

## Commentary and Debate

To conserve space for the publication of original contributions to scholarship, the comments in this section must be limited to brief critiques; author replies must be concise as well. Comments are expected to address specific substantive errors or flaws in articles published in *AJS*. They are subject to editorial board approval and peer review. Only succinct and substantive commentary will be considered; longer or less focused papers should be submitted as articles in their own right. *AJS* does not publish rebuttals to author replies.

### RACIAL RIGIDITY IN THE UNITED STATES: COMMENT ON SAPERSTEIN AND PENNER<sup>1</sup>

As many sociologists have pointed out, the system of racial categorization in the United States is historically unique (Davis 2001; Morning 2008). One oft-discussed example of U.S. exceptionalism is the rigidity surrounding the infamous “one-drop rule” and the notion that all persons with any African ancestry are black. In countries without that historical legacy there is much less formality in racial categorization and thus a greater potential for individuals to change racial classifications as they move through the life course.<sup>2</sup>

In “Racial Fluidity and Inequality in the United States” (*American Journal of Sociology* 118 [3]: 676–727) Saperstein and Penner (2012; henceforward S&P) call into question the assumed high level of rigidity in racial categorization in the United States. They assert instead that the racial rigidity that social science has long documented and analyzed is mistaken (or has changed) and that, regardless of phenotype, the probability that an individual will identify or be seen as white or black can be significantly altered by undergoing raced status experiences such as going to prison, falling into poverty, or utilizing public assistance.

<sup>1</sup> All authors contributed equally. Direct correspondence to Rory Kramer, Department of Sociology and Criminology, Villanova University, 800 East Lancaster Avenue, Villanova, Pennsylvania 19085. E-mail: rory.kramer@villanova.edu

<sup>2</sup> As Telles and Paschel (2012) point out, the degree of racial fluidity in other countries should not be exaggerated. Even in countries where race is very ambiguously defined, skin tone remains the most significant determinant of racial classification and identity.

One implication of S&P's findings is that white privilege in the United States is not as durable as commonly believed, since an individual's likelihood of being seen as white, and of enjoying all of the privileges that go along with that racial status, can be significantly harmed by one's missteps and misfortunes. Conversely, members of racial minority groups can increase the likelihood that they will be classified and treated as white by avoiding such things as involvement with the criminal justice or public welfare systems. While for most of sociology's history there has been clear agreement that race in the United States is best thought of as an ascribed master status, S&P challenge this common wisdom and claim that fluidity is "an integral part of the 'social invention' . . . that is race in the United States" (p. 680). Moreover, S&P suggest that their results are not "driven by the small minority of respondents for whom racial perceptions are thought to be more fluid and complex" (p. 686) but rather can be generalized to the average person.

We appreciate S&P's basic point that racial categorization is a social construction and thus theoretically subject to change. However, S&P's article presents a provocative thesis regarding widespread racial fluidity across the life course, but on close inspection it does not provide evidence that racial change is a common experience across all racial and ethnic groups nor that social status experiences cause individual racial change. Contrary to S&P's claim that their modeling strategies "demonstrate consistent evidence of the reciprocal relationship between racial fluidity and inequality in the United States" (p. 687), we argue that none of their modeling strategies disentangle the well-documented effect of race on social status from their newly proposed effect of social status on racial categorization. While S&P argue that the robust correlations between race and social status variables reflect how racial groups are stereotyped, these relationships could simply reflect the enduring structural reality of racism's impact on life chances. For example, the observed statistical associations between the status variables and black self-identification do not necessarily tell us that people are more likely to see themselves as black if they experience a disfavored social status. They only reveal what sociologists have long known—that black people are more likely to suffer poverty, incarceration, unemployment, and other negative social outcomes.

S&P incorrectly assume that random error in the survey would make it "more difficult, perhaps even impossible, to find evidence of the expected relationship between social position and racial classification" (p. 689). In fact, that error is exactly what enables their results, as a flawed control variable can lead to unfounded conclusions regarding the theoretically important independent variables in the model. In this case, even slight error in the measurement of prior racial classification/identification can lead to unwarranted conclusions about the causal impact of social position on racial categorization.

