Research Article

Does race response shift impact racial inequality?

Jerônimo O. Muniz

Stanley R. Bailey


This open-access work is published under the terms of the Creative Commons Attribution 3.0 Germany (CC BY 3.0 DE), which permits use, reproduction, and distribution in any medium, provided the original author(s) and source are given credit.
See https://creativecommons.org/licenses/by/3.0/de/legalcode.
Contents

1 Introduction 936
2 Racial fluidity and response shift 937
3 Racial inequality and response shift 938
4 Data 939
5 Measures and methods 940
5.1 Racial classification measure 940
5.2 Income measure 941
5.3 Racial inequality measure: ratio of earnings 941
5.4 Analytic approach 942
6 Results 944
6.1 Racial earnings inequality in Brazil at time t 944
6.2 Racial response shift between times t and t+9 948
6.3 Racial response shift and estimates of earnings inequality 948
6.4 Explaining stagnant inequality estimates amid shifting population composition 949
6.5 Observed and simulated racial response trends and their impact on inequality estimates 952
7 Discussion and conclusion 956
8 Acknowledgments 958
References 959
Does race response shift impact racial inequality?

Jerônimo O. Muniz¹
Stanley R. Bailey²

Abstract

BACKGROUND
Previous research posits that racial reclassification, or response shift, may confound measures of racial earnings inequality. However, this claim has not been systematically tested.

OBJECTIVE
We measure racial response shift in Brazil and examine its impact on white-to-nonwhite earnings inequality between survey waves over ten years at nine-month intervals.

METHODS
We use individual-level linked data from the 2002–2012 Monthly Employment Survey, involving Brazil’s six largest metropolitan areas (n = 400,046). We describe the level and pattern of racial reclassification across time and by income rank. We then decompose racial inequality into two components (income and population ratios) to examine the impact of racial response shift on estimates of racial inequality and to construct analytic counterfactuals.

RESULTS
Results reveal that 16% of our sample shifted their racial responses between survey waves. Nonetheless, we show that this level of response shift had no substantial impact on estimates of income inequality. We explain the counterintuitive results by demonstrating how bidirectional racial classification flows – lightening and darkening – countervail each other due to their similar income profiles and racial reclassification rates.

CONTRIBUTION
We offer a first empirical analysis of how racial response shift impacts estimates of racial earnings inequality using individual-level linked data from a large-sample survey.

¹ Corresponding author. Department of Sociology, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil. Email: jeronimo@ufmg.br.
² Department of Sociology, University of California, Irvine, CA, United States.
1. Introduction

The persistence of racial income inequality in the United States during the last 60 years, despite progressive intervention, is a central paradox in the field of social stratification (Morris and Western 1999; Reskin 2012; Bloome 2014). Scholars endeavor to understand how various established economic and compositional factors combine and interact to explain stagnant estimates across dynamic contexts (Wilson 2009; Bayer and Charles 2018). Moreover, recent research suggests that a neglected compositional factor, racial response shift, merits serious scrutiny for its possible role in complicating racial inequality estimates. Using individual-level linked census data, Liebler and her colleagues demonstrated that fully 6% of the US population changed racial responses in 2010 compared to 2000 (Liebler, Bhaskar, and Porter 2016; Liebler and Ortyl 2014; Liebler et al. 2017). Though examining only reclassification, these scholars challenge stratification scholars to examine the possible impact of racial response shift on estimates of racial inequality. Indeed, existing scholarship using alternative longitudinal and quasi-experimental designs suggests the impact of racial reclassification on racial disparities (Saperstein and Gullickson 2013; Saperstein and Penner 2010, 2012; Villareal and Bailey 2020). We seek to contribute to disentangling the forces behind complexing racial inequality gaps by examining the impact of racial response shift on estimates of racial earnings inequality in Brazil.

Hypothetically, how might racial response shift impact estimates of racial earnings inequality? Using the example of educational disparity, Liebler et al. (2017: 261) explain: “If a highly educated person changes her [racial] response from X to Y [between times \( t \) and \( t+1 \)], then group Y’s mean education rises and group X’s falls.” Hence a joiners/leavers dynamic may impact inequality estimates at time \( t+1 \) compared to time \( t \) (Liebler, Bhaskar, and Porter 2016). Regarding earnings inequality, it follows that individuals who shift their racial response between two survey periods carry with them not only their ‘race’ but also their income, thereby producing simultaneous, cross-category effects on the composition and mean income of joining and leaving subpopulations.

Despite hypothetical expectations, no research demonstrates the impact of response change on estimates of racial income inequality using individual-linked data. Brazil is a strategic research site in which to examine this question due to its noted combination of salient racial fluidity and persistent racial inequality (Muniz and Bastos 2017; Telles 2004). We analyze rare, individual-level linked data across two survey waves at multiple nine-month intervals between 2002 and 2012 (\( n = 400,046 \)). Our results reveal substantial racial response shift (fully 16% of our sample) in Brazil’s six largest metropolitan areas. Counterintuitively, this level of response shift had no meaningful impact on estimates of racial income inequality across survey waves. We explain these surprising results by
showing that bidirectional racial response shifts – lightening and darkening – symmetrically countervail each other during the period under study in Brazil (2002–2012), offsetting their impact on earnings inequality.

