DOES IT MATTER WHO ANSWERS THE RACE QUESTION? RACIAL CLASSIFICATION AND INCOME INEQUALITY IN BRAZIL*

EDWARD E. TELLES AND NELSON LIM

Previous studies of racial inequality have relied on official statistics that presumably use self-classification of race. Using novel data from a 1995 national survey in Brazil, we find that the estimates of racial income inequality based on self-classification are lower than those based on interviewer classification. After human capital and labor market controls, whites earn 26% more than browns with interviewer classification but earn only 17% more than browns with self-classification. Black-brown differences hardly change: Blacks earn 13% and 12% less than browns with interviewer classification and self-classification, respectively. We contend that interviewer classification of race is more appropriate because analysts of racial inequality are interested in the effects of racial discrimination, which depends on how others classify one's race.

Brazil, with the largest African origin population outside Nigeria, has substantial racial inequality. Official statistics consistently show a large difference between the incomes of whites and nonwhites and a relatively small difference between browns and blacks.[1] Data from the 1991 census reveal that, among male workers in Brazil, browns earned an average of 68% of whites' income, and blacks earned an average of 63% of whites' income (Barros, Mendonça, and Velazco 1996). Such estimates depend on official statistics, which presumably are collected using self-classification.

These estimates may be deficient if we are interested in measuring racial discrimination. Because racism is socially defined, a more appropriate estimation should rely on racial classification by others. Income differences by race are at least partly products of discrimination, in which discriminators, such as employers or consumers, reward or punish a person based on their perceptions of the person's race. Perceptions of race in Brazil depend primarily on phenotype, although they may be influenced by perceptions of social class or social context. Self-classification is subject to similar ambiguity, but may be an especially poor proxy for how one is treated in the labor market. Consciousness about one's identity, social networks, and cultural practices may further affect one's self-classification, although these factors often are not perceived by others and thus are not important criteria in the calculus of social classification. Such ambiguity may be especially great in countries like Brazil, where race has never been defined under law, unlike multiracial countries like the United States and South Africa.[2]

In this paper, we examine the extent to which there are white-brown, white-black, and brown-black income gaps when race is based on both self-classification and on interviewer classification. Do alternative definitions of race affect the level of estimated racial income inequality? We are particularly interested in the position of the brown population. In terms of income, are browns more similar to blacks than to whites, halfway between blacks and whites, or closer to whites than to blacks? To answer these questions, we examine data from a national survey conducted in Brazil in 1995.

BACKGROUND

Until the 1970s, scholars claimed that contemporary racial income differences in Brazil had little or nothing to do with racial discrimination; rather, they argued, these differences resulted primarily from the recent emergence of Afro-Brazilians from slavery (Pierson 1942; Wagley 1969) or from the deficient culture that they inherited from slavery (Fernandes 1965). These authors expected that racial differences would eventually disappear, as nonwhites would gradually acquire the necessary human and cultural capital to compete with whites. Studies using recent census and household survey data, however, show that as much as one third of the difference in income between whites and nonwhites cannot be explained by racial differences in variables like education, work experience, social origins, and region. This suggests that a substantial part of racial inequality is caused by racial discrimination in the labor market (Barros, Medonça, and Velazco 1996; Lovell 1989; Silva 1985).

Given the strong evidence for the persistence of unexplained white-nonwhite income differences, the scholarly

*Edward E. Telles, The Ford Foundation, Rio de Janeiro Office, 320 E. 43rd St., New York, NY 10017; e-mail: e.telles@fordfound.org. Nelson Lim, Department of Sociology, UCLA. We thank William Darity, David McFarland, Antonio Sergio Guimarães, and *Demography* reviewers for comments on earlier drafts. This is a revised version of a paper presented at the 1997 annual meeting of the Population Association of America, Washington DC. The research was supported by NSF Grant SBR-9710366.

1. In this paper, the term *nonwhite* includes both the census categories of brown (*pardo*) and black (*preto*). *Brown* is roughly equivalent to *mixed race*. Rather than include the many terms Brazilians use to describe their race, the Brazilian Census Bureau simplifies the intermediate categories between white, black, and Indian into a single category. The intermediate categories are composed of both ancestry (e.g., *caboclo*, *mestizo*, *cafuzo*, *mulatto*) and appearance (e.g., *escuro*, *marron*) categories.

2. There is also racial ambiguity in the United States among the black and white populations, despite the historical rigidity of the U.S. racial system in which any child of a black person is considered black (Davis 1991). Because such laws in the United States were based on ancestry rather than appearance, a sector of the self-defined U.S. black population might be classified by others as of another racial group (Hahn, Mulinare, and Teutsch 1992). Conversely, persons may be classified by others as black, but self-classify as another race or as multiracial (Twine 1995).

discussion has turned to the extent of black-brown differences. Degler's (1986 [1971]) well-known "mulatto escape hatch" theory claimed that the basic difference between race relations in Brazil and the United States is that mulattos in Brazil enjoy a favored status vis-à-vis blacks, whereas mulattos in the United States are treated like blacks.[3] According to Degler, the status of the mulatto in Brazil reflects a national belief in whitening, where whiteness is desirable and blackness is to be escaped.