After demonstrating how their findings could be driven by measurement error, we more directly test S&P's causal mechanism by exploiting variability in an interviewer's knowledge about a respondent and find no evidence of a reciprocal relationship between an individual's social status and an observer's assessment of that individual's race. In the United States, race continues to be best conceived of as an ascribed master status.

#### RESULTS DESPITE OR BECAUSE OF MEASUREMENT ERROR

S&P present two alternative readings of the observed discrepancy in racial identification over time—either a social constructivist view in which fluidity is part of the unstable equilibrium of U.S. racial divisions or a “primordialist” view in which fluctuation is “generated by poor question wording and limited answer options or is an issue of comprehension. . . . This [primordialist view] implies that clearer questions or better categories would eliminate the fluidity we observe and that eliminating these inconsistencies is desirable” (pp. 681–82). Such framing presents a false dichotomy. Acknowledging that the NLSY survey includes flawed measures of racial identity that inaccurately map to common understandings of racial categories does not imply a belief in primordialism; rather, it rightly recognizes that question wording matters. Social constructivism does not negate the need for clear racial categories on surveys, nor does everything social scientists observe have deep theoretical implications. Sometimes, observed data variation is just measurement error.

A closer inspection of the NLSY data raises important issues about its suitability for S&P's theoretical interests. They claim that fluidity is surprisingly widespread because 20% of the NLSY sample experiences at least one change in white/black/other racial classification over two decades. We have serious concerns about presenting an analytic model based on those survey options because “other” is not a racial identity; it is the absence of measured identity, and it is disproportionately involved in the classification changes observed by S&P.<sup>3</sup> While the vast majority of those initially classified as white or black were consistently classified that way across the full 17 survey years, only 1% of the respondents who were classified as “other” in 1979 consistently received that designation. Furthermore, the 20% overall rate of change identified by S&P is more accurately described as over 85% for the Hispanic population and 8% for the non-Hispanic population. We believe this is not due to racial fluidity, per se, but rather a flawed classification scheme given to interviewers with inadequate

<sup>3</sup> Given S&P's emphasis on the causal role of stereotypes in racial identification and classification, the “other” category is problematic because it lumps together disparate racial groups (e.g. Asians and Native Americans) as similarly stereotyped regarding social status.

directions on its use. In sum, we believe that the bulk of racial change identified by S&P is driven by unclear survey categories (“other”) or alternative forms of measurement error (about two-thirds of the non-Hispanic changers experience only one change across the 17 survey years).

S&P could argue that their appendix table A2 (p. 718) demonstrates that the effects of the intervening social status variables on racial classification were significant even when filtering out respondents that self-identified as Hispanic, Native American, or multiracial in 1979 (“populations with high, but theoretically fixed, propensities toward ambiguity”; p. 707).<sup>4</sup> Moreover, it could be that data noise, whether due to categorical ambiguity or human error, would only serve to make it harder for them to find statistically significant status effects were the noise randomly distributed (p. 689). We argue that while it is true that noise in the status variables decreases the likelihood that the status effects will be statistically significant, noise in the key control variable, the racial category selected on the prior survey, does the opposite.

To illustrate this point, the first column in table 1 provides estimates from a logistic regression analysis of black self-identification of the type used by S&P. We only include the social status variables that we were able to accurately match with the values in S&P’s descriptive statistics table (p. 694). The point of our analyses is not to exactly replicate S&P’s coefficients but rather to illustrate how the modeling strategy of controlling for prior racial categorization cannot distinguish between random error and status driven fluidity. This is true for all of S&P’s models in the discussion under scrutiny here, regardless of the specific status variables in the model, the use of demographic controls, and adjustments for missing data.

Consistent with S&P’s results (p. 699), we find highly statistically significant independent effects for poverty, long-term unemployment, ever married as a teen, ever a teen parent, and currently living in an inner city on racial identity (table 1, col. 1). However, these results do not demonstrate the existence of theoretically meaningful fluidity where social status is systematically related to changes in racial identity. Instead, as long as there is a small amount of random error in the 1979 racial measurement that differs from the random error in the 2002 racial measurement, the social status variables will simply exhibit their familiar structural relationships with race (e.g., white people are less likely to experience incarceration) that S&P interpret in the reverse direction (incarcerated individuals are less likely to be seen as white).

<sup>4</sup>However, S&P did not report the results of any fixed-effects analyses for the non-Hispanic population separately. When we estimate racial classification fixed-effects models for the non-Hispanic population we find that none of the social status variables are significant (using only the ever-status variables that we could successfully reproduce; see table 1). Results are available on request.

## Racial Rigidity in the United States

TABLE 1  
LOGISTIC REGRESSION OF STATUS EFFECTS FOR BLACK SELF-IDENTIFICATION IN 2002  
UNDER DIFFERENT ASSUMPTIONS ABOUT FLUIDITY VERSUS RANDOM ERROR

Variable	Estimates controlling for the original <i>black in 1979</i> measure used by S&P to model status-driven fluidity	Estimates controlling for a created <i>black in 1979</i> measure that is different from <i>black in 2002</i> only via induced random error
Ever incarcerated . . . . .	.28 (.35)	.33 (.27)
Ever unemployed > 4 months . . . . .	.75*** (.20)	.77*** (.15)
Ever below poverty line . . . . .	.95*** (.24)	.91*** (.17)
Ever received welfare . . . . .	.29 (.21)	.30 (.16)
Ever graduated college . . . . .	-.27 (.25)	-.20 (.18)
Ever married as a teen . . . . .	-1.07** (.35)	-1.46*** (.26)
Ever teen parent . . . . .	.89*** (.29)	.97*** (.23)
Lives in an inner city . . . . .	1.00*** (.27)	1.04*** (.21)
Lives in a suburb . . . . .	.074 (.26)	.047 (.19)
Black Self-ID, 1979 . . . . .	8.15*** (.22)	6.73*** (.15)
N . . . . .	7,718	7,718

NOTE.—Numbers in parentheses are SEs. In both the original and the simulation data the same percentage of the sample (1.6%) has a discrepancy in black racial identification for 1979 and 2002. All variables follow S&P’s coding instructions and match the approximate means S&P provide in their table 1 (p. 694).

- \*  $P < .05$ .
- \*\*  $P < .01$ .
- \*\*\*  $P < .001$ .

*Ed. note.*—AJS received updated estimates for the second column of table 1 after the contributions to this exchange were finalized for print. In the interests of completeness and accuracy, this updated information is available as an online appendix to this comment; these estimates do not represent a substantive change in the authors’ analysis but do alter the magnitude of the output slightly.

To show that S&P’s results can be reproduced without theoretically meaningful fluidity we present the results of a simulation analysis in the second column of table 1. We first assume that there is no individual racial change and set each individual’s racial identity in 1979 equal to the corresponding 2002 self-identity. At this point, by definition, if we were to regress the 2002 identity on the 1979 identity it would explain 100% of the variation. We then randomly change the 1979 black/non-black status of 1.6% of the sample, the same percentage of individuals who actually

changed in the NLSY. We then replicate this procedure 100 times, estimating the baseline logistic regression model in table 1 after each random assignment of 1979 racial self-identities.

The results of the simulation are striking: averaging the 100 sets of estimated coefficients and standard errors for each of the social status variables reveals that we can reproduce S&P's basic pattern of results (shown in table 1) even though by construction the changes or nonchanges in race between 1979 and 2002 are random. As such, S&P's methodology is fundamentally incapable of distinguishing between the impact of racial categorization on social status and the reverse. Their inclusion of prior racial categorization as a key control variable is, in practice, irrelevant for their theoretical concerns, and the residual correlations tell us nothing about their central causal claim. In sum, S&P's results do not depend on the existence of theoretically meaningful fluidity; they can be achieved simply by introducing a sufficient amount of random noise in their primary control variable (here, the respondents' self-reported race when they were 14–22 years old).<sup>5</sup>

We do not contend that the difference between the two racial measurements is always the product of random error.<sup>6</sup> In fact, we expect that *nonrandom measurement error* plays the same role. The point of the simulation analysis is that, consistent with the logic behind statistical significance testing, it is important to first rule out random error. Because S&P do not, the models merely report well-established correlations between race and social status with the variables flipped from one side of the equation to the other. For example, when we simply move the *ever below the poverty line* variable to the other side of the equation and use the race the respondent self reported as an adult as the predictor, we unsurprisingly find the same significant association (see table 2). In interpreting this significant relationship, it is important to keep in mind that just because an individual has his or her racial identity officially recorded after an event has occurred does not mean that the person lacked that racial identity in the years leading up to the event; race existed and influenced these respondents *before* they entered the survey and anyone asked them about it.

Presumably to buttress their causal claim, S&P report the results of Granger causality estimates in their appendix (table A6).<sup>7</sup> However, it is

<sup>5</sup> Again, this is a general outcome that results from the basic structure of S&P's empirical models and applies to whatever data are used, either self-identified or interviewer classified race.

<sup>6</sup> As S&P note, in 1979, respondents chose from a list of 28 ethnicities/nationalities, but in 2002 they chose from only five racial groups and had a separate question about Hispanic ethnicity. This dramatic change in survey questions could obviously introduce significant error in longitudinal social science models. Such error need not be randomly distributed.

<sup>7</sup> We additionally note that when S&P include fixed effects, only 3 of 13 coefficients are significant in each of the classification models, even with *N*'s over 160,000 (S&P's table

## Racial Rigidity in the United States

TABLE 2  
LOGISTIC REGRESSIONS UNDER DIFFERENT ASSUMPTIONS ABOUT CAUSAL DIRECTION

Variable	Black Self-ID 2002	Ever in Poverty 1979–2002
Ever in poverty, 1979–2002 . . . . .	1.42*** (.23)	. . .
Black Self-ID, 2002 . . . . .	. . .	1.42*** (.23)
Black Self-ID, 1979 . . . . .	8.22*** (.21)	-.07 (.23)
N . . . . .	7,718	7,718

NOTE.—Numbers in parentheses are SEs. All variables follow S&P’s coding instructions and match the approximate means S&P provide in their table 1 (p. 694).

- \*  $P < .05$ .
- \*\*  $P < .01$ .
- \*\*\*  $P < .001$ .

widely accepted that Granger causality tests cannot reliably reveal true structural causality.<sup>8</sup> For example, they cannot overcome a basic concern in gauging causality, whereby a third factor could be driving both the dependent variable and the variable in question. In this case, a likely third factor is the racial identity and raced experiences that an individual had before he or she was ever contacted by the NLSY. More fundamentally, S&P delineate a very particular causal mechanism: social status alters racial self-perception and the perceptions of others via internalized racial stereotypes. To establish causality, S&P must provide evidence directly related to that specific causal mechanism. In the following section, we present two direct tests of their causal mechanism using NLSY data.

### POSSIBLE WAYS TO TEST S&P’S CAUSAL MECHANISM

Showing that random error can create similar findings to S&P and that Granger causality is very different from actual causality does not mean that S&P’s thesis is untestable with the NLSY data. Building on an idea

A3). Moreover, the three are different for black compared to white. For example, completing college is associated with a higher probability of being classified as white but not a lower probability of being classified as black. Contrary to their theory, completing college is associated with a lower probability of self-identifying as white. The lack of consistency across racial categories and between self-identified and observer-classified race is unexplained by their theory or discussion.

<sup>8</sup>For example, Stock and Watson (2003) explain that “Granger causality has little to do with causality in the sense that it was used elsewhere in this book [where] causality was defined in terms of an ideal randomized controlled experiment, in which different values of  $X$  are applied experimentally and we observe the subsequent effect on  $Y$ . . . . While ‘Granger predictability’ is a more accurate term than ‘Granger causality,’ the latter term has become part of the jargon of econometrics” (p. 449).

S&P present in a footnote, we offer below two novel approaches to testing S&P's causal argument directly. For the specific variables we examined with these approaches (illicit drug use, illicit drug sales, and prison record), none of the analyses support S&P's thesis.

S&P provide only one direct test of their argument that an interviewer's racial classification could be "colored by the respondent's answers" (p. 688). They point out that interviewers heard the respondent racially self-identify in 1979 before classifying that respondent as white/black/other. To test whether hearing the respondent's answers to the racial identification question leads to bias in an interviewer's classification S&P compare the level of consistency between racial identification and classification in 1979 with the level of consistency between 1979 racial identification and 1980 racial classification. The logic underlying this test is that in 1980 the racial self-identity question was not asked, and, presumably, the interviewer would not know or remember the youth's responses from the previous year. Noting that statistical tests fail to find any difference, S&P concluded that "the interviewers' hearing the respondents' self-identification in 1979 did not significantly influence their classification" (p. 688). An interesting aspect of this conclusion is that it implies that interviewers' racial classifications will not be influenced by hearing a respondent specify a black identity, yet will be influenced by a respondent mentioning a status indirectly associated with black identity, such as having one's first child as a teenager.<sup>9</sup> For the sake of this comment, the most important aspect of this conclusion is that it points to the possibility of better tests of S&P's causal mechanism that exploit myriad variation in the data concerning the interviewer's level of knowledge about the respondent's answers.

#### Question Variation in Different Survey Waves

In 1980 (but not in 1981) the NLSY asks a series of questions about the respondent's criminal behavior. Thus, extending the logic of S&P's test of the influence of the 1979 racial identification question, one could examine whether survey items about a respondent's criminal behavior matter for interviewer racial classification, relative to years when the questions were not asked. Here, we would be comparing the impact of reported criminal behavior in 1980 on racial classification in 1980 with the impact of reported criminal behavior in 1980 on racial classification in 1981. While S&P only include variables in their models where racial stereotypes reflect and ex-

<sup>9</sup> Arguably, this finding is inconsistent with much of the psychological literature on priming, where the word "black" is shown to invoke a variety of subconscious stereotypical associations (e.g., Wittenbrink, Judd, and Park 1997). More important, we believe that this finding is inconsistent with S&P's own theorizing.

aggerate an empirical reality (i.e., black people are more likely to be *perceived* as having a prison record and black people are more likely to *have* a prison record), an alternative approach would be to examine prominent stereotypes with absolutely no basis in reality. Doing so would help isolate the effects of racial stereotypes on an interviewer's classification from the well-known effects of structural inequality. A respondent's self-reporting of drug crimes provides an excellent opportunity to test empirically the theorized importance of stereotypes for racial classification. While there are pervasive stereotypes of black Americans as more likely than white Americans to use illegal drugs, survey evidence consistently contradicts this widespread belief (Wallace and Bachman 1991; Wu et al. 2011). Additionally, while black Americans are stereotyped as being much more involved in the sale of illegal drugs than whites, the reality is very different in the case of marijuana, the most widely used illicit substance (Mohamed and Fritsvold 2011).

We conducted logistic regression analyses of the relationship between self-reported drug crimes and interviewer racial classification for 1980 (the survey wave when the drug questions were asked) and 1981 (when the drug questions were not asked). If interviewers were more likely to classify an individual as black after hearing reports of illicit drug use or selling, that would support S&P's theory. Alternatively, and following S&P's reasoning, if there is no discernible difference between the coefficients for 1980 and 1981, one can conclude that hearing a respondent's self-reported drug activity does not significantly influence an interviewer's classification.

The results are presented in table 3. For each model, the coefficients associated with racial classification are statistically indistinguishable across both years. In particular, and contrary to stereotypes, respondent marijuana use actually increased the likelihood of white classification, regardless of whether the interviewer was informed about the offense. In fact, drug use of any type increased the likelihood of white classification across both years in statistically identical manners, while marijuana use or distribution was uniformly associated with a lower likelihood of black classification. The association between black classification and nonmarijuana drug selling was positive in both years, but the difference between the two coefficients is not statistically significant. Thus, it appears that a respondent's answers to the drug crime questions do not significantly affect the interviewer's racial classification.<sup>10</sup>

<sup>10</sup> Additionally, controlling for racial classification in 1979 did not change the significant positive relationship between selling marijuana and white classification in 1980. Thus, if we were to employ S&P's analytic models and follow their interpretation of results in terms of status driven fluidity, one would have to conclude that marijuana dealing was stereotypically associated with whiteness in 1980 and this stereotype had filtered into the subconscious of interviewers.

TABLE 3  
 COMPARISON OF SELF-REPORTED DRUG CRIME EFFECTS ON RACIAL CLASSIFICATION BY  
 WHETHER THE INTERVIEWER DIRECTLY HEARD ABOUT THE CRIMES (1980)  
 OR DID NOT (1981)

Drug Crime	WHITE		BLACK	
	1980	1981	1980	1981
Model 1:				
Used marijuana in past year . . . . .	.31*** (.06)	.38*** (.057)	-.33*** (.065)	-.34*** (.065)
Sold marijuana in past year . . . . .	.24** (.08)	.19* (.08)	-.23** (.09)	-.24** (.09)
N. . . . .	10,190	10,190	10,190	10,190
Model 2:				
Used other drugs in past year . . . . .	.94*** (.08)	.91*** (.08)	-1.04*** (.09)	-1.02*** (.09)
Sold other drugs in past year . . . . .	-.34* (.16)	-.37* (.16)	.45** (.17)	.38* (.19)
N. . . . .	10,240	10,240	10,240	10,240

NOTE.—Numbers in parentheses are interviewer clustered SEs. All variables are coded 0-1, and all models are estimated with logistic regression. Missing data are excluded using casewise deletion.

- \*  $P < .05$ .
- \*\*  $P < .01$ .
- \*\*\*  $P < .001$

Variation in Interviewer/Respondent History

For their analytic models, S&P constructed several of their own “ever” measures, such as ever incarcerated or ever poor, by cumulating responses across survey waves from questions that were assessing events only in a particular year. This type of variable construction has implications for testing S&P’s causal mechanism. While in the example cited earlier regarding racial self-identification’s impact on classification S&P appeared to assume that the interviewer in 1980 would not have access to 1979 data, the opposite assumption appears to have been made for many of S&P’s ever-status variables. The assumption, that, for example, an interviewer in 1994 would know something about a respondent’s poverty status from 1984, when it was not on the questionnaire in front of them, is inconsistent with both common survey practice and statements from NLSY representatives (e-mail, NLS archivist and NLS User Services 2013).

In general, interviewers did not have any access to information about a respondent’s previous social experiences from answers given in earlier surveys.<sup>11</sup> However, it is possible that in cases where an interviewer had the

<sup>11</sup> Interviewers had access only to limited information from the very last survey on a need to know basis. Such information was included on an interviewer information sheet

respondent in previous survey rounds, interviewers might remember the respondent's earlier answers. While such a possibility seems remote to us for almost all of S&P's ever measures (e.g., an interviewer remembering a respondent being unemployed for 17 or more weeks 10 years ago), we could certainly see the possibility for the durable stigma of an incarceration record. Indeed, S&P explicitly argue that "a stint in prison" decreases the respondent's odds of being seen as white and increases the respondent's odds of being classified as black by the interviewer "in any future encounter" (p. 707).<sup>12</sup>

The fact that about half of the respondents in the NLSY were assigned a brand new interviewer in any given year provides yet another way to test S&P's argument. If S&P are correct, the relationship between ever incarcerated and racial classification should be significantly weaker when the interviewer is brand new and thus had no possible access to information about imprisonment several years ago. Conversely, following S&P's logic, the ever-incarcerated effect should be most pronounced when interviewers have familiarity with a respondent's past.

To test S&P's causal mechanism, we follow the basic structure of S&P's models and estimate logistic regressions in which a respondent's current classification as either black or white depends on the previous year's racial classification and whether a respondent was ever incarcerated. We estimate these basic models for two samples: one in which the respondent has been interviewed by the same interviewer before and one in which the interviewer has never met the respondent and thus has no information about a respondent's incarceration history beyond the current/last year items on the questionnaire in front of the interviewer.<sup>13</sup> In both cases, the sample is restricted to those respondents who were not incarcerated during the current year.<sup>14</sup>

---

for the purpose of framing current questions (e.g., "Are you still employed with [insert name of company]p?"). The interviewer sheet had no information about, e.g., a respondent's incarceration experience several years ago.

<sup>12</sup> A close look at the data reveals that among those "ever incarcerated," the slight majority of people whose classification changed from white to black had this happen *before* they were incarcerated. This directly contradicts S&P's assertion.

<sup>13</sup> We also estimated the basic models for the full sample and included interaction terms between the status variables and the continuous measure of the number of times the respondent had been seen by that interviewer. The interaction was statistically insignificant and minuscule in all models, suggesting that the incarceration effect is not contingent on an interviewer's potential familiarity with a respondent's past. Additional analyses with alternative specifications (e.g., unit fixed-effects models) and other status variables in S&P's models (e.g., ever below the poverty line) also support these same conclusions.

<sup>14</sup> To be thorough, we also exclude 1980 because of the previously mentioned series of criminal history questions. Additionally, to ensure that items concerning reasons for unemployment since the last interview were not affecting the results the models were

TABLE 4  
 LOGISTIC REGRESSION FOR THE LIKELIHOOD OF BEING CLASSIFIED AS WHITE OR BLACK

Status Variable	RESPONDENT	
	Previously Had Interviewer	Did Not Previously Have Interviewer
Ever incarcerated:		
White . . . . .	-.62 *** (.15)	-.62 *** (.10)
Black . . . . .	.59 * (.24)	.66 ** (.23)
N . . . . .	64,955	64,438

NOTE.—Data are from NLSY79. Numbers in parentheses are interviewer clustered SEs. N is given in person-years. All models include lagged race of respondent and year fixed effects. Missing data are excluded using casewise deletion.

- \*  $P < .05$ .
- \*\*  $P < .01$ .
- \*\*\*  $P < .001$ .

Results from the logistic regressions are displayed in table 4. As in S&P’s analysis, the effect of ever being incarcerated was highly statistically significant and in the expected direction for the sample of respondents that had the same interviewer previously. But the effect was also present for the sample of respondents that never had the interviewer before. More important, the effects were statistically indistinguishable between the two samples.

Both in terms of an interviewer hearing a respondent’s admitted drug crimes or potentially remembering that a respondent was interviewed in prison, our results offer no support for the notion that the observed associations with racial classification are due to the causal mechanism described by S&P, where an interviewer’s “classification is colored by the respondent’s answers” (p. 688). While the limits of comment space prevent us from significantly detailing other analyses, the data offer more possibilities for investigating S&P’s thesis. For example, in contrast to the full interview in which the respondent had their race observed at the end, in the short pre-interview to get into the NLSY sample the respondent’s race was classified early on, before any questions were asked about income and public assistance. Therefore, one could exploit the variability in question order between the two surveys to test whether hearing a respondent declare income from welfare matters for the observer’s classification of the respondent as black (our own analysis suggested that it did not; results are available on request). To test arguments about racial fluidity we encourage future research to use the fact that interviewers cannot be influenced by knowledge that, realistically, they do not possess.

---

reestimated to exclude those who were incarcerated either in the current interview cycle or the last interview cycle. The conclusions were unchanged (results available on request).

Our attempts to directly test S&P's causal mechanism fail to support their claim that social status variables exhibit their significant associations with race because of the internalization of powerful stereotypes. We conclude that the correlations reflect well-known structural relationships where racial categorization probabilistically determines social outcomes.

### CONCLUSION

As S&P highlight, researchers unquestioningly rely upon racial classifications provided by survey research too frequently without interrogating whether and how those classifications then impact the analytic results (Zuberi 2001; Zuberi and Bonilla Silva 2008). S&P ask an important question by considering how racial classification is related to socioeconomic status and inequality, but they do not take the steps necessary to answer that question. Unfortunately, while criticizing other sociologists for rarely questioning the presumed rigidity of racial classifications, they did not test their own causal mechanism adequately, nor interrogate the actual breadth and depth of racial fluidity.

S&P "suggest that taking racial fluidity into account in studies of inequality is less a matter of reestimation than reinterpretation" (p. 685). Consistent with this statement, we also find robust correlations between the same social status variables and race that can be interpreted in either causal direction. However, S&P go on to conclude that they "offer new evidence" (p. 676) and that their models "show that racial fluidity is not simply a sign of confusion or misapprehension among particular respondents but a function of the relationship between race and inequality" (p. 716) and "demonstrate . . . a reciprocal relationship between racial fluidity and inequality" (p. 687). Our reanalysis indicates that S&P do not offer, show, or demonstrate these things. More specifically, while we appreciate the innovation of S&P's provocative thesis, their empirical results do not actually disentangle potential reciprocal relationships, and they do not rule out the possibility of measurement error, except by assumption.

We provide evidence showing that one obtains the basic pattern of S&P's results even when variation in racial identity from one period to the next is random. Their empirical methodology is incapable of testing what they set out to test. By contrast, variation in the knowledge actually possessed by interviewers offers a direct test of their specific causal mechanism. Results from such tests fail to support their conclusion that social status variables exhibit their significant associations with race because of the internalization of powerful stereotypes by both the observed and the observer. Instead, we argue that social status effects indicate the enduring power of structural inequality in the United States. In the United States, the average black person does not increase his or her chance of being seen as white by avoiding

poverty or unemployment and the average white person does not lose his or her racial privilege by going to jail or receiving welfare. Acknowledging that reality does not make us primordialists but rather realists.

Rory Kramer, Robert DeFina, and Lance Hannon  
*Villanova University*

REFERENCES

- Davis, F. James. 2001. *Who Is Black? One Nation's Definition*. University Park: Penn State Press.
- Mohamed, Rafik, and Erik D. Fritsvold. 2011. *Dorm Room Dealers: Drugs and the Privileges of Race and Class*. Boulder, Colo.: Lynne Rienner.
- Morning, Ann. 2008. "Reconstructing Race in Science and Society: Biology Textbooks, 1952–2002." *American Journal of Sociology* 114 (suppl.): S106–S137.
- Saperstein, Aliya, and Andrew Penner. 2012. "Racial Fluidity and Inequality in the United States." *American Journal of Sociology* 118 (3): 676–727.
- Stock, James, and Mark Watson. 2003. *Introduction to Econometrics*. Boston: Addison-Wesley.
- Telles, Edward, and Tianna Paschall. 2012. "Beyond Fixed or Fluid: Degrees of Fluidity in Racial Identification in Latin America." Presented at Latin American Studies Association Meetings, San Francisco, Calif. May 25.
- Wallace, John M., and Jerald G. Bachman. 1991. "Explaining Racial/Ethnic Differences in Adolescent Drug Use: The Impact of Background and Lifestyle." *Social Problems* 38 (3): 333–57.
- Wittenbrink, Bernd, Charles M. Judd, and Bernadette Park. 1997. "Evidence for Racial Prejudice at the Implicit Level and Its Relationship with Questionnaire Measures." *Journal of Personality and Social Psychology* 72 (2): 262–74.
- Wu, Li-Tzy, George E. Woody, Congming Yang, Jeng-Jong Pan, and Dan G. Blazer. 2011. "Racial/Ethnic Variations in Substance-Related Disorders Among Adolescents in the United States." *Archives of General Psychiatry* 68 (11): 1176–85.
- Zuberi, Tukufu. 2001. *Thicker Than Blood: An Essay on How Racial Statistics Lie*. Minneapolis: University of Minnesota Press.
- Zuberi, Tukufu, and Eduardo Bonilla Silva. 2008. *White Logic, White Methods: Racism and Methodology*. New York: Rowman & Littlefield Publishers.