What You See and What She Gets:
Isolating the Effect of Inconsistent Racial Classification on Women's Earnings and Income

Aliya Saperstein, *University of California, Berkeley*
Bryan L. Sykes, *University of Washington, Seattle*

Abstract

Recent studies have shown that using different measures of race results in different estimates of everything from vital rates to racial disparities in income and medical treatment. Here, instead of comparing conclusions based on different measures, we combine measures of interviewer-classified and self-reported race to examine whether differences exist in the earnings and income of people who are consistently classified compared to those who are not. Because inconsistent racial classification is not randomly distributed across the population, we use propensity-score matching techniques to ensure comparison between individuals who are similar on all other characteristics, except their racial identity and classification. Drawing on data from the 1988 National Survey of Family Growth, we find that women who self-report a race that they are not perceived to be have earnings and income that are significantly lower than their consistently classified counterparts. This finding has implications for understanding how racial inequalities are perpetuated in the United States.

Paper prepared for the 2008 Population Association of American Annual Meeting, New Orleans, April 17-19.
Please do not quote or cite without permission.

Whiteness has been synonymous with social and economic privilege in the United States since the country's founding. Until the mid-20th century, being "white" was a requirement for citizenship (Lopez 1996), a litmus test for who could marry whom (Golden 1958), and carried with it a whole host of other "rights" (Harris 1993). Even today, whites remain ahead of most other racial populations – but especially blacks – in "life, liberty and the pursuit of happiness" (Fischer and Hout 2006).

But what is "whiteness" exactly? Is it enough to be perceived by others as white (or conversely, to self-identify as white)? Given the complexities in identifying and defining race, what is the mechanism that perpetuates racial inequality in post-civil rights era America? These questions motivate the analyses that follow, and highlight the links between racial identity, racial classification, stratification and mobility that remain undertheorized in most work on racial inequality in the United States.

We assess whether being perceived as white is rewarded in either the labor market or the marriage market, as theories of categorical discrimination would suggest (e.g., Tilly 1998), drawing on data from the 1988 National Survey of Family Growth that includes measures of both interviewer-classified and self-identified race. We use cases of inconsistent racial classification, where the perception of one's race and one's self-identification do not match, to isolate the causal link between race and economic inequality. Further, we do so in a counterfactual modeling framework, which takes into account that characteristics, such as educational attainment or urban residence, that affect one's level of income may also affect, or be affected by, the perception and identification of one's race. These methods represent significant improvements over previous attempts to estimate the effects of discrimination on racial differences in earnings and income, which tend to assume that race is a static (or exogenous) characteristic and that the perception of one's race and one's self-identification capture the same information.

Our results show that being perceived as white by itself does not "pay off" in either the labor market, as measured by the respondent's earnings, or the marriage market, as measured by family income (which combines the respondent's earnings with that of a spouse, if one is present). For example, we find that the average woman who is seen as white but identifies as black is doing no better economically, net of all other factors, than the average woman who is both seen as black and identifies as black. This suggests that the mechanism that perpetuates

racial inequality in the labor market and the marriage market is not solely discrimination based on appearance (by either employers or potential spouses), but a more complex interaction of preferences and social background characteristics that are not measured in our data. Exploring these mechanisms more fully awaits both explicit theorizing about the relationships between racial identification, racial classification and inequality and the incorporation of multiple measures of race into standard surveys.

Race, inequality and the confounding of the two

In the legislation and legal decisions of the past, one's race was determined by descent, specifically the proportion of white (or nonwhite) ancestors in one's bloodline (Davis 2001). This definition of race as ancestry remains in contemporary conceptions of racial identification, and hides a largely unexamined tautology: one is white because all (or most) of one's ancestors were white. But what made one's ancestors white? Was it their skin tone, their physical features, their national origin?

Lopez (1996) has shown that descent-based arguments were not the only ones made to determine "whiteness" for the purposes of acquiring citizenship in the United States. Judges also relied on the criteria of how one was perceived by one's community; a kind of everyone-knows-someone-of-a-different-race-when-they-see-one claim akin to later Supreme Court decision regarding pornography.[1] In the early 20th century, the U.S. census bureau relied on similar logic in its instructions to enumerators regarding how people of mixed American Indian and white or American Indian and black ancestry should be counted (Nobles 2000). This suggests that, even in the past, descent was a necessary but not sufficient condition for determining one's race. One's appearance – or perhaps one's standing in the community – also played an important role.

This definitional complexity, combined with recent increases in immigration and procreation across conventional defined racial lines, makes for significant challenges in measuring race in standard surveys. Several recent studies show that using different measures of race results in different estimates of fertility, mortality and injury rates (Arias et al. 2007; Morgan et al. 1999; Sugarman et al. 1993; Hahn et al. 1992), racial disparities in income

[1] For a summary, see "Movie Day at the Supreme Court or "I Know It When I See It": A History of the Definition of Obscenity," accessed online March 19, 2008, http://library.findlaw.com/2003/May/15/132747.html

(Saperstein 2006; Telles and Lim 1998), and medical treatment (e.g., Franks et al. 2005). These differences arise when self-identified measures of race are compared with racial classifications made by others, such as nurses, doctors, funeral directors or survey interviewers.

Most of these studies are concerned with quantifying the over- or under-estimates of the outcome of interest, based on the assumption that one or the other measures of race is the more "correct" one. Since at least 1970, self-identification has been considered the gold-standard of racial reporting in the United States, though Telles and Lim (1998) note that the racial classification of the survey interviewer may be a better proxy for the perceptions (and therefore the discrimination) of employers and other persons in positions of authority. This raises the issue of mechanisms: how are different measures of race related to inequality? We know different measures of race yield different empirical results; the more interesting question is why?