Degler's belief in the intermediary status of the mulatto, however, has not held up to empirical investigation. Based on a human capital model, estimated using 1976 national household survey data, Silva (1985) refuted Degler's argument. He found that the unexplained income difference between browns and whites was similar to that between blacks and whites, and concluded that blacks and browns suffer similar levels of discrimination. Since then, other empirical studies have also found that the income gap between browns and blacks is small compared with that between whites and browns. Unlike Silva (1985), others using similar human-capital models found that browns generally earn more than blacks, but the brown-black income difference is smaller than the white-brown gap (Barros, Mendonça, and Velazco 1996; Lovell 1989).

Some analysts question the usefulness of census data because they believe that "money whitens," as the popular Brazilian saying goes. Based on his study of a village in Brazil, Harris (1964) reported a tendency for race-color identity to shift toward white among wealthier and better-educated nonwhites. In other words, race is based on a combination of phenotype and class factors. Although Harris did not distinguish between self-classification and classification by others, we assume he was referring to both. Thus, many better-off blacks classify themselves and are classified by others as brown or even white, and better-off browns are reclassified as white. Harris also found that the extent to which class and phenotype contribute to race is ambiguous and variable across observers. Based on his analysis of census data, Wood (1991) similarly claimed that widespread economic mobility between 1950 and 1980 allowed many persons classified as black in 1950 to reclassify as brown in 1980. A smaller proportion of browns also seem to have reclassified as white. On the other hand, Wagley (1968) doubted that money significantly "whitens" one's race, but suggested that money merely makes nonwhites more socially acceptable to whites. He refers to this as *social race*.

If money whitens one's racial classification, as Harris claimed, then the studies based on official data are likely to have overestimated white-black inequality: The income of blacks would be deflated because higher income blacks move out of the black category with self-identification. The income of whites is unlikely to be affected because better-off blacks would have income more similar to that of average whites. White-brown and brown-black inequality may also be biased, although its direction of error is not clear, as the brown category would be inflated by the movement of better-off blacks into it and better-off browns out of it.

For different reasons, Wade (1995) was also critical of studies based on official data. Based on his fieldwork in Colombia, Wade argued that in many South American countries, including Brazil, the primary racial cleavage is between blacks and nonblacks because discrimination is much harsher against blacks than against browns. Because many whites identify and treat as black, persons who identify themselves as brown, data using self-classification may overestimate discrimination against browns. Wade found that blacks often identify as brown when they migrate to non-black communities or leave the social networks of black communities. Thus, Wade supported Degler's argument about the mulatto escape hatch, suggesting that Degler's detractors erroneously relied on estimates that are flawed because they are based on self-classification. He claimed that official statistics deflate browns' income, leading to overestimates of brown-white inequality and to underestimates of brown-black inequality. The income of browns, according to Wade, is indeed between the incomes of blacks and whites and may be closer to that of whites, especially after human-capital and labor market variables are controlled.

The Brazilian Census Bureau instructs interviewers to collect race data based on respondents' self-reports, as the United Nations recommends and is currently international practice (Goyer and Domschke 1983; Pinto 1996). However, interviewers often do not ask respondents their race as they are instructed, but rather classify respondents themselves. This may occur because (1) interviewers feel certain about respondent's race, (2) automation creeps into interviewer routines, and (3) interviewers feel uncomfortable asking about race (Pinto 1996). Thus, the Brazilian census data on race are collected using a combination of self-classification and interviewer classification, making previous inequality estimates based on these data subject to an unknown mix of collection methods.

A separate analysis by Telles (1996) supports Wade's contention that the racial composition of social networks affect how one self-classifies. Telles found that darkening by self-classification (or whitening by the interviewer) occurs at the same rate as inconsistent classification in the other direction. Although income and social class have no predictive effects on whitening or darkening, schooling does. Inconsistent classification between interviewer and respondent are particularly common among the least-educated Brazilians, whereas the most educated are most likely to self-classify consistently with interviewer assigned classifications. Thus, the effects of social factors on self-classification compared with interviewer identification are complex, and the direc-

3. Despite Degler's claim, studies have shown that mulattos or lighter skin-toned blacks in the United States have significantly higher incomes and life chances, in general, than their darker counterparts (Keith and Herring 1991; Ransford 1970). Indeed, Telles (forthcoming) found that skin-tone differences among the African-origin population in the United States are greater than between blacks and browns in Brazil. The more important difference between the two countries is that persons with only partial black ancestry are *considered* black in the United States (Davis 1991) but as distinct from blacks in Brazil.

tion or extent of differences between the two forms of racial classification are not easily predictable.

DATA

We analyze data from a national face-to-face survey conducted by the Data Folha Instituto de Pesquisas, the survey unit of the *Folha de São Paulo*, one of Brazil's major daily newspapers. The survey was conducted in April 1995 and is officially called "300 Anos de Zumbi: Os Brasileiros e o Preconceito de Cor" (300 Years of Zumbi[4]: Brazilians and Racial Prejudice). For the first time in a national survey, each respondent's race was classified by both the respondent and the interviewer, providing a unique opportunity to examine the effect of alternative classifications on racial inequality.

The data are based on a national random sample of the urban population aged 16 and over. Municipalities were selected at random from within representative socioeconomic level, region, and size strata. Successive random samples were then taken of neighborhoods, then streets, and then individuals. The complete sample consists of 5,014 persons sampled across 121 municipalities.

Based on interviewer classification and self-classification, respondent's race was coded according to the five census racial categories: white (*branco*), brown (*pardo*), black (*preto*), yellow or Asian (*amarello*), and indigenous (*indigena*). Because the debate on racial inequality focuses on the black-to-white racial continuum, we exclude respondents that self-classified or were classified by interviewers as Asian or indigenous. Thus, we limit the sample to white, brown, and black persons, under both self-classification and interviewer classification. We further limit the analysis to those providing information about their income. This brings the final sample size to 4,000.