2. Racial fluidity and response shift

Core constructivist principles maintain that racial and ethnic categories are not fixed but fluid and can vary across time and place (Omi and Winant 1994; Nagel 1994; Alba 2020; Alba and Islam 2019; Dahis, Nix, and Qian 2019; Eschbach, Khalil, and Snipp 1998). The case of Brazil is often recognized in the literature as illustrative of racial fluidity (Winant 1992). Racial identification in Brazil is constructed from a dynamic combination of perceived behavioral, social, and physical traits (skin color, facial structure, the shape of the nose and lips, and hair texture). Scholars frequently contrast Brazil’s racial identification fluidity to the rigidity of racial identification in the United States, where perceived ancestry is central and may produce sharper boundary distinctions (Nobles 2000; Nogueira 1985; Telles 2004; Daniel 2006). The seemingly more dynamic process of inferring ‘race’ in Brazil leads to presumed higher levels of ambiguity, fluidity, and subjectivity relative to the United States (Davis 1991; Skidmore 1993).

Racial identification in everyday Brazilian life further hinges on individuals’ family, neighborhood, and broader social interactions and may be less salient in specific contexts or activities than in others (Moraes Silva and Leão 2012; Sansone 2003). This situation-contingent character in the lived racial dynamics in Brazil exemplifies the multidimensionality of ‘race’ (Roth 2016). As Roth explains, in everyday life and across contexts, individuals “experience ‘race’ not as a single, consistent identity but as a number of conflicting dimensions” (2016: 1310). The effect of multidimensionality may appear as racial inconsistency across classification formats (or dimensions), between self- versus interviewer classification, and in open- versus closed-format questions (Bailey, Loveman, and Muniz 2013; Loveman, Bailey, and Muniz 2012; Muniz 2010, 2012, 2016; Telles and Lim 1998).

Notably, the fluid and multidimensional character of racial classification in Brazil produces variation in the racial composition of populations across time, as Liebler et al. (2017) demonstrated in the United States. For example, research using projection methods on Brazilian census data provides evidence of racial composition shift across time above and beyond the demographic effects of birth, death, and immigration (Carvalho, Wood, and Andrade 2004; Miranda 2015; Wood 1991). Moreover, using linked data from the Monthly Employment Survey (Pesquisa Mensal de Emprego, or PME), Miranda (2014) shows that 23% of respondents shifted their racial classification across waves. Of the self-reported brown (parda) category, 20% reclassified as white
Of the self-reported black category, 31% shifted to brown and 8% to white. And from the white category, 14.6% reclassified as brown and 1.4% as black. Senkevics (2017) also reports that 22% of those who enrolled in the National High School Exam (ENEM) revised their racial self-classification between 2010 and 2014. Lastly, from 2008 to 2015, Silveira (2019) reports that 11.5% of individuals in the Annual Social Information Survey (RAIS) were racially reclassified. Overall, this evidence suggests racial response inconsistency among at least one-fifth of surveyed populations in Brazil, potentially remaking the population’s racial composition over time through response change.

Regarding the direction of racial response flows in Brazil, scholarship has long noted the dominance of whitening dynamics (Hasenbalg 2005; Schwartzman 2007). Nonetheless, more recent research suggests that whitening may be waning and that the increased salience of movements and discourses of black affirmation, or negritude, favors racial darkening (Marteleto 2012; Miranda 2014, 2015; Telles and Paschel 2014). This darkening response shift may be most salient among younger and more educated individuals, as well as those subject to the impact of wide-ranging race-targeted affirmative action beginning in the early 2000s (Francis-Tan and Tannuri-Pianto 2015; Marteleto 2012; Bailey and Fialho 2018). In sum, the increasing destigmatization and valuation of blackness through racially affirmative discourse, social movements, and public policy distinguish 21st-century Brazil’s darkening dynamics from the whitening dynamics of the 20th century (Bailey, Fialho, and Loveman 2018; Marteleto 2012; Miranda 2015; Bailey 2009).

3. Racial inequality and response shift

Brazil is equally known for its persistent racial inequality and its racial fluidity. In this way, deep racial hierarchy is a core characteristic of both the United States and Brazil (e.g., Telles 2004). Regardless of the measure – Gini, Theil, dissimilarity index, or ratio – and irrespective of the outcome of interest (life expectancy, education, occupational prestige, or income), the white population in Brazil consistently has more resources and occupies higher hierarchical positions than the nonwhite population (Chaderevian 2011; Hasenbalg 2005; Henriques 2001; Marteleto 2012; Osório 2009; Ribeiro 2006; Soares 2000). In addition, though the Brazilian census has long used an intermediate (or mixed race) category, scholarship has shown that the racial inequality divide may be best conceptualized as binary, constituted by a white versus nonwhite cleavage (Hasenbalg 1985; Marteleto 2012; Silva 1985; Telles 2004).

Research often links racial response shift to socioeconomic status in racially unequal though fluid settings. Even in the constrained US context, research suggests an
endogenous relationship between social mobility and racial classification (Saperstein and Gullickson 2013; Saperstein and Penner 2012). This relationship may be more salient in Brazil. When Degler (1971) coined the term ‘mulatto escape hatch,’ he argued that upward social mobility could alter perceived racial lines in the direction of whiteness. Telles and Lim (1998: 473) provide a unique empirical demonstration of Degler’s thesis by showing that interviewers in a survey setting often “whiten those with higher status and darken those of lower status.” From another empirical approximation regarding racial whitening’s connection to social status, Schwartzman (2007) reports that higher social status raises the likelihood that nonwhite parents will classify their children as white and lowers the likelihood that white parents will classify their children as nonwhite. In sum, regarding the hypothetical effect of racial whitening on racial income inequality, Bailey, Loveman, and Muniz (2013: 109) posit, “To the extent that racial mobility is occurring, it likely contributes to overestimates of racial inequality in income because higher-income individuals tend to be reclassified from darker categories to lighter ones.”