Theories of racial discrimination

There is a long-standing debate in both sociology and economics regarding how much of the observed racial disparities in income, and other socioeconomic outcomes, can be explained by characteristics of the individual, such as verbal or cognitive skills, educational attainment and job-related experience, and how much might be explained by discrimination in the labor market (see Darity and Mason 1998 for a review). The former are often called "human capital" theories; these assume that racial disparities in wages, for example, are the result of pre-market forces. People arrive at the labor market with differing skills and experience and those characteristics are rewarded differently. Theories of discrimination claim that even people with the same skills are treated differently or that one group's lack of human capital may reflect a rational adaptation in the face of repeated or past discrimination.

Each of these theories relies on one of two mechanisms, being rendered non-competitive during the process of accruing human capital (this mechanism is the same for some theories of discrimination, such as accumulated disadvantage, as well as theories that claim different races also have different "cultural" dispositions) or having employers act on their preferences for some workers over others. To account for individual-level or human capital explanations in statistical analyses requires that researchers use a standard set of controls, such as years of education, hours of work, age, marital status and number of children. Generally speaking, if being a member of a

given racial population is still associated with negative outcomes after introducing such controls, the remainder of the effect is attributed to discrimination.

Theories of "discrimination" or social closure in the marriage market posit similar mechanisms. Either people of different racial populations are actively kept apart by social norms and/or active pressure from family members, or they are kept apart because they do not travel in the same circles – they don't go to the same schools, live in the same neighborhoods or work in the same jobs (Kalmijn 1998). Further, in poor communities, a dearth of "marriageable men" may lead women to forgo marriage entirely (e.g., Wilson 1987). If members of a disadvantaged racial population only marry each other, or forgo marriage entirely, not only their individual earnings but their family incomes will be lower on average than the incomes of members of more advantaged groups.

However, what all of these theories are missing is an explicit statement of how active discrimination occurs. In the United States, one's race is not recorded on most official documents, and applications for schools and jobs ask individuals to self-identify. It seems to be understood then that employers, or potential spouses, can recognize someone of a different – and disadvantaged – race when they see them. That suggests it is not one's self-identification that is the most detrimental characteristic in a race-conscious society, but the perception of one's race by others.[2]

Certainly there is reason to believe that one's self-identity, or at least the expression of it on a survey, is affected by how one is perceived and treated by others (Doyle and Kao 2007; Harris and Sim 2002; Nagel 1994). This would lead most people to have consistent racial classifications: they would claim the same race that a stranger would use to classify them. However, in cases where the two classifications differ, theories of discrimination would lead one to expect that people who are perceived as white will be better off economically than people who

[2] This perspective is supported by the literature on skin tone stratification among African Americans and Hispanics in the United States (e.g., Hunter 2002; Keith and Herring 1991; Telles and Murguia 1990; Arce et al. 1987). However, Gullickson (2005) finds that the effects of skin tone on socioeconomic status, net of family background characteristics, is nonsignificant for cohorts of African Americans born after the mid 20th century (i.e., those who came of age during, or after, the Civil Rights Movement). To the extent that perceived race is a proxy for skin tone (or vice versa), this finding could be reflected in our sample, as well. Though the oldest women in the 1988 NSFG were born in 1944, the majority of the women are members of the birth cohorts whose outcomes were not shaped by skin tone stratification, according to Gullickson (2005).

are perceived as nonwhite, regardless of how the individuals choose to self-identify. This is the hypothesis we seek to test in the analyses that follow.

Money "whitens" – a cautionary tale of causality

But what if people change their race as they improve their socioeconomic status, as has been observed in Brazil? This is generally thought to occur because people choose to "whiten" themselves, or their children, as both the result of previous status improvement and a desire for continued mobility (Schwartzman 2007). However, if money "whitens," then the effect of race on one's income, for example, is confounded by the effect of one's income on one's race. It also begs the question of whether and how the perception of one's race by others matters.

If one "whitens" oneself after experiencing some mobility, then whiteness may be as much a status symbol as a barrier to achievement. The assumption that if one has "arrived" then one must be white is different than assuming one must be white in order to attain higher levels of socioeconomic status in the first place. The latter presumably requires active discrimination based on appearance; the former does not (though it suggests it likely occurred in the past). This is a very important distinction because assuming individuals "whiten" themselves ignores the perception of employers which produce the very rewards that are endogenous to the whitening process. Our framework shifts the mechanism from an individualistic determination to a relational one (cf. Telles and Lim 1998).

To truly test the hypothesis that money "whitens" one would need either longitudinal or intergenerational data (or both). However, cross-sectional data that includes measures of perceived and self-identified race can also provide some clue as to whether individuals who are perceived as white have higher status on average than individuals who self-identify as white, or vice versa.

Data and Measures

The National Survey of Family Growth (NSFG) is based on in-person interviews with women aged 15-44 and is typically used for studies of pregnancy, childbearing, contraception,

and related aspects of maternal and infant health.[3] However, the survey also includes detailed background information about the respondent and her husband (if relevant), such as education, religion, ethnic origin, occupation, and earnings.[4]

In 1988, the NSFG coded the respondent's race in two ways. First, the interviewer made her observation of the respondent's race, recording it as one of three categories: Black, White and Other. Then, amidst a series of demographic questions, the respondent was asked "Which of these groups best describe your racial background?" The category options were: American Indian, Asian or Pacific Islander, Black and White. Respondents could choose as all four racial categories if they wished. To establish whether respondents were consistently classified or not, we collapsed self-reports of American Indian or Asian or Pacific Islander into one "other" category to better match the coding of perceived race. Table 1 shows the cross-tabulation of perceived and self-reported race. The vast majority of cases are consistently classified but previous analysis indicates that, despite the relatively small numbers, the remainder cannot be characterized as the result of measurement error (Saperstein 2008).