Two sources of bias contained in the survey may limit generalizations of our findings to all of Brazil. First, the sample includes only urban areas, but these accounted for fully 76% of the Brazilian population according to the 1991 census (Associação Brasileira de Estudos Populacionais 1996). Second, the survey overestimated the size of the black population, which is a problem for describing overall racial distributions, but is not a problem for most of the analyses conducted here. According to the 1991 census, the population of Brazil is 52% white, 42% brown, 5% black, 0.4% yellow (Asian), and 0.2% indigenous. The entire sample for the survey, according to self-classification of race, is 53% white, 36% brown, 10% black, 0.6% yellow, and 1.1% indigenous.

For the two race variables, we rely on a question about race and the interviewer's assessment of respondent's race. The wording of the close-ended survey question was, "Considering the following categories, what is your race: white, black, brown, yellow, or indigenous?"[5] These categories are the same as those used in the 1991 census of Brazil.

Interviewers were instructed to note the race of the respondent using the close-ended census categories before asking questions from the survey. According to the survey director, in most cases, interviewers readily classified respondents, and there was little doubt about the respondent's race. Clearly, racial classification depended largely on the tastes of the interviewers, even if they claimed classification was straightforward. We have no alternative evidence, however, on the level of racial ambiguity in Brazil that would permit any sensitivity analysis. One advantage of this survey is that interviewers resided in the same region as interviewees, diminishing errors in classification from variation in regional conceptions of racial classification.

In a few cases in which interviewers had doubts about racial classification, they met with the central survey committee to decide about classification. The final decision usually confirmed the interviewer's initial impression. Although a more objective classification of race might have been made by a panel of interviewers established for each region or by an interview with a person outside of his or her social context, this was not possible because of the extraordinary costs and efforts that would have entailed in a country as large as Brazil. The concern about the effect of social context is based on the assumption that money may whiten one's classification. As we mentioned previously, however, Telles (1996) found that persons with inconsistent racial classification tend to be poor and less educated, whereas those with the greatest consistency in racial identity are middle class and well educated.

We have no data on the characteristics of the interviewer. We know that most of the interviewers were white and relatively well educated, reflecting the correlation between being white and having higher status. Although this might be viewed as a bias, classification by lighter and more-educated persons may be an advantage here: Such persons are especially likely to be in social positions in which decisions about racial classification affect the incomes of the persons being classified.

Monthly income, the dependent variable, was collected for only five categories: 0–150 *Reais*, 151–375 *Reais*, 376–750 *Reais*, 751–1500 *Reais*, and 1501 and more *Reais*.[6] Unfortunately, the survey designers were interested in income only to be able to identify major social strata. The few income categories might generally preclude us from running standard human-capital models using the ordinary least square estimation procedure. However, we use a methodology that overcomes the deficiencies of categorical data when continuous data are preferred.

Independent variables include standard human-capital and labor market variables. Human-capital variables are age, age squared, and three categories of schooling: primary schooling (less than eight years of education), secondary

4. The title refers to 300 years since the birth of Zumbi, the leader of a runaway slave colony (Quilombo de Palmares) that lasted nearly 100 years.

5. The Portuguese wording of the questionnaire uses the word "cor," which can be literally translated as *color*. This concept, however, refers not merely to skin color but to a range of phenotypic characteristics (Ribeiro 1996). Thus we translate cor as race.

6. At the time of the interview, one *real* (plural *reais*) was equal to 0.90 U.S. dollar.

schooling (completion of primary school and at least some secondary education), and college (at least some post-secondary schooling). We also include a control for sex because human capital in Brazil varies by sex.

We include two controls for local labor markets. The first, Northeast, controls for the lower wages generally found in the Northeast region compared with the rest of Brazil. Throughout this century, the Northeast has been poor and economically underdeveloped, standing in sharp contrast to other, often industrialized areas of Brazil (Merrick and Graham 1979). Although the survey allows us to distinguish four regions (Northeast, Southeast, South, and North/Central East), we find little regional variation in income, except between the Northeast and all others. We also control for the size of the urban area, as income and costs of living tend to be higher in large cities. We control for large urban areas, those with over 500,000 persons. Although data are available, we do not control for employment sector (e.g., government) or type of employment (e.g., formal, self-employed) because we believe that, like income, these are outcomes of human capital and thus belong on the right-hand rather than left-hand side of a human-capital equation.

METHOD

Because income is reported only by category, the exact amount of a respondent's income is unobserved; we know only that income is within one of five intervals and the income thresholds that bound the categories. Thus, the dependent variable is completely censored and takes the following form:

$y = 1$ if income ≤ 150
$y = 2$ if $150 <$ income ≤ 375
$y = 3$ if $375 <$ income ≤ 750
$y = 4$ if $750 <$ income $\leq 1{,}500$
$y = 5$ if $1{,}500 <$ income.

We could have chosen from among several alternative strategies to analyze the determinants of income with this type of limitation. The first is to code income into five categories. Because the coded values of the dependent variable are no longer of the same scale as income, however, OLS regression is inappropriate (Green 1990:738; Stewart 1983). A common approach is to assign the midpoint value to the observations in any given closed interval, to provide some value for the open-ended interval, and then to proceed with ordinary least squares regression.[7] Stewart (1983:740–41), however, found this approach to yield inconsistent estimates. Another alternative is to use ordered logit or probit models to estimate the effects of independent variables on the probability of a respondent being in a given category. However, this ignores the values of the thresholds of income intervals, which provide the information about the scale of the dependent variable. A related disadvantage of this strategy is that one cannot interpret the results in terms of the natural unit of income.