We focus our analytic lens on both Brazil’s racial fluidity and its equally well-documented racial inequality. Significantly, though, we go beyond the ordinary course to a research design that permits a disentangling of the effects of Brazil’s racial fluidity on estimates of racial inequality.

4. Data

To examine the impact of racial response change on racial earnings inequality in Brazil, we use linked individual-level data from the PME collected by the Brazilian Institute of Geography and Statistics (IBGE 2007). The PME was fielded between March 2002 and December 2012 in the six largest metropolitan areas of the country (São Paulo, Belo Horizonte, Rio de Janeiro, Porto Alegre, Recife, and Salvador), which account for about 25% of the national population. The PME longitudinal design follows the Current Population Survey conducted by the US Census Bureau and the US Bureau of Labor Statistics. Sampled households are in the survey for four months, out for eight, and then back for another four months. For example, some individuals interviewed in March, April, May, or June of a given year were questioned again in December, January, February, or March (Ribas and Soares 2009). Unique identifiers allowed the linkage of

---

3 We retrieved PME linked data using datazoom_social, a Stata program developed by the Pontifical Catholic University of Rio de Janeiro, to extract original microdata from household surveys produced by IBGE (Data Zoom 2022).

4 The generalizability of all results is constrained by the urban geography of our sample, as represented by the data’s six major metropolitan areas; results may vary within and between geographical areas of the country, including rural areas.
households over consecutive survey waves, and data on sex, date of birth, and years of education allowed the identification of the same person across waves (Miranda 2014; Queiroz 2007; Ribas and Soares 2009).\(^5\) We focus on survey waves 4 (time \(t\)) and 5 (time \(t+9\)), which were nine months apart but distributed across a ten-year survey period (2002–2012).\(^6\) Our pooled analytical sample contains 400,046 individuals. All analyses use complex survey weights to account for the PME sampling design.

5. Measures and methods

The measurement of racial earnings inequality rests on a core duality: the relative distribution of income mass and individuals between subpopulations (Lam 1986; Lam and Levison 1992; Robinson 1976). Conclusive arguments about what drives inequality across time must thus disentangle the relative allocation of income via economic factors from the relative distribution of persons among subpopulations via compositional characteristics. Antidiscrimination and income redistribution policies, for example, impact the income mass of specific subpopulations, while fertility, mortality, and migration affect their relative size (Bloome 2014; Chu and Jiang 1997). Our analytic approach, however, examines a neglected force possibly impacting relative population size and hence estimates of racial inequality: racial reclassification, or response change (Liebler et al. 2017; Saperstein and Gullickson 2013; Saperstein and Penner 2012).

5.1 Racial classification measure

The IBGE, Brazil’s official statistics bureau, tasked with the census and the PME, asks individuals to self-classify by raça ou cor (race or color) into one of five categories. More than 99% of our sample classified as either white (branca), brown (parda), or black (preta), and we excluded Asian (amarela) and indigenous (indígena) origin respondents (0.62% of the sample). One person (usually the head of the household) responded to the questionnaire for household members. Reported racial category classification, or racial

\(^5\) Because the PME is an unbalanced panel, 25% of the sample is replaced between consecutive waves while 75% overlaps from one month to another. In the data under scrutiny, the average attrition rate was 35%. Similar mean values among those who left, stayed, or entered the sample between rounds 4 and 5 provide evidence against sample selection bias (see, however, Ribas and Soares 2010). In addition, we investigated the sensitivity and robustness of our conclusions to missing data using multivariate imputation by chained equations (MICE). A systematic comparison between imputed and observed analytic samples provided similar results regarding patterns of racial reclassification and inequality. Analysis available upon request.

\(^6\) See supplementary Table S1 for further information on the composition and structure of wave 4 and 5 interviews.
response, is thus a combination of self-classification and other classification typical in national censuses.

We focus on a core binary categorical divide in Brazil: branca and não-branca – henceforward white and nonwhite – by collapsing parda and preta categories into a single nonwhite statistical grouping (see Marteleto 2012). This operationalization is supported in the lion’s share of research that documents similar socioeconomic profiles for the parda and preta population categories (Hasenbalg 1985; Marteleto 2012; Silva 1985; Telles 2004). Moreover, Brazilian black (negro) movements, state affirmative action policy, and numerous area scholars embrace a binary division regarding the three census categories we engage (Machado, Eurístenes and Feres Júnior 2017; Paschel 2016). This approach has the additional benefit of cognitive consonance with the white versus black divide in the United States. Nonetheless, we provide supplementary analyses that maintain the ternary distinction – branca, parda, and preta.