To ensure that the inconsistently and consistently classified women are comparable on as many other characteristics as possible, we selected both characteristics that would predict income (e.g., educational attainment) and information that the interviewer learned about the respondent prior to coding her race or could have assessed through direct observation of the respondent's home and neighborhood (e.g., urban residence). The interviewer heard a range of information about the respondents' health history, parity and marital status, and conceptive use prior to coding her race. From the many potential responses we selected characteristics that would be commonly known to vary by race (e.g., sickle cell anemia, high blood pressure, number of pregnancies, growing up in an intact family), under the assumption that if the respondent's race were otherwise ambiguous the interviewer might rely on these types of cues to make her

[3] The fact that the NSFG only samples women is a limitation, but it is the only national survey of adults that includes multiple measures of race in the same year. The 1979 National Longitudinal Survey of Youth does include multiple measures, but the two measures of race only occur together in the first year, when the respondents are 14-23 years of age. The National Longitudinal Survey of Youth also has multiple measures of race, but has the same age range bias as the NLSY. The 1996 and 2000 General Social Surveys includes multiple measures of race in the same year, but it has a much smaller sample size that makes identifying significant differences between classification consistency groups difficult.

[4] For additional details on the survey, see the National Center for Health Statistics website (http://www.cdc.gov/nchs/products/elec_prods/subject/nsfg.htm) or the NSFG webpage from the Office of Population Research at Princeton University (http://opr.princeton.edu/archive/nsfg/).

decision. Table 2 presents descriptive statistics of how inconsistently classified women differ in earnings and family income, on average, from other women who are either perceived to be or identify as the same race, as well as how they differ on some of the demographic characteristics described above.

Based on these descriptive statistics it seems that women who are seen as white but identify as nonwhite are generally not as well of economically as other women who are perceived to be white; nor are they better off than other women who identify as nonwhite. However, all of these women vary on other characteristics such as region of residence and marital status that may be related to both their earnings and incomes and their likelihood of having inconsistent racial classifications. This makes it all the more important to use a counterfactual framework, such as propensity-score matching, rather than ordinary least squares regression to assess differences in socioeconomic status between these racial populations. We describe the technique and its justification in more detail below.

The NSFG does not make individual-level information about interviewers publicly available, so we cannot assess whether or not characteristics of the interviewer are also associated with inconsistent racial classification of the respondents. Aggregate data provided by the NSFG indicates that the vast majority of its interviewers are relatively highly educated white women (Abma, personal communication). Previous studies do find race of interviewer effects in the perception of skin tone (Hill 2002a) and the reporting of political attitudes (Krysan and Couper 2003), but studies like this one that examine the racial classification of respondents do not find that interviewer characteristics bias the results (Penner and Saperstein 2008; Campbell and Troyer 2007). Thus, we assume that the racial classifications of NSFG interviewers are good proxies for the perceptions of both employers and potential spouses.

Methods

We are interested in understanding whether there are systematic and significant earnings differences between two groups of women: women who are perceived to be white, but identify as nonwhite and women who are both seen as and identify as the same race. Because inconsistent racial classification is not randomly distributed across the population, we use propensity-score matching techniques to ensure comparison between individuals who are similar on all other

characteristics, except their racial identity and classification. Essentially, propensity-score matching turns observational data into experimental data by creating treatment and control groups after the fact. Our "treatment group" includes women whose perceived race is inconsistent with their self-identity, while our baseline or control group includes women whose self-reported racial backgrounds are consistent with how the interviewer classified them. This extra step, above and beyond standard regression models, is necessary for several reasons.

First, by estimating the propensity score, we test for pretreatment differences in social background indicators between the two groups of women. If there are significant differences for any of the covariates (as Table 2 suggests), then the inconsistent racial classifications are not random on that dimension. To rectify this, the propensity score is then balanced by constructing groups of respondents where there are no systematic differences in the pretreatment characteristics, which ensures the randomness of the inconsistent racial classifications (Rubin 1973a, 1973b, 1976a, 1976b, 1977, 1978, 1980, 1986, 1991, 2000; Rubin and Thomas 1991, 1996, 2000; D'Agostino and Rubin 2000).[5]

Second, by estimating the propensity score, we reduce the dimensionality of including a great number of regressors into the earnings equation (Rosenbaum and Rubin 1983, 1984, 1985a, 1985b). The propensity score captures and summarizes the overall effect of all covariates on the likelihood of having one's race inconsistently observed between respondents and interviewers. This leaves just three effects to be estimated in the income models: the effect of being inconsistently classified, the effect of having characteristics associated with being inconsistently classified (i.e., the propensity score) and an interaction effect between the two. If only the coefficient on the propensity score is statistically significant, then differences in income are caused by characteristics associated with inconsistent classification; however, being inconsistently classified in and of itself does not necessarily translate into measurable earnings differences. Statistically significant effects for being inconsistently classified, the propensity

[5] We do not present estimates from the propensity score model or the balancing results because of the large number of models being estimated. We did find several significant "pretreatment" differences between groups that were related to inconsistent racial classification, including living in the western United States, living in central cities, not growing up in an intact family and, in some cases, being foreign born. These are not surprising given the geographic racial diversity of the United States (Jones and Smith 2001), the link between self-identity and family or community socialization (e.g., Harris and Sim 2002; Liebler 2004; Xie and Goyette 1997) and that definitions of race differ across countries (e.g., Davis 2001; Rodriguez 2000).

score, and the two-way interaction indicate that being perceived as a different race than that with which one self-identifies is a causal link between race and income inequality.

Lastly, little is known about the distribution from which inconsistently classified women are likely to be drawn. While it is possible that this distribution is normal, there is no evidence or literature to suggest normality, particularly along certain social background characteristics. To address this issue and ensure confidence in our inferences about income disparities, we augment the propensity score matching method by bootstrapping (or resampling) estimates to create a likelihood distribution from which our standard errors (and confidence intervals) are more robust and representative without making any distributional assumptions.[6]

Models

Below, we estimate models for three different "treatment" groups: women who are perceived as white but identify as nonwhite, women who are perceived as white but identify as "other" and women who are perceived as white but identify as black. We run models separately for the last two groups to make sure that the effects are not operating in opposite directions, thus cancelling each other out in the more general nonwhite models.