Given these limitations, we choose a maximum likelihood approach to estimate our income regression (Green 1990:738–39; Stata Corp. 1997a:141–45; Stewart 1983). The procedure is quite similar to the estimation of the Tobit model in which only some of the data are observed (Breen 1996). The main difference is that we have no observed data for the dependent variable, although we know the category thresholds.

In this approach, the latent dependent variable is assumed to be given by, $y_i^* = \beta'\mathbf{x}_i + e, (i = 1,\Lambda,N)$, where y_i^* is the unobserved dependent variable, and $\mathbf{x}_i$ and β are vectors representing the independent variables and the unknown coefficients, respectively. The errors ε_I are assumed to be independent, identical, and normally distributed random variables; to have a mean and variance σ^2 equal to 0; and to be independent of $\mathbf{x}_i$. Because these assumptions are quite stringent, we use Huber-White corrections for the biases caused by the potential clustering among the errors (Stata Corp. 1997b:145–47). The underlying latent variable in this study is, like income variables in general, continuous and positively skewed. Thus, we transform the dependent variable into log income, giving it the following values:

$y = 1$ if $y^* \leq 5.01$
$y = 2$ if $5.01 < y^* \leq 5.93$
$y = 3$ if $5.93 < y^* \leq 6.62$
$y = 4$ if $6.62 < y^* \leq 7.31$
$y = 5$ if $7.31 < y^*$.

The next step in the procedure is to obtain a maximum-likelihood estimator by maximizing

$$\ln L = \sum_0 \ln \text{Prob}[y = 0] + \sum_1 \ln \text{Prob}[y = 1] + K$$
$$+ \sum_j \ln \text{Prob}[y = j].$$

The probability of a respondent taking the value of j can be given by

$$\text{Prob}[y_i = j] = \Phi\left(\frac{a_{j+1} - \beta' x}{\sigma}\right) - \Phi\left(\frac{a_j - \beta' x}{\sigma}\right).$$

For example, the probability of a respondent being in the third category of the income variable can be obtained by

$$\text{Prob}[y_i = j = 3] = \Phi\left(\frac{7.31 - \beta' x}{\sigma}\right) - \Phi\left(\frac{6.62 - \beta' x}{\sigma}\right).$$

Summing the log of these probabilities gives the maximand for the estimation. The resulting estimates from the maximum-likelihood method can be interpreted in the same way as coefficients from any log-income regression.

FINDINGS

The claim that income inequality may be quite different when interviewer classification rather than self-classification of race is used assumes that many individuals do not classify themselves in the same way they are classified by interviewers. The last column in Table 1 shows that the racial composition of our sample is 56.0% white, 33.3%

7. A common procedure is to estimate the mean for the highest category using the Lorenz criterion (Shryock and Siegel 1976).

TABLE 1. SELF-CLASSIFICATION BY INTERVIEWER CLASSIFICATION OF RACE

	Interviewer Classification				
Self-Classification	White (%)	Brown (%)	Black (%)	Total (%)	Percentage Distribution
White (%)	88.6	11.0	0.4	100.0	56.0
Brown (%)	20.2	71.0	8.8	100.0	33.3
Black (%)	2.2	39.8	57.9	99.9[a]	10.7
Total (%)	55.9	30.7	13.4	100.0	100.0

Note: $N = 4{,}000$.

[a]Row does not sum to 100.0 because of rounding.

brown, and 10.7% black using self-classification; the bottom row shows that racial composition is 55.9% white, 30.7% brown, and 13.4% black using interviewer classification. Thus, *net* shifts in racial classification occur from brown to black when changing from self-classification to interviewer classification, whereas the percentage white is stable across alternative forms of classification. Based on these data, 20% ((13.4 – 10.7) / 13.4) of persons who appeared black to interviewers classified themselves as brown.

Net change, however, may hide much inconsistency in racial classification for individuals. The first three rows of the first three columns in Table 1 show how self-classified persons in each racial group were classified by interviewers. Inconsistencies in racial classification are particularly great for nonwhites. Among those who self-classified as brown, interviewers classified 20.2% as white and 8.8% as black. Among those identifying themselves as black, interviewers classified 39.8% as brown and a surprising 2.2% as white. Finally, among self-classified whites, interviewers categorized 11.0% as brown and 0.4% as black. Thus, there are inconsistencies in racial categorization not only between proximate color categories but, in a few cases, between black and white racial categories, skipping the brown category altogether. Based on data from Table 1, we calculate that 21% of the sample was classified inconsistently across the two types of racial categorization.

Given this much inconsistency in classification, we expect that the income distributions of racial groups using alternative categorizations will differ. Table 2 shows the distribution of incomes using self-classification compared with interviewer classification. With interviewer classification, whites shift to higher income categories, whereas nonwhites shift toward lower income categories. This is apparent from comparing the percentage of persons in the poorest category (≤ 150 *reais*), which includes nearly half (48.3%) of the sample. In the poorest category, the percentage white drops from 44.4% to 42.6%, whereas the percent brown increases from 52.5% to 55.0%, and the percentage black increases from 55.0% to 57.4%. These findings suggest that, compared with interviewer classification, self-classification underestimates whites' income and overestimates nonwhites' income.

The means of independent variables in Table 3 reveal the higher socioeconomic status of whites and particularly

TABLE 2. MONTHLY PERSONAL INCOME BY RACE AND BY SELF-CLASSIFICATION AND INTERVIEWER CLASSIFICATION

Classification	≤ 150	151–375	376–750	751–1,500	> 1,500	Total (%)
Self-Classification (%)						
White	44.4	21.6	17.0	10.7	6.3	100.0
Brown	52.5	22.9	14.6	7.3	2.8	100.1[a]
Black	55.0	25.2	13.8	4.3	1.7	100.0
Interviewer Classification (%)						
White	42.6	21.7	17.5	11.7	6.5	100.0
Brown	55.0	23.4	13.7	5.6	2.3	100.0
Black	57.4	23.4	13.6	3.7	1.9	100.0
Total	48.3	22.5	15.8	8.8	4.6	100.0

[a]Row does not sum to 100.0 because of rounding.

TABLE 3. MEANS OF EXPLANATORY VARIABLES FOR THE TOTAL SAMPLE AND BY SELF-CLASSIFICATION AND INTERVIEWER CLASSIFICATION OF RACE

Variable	Total Sample	White		Brown		Black	
		Self-Classification	Interviewer Classification	Self-Classification	Interviewer Classification	Self-Classification	Interviewer Classification
Male (%)	52.5	52.2	52.3	53.3	53.2	51.8	51.5
Age	35.6	36.1	36.2	34.8	34.4	34.8	36.1
Education (%)							
Elementary	60.0	55.4	53.3	64.1	67.6	69.8	71.4
Secondary	29.7	31.0	32.3	29.3	27.0	25.1	24.2
College	10.3	13.6	14.3	6.6	5.4	5.1	4.5
Northeast Region (%)	22.7	16.0	14.8	32.3	35.6	28.4	23.7
Large Urban Area (%)	39.8	37.7	37.5	40.1	41.0	47.7	47.8

stark education and regional differences. For example, whereas 14.3% of interviewer-classified whites have attended at least some college, only 5.4% of browns and 4.5% of blacks have. Further, only 14.8% of whites reside in the Northeast region compared with 23.7% of blacks and 35.6% of browns. On the other hand, blacks are most likely and whites are least likely to reside in large urban areas. Similar differences hold using self-classification.

The comparison of self-classified with interviewer-classified racial groups in Table 3 shows that education is slightly higher for interviewer-classified whites and lower for interviewer-classified browns and blacks. The geographical variables present the biggest difference between forms of classification for nonwhites. The percentage of browns living in the Northeast is 35.6% with interviewer classification, 3.3 percentage points more than with self-classification. By contrast, 33.3% of interviewer-classified blacks reside in the Northeast, 4.7 percentage points less than for self-classified blacks.

We estimate racial differences in income while controlling for the human-capital and labor market characteristics using maximum-likelihood regression. Specifically we regress log income on sex, age, age squared, education, region, size of urban area, and race. We present the coefficients of the regressions in Table 4, using both self-classification (column 1) and interviewer classification (column 2) of race. The coefficients for race represent the difference in log income of blacks and whites compared with browns, the omitted race category. Despite our expectations, we find no interactions between race and the other variables. Thus, we present only a main-effects model with race as an independent variable.

Compared with browns' income, the log income of whites is greater when race is interviewer classified than when it is self-classified (.234 versus .155), and black income is somewhat lower with self-classification than with interviewer classification (–.125 versus –.145). These results reveal greater racial inequality net of the human-capital and labor market variables among the three racial groups with interviewer classification. The human-capital and labor market variables have the same or similar effects on (logged) income, regardless of the form of racial classification.

We exponentiate the results in Table 4 and, in Figure 1, illustrate the extent of racial inequality using actual income. The income of whites is 17% higher than for the income of browns with self-classification, but white-brown inequality increases to 26% with interviewer-classified race. On the other hand, brown-black inequality hardly changes under alternative categorizations. The income of blacks is 12% lower than the income of browns with self-classification compared with 13% with interviewer classification. Thus, white-brown and white-black inequalities are greater with interviewer classification, but brown-black inequality is roughly the same under both forms of classification.

To examine whether changes from inconsistency in classification are statistically significant, we include both classifications in the same model[8] (column 3 of Table 4). Because most respondents in the sample have consistent racial classification, we cannot simply enter both self-classification and interviewer-classification race variables in the model at the same time: This would lead to multicollinearity between variables. Rather, we first include the dummy variables for racial self-classification in the model and six additional dummy variables indicating all possible inconsistent forms of self-classification (i.e., all off-diagonal cells).

We find that changes in individual coefficients for race are consistent with results from the previous analyses. After we control for inconsistent classifications with the six dummy variables, the positive effect of being white increases and the negative effect of being black or brown increases compared with the model with only self-classified race variables (column 1). Moreover, the likelihood ratio

8. We thank the editor and an anonymous reviewer for suggesting this analysis.

TABLE 4. MAXIMUM-LIKELIHOOD REGRESSION ESTIMATES OF LOG INCOME ON SELECTED VARIABLES WITH HUBER-WHITE CORRECTIONS AND ROBUST STANDARD ERRORS