5.2 Income measure

Our income measure is mean hourly wage (Gould 2020). We exclude individuals missing income information and those with extreme earnings. We standardize employment wages from all jobs by the number of hours worked in a month and adjust for inflation to express real values in August 2014 (Courseuil and Foguel 2002). We also use a measure of income rank operationalized as an individual’s position in the income distribution in 20 quantiles (Bayer and Charles 2018; Manduca 2018).

5.3 Racial inequality measure: ratio of earnings

No measure of racial income inequality satisfies all the desirable properties of a sufficient indicator. The choice of one metric over another frequently involves trade-offs. We use the ratio of mean hourly income between racial categories, which we lay out in Equation (1). We choose this measure over several other metrics (Gini, Theil, Hoover, and Atkinson indexes; coefficient of variation; general entropy measures) for practical and substantive reasons. The ratio of mean income is practical because it requires information

---

7 Sensitivity analysis shows that utilizing the median or the mean income (after excluding extreme values) does not influence the results.
8 Fully 42% of the linked sample reported wages and the number of hours worked in a month. The detection of outliers is based on the Mahalanobis distance metric with a 0.15 percentile criterion, which marks about 1% of respondents as extreme values (Weber 2010).
9 The use of racial population–specific income distributions provided similar results.
on only individuals’ racial classification and reported earnings, through which we obtain information on population size and income mass. This measure also possesses an intuitive meaning (that is, one population earns twice or \(x\%) more than the other) and is of substantive importance for capturing the overall economic distance between racial populations. Moreover, the chosen inequality measure facilitates decomposition and the construction of informative counterfactual scenarios, which we use to disentangle the relative impact of racial response change and income redistribution on racial earnings inequality.

The ratio of mean hourly income between whites and nonwhites decomposes into two multiplicative components: (1) the income ratio of whites to nonwhites and (2) the population ratio of nonwhites to whites. Mathematically, it is possible to express these two components at time \(t\) as:

\[
\text{Racial inequality at time } t = \frac{\text{Whites' hourly mean income at time } t}{\text{Nonwhites' hourly mean income at time } t}
\]

\[
= \frac{\sum_{i} y^{i,t}_w}{n^t_w} = \frac{\sum_{i} y^{i,t}_w}{\sum_{i} y^{i,t}_{nw}} \times \frac{n^t_{nw}}{n^t_w},
\]

where \(y\) represents the remuneration of an individual \(i\) at time \(t\), \(n\) is the absolute population size, and the subscripts indicate if reported racial response is white (\(w\)) or nonwhite (\(nw\)). The first component, \(\frac{\sum_{i} y^{i,t}_w}{\sum_{i} y^{i,t}_{nw}}\), shows how large the amount of income held by whites is in comparison to the sum of nonwhites’ earnings. The second component, \(\frac{n^t_{nw}}{n^t_w}\), describes the population ratio of nonwhites to whites. Equation (1) demonstrates that racial income inequality is a combination of the relative distribution of income and individuals between subcategories of a population.

### 5.4 Analytic approach

It follows that racial earnings inequality varies between two points in time whenever there is a change in relative racial income mass or racial population size. Previous research on
compositional effects (that is, those affecting relative population size) demonstrates, for example, how differential migration and fertility can impact inequality estimates (Lam 1986; Muniz 2012). These two compositional factors produce their effect on relative population size by adding new ‘exogenous’ members to existing subpopulations. We examine something novel: how a compositional impact wrought by racial response change might cause variation in inequality estimates. A critical difference, then, in our research’s question and design: While the impact of differential migration and fertility results from adding new members to existing subpopulations, we examine an internal dynamic resulting from switching between existing subpopulations.

Liebler and colleagues (2016, 2017) conceptualize this dynamic as involving exiting and joining dimensions. That is, the reallocation of people and income mass across time through shifts in racial response simultaneously impacts the size and income mass of both the origin population (that individuals exit) and the destination population (that they join). For example, if nonwhite respondents (at time $t$) change their racial response to white nine months later, at time $t+9$, they carry with them their ‘race’ and their income; hence these shifts not only impact the size and income mass of the nonwhite population they exited but also automatically impact the size and income mass of the white population they joined.

To flesh out the effects of racial response shift on the measurement of racial earnings inequality across two periods of time, our analysis proceeds in five core stages:

1. We begin with a cross-sectional lens on racial earnings inequality. We use Equation (1) to estimate observed inequality at time $t$ (wave 4). In addition, we present a novel visualization of the dynamic relationship of changing income mass and population size on ratios of earnings inequality at time $t$ by survey year. Lastly, we use counterfactuals to gauge the relative weight of changing income mass and population size on our cross-sectional time $t$ estimates by survey year.

2. We then turn to a longitudinal design. We estimate the observed percentage of the population that changes their racial response between survey waves 4 (time $t$) and 5 (time $t+9$). Response shift can be in two directions in our research design. On the one hand, white respondents at time $t$ might change their racial response to nonwhite at time $t+9$. We refer to this recategorization shift as darkening. On the other hand, nonwhite respondents at time $t$ might change their racial response to white at time $t+9$, which we label lightening.

3. Next we calculate the observed racial earnings inequality at time $t+9$ to compare with observed racial earnings inequality at time $t$. The impact of racial response change – lightening and darkening – on estimates of racial income inequality is evidenced by the difference between observed inequality at time $t$ and observed inequality at time $t+9$ (Bayer and Charles 2018; Manduca 2018).
4. To explain the results of our analysis regarding the impact of racial response shift on estimates of racial earnings inequality, we examine racial response shift – lightening and darkening – by income and along the income distribution.