We also estimate two types of income models: family income and potential earnings. Family income refers to income for the 12 months prior to the survey date, and was collected by using a card with 17 income categories that ranged from $1,500 to $3,000 a year to $50,000 and up. Responses were recoded to represent the mid-point of these ranges with the open-ended top category set at $62,500.[7] Potential earnings represents the amount of money a woman would make if she were not removed from the labor force for childrearing and other reasons, compared to other women who may not have children.[8] Both sets of models are restricted to women who are 25 and older to allow for educational completion and labor force experience.

[6] We performed 1000 bootstrap replications for each model presented.

[7] To the extent that there is racial inequality within these income ranges (such that "whites," for example, have incomes at the top end of the category while "blacks" have incomes at the bottom of the same range), this coding scheme will underestimate racial differences in income.

[8] The NSFG asks women what their earnings are for their current job or were for their most recent job. We also have data on when the respondent last worked at their previous job. Using the consumer price index we adjust former earning to 1988 dollars giving us a measure of earnings potential for all ever-employed women in the sample.

The average income of each of these treatment groups is compared to the distribution of income among its two possible reference groups: other women who are perceived as white and other women who identify as nonwhite. Thus, we estimate a total of twelve models; two for each income type and reference group for each of the three treatment groups.[9]

Coefficients in all the models are interpretable in dollars. For example, an estimate of 1,000 corresponding to the indicator for inconsistent racial classification would mean that women who are perceived to be a race with which they do not identify earn $1,000 more on average than otherwise identical women with consistent racial classifications. The propensity score ranges from zero to 100 percent; so a one percentage-point increase in a woman's propensity to have an inconsistent racial classification results in an increase (or decrease) in the woman's income of the amount specified by the coefficient.

Hypotheses

If whiteness is rewarded in either the labor market or the marriage market, through the perceptions of employers and potential spouses, then, on average, self-identified nonwhites who are perceived to be white should be:

1. No worse off economically than perceived whites who also identified as white, and/or
2. Better off economically than self-identified nonwhites who are also seen as nonwhite (net of all other factors)

Thus, in models referencing the distribution of income among perceived whites, the coefficient for inconsistent racial classification should be small and nonsignificant (or possibly significant and positive). In models comparing the distribution of income among self-identified nonwhites, the coefficient on inconsistent classification should be both positive and significant.

Results

We discuss the models in order of specificity, starting with the more general treatment group of self-identified nonwhites who are perceived as white. In comparing each treatment

[9] Actually, we estimate 24 models. Each "final" model is paired with a model that includes only an indicator for inconsistent racial classification (see Table 3). Comparing the results between these two models helps to illustrate the improvement provided by using propensity-score matching techniques.

group to its two reference distributions, we follow the order of the hypotheses described above. We also present models of different income types starting with family income because it represents a combination of effects from the labor market and the marriage market.

Seen as white, identifies as nonwhite

Our results indicate that women who are seen as white but self-identify as nonwhite have significantly lower family incomes and potential earnings than women who are both seen as white and identify as white. The total effect of being inconsistently classified, with a one-percentage-point increase in the propensity score, results in an average family income that is roughly $7,300 lower than an otherwise identical woman who is consistently classified as white (see Table 3, Panel A). The potential earnings of women who appear white but self-identify as nonwhite are also significantly reduced relative to women who both appear white and self-identify as white (by $3,300 per year in 1988 dollars). It seems, then, that appearing white alone does not equalize the incomes of otherwise similar women of different self-reported races.

Of course, it is possible that the socioeconomic status of women who appear white but self-identify as nonwhite falls in between their consistently classified white and consistently classified nonwhite counterparts. In this sense, their perceived whiteness would still provide some benefit relative to other self-identified nonwhites. However, our results indicate this is not the case either. When we compare the family incomes and potential earnings of all self-identified nonwhites, the only coefficients that are statistically significant are those for the propensity score – the characteristics associated with being inconsistently classified (and these are both negative rather than positive as a perceived white advantage would suggest). The effect of being inconsistently classified itself is estimated to be positive (at $3,522 and $1,696 for family income and earnings, respectively) but statistically nonsignificant. If women who self-identify as nonwhite and appear white are doing better than their otherwise identical nonwhite counterparts, the difference is not a measurable one.

Seen as white, identifies as "other"

Research shows that Americans with Asian or American Indian ancestry are more likely to "become white" or (be allowed to) assimilate and intermarry than Americans with African ancestry (Zhou 2004; Liebler 2004). Similarly, some scholars have argued that the color line

between the haves and have-nots in the United States is shifting from a white-nonwhite to a black-nonblack divide (Fischer and Hout 2006; Lee and Bean 2004). This suggests that the experiences of women who are seen as white and identify as "other" (i.e., American Indian or Asian or Pacific Islander) may be very different from women who are seen as white and identify as black. We take that difference into account in our analyses below.

Our results indicate that women who are seen as white but self-identify as "other" do not experience a significant reduction in family income or personal earnings compared to all women who are seen as white (see Table 3, Panel B). The characteristics associated with inconsistent racial classification do result in significantly lower average family income and earnings ($1,324 and $1,576 less, respectively, for a one-percentage point increase in the propensity score), but inconsistent racial classification itself does not produce significant differences. Statistically speaking, women who are seen as white but identify as other are doing just as well as women who are both seen as white and identify as white.[10]

However, they are also not doing measurably better than otherwise identical women who identify as "other" and are seen as "other." Though the effects of being seen as white, among self-identified "others," is positive for both family income ($8,861) and earnings ($995) the estimates are not statistically significant. In fact, the only significant effect is for the characteristics that are associated with inconsistent racial classification (the propensity score) in the earnings model, and it is negative. This could reflect the fact that, on average, the earnings and income of otherwise similar Asians and whites are already roughly equal (e.g., Fischer and Hout 2006) so there is little or no additional economic advantage to be gained in either the labor market or the marriage market by being seen as white for most women who identify as "other."