Independent Variables	Self-Classification	Interviewer Classification	Both Classifications
Male	.900	.900	.900
	(.039)	(.039)	(.039)
Age	.098	.098	.097
	(.008)	(.008)	(.008)
Age^2 (× 100)	–.103	–.103	–.103[a]
	(.009)	(.009)	(.009)
Secondary Schooling	.706	.688	.687
	(.043)	(.043)	(.043)
College	1.545	1.512	1.512
	(.061)	(.064)	(.061)
Northeast	–.449	–.426	–.426
	(.050)	(.050)	(.050)
Large Urban Area	.305	.314	.314
	(.039)	(.043)	(.039)
Race			
White	.155	.234	.228[a]
	(.049)	(.043)	(.051)
Black	–.125	–.145	–.145[a,†]
	(.062)	(.068)	(.081)
Inconsistent Racial Classification			
Self-classified white, interviewer-classified brown			–.245
			(.085)
Self-classified white, interviewer-classified black			–1.070
			(.521)
Self-classified brown, interviewer-classified white			.233
			(.127)
Self-classified brown, interviewer-classified black			–.131[†]
			(.126)
Self-classified black, interviewer-classified white			.256[†]
			(.366)
Self-classified black, interviewer-classified brown			.131[†]
			(.103)
Constant	2.234	2.197	2.211
Log-Likelihood	–4,550	–4,539	–4,538
Model χ^2	1,605	1,627	1,629
Degrees of Freedom	9	9	15
Number of Cases	3,993	3,993	3,993
Likelihood Ratio Test χ^2 (6) for Difference Between Columns 1 and 3.			23.31***

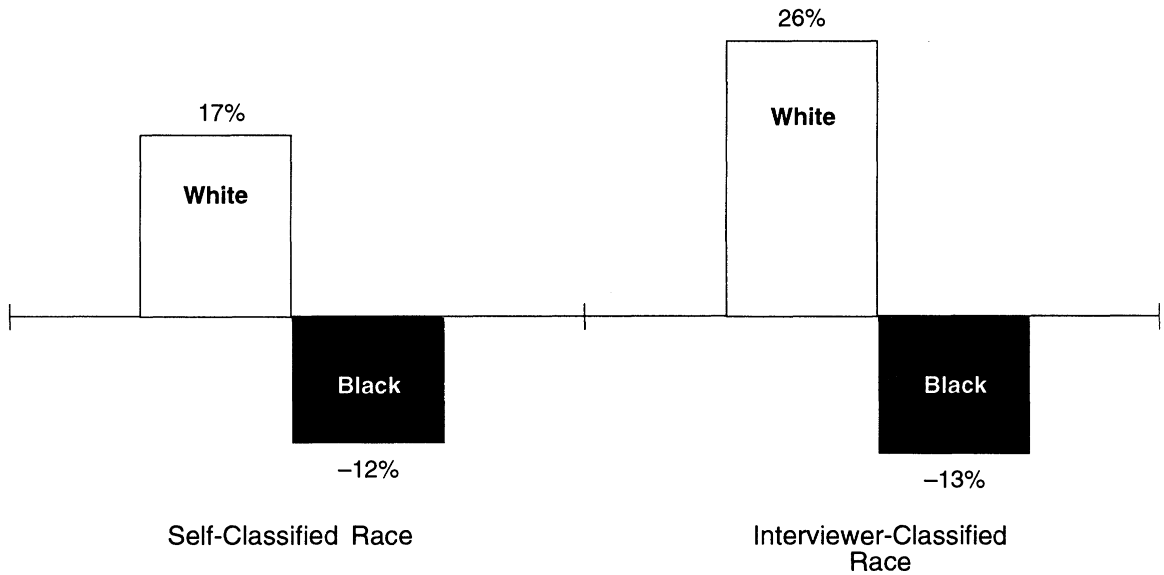
Note: Numbers in parentheses are standard errors.

[a]Self-classification.

[†] Not significant at the .05 level. (All other coefficients are significant at $p < .05$.)

***$p < .001$.

FIGURE 1. PERCENTAGE DIFFERENCES IN INCOME OF SELF- AND INTERVIEWER-CLASSIFIED WHITES AND BLACKS COMPARED WITH BROWNS, AFTER HUMAN-CAPITAL AND LABOR MARKET CHARACTERISTICS ARE CONTROLLED: BRAZIL, 1995



tests between these models (columns 1 and 3) indicate that the addition of the six dummy variables improves the log-likelihood of the models, and this reduction is highly statistically significant, indicating greater goodness of fit. Thus, we reject the null hypothesis that the collective effect of these six dummy variables is zero.

The results in Table 4 also show that those with lower incomes were interviewer-classified in darker categories than they self-classified (all three above diagonal cells). Alternatively, those with higher incomes were interviewer-classified in lighter categories than they self-classified (all three below diagonal cells).

Finally, Table 5 presents percentage differences in income from consistently classified browns and sample sizes for each cell. This allows for the comparison of all cells with one another. These values are calculated from column 3 of Table 4, where consistently classified browns is the comparison category. For example, consistently classified whites have 25% higher income than do consistently classified browns.

Table 5 reveals that income is similar across self-classified cells that are within interviewer-classified racial categories.[9] Conversely, income varies widely across self-classified racial categories. Thus, interviewer classification more reliably accounts for variations in income by race than does self-classification. Incidentally, this finding is confirmed by the model fits in Table 4, where the interviewer-classification model (column 2) fits better than the self-classification model (column 1) and fits about as well as the model with self-classification plus inconsistent cells (column 3). This strongly suggests that interviewer classification is preferable for measuring racial discrimination if we believe, as many economists do, that unexplained racial income differences, after human-capital is controlled, are due primarily to discrimination.