5. Lastly we explore overall racial classification trends – lightening, darkening, and stability – in observed and simulated scenarios and their relationship to racial earnings inequality in Brazil.

6. Results

6.1 Racial earnings inequality in Brazil at time $t$

We begin with a benchmark, cross-sectional estimation of observed racial earnings inequality in Brazil at time $t$ (survey wave 4). We use Equation (1) and data on racial response (at time $t$) and income mass (at time $t$) aggregated from the survey’s ten-year period (see time $t$ row marginals in Table 1). Hence

$$
\frac{y_{11} + y_{12}}{y_{22} + y_{21}} \times \frac{n_{22} + n_{21}}{n_{11} + n_{12}} = \frac{122,007,856}{51,351,687} \times \frac{7,879,764}{10,577,740} = 1.77,
$$

where the terms in this equation now represent aggregated measures of income ($y$) and individuals ($n$) and the subscripts index the rows and columns of cells in Table 1. The result from Equation (2) shows that whites had hourly earnings 77% higher than nonwhites at time $t$ – a racial inequality of 1.77.

<table>
<thead>
<tr>
<th>Time $t$ plus nine months</th>
<th>Whites</th>
<th>Nonwhites</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time $t$</td>
<td>$9,113,854$</td>
<td>$1,463,886$</td>
<td>$10,577,740$</td>
</tr>
<tr>
<td></td>
<td>($110,805,056$)</td>
<td>($11,202,800$)</td>
<td>($122,007,856$)</td>
</tr>
<tr>
<td>Nonwhites</td>
<td>$1,492,063$</td>
<td>$6,387,701$</td>
<td>$7,879,764$</td>
</tr>
<tr>
<td></td>
<td>($11,637,187$)</td>
<td>($39,714,500$)</td>
<td>($51,351,687$)</td>
</tr>
<tr>
<td>Total</td>
<td>$10,605,917$</td>
<td>$7,851,587$</td>
<td>$18,457,504$</td>
</tr>
<tr>
<td></td>
<td>($122,442,244$)</td>
<td>($50,917,300$)</td>
<td>($173,359,544$)</td>
</tr>
</tbody>
</table>

Note: Sums of mean individual incomes are in parentheses. These figures refer to aggregate flows of racial reclassification, 2002 to 2012.

For an alternative, cross-sectional view of racial earnings inequality in Brazil, we disaggregate time $t$ data by survey year (as opposed to aggregating data from all time $t$...
interviews across the ten-year period). These findings suggest that racial inequality, though substantial, decreased in Brazil during the survey’s ten-year period (Figure 1, solid-dotted line). For example, in 2002 the observed ratio of white-to-nonwhite earnings was 1.9; in 2012 that ratio was 1.6. These observed ratios represent cross-sectional snapshots, and our finding of decreasing inequality during this period is supported by the existing research literature (Augusto, Roselino, and Ferro 2015; Paiva 2016). Interestingly, Brazil’s decreasing racial income inequality contrasts with that of the United States; in the latter context, black-to-white earnings inequality increased during the same period (Akee, Jones and Porter 2019; Bayer and Charles 2018; Gould 2020; Manduca 2018).

Figure 1: Ratios of white-to-nonwhite earnings inequality, income mass, and population size, 2002–2012

<table>
<thead>
<tr>
<th>Year</th>
<th>Income share ratio</th>
<th>Population share ratio</th>
<th>Observed racial inequality at the baseline interview</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Income share ratio refers to the first component of Equation (1). Population share ratio refers to the second component of that equation. Both ratios refer to values at the baseline interview, at time t.

Figure 1 also offers a novel visualization of the dynamic impact of changing income mass and population size on estimates of racial income. The top (solid) line plots income mass ratios by survey year. The bottom (dotted) line plots population size ratios by survey year. The middle (solid-dotted) line plots earnings inequality ratios by survey year, as mentioned above. The downward slope of inequality ratios results from the incremental
convergence of relative income mass and population sizes across the ten-year period. Both the decrease in racial inequality and the increase in the nonwhite-to-white population share that our analysis reveals are supported in the literature (Athias 2018; IBGE 2019; Jesus and Hoffmann 2020; Soares 2008a, 2008b).

The period’s decrease in racial earnings inequality, however, was not even across the earnings distribution. To illustrate, in Figure 2 we present estimates of racial income inequality across survey years by 20 quantiles of the earnings distribution – that is, income rank. We show that white to nonwhite inequality varies substantially by income rank. For example, whites in the lowest 10% (q2) of the income distribution earned 40% more than nonwhites in 2002, whereas whites in the highest 10% of the income distribution (q19) earned hourly wages about 130% higher than nonwhites. In 2012, in contrast, whites at the bottom of the income distribution (q2) had hourly wages only 20% higher than nonwhites, and those at the highest 10% had wages 90% higher than nonwhites at the upper-income quantile (q19). In sum, overall racial inequality decreases across survey years while it increases with upward wage mobility, in tandem with existing literature (Bailey, Loveman, and Muniz 2013).