Seen as white, identifies as black

The results above indicate that the negative effects of inconsistent racial classification on family income and potential earnings we find in the general models for self-identifying as

[10] The magnitude of the coefficients for the earnings model compared to all other women who are seen as white is roughly comparable to the magnitude of the coefficients in the more general nonwhite earnings model presented above. It is possible that the lack of statistical significance can be explained by cutting the size of the inconsistently classified group in half. If so, it suggests that the earnings of women who are seen as white but identify as other are *substantively* different – and at least $1,700 lower per year (in 1988 dollars) judging by the total effect – than the earnings of other women who are seen as white, even if they are not statistically significantly different.

nonwhite are driven largely by the experiences of women who appear white but identify as black. This is borne out by the models that estimate their incomes separately, as described below.

Women who are seen as white but identify as black not only reap no advantage from their perceived whiteness in the labor market or the marriage market, compared to either other perceived white women or other self-identified black women, they experience a significant disadvantage in the marriage market. That is, compared to other women who appear white, women who appear white but identify as black have no measurable gap in earnings but a significantly lower average family income (see Table 3, Panel C).[11] The effect of inconsistent classification on family income, by itself, results in roughly $11,000 less per year, and the total effect, for each percentage-point increase in the propensity score, results in an additional loss of at least $1,000.

Compared to other women who self-identify as black, women who appear white but identify as black only have an advantage in family income if they also have a high propensity to be inconsistently classified. Though the main effect of inconsistent racial classification on family income is negative (-$6,596), the interaction effect is relatively large and positive ($1,816), indicating that women who identify as black but are seen as white, and have a propensity score of more than 5 percent also have family incomes than are higher, on average, than otherwise identical women who are both seen as and identify as black. Again, however, women who appear white but identify as black received no higher earnings than their otherwise identical self-identified black counterparts.

Discussion

Previous studies of racial discrimination in the labor market and marriage market claim, either explicitly or implicitly, that being seen as white – or at least having lighter skin tone – provides significant advantages. However, these studies do not have direct measures of how the respondents are racially classified by others. When we include this information along with the respondent's self-identity in a counterfactual framework, which allows us to isolate the effect of

[11] The earnings model for women who identify as black, compared to all other women who are seen as white, should be treated with the same caution as the one for women who identify as other, as noted above (footnote 8). It is interesting to note, though, that in this case the estimates indicate that for each percentage-point increase in their propensity to be inconsistently classified women who are seen as white but identify as black actually gain $250 in earnings per year.

being seen as a member of a race with which one does not identify from the effect of other characteristics that may both affect income and inconsistent racial classification, we find that "whiteness," in and of itself, does not result in higher family income or earnings. In fact, not only are people who are seen as white but identify as nonwhite not as well off as other perceived whites, they are little or no better off than otherwise identical self-reported nonwhites.

Our results also speak to whether or not money "whitens" in the United States. The fact that women who appear white to others but identify as nonwhite are not better off economically than consistently classified nonwhites suggests that interviewers are not "whitening" people with higher socioeconomic status, as Telles and Lim (1998) found in Brazil. We did not present models examining the opposite kind of inconsistent racial classification, where women identify as white but are seen as nonwhite by the interviewers. The number of such cases is even smaller in the NSFG than for the groups we analyzed above, making it even more challenging to identify statistically significant differences.[12] However, preliminary models comparing women who self-identify as white but appear nonwhite to all other women who self-identify as white suggest that there is no statistically significant income or earnings difference between the two groups (results not shown). This could indicate that Americans are more likely to "whiten" themselves once they achieve a certain level of socioeconomic status, but further analysis is necessary to solidify the results.

There are limitations to our analyses, but it is not clear that they would significantly alter our conclusions about whether whiteness was being rewarded in the labor market or the marriage market during the late 1980s in the United States. The primary limitation is that, in 1988, the NSFG only sampled women of childbearing age. From the late 1970s to the early 1990s, women of different races but otherwise similar human capital characteristics had achieved rough parity in earnings in the United States (Darity and Mason 1998). It is possible that we do not see a "discrimination" effect among women because of their convergence in income. However, some have argued that phenotypic characteristics, such as skin tone, and perceptions of beauty more generally affect the socioeconomic outcomes of women more than they do men (e.g., Hill 2002b; Hunter 2002). Further, evidence from the 2004 and 2005 waves of the Behavioral Risk Factor Surveillance System, an annual health survey designed by the Centers for Disease Control that

[12] There are 20 women who are seen as black but identify as white, and 30 women who are seen as "other" but identify as white (see Table 1).

also includes multiple measures of race, suggests that women are overrepresented among Americans who identify as nonwhite but are classified by others as white, while men are overrepresented among Americans who identify as white but are classified by others as nonwhite.[13] While this opens up the intriguing possibility of a gender difference in the likelihood of "whitening" oneself, it also makes it unlikely that being perceived as white by others is a significant mechanism affecting the income or earnings of men. A secondary limitation is that we cannot account for the possibility that women who identify as nonwhite but are seen as white have other characteristics that serves as cues to employers or spouses of their racial identity, such as a distinctive name.[14] But it seems safe to assume that if the characteristic would be racialized by employers, it would also have been picked up on during the first half of the survey interview before the interviewer was instructed to code the respondent's race.

In many ways, these results leave us asking the same questions with which we began. If the advantage of "whiteness" cannot be attributed to racial classification or appearance alone, then what is it? What characteristics, above and beyond "human capital," are being rewarded in the labor market or the marriage market? The traditional counterpoint to theories of active discrimination is theories of cultural (or biological) differences between races (e.g., Herrnstein and Murray 1994, Moynihan 1965). However, all of the theories operate by reducing the human capital of members of the disadvantaged group or affect other characteristics, such as marital status and number of children, for which our models control.