Because our estimates of white-brown inequality vary depending on the method of classification, we expect an examination of the two white-brown cells in Table 5 to reveal the source of the difference. Interviewers classified many (248) self-classified browns as white, and these persons tended to have incomes similar to those of consistently classified whites (26% versus 25% greater than consistently classified browns). Because these persons have incomes similar to those of whites, the inclusion of this sizable group (20.2% of all self-classified browns) in the brown category increases the average income of browns based on self-classification, making it more similar to that of whites and thus decreasing white-brown inequality compared with that based on interviewer classification. On the other hand, the inclusion of self-classified browns who were interviewer-classified as white reduces the average income of browns based on self-classification; the degree of the effect (–12%) and the size of the cell (8.8% of all self-classified browns), however, are substantially outweighed by the increase.

Interviewers classified a similar number (247) of self-classified whites as brown, and the average income of these

9. The exceptions are the two black-white categories that have exceptionally small cell counts, leading to possibly unstable values.

TABLE 5. PERCENTAGE DIFFERENCE IN INCOME FOR ALL COMBINATIONS OF SELF-CLASSIFICATION AND INTERVIEWER CLASSIFICATION, COMPARED WITH CONSISTENTLY CLASSIFIED BROWNS

	Interviewer Classification		
Self-Classification	White (%)	Brown (%)	Black (%)
White (%)	25 (1,981)	4 (247)	–40 (9)
Brown (%)	26 (248)	— (872)	–12 (108)
Black (%)	16 (12)	0 (213)	–13 (310)

Notes: Total N = 4,000; N's for individual cells are shown in parentheses.

persons was only 4% greater than that of consistently classified browns. These persons have an average income well below that of consistently classified whites (25%). Thus, their inclusion in the self-classified white category decreases the average income of whites, further decreasing white-brown inequality when self-classification rather than interviewer classification is used.

Similarly, interviewers tended to classify better-off self-classified blacks as brown and worse-off self-classified browns as black. This leads to greater white-black inequality with interviewer classification. Brown-black inequality remains roughly the same, however, because interviewer classification deflates the incomes of blacks and browns to similar degrees.

SUMMARY AND DISCUSSION

We find that the estimate of white-nonwhite income inequality in Brazil is greater when interviewer classification is used than when self-classification is used. We believe that interviewer classification is a more appropriate method for determining racial inequality because the perceptions of others about one's race weigh more heavily than self-classification in determining labor market outcomes. Our results also demonstrate that interviewer-classified race explained racial differences in income much better than did self-classified race, further suggesting that interviewer classification is preferable for measuring racial discrimination. Thus, we believe that previous studies have underestimated racial inequality because they have relied solely on official statistics, in which race is based on self-classification, or on an unknown mix of self-classification and interviewer classification.

Our estimates provide some support for Degler's (1986) mulatto escape hatch theory, thus refuting Silva's (1985) well-known challenge and Wade's (1995) claim that the primary racial cleavage in countries like Brazil is between blacks and nonblacks. The income of browns, based on interviewer classification, is one third of the way between the incomes of blacks and whites and is closer to that of blacks. Clearly, the actual income of browns is much closer to the income of blacks than to the income of whites, but this is because of the cumulative disadvantages of nonwhites (which we controlled), especially low levels of education, and concentration in less-developed regions. Given the especially high income concentration and highly skewed returns to education found in Brazil (Lam and Levinson 1987), the large differences in actual income are not surprising. It is interesting that estimates based on self-classification show that the income of browns is closer than Degler expects to the halfway point between the income of whites and blacks.

We also find that persons who interviewers classified as brown but self-classified as white, and persons who interviewers classified as white but self-classified as brown, account for the greater white-brown inequality when interviewer-classification estimates are used. Interviewers classified as white 20% of all individuals who classified themselves as brown, perhaps because of their greater socioeconomic status. Similarly, interviewers darkened 11% of self-classified whites as brown, and these persons tended to have incomes similar to those of average browns. The evidence in this paper suggests a similar pattern for black-brown inconsistencies. We find support for a "money whitens" argument: Interviewers whiten those with higher status and darken those of lower status. The inconsistent cells may include many persons who can physically pass as either white or brown (or as brown or black), leaving interviewers to rely on their social status in classifying race.

We certainly need to be cautious about generalizing our findings to the entire urban Brazilian population. The direction of the change in our sample, however, is likely to be similar for the entire population, although the magnitude of the change may not be the same. Furthermore, ours is the only evidence so far that is based on a national sample showing how such changes in racial classification might affect inequality in Brazil. Previous findings on the subject were based on ethnographical studies of single towns. These towns can represent only one region in this regionally heterogeneous country, and towns compose only a small proportion of the national population.

The implications of these findings may extend beyond Brazil to other Latin American countries and, to a lesser extent, to countries outside the region such as the United States and South Africa. Race is similarly ambiguous throughout Latin America (Graham 1991) and in some countries of the region where racial dynamics might be quite similar to those in Brazil. Although the Brazilian case is sometimes considered exceptional, ambiguity and subjectivity about racial classification are likely to increase in places with stricter classification systems, like the United States and South Africa, as legal definitions of race fade further into the past. Continued miscegenation and immigration from countries like Brazil are also likely to promote greater racial ambiguity.