Figure 2: Ratio of mean incomes at the baseline interview across quantiles of the earnings distribution, 2002–2012

Source: Based on 2002–2012 PME.
Lastly, we use counterfactuals with our cross-sectional data to disentangle the relative importance of economic (income mass) and demographic or compositional (racial population size) components in producing decreasing income inequality across this ten-year survey period in Brazil (Figure 3). We present income inequality by survey year from two simulated (unobserved) contexts alongside observed inequality at time \( t \). The top (black) line simulates what the racial inequality trend would be if income mass were held constant at its 2002 level but population size was allowed to shift across years. This counterfactual scenario reveals that racial earnings disparity would have gradually increased across time. The bottom (dotted) line represents hypothetical inequality levels if the population ratio of nonwhites to whites had remained as in 2002 but income mass was allowed to shift across years. The trend line shows that inequality would have declined significantly. The fact that the top simulation (which negates income variations but allows population shifts) results in inequality in the opposite trend direction of what was indeed observed (as plotted in the solid-dotted line) indicates that the income component outweighed the population component in driving the white/nonwhite earnings ratio.

**Figure 3: Counterfactual simulations of racial inequality, 2002–2012**

*Note:* Simulations are based on the income and population components introduced in Equation (1). The solid black line keeps the income share ratio as observed in 2002 while varying the population share ratio over time, as observed at the baseline interview. The dotted gray line does the opposite: It keeps the population share ratio at its 2002 values while varying the income share ratio over time.
6.2 Racial response shift between times $t$ and $t+9$

We now turn to a longitudinal perspective to examine racial reclassification – response shift – at wave 5 (time $t+9$) relative to wave 4 (time $t$). Our cross-tabulation of racial responses at time $t$ and time $t+9$ shows that 16% of our sample changed their racial response across the two waves (Table 1). Through a comparative lens using the US example, Liebler et al. (2017) find that ‘only’ 3% of non-Hispanic whites and 6% of non-Hispanic blacks changed their racial response between 2000 and 2010. Importantly, then, the substantially larger size of racial response shift in Brazil compared to the United States clearly lends support to our empirical question of whether racial response shift could be consequential for measures of income inequality.

In terms of the direction or flow of racial response shift, what did we expect and what did we find? Research on Brazilian racial identification dynamics in the 20th century assumed the dominance of racial lightening (reclassification from nonwhite to white) (Degler 1971; Schwartzman 2007). In contrast, contemporary research has begun to explore the growing salience of racial darkening in Brazil (Marteleto 2012; Telles and Paschel 2014). We found evidence of both flows in our analysis (Table 1). On the one hand, 7,879,764 individuals classified as nonwhite at time $t$, and 1,492,063 of these individuals changed their response to white (or lightened) at time $t+9$ (8.08% of the overall sample). On the other hand, 10,577,740 individuals classified as white in the fourth wave, and 1,463,886 reclassified as nonwhite (or darkened) at time $t+9$ (7.93% of the overall sample). We found, then, that racial response shift was bidirectional and almost identical in size.\(^\text{10}\)

6.3 Racial response shift and estimates of earnings inequality

The magnitude of racial response shift, in essence, reconfigures the membership composition of white and nonwhite populations in Brazil. The substantial compositional effect of racial response shift could thus impact the estimation of racial earnings inequality, our core research question. To engage the possibility of response shift’s compositional effect on racial earnings inequality, we use linked individual-level data and calculate observed earnings inequality at time $t+9$, again using Equation (1), as

$$
\frac{y_1 + y_2}{y_2 + y_1} \times \frac{n_{22} + n_{12}}{n_{11} + n_{21}} = \frac{122,442,244}{50,917,300} \times \frac{7,851,587}{10,605,917} = 1.78,
$$

\(^{10}\) See Table S2 for replication of the analysis presented in Table 1 but using the ternary classification scheme.
where the subscripts of income mass $y$ and population $n$ identify the cells in Table 1.

Equation (3) shows that the ratio of white to nonwhite earnings at time $t+9$ is 1.78. As mentioned, the impact of racial response change on estimates of racial income inequality across two periods of time is evidenced by the difference between observed inequality at time $t$ and observed inequality at time $t+1$ (Bayer and Charles 2018; Manduca 2018). Comparing the ratio at time $t+9$ to the previously calculated racial inequality at time $t$ through Equation (2) reveals that the difference between these estimates appears negligible (1.78 versus 1.77).

6.4 Explaining stagnant inequality estimates amid shifting population composition

Our results are clearly counterintuitive and even paradoxical: We find no substantial change in estimates of racial earnings inequality between time $t$ and time $t+9$, in aggregate and by survey year, despite substantial racial response shifts across the survey’s ten-year period. In essence, the racial response shifts produced the recomposition of the membership of white and nonwhite population segments, but our analysis suggested no substantial compositional effects. How can we explain this?