Perhaps, then, employers and spouses also recognize a different kind of capital. Cultural capital: the level of eloquence, the omnivorous knowledge of arts and leisure, that *je ne sais quoi* that is rewarded in school, but not explicitly taught (Dimaggio 1982; Bourdieu 1977). Without it, a woman who identifies as nonwhite but appears white is perceived as simply a poor, white

[13] Results available from the first author upon request, or see the BRFSS website to access the public-use data files. http://www.cdc.gov/brfss/technical_infodata/surveydata.htm

[14] Research has shown that both "ethnic" (Riach and Rich 1991, cited in Darity and Mason 1998) and racialized (Bertrand and Mullainathan 2004) names reduce the likelihood of being selected for interviews when applying for jobs by mail.

person.[15] This suggests that one's perceived race is a necessary but not sufficient condition for achieving socioeconomic mobility.

Another possible, and potentially complementary, mechanism is revealed by the fact that the coefficients on the propensity scores are negative and statistically significant in eight of the 12 models presented above. If the characteristics that predict inconsistent racial classification are also associated with negative socioeconomic outcomes, then perhaps women who are less well-off economically do not "feel" white, even if this is how they are often seen. Like the mechanisms through which money "whitens," but in reverse, this suggests that Americans consider whiteness an achieved status as well as an ascribed one.

In summary, it seems that categorical distinctions based on appearance alone do not drive economic inequalities in the contemporary United States. Though it is difficult from our models to draw conclusions about what exactly is going on, it is quite clear from our results what is *not* going on: being perceived as white is not rewarded in either the labor market or the marriage market. Instead, the mechanisms that produce racial disparities in earnings and income are far more complex. This makes it all the more important for future research to use both multiple measures of race and a counterfactual modeling framework to sort out the causal relationships between racial identities, racial classifications and racial inequality in America.

[15] This claim is supported by evidence from social psychological studies of "subtyping." That is, within large, global stereotypes of "whites" or "blacks," people also recognize smaller subtypes that may have completely unrelated (stereotypical) characteristics, such as "black businessmen" (see Kunda and Thagard 1996 for a review).

References

Arce, Carlos H., Edward Murguia, and W. Parker Frisbie. 1987. "Phenotype and Life Chances Among Chicanos." *Hispanic Journal of Behavioral Sciences* 9(1): 19-32.

Arias, Elizabeth, William S. Schaumann and Paul Sorlie. 2007. "Race and Hispanic Origin Reporting on Death Certificates in the United States: Status and Effects." Paper presented at the Population Association of America Annual Meeting, New York. http://paa2007.princeton.edu/download.aspx?submissionId=71842

Bertrand, Marianne and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *The American Economic Review* 94: 991-1013.

Bourdieu, Pierre. 1977. "Cultural Reproduction and Social Reproduction," Pp. 487-511 in *Power and Ideology*, Jerome Karabel and A.H. Halsey, editors. Oxford University Press.

Campbell, Mary E. and Lisa Troyer. 2007. "The Implications of Racial Misclassification by Observers." *American Sociological Review* 72:750-65

D'Agostino, Ralph and Donald Rubin. 2000 "Estimating and Using Propensity Scores with Partially Missing Data (in Applications and Case Studies)" *Journal of the American Statistical Association*, Vol. 95, No. 451. pp. 749-759.

Darity, William and Patrick Mason. 1998. "Evidence on Discrimination in Employment: Codes of Color, Codes of Gender." *Journal of Economic Perspectives* 12(2):63-90.

Davis, F. James. 2001 [1991]. *Who is Black?: One Nation's Definition*. University Park, PA: Pennsylvania State University Press.

Dimaggio, Paul. 1982. "Cultural Capital and School Success: The Impact of Status Culture Participation on the Grades of United States High-School Students." *American Sociological Review 47(2):189-201.*

Doyle, Jamie M. and Grace Kao. 2007. "Are Racial Identities of Multiracials Stable? Changing Self-Identification Among Single and Multiple Race Individuals." *Social Psychology Quarterly* 70(4): 405-23.

Fischer, Claude S. and Michael Hout. 2006. *Century of Difference: How America Changed in Last One Hundred Years*. New York: Russell Sage.

Franks, Peter, Kevin Fiscella and Sean Meldrum. 2005. "Racial Disparities in the Content of Primary Care Office Visits." *Journal of General Internal Medicine* 20: 599-603.

Golden, Joseph. 1958. "Social Control of Negro-White Intermarriage." *Social Forces* 36(3):267-9.

Gullickson, Aaron. 2005. "The Significant of Color Declines: A Re-Analysis of Skin Tone Differentials in Post-Civil Rights America." *Social Forces* 84(1): 157-80.

Hahn, R.A., J. Mulinare, and S.M. Teutsch. 1992. "Inconsistencies in Coding of Race and Ethnicity between Birth and Death in US Infants: A New Look at Infant Mortality, 1983 through 1985." *Journal of the American Medical Association* 267:259-63.

Harris, Cheryl I. 1993. "Whiteness as Property," Harvard Law Review 106: 1707-91.

Harris, David R. and Jeremiah Joseph Sim. 2002. "Who is Multiracial? Assessing the Complexity of Lived Race." *American Sociological Review* 67:614-27.

Herrnstein, Richard and Charles Murray. 1994. *The Bell Curve: Intelligence and Class Structure in American Life*. Basic Books.

Hill, Mark. 2002a. “Race of the Interviewer and Perception of Skin Color: Evidence from the Multi-City Study of Urban Inequality.” *American Sociological Review* 67(1): 99-108.

Hill, Mark. 2002b. “Skin Color and the Perception of Attractiveness Among African Americans: Does Gender Make a Difference?” *Social Psychology Quarterly* 65(1): 77-91.

Hunter, Margaret. 2002. “’If You’re Light You’re Alright’: Light Skin Color as Social Capital for Women of Color.” *Gender and Society*16(2):175-93.

Jones, Nicholas and Amy Symens Smith. 2001. “The Two or More Races Population: 2000.” *Census 2000 Brief.* U.S. Census Bureau.

Kalmijn, Matthijs. 1998. “Intermarriage and Homogamy: Causes, Patterns, Trends.” *Annual Review of Sociology* 24:395-421.

Keith, Verna M. and Cedric Herring. 1991. “Skin Tone and Stratification in the Black Community.” *American Journal of Sociology* 97(3):760-78,

Krysan, Maria and Mike P. Couper. 2003. “Race in the Live and Virtual Interview: Racial Deference, Social Desirability and Activation Effects in Attitude Surveys.” *Social Psychology Quarterly* 66:364-83.

Kunda, Z. and P. Thagard. 1996. “Forming Impressions from Stereotypes, Traits and Behaviors: A Parallel-Constraint-Satisfaction Theory.” *Psychological Review* 103(2):284-308.

Liebler, Carolyn A. 2004. “Ties on the fringes of identity.” *Social Science Research* 33(4): 702-23.

Lee, Jennifer and Frank D. Bean. 2004. “America’s Changing Colors Lines: Immigration, Race/ethnicity and Multiracial Identification,” *Annual Review of Sociology* 30:221-42.

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

Morgan, S. Philip, N. Botev, R.B. Chen and J.P. Huang. 1999. “White and Nonwhite Trends in First Birth Timing: Comparisons using Vital Registration and Current Population Surveys.” *Population Research and Policy Review* 18: 339-356.

Moynihan, Daniel Patrick. 1965. “The Negro Family, the Case for National Action.” Prepared for the United States Department of Labor, Office of Policy Planning and Research. U.S. Government Printing Office.

Nagel, Joanne. 1994. “Constructing Ethnicity: Creating and Recreating Ethnic Identity and Culture.” *Social Problems* 41: 152-176.

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

Penner, Andrew and Aliya Saperstein. 2008. “The Race of a Criminal Record: How Incarceration Colors Racial Perception.” Paper presented at the Population Association of America Annual Meeting, New Orleans, April 17.

Rodriguez, Clara E. 2000. *Changing Race*: *Latinos, the Census, and the History of Ethnicity in the United States.* New York University Press.

Rosenbaum, Paul and Donald Rubin. 1983 “The Central Role of the Propensity Score in Observational Studies for Causal Effects” *Biometrika*, Vol. 70, No. 1. pp. 41-55.

Rosenbaum, Paul and Donald Rubin. 1984 “Reducing Bias in Observational Studies Using Subclassification on the Propensity Score (in Applications)” *Journal of the American Statistical Association*, Vol. 79, No. 387. pp. 516-524.

Rosenbaum, Paul and Donald Rubin. 1985a. “The Bias Due to Incomplete Matching” *Biometrics*, Vol. 41, No. 1. pp. 103-116.

Rosenbaum, Paul and Donald Rubin. 1985b. “Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score” *The American Statistician*, Vol. 39, No. 1. pp. 33-38.

Rubin, Donald. 1973a. “Matching to Remove Bias in Observational Studies” *Biometrics*, Vol. 29, No. 1. pp. 159-183.

Rubin, Donald. 1973b. “The Use of Matched Sampling and Regression Adjustment to Remove Bias in Observational Studies” *Biometrics*, Vol. 29, No. 1. pp. 185-203.

Rubin, Donald. 1976a. “Multivariate Matching Methods That are Equal Percent Bias Reducing, I: Some Examples” *Biometrics*, Vol. 32, No. 1. pp. 109-120.

Rubin, Donald. 1976b. “Multivariate Matching Methods That are Equal Percent Bias Reducing, II: Maximums on Bias Reduction for Fixed Sample Sizes” *Biometrics*, Vol. 32, No. 1. pp. 121-132.

Rubin, Donald. 1977. “Assignment to Treatment Group on the Basis of a Covariate” *Journal of Educational Statistics*, Vol. 2, No. 1. pp. 1-26.

Rubin, Donald. 1978. “Bayesian Inference for Causal Effects: The Role of Randomization” *The Annals of Statistics*, Vol. 6, No. 1. pp. 34-58.

Rubin, Donald. 1980 “Bias Reduction Using Mahalanobis-Metric Matching (in Shorter Communications)” *Biometrics*, Vol. 36, No. 2. pp. 293-298.

Rubin, Donald. 1986. “Statistics and Causal Inference: Comment: Which Ifs Have Causal Answers (in Theory and Methods)” *Journal of the American Statistical Association*, Vol. 81, No. 396. pp. 961-962.

Rubin, Donald. 1991 “Practical Implications of Modes of Statistical Inference for Causal Effects and the Critical Role of the Assignment Mechanism” *Biometrics*, Vol. 47, No. 4. pp. 1213-1234.

Rubin, Donald. 2000 “Causal Inference Without Counterfactuals: Comment (in Theory and Methods)” *Journal of the American Statistical Association*, Vol. 95, No. 450. pp. 435-438.

Rubin, Donald and Neal Thomas. 1992 “Characterizing the Effect of Matching Using Linear Propensity Score Methods with Normal Distributions” *Biometrika*, Vol. 79, No. 4. pp. 797-809.

Rubin, Donald and Neal Thomas. 1996 “Matching Using Estimated Propensity Scores: Relating Theory to Practice” *Biometrics*, Vol. 52, No. 1. pp. 249-264.

Rubin, Donald and Neal Thomas. 2000. “Combining Propensity Score Matching with Additional Adjustments for Prognostic Covariates (in Theory and Methods)” *Journal of the American Statistical Association*, Vol. 95, No. 450. pp. 573-585.

Saperstein, Aliya. 2008. “(Re)Modeling Race: Moving from Individual Characteristic to Multidimensional Marker of Status,” Forthcoming in *New Directions in Race Research*, ed. Charles Gallagher. University of North Carolina Press.

_____ 2006. “Double-checking the Race Box: Examining Inconsistency between Survey Measures of Observed and Self-Reported Race.” *Social Forces* 85(1): 57-74.

Schwartzman, Luisa F. 2007. “Does Money Whiten? Intergenerational Changes in Racial Classification in Brazil.” *American Sociological Review* 72(6):940-963.

Sugarman, Jonathan R., Robert Soderberg, Jane E. Gordon, and Frederick P. Rivara. 1993. "Racial misclassification of American Indians: Its effect on injury rates in Oregon, 1989 through 1990." *American Journal of Public Health* 83(5):681-684.

Telles, Edward and Nelson Lim. 1998. “Does it Matter Who Answers the Race Question? Racial Classification and Income Inequality in Brazil.” *Demography* 35: 465-74.

_____ and Edward Murguia. 1990. “Phenotypic Discrimination and Income Differences Among Mexican Americans.” *Social Science Quarterly* 71(4): 682-94.

Tilly, Charles. 1998. *Durable Inequalities*. University of California Press.

Wilson, William Julius. 1987. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy.* University of Chicago 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.

Zhou, Min. 2004. “Are Asian Americans Becoming ‘White’?” *Contexts 3, 1: 27-35.*

Table 1: Comparison of Observed and Self-reported Race in the 1988 NSFG

		Perceived race			
		Black	**White**	**Other**	**Row total**
Self-reported race	**Black**	2634	41	5	2680
	White	20	5118	30	5168
	Other	15	58	161	234
	Black-White	11	1	1	13
	Black-Other	54	0	2	56
	White-Other	1	108	3	112
	All three	12	0	0	12
	DK/Refused	1	2	1	4
	Column total	2748	5328	203	8,279

Note: Unweighted counts. Self-reported “Other” is a combination of American Indian and Asian or Pacific Islander race responses.

Table 2. Descriptive statistics, 1988 NSFG

	Perceived whites	Seen as white, identifies nonwhite	Seen as white, identifies as other	Self-Reported Others	Seen as white, identifies as black	Self-Reported Blacks
R's earnings (1988$)	$14,792	$12,217	$10,742	$13,795	$14,010	$13,476
Family income	$34,748	$28,213	$33,102	$30,212	$22,071	$22,033
Age (years)	30	28	26	29	29	29
Education (years)	13.1	12.5	12.6	13.0	12.4	12.7
Hours worked	36.2	35.5	31.9	36.4	39.7	37.1
Foreign born	6%	17%	21%	31%	12%	4%
Married	57%	40%	47%	47%	32%	29%
Cohabiting	5%	4%	7%	5%	0%	5%
Never married	28%	42%	41%	38%	44%	48%
Number of kids	1.0	1.1	1.0	0.9	1.2	1.2
Ever pregnant	66%	68%	59%	65%	80%	76%
Ever used the pill	20%	22%	20%	17%	26%	22%
Lives in central city	18%	29%	21%	29%	41%	52%
Northeast	21%	12%	10%	17%	15%	16%
South	32%	29%	21%	24%	41%	54%
Midwest	28%	25%	17%	17%	37%	22%
West	20%	33%	52%	42%	7%	8%
Lived with both parents	78%	71%	79%	69%	59%	51%
Mother's education	11.8	10.8	11.0	11.5	10.4	10.9
Mother's age at 1st birth	27	26	22	27	33	34
History of hypertension	11%	17%	17%	15%	17%	18%
History of diabetes	3%	8%	9%	5%	7%	3%
Sickle cell anemia	0%	1%	0%	1%	2%	2%

Table 3. Estimating Income Differences by Inconsistent Racial Classification and Propensity Score, 1988 NSFG

Panel A. Seen as white, self-identifies as nonwhite

	Comparison with perceived whites				*Comparison with self-reported nonwhites*			
	Family income		Earnings		Family income		Earnings	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Inconsistently classified	-7252 *	-6312 *	-2377 +	-2850	5182 *	3530	-771	1696
	(2459)	(2826)	(1257)	(1827)	(2442)	(3522)	(1286)	(2051)
Propensity score		-2923 ***		-1486 ***		514 ***		-273 ***
		(248)		(141)		(147)		(69)
Two-way interaction		1896 ***		1016 ***		-81		-153
		(538)		(257)		(455)		(137)

Panel B. Seen as white, self-identifies as other

	Comparison with perceived whites				*Comparison with self-reported others*			
	Family income		Earnings		Family income		Earnings	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Inconsistently classified	377	2260	-3945 *	-2543	5667	8861	-3748 +	995
	(3606)	(4249)	(1858)	(3106)	(3709)	(7041)	(2038)	(4467)
Propensity score		-1324 ***		-1576 ***		-93		-300 ***
		(254)		(176)		(158)		(59)
Two-way interaction		541		897		-123		-23
		(909)		(595)		(286)		(130)

Panel C. Seen as white, self-identifies as black

	Comparison with perceived whites				*Comparison with self-reported blacks*			
	Family income		Earnings		Family income		Earnings	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Inconsistently classified	-15033 ***	-11153 **	-653	-2207	-1589	-6596 *	1109	-112
	(2332)	(3876)	(1566)	(2769)	(2450)	(3336)	(1573)	(2590)
Propensity score		-2686 ***		-226		-275		-118
		(462)		(153)		(322)		(182)
Two-way interaction		1817 +		582		1816 *		309
		(1020)		(405)		(735)		(626)

Note: + p<.10 * p<.05 ** p<.01 *** p<.001