Finally, our results demonstrate that racial characteristics are not incontrovertible or objective facts but are often ambiguous, subject to variation according to the classifier, and affected by nonphysical criteria. Analysts of data with such variables should acknowledge their subjectivity. A finding that race is categorized inconsistently need not indicate response error bias, but instead that racial classification is subject to differences in social perception. These differences in perceptions of race are important because categorizing persons and treating them accordingly often has harmful consequences for individuals.

REFERENCES

Associação Brasileira de Estudos Populacionais. 1996. *Diversidades Brasileiras: Um Olhor Demografico* (Brazilian Diversity: A Demographic View). Information poster. Brasilia: United Nations Fund for Population Activities.

Barros, R.P. de, R. Mendonça, and T.M. Velazco. 1996. "O Papel da Cor no Processo de Estratificação Social Brasileiro" (The Role of Color in Social Stratification in Brazil). Unpublished manuscript, Institute de Pesquisas Economicas Aplicadas, Rio de Janeiro.

Breen, R. 1996. *Regression Models: Censored, Sample Selected, or Truncated Data*. Thousand Oaks, CA: Sage Publications.

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

Degler, C.N. 1986 [1971]. *Neither Black nor White: Slavery and Race Relations in Brazil and the United States*. Madison, WI: University of Wisconsin Press.

Fernandes, F. 1965. *A Integração do Negro na Sociedade de Classes* (Integration of Blacks in Class Society). São Paulo: Dominus.

Goyer, D. and E. Domschke. 1983. *The Handbook of National Population Censuses: Latin America and the Caribbean, North American and Oceania*. Westport, CT: Greenwood Press.

Graham, L., ed. 1991. *The Idea of Race in Latin America*. Austin, TX: University of Texas.

Green, W.H. 1990. *Econometric Analysis*. New York: Macmillan.

Hahn, R.A., J. Mulinare, and S. M. Teutsch. 1992. "Inconsistencies in Coding of Race and Ethnicity between Birth and Death in U.S. Infants" *Journal of the American Medical Association* 267:259–63.

Harris, M. 1964. "Racial Identity in Brazil." *Luso-Brazilian Review* 1:21–28.

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

Lam, D. and D. Levinson. 1987. "Age, Experience and Schooling: Decomposing Earnings Inequality in the U.S. and Brazil." Unpublished manuscript, Population Studies Center, University of Michigan.

Lovell, P.A. 1989. *Racial Inequality and the Brazilian Labor Market*. Unpublished doctoral dissertation, Department of Sociology, University of Florida.

Merrick, T.W. and R. Graham. 1979. *Population and Economic Development in Brazil: 1800 to the Present*. Baltimore: Johns Hopkins University.

Pierson, D. 1942. *Negroes in Brazil: A Study of Race Contact at Bahia*. Chicago: University of Chicago Press.

Pinto, R.P. 1996. "Classifying the Brazilian Population by Color: Underlying Problems." Pp. 189–214 in *Brazilian Issues on Education, Gender and Race*, edited by E.S. de Sá Barreto and D.M.L. Zibas. São Paulo: The Carlos Chagas Foundation.

Ransford, H.E. 1970. "Skin Color, Life Chances and Anti-White Attitude." *Social Problems* 18:164–78.

Ribeiro, F.R. 1996. "Diversity and Assimilation in Brazil." Pp. 125–29 in *Now That We Are Free: Coloured Communities in a Democratic South Africa*, edited by W. James, D. Caliguire, and K. Cullinan. London: Lynne Rienner Publishers.

Shryock, H.S. and J.S. Siegel. 1976. *The Methods and Materials of Demography*. New York: Academic Press.

Silva, N. do Valle. 1985 "Updating the Cost of Not Being White in Brazil." Pp. 42–55 in *Race, Class and Power in Brazil*, edited by P.-M. Fontaine. Los Angeles: UCLA Center for Afro-American Studies.

Stata Corp. 1997a. *Stata Reference Manual Release 5 Vol. 1 A–F*. College Station, TX: Stata Corp.

———. 1997b. *Stata Reference Manual Release 5 Vol. 3 P–Z*. College Station, TX: Stata Corp.

Stewart, M. 1983. "On Least Squares Estimation When the Dependent Variable Is Grouped." *Review of Economic Studies* 50:141–49.

Telles, E.E. 1996. "Self vs. Social Classification of Race: Inconsistency, Category, Ambiguity and Affirmation in Brazil." Unpublished manuscript, The Ford Foundation.

———. Forthcoming. "Ethnic Boundaries and Political Mobilization among African Brazilians: Comparisons With the U.S. Case." In *Racial Politics in Brazil*, edited by M.G. Hanchard. Durham, NC: Duke University Press.

Twine, F.W. 1995. "Brown Skinned White Girls: Racial Classification among Multiracial Surburban High Schoolers." Unpubished manuscript, Department of Sociology, University of California, Santa Barbara.

Wade, P. 1995. *Blackness and Race Mixture: The Dynamics of Racial Identity in Colombia*. Baltimore: Johns Hopkins University.

Wagley, C. 1968. *Race and Class in Rural Brazil*. Paris: UNESCO.

———. 1969. "From Caste to Class in North Brazil." Pp. 60–75 in *Comparative Perspectives in Race Relations*, edited by M. Tuman. Boston: Little, Brown and Company.

Wood, C.H. 1991. "Categorias Censitárias e Classificações Subjetivas de Raça no Brasil" (Census Categories and Subjective Classifications of Race in Brazil). Pp. 93–111 in *Desigualdade Racial no Brasil Contemporâneo*, edited by P.A. Lovell. Belo Horizonte: UFMG/CEDEPLAR.