To engage this puzzle, we first explore the relationship of racial response shift to income. Figure 4 maps the conditional probability of lightening and darkening racial response flows at time $t+9$ by income rank across survey years (2002 versus 2012). The analysis shows that the probability of lightening is positively associated with income level in both 2002 and 2012; in contrast, the probability of darkening is negatively associated with income level during those two survey years. For example, the results show that the lower the income rank of a white respondent, the higher the probability of that respondent shifting to the nonwhite category (see Saperstein and Penner 2012). Although these results affirm the need to consider someone’s relative income position to better understand the process of racial response shift in relation to earnings inequality, they do not elucidate how racial response shift resulted in unchanging estimates of earning inequality in our study.11

11 Our findings do not speak to the issue of causality. For resolving that question, the influence of other factors (education, age, sex, region, and so on) would have to be controlled, and the temporal order of events—whether racial or income mobility comes first—would have to be precisely established. Moreover, the advantage of our research design is that it is able to uniquely isolate the effect of response shift because the time between survey waves (nine months) was too short for significant shifts in the types of capital that produce social mobility.
To better understand the persistence of consistent ratios of white-to-nonwhite earnings inequality across waves, despite substantial racial response shift, we present a visualization of mean hourly income by racial response type at time $t$ across survey years (see Figure 5). First, those who darken at time $t+9$ – shifting from white to nonwhite (w-nw) – have mean incomes higher than those in their (nonwhite) destination (shift-to) category, as represented by the nw-nw (black) line. In addition, those who lighten – shifting from nonwhite to white (nw-w) – have lower mean earnings than those of their (white) destination category, represented by the w-w (gray) line. One would expect, then, that these racial response shifts would lower the mean income (income mass) of the white population at time $t+9$ and raise that of the nonwhite population at time $t+9$. Counterintuitively, these racial response shifts did not have that impact because the exiting whites were replaced by entering (‘new’) whites, whose aggregate mean income was equal to that of exiting whites, and vice versa for exiting and entering (‘new’) nonwhites. That is, the two lines representing income shares of those who lighten and

---

**Figure 4:** Probabilities of racial lightening or darkening across quantiles of the income distribution in the six largest metropolitan areas of Brazil, 2002–2012

![Graph showing probabilities of racial lightening or darkening across quantiles of the income distribution.](image-url)
those who darken completely overlap. The impact of population reallocation via racial response shift on estimates of earnings inequality gaps is therefore neutralized by a unique double symmetry in the flow of income mass. That is, even with the suggested association between income and racial response (Figure 4), the mean income of the population of individuals who lighten is approximately the same as the mean income of those who darken. And these bidirectional flows countervail one another with regard to their impact on estimates of racial earnings inequality during this survey’s ten-year period.\textsuperscript{12}

\textbf{Figure 5:} \textit{Mean hourly income at time} $t$ \textit{by racial response across survey waves}

\begin{figure}[h]
\centering
\includegraphics[width=\linewidth]{figure5.png}
\caption{Mean hourly income at time $t$ by racial response across survey waves.}
\end{figure}

Note: The first letter or letters in the legend (w means white; nw means nonwhite) identify the racial response reported at time $t$, and the second letter or letters refer to the response reported at time $t$ plus nine months. 
Source: Based on 2002-2012 PME.

\cite{12} See Table S3 for descriptive statistics, including hourly income, by possible response types using the ternary classification scheme.
6.5 Observed and simulated racial response trends and their impact on inequality estimates

While our results reveal that the paradox of stagnant estimates results from countervailing flows of individuals and income mass, can our data speak to racial identification trends and their relationship to racial earnings inequality? We address this question using both observed and simulated racial trend scenarios. Beginning with observed data, we return to Figure 4, comparing response shift types – lightening and darkening – in 2002 and 2012 by quantiles of hourly earnings. Results reveal that, on the one hand, there was clearly a substantial decrease in racial response lightening in 2012 compared to 2002, regardless of income rank: The probability of nonwhites reclassifying as whites declined from between 20% and 50% in 2002 to between 10% and 30% in 2012. On the other hand, we also see a lower probability of darkening (white to nonwhite) in 2012 compared to 2002, but the trend is comparatively modest vis-à-vis the decreasing lightening trend.

In line with the decreasing trends in racial fluidity in Brazil via both lightening and darkening, our results also reveal its mirror image in increasing racial classification stability across the ten-year survey period (Figure 6). In 2002, about 70% of nonwhites in median income quantiles (q9–q11) reported the same racial response across waves. Ten years later, this figure had increased to 80%. The clearest illustration of the trend toward increasing racial classification stability among nonwhite Brazilians is the substantial and near-consistent distance between the decreased probability of lightening in 2002 compared to 2012. Our results also show an increase in the stability of whiteness over the same period, though the trend of racial identification stability is most pronounced for nonwhites.13 These classification trend dynamics are consistent with research that suggests the waning of racial classification fluidity in contemporary Brazil (Bailey, Fialho, and Loveman 2018; Miranda 2014).

13 In addition, the income rank of those most likely to maintain stable white and nonwhite racial responses is strikingly dissimilar: The probability of white response stability is highest at the top income quantile while the exact opposite is true for nonwhite response stability.
Figure 6: Probabilities of racial response stability across quantiles of income distribution in the six largest metropolitan areas of Brazil, 2002–2012

*Note:* Income quantiles map the relative income distribution for the entire population in each specific year but not for each reported racial response. Response-specific quantiles, however, provide similar trends.

*Source:* Based on 2002–2012 PME.

The more accelerated decline in the salience of racial lightening compared to the modest decline in darkening, though not impactful during the ten-year period under study, could lead to future reconfigurations of the composition of whites and nonwhites in Brazil that could hypothetically impact estimates of racial earnings inequality. To explore this possibility, we leverage two counterfactual scenarios.

**Scenario 1: lightening.** If the impact of racial response change is blunted due to bidirectional, countervailing shifts, what would happen if racial recategorization was unidirectional – that is, what would happen if only lightening was salient among our sample population? Such a scenario proxies 20th-century racial dynamics in Brazil, a historical period in which scholarship posited the hegemony of racial whitening and largely ignored racial darkening. We simulate racial lightening by reallocating the 8% of whites (at time t) who darkened (at time t+9) back to their original racial response (white at time t). At the same time, we allow nonwhites (at time t) to lighten their racial response to white (at time t+9). Under this lightening counterfactual, using Equation (1), income inequality becomes
\[
\frac{y_{11} + y_{21} + y_{12}}{y_{22}} \times \frac{n_{22}}{n_{11} + n_{21} + n_{12}} = 1.78,
\]

which does not differ significantly from the ratio of 1.77 at time \( t \) calculated in Equation (2). This result suggests, then, that there would be no effect of exclusively whitening response shift on the estimate of racial earnings inequality for the ten-year survey period in aggregate.

Nonetheless, could disaggregating by survey year reveal more nuance? Figure 7 compares observed ratios of racial earnings (at times \( t \) and \( t+9 \)) with simulated scenarios of racial response by survey year (2002 versus 2012) and confidence intervals. The results show that under the simulated conditions of scenario 1, whites would earn 91% more than nonwhites in 2002 and 63% more in 2012. At time \( t \), these figures are 88% and 62%, respectively. The overlapping confidence intervals of these data points suggest that the exclusive lightening response shift simulation did not substantially impact estimates of racial income inequality. This finding may suggest the need to reconsider past scholarly assumptions on the distorting effect of racial whitening on racial inequality in Brazil.
Muniz & Bailey: Does race response shift impact racial inequality?

**Figure 7:** Racial inequality in alternative racial response scenarios, 2002–2012

Note: Vertical error bars represent 95% confidence intervals, assuming Gaussian distribution and independence between white and nonwhite mean incomes (Fieller 1940).

**Scenario 2: darkening.** Alternatively, what if racial response shift was exclusively of the darkening type? This counterfactual is of particular interest to Brazil’s black (negro) movement because it may proxy the goal of racial consciousness-raising (Miranda 2015). The movement’s racial formation project seeks to affirm and expand black/negro/nonwhite identification among individuals of some perceived African ancestry. To simulate this scenario, we hold stable the pre-shift racial response of nonwhites (at time $t$) who lightened (at time $t+9$) while allowing whites (at time $t$) to darken (at time $t+9$). Based on data from Table 1, the income inequality resulting from this darkening counterfactual is
\[
\frac{y_{11}}{y_{22} + y_{21} + y_{12}} \times \frac{n_{22} + n_{21} + n_{12}}{n_{11}} \quad (5)
\]

\[
= \frac{110,805,056}{50,917,300 + 11,637,187} \times \frac{7,851,587 + 1,492,063}{9,113,854} = 1.82,
\]

which is slightly higher than the inequality ratio calculated in Equation (2) at time \( t \) (1.77). By survey year, inequality would have been 5% greater in 2002 and 3% higher in 2012 compared to that observed at time \( t \). While higher, these are not significant differences, as suggested by the overlapping confidence intervals in Figure 7. The results suggest that simulated hegemonic darkening also does not substantially impact estimates of racial income inequality.

These two counterfactual demonstrations further suggest a high threshold for compositional change via racial response shift to impact measures of racial earnings inequality. That is, our research suggests a certain imperviousness of the structure of white to nonwhite inequality in Brazil to internal exiting and entering dynamics or racial response shift (see Liebler, Bhaskar, and Porter 2016; Liebler et al. 2017).

7. Discussion and conclusion

We began this study by noting a core paradox in the field of social stratification – nearly unchanging estimates of racial inequality across time despite progressive intervention – and the ongoing endeavor of scholars to disentangle the combination of economic and compositional forces behind these estimates (Bayer and Charles 2018; Bloome 2014; Morris and Western 1999; Reskin 2012). Moreover, we reported that recent scholarship raises questions regarding an understudied compositional factor possibly impacting estimates of racial inequality across time: racial response shift (Liebler, Bhaskar, and Porter 2016; Liebler et al. 2017; Saperstein and Gullickson 2013; Saperstein and Penner 2012). In response, we engaged the case of Brazil to examine how racial response shift may impact estimates of white to nonwhite earnings gaps.

14 See Table S4 for information on the impact of racial response shift on measures of earnings inequality using \textit{branca}, \textit{parda}, and \textit{preta} categories. For example, results show a higher white/\textit{preto} ratio than white/\textit{pardo} racial inequality.
Our initial cross-sectional analysis of observed inequality at wave 4 (time $t$) revealed significant overall levels of income inequality between white and nonwhite populations in the country’s six largest metropolitan areas across the survey period (2002–2012). To gauge how racial response shift can impact the measurement of racial inequality in that context, we used individual-level linked data from survey wave 5 (time $t+9$), fielded nine months after the baseline wave 4 interviews (time $t$). We found that fully 16% of respondents shifted their racial response across waves. Counterintuitively, though, and in direct response to our research question, our comparative estimation of observed income inequality at times $t$ and $t+9$ showed no substantial response shift effects.

Our findings resonate with existing research on racial income inequality and racial reclassification dynamics in contemporary Brazil. Regarding racial classification and fluidity, scholars note the varying salience of coexisting yet opposing racial ideologies and discourses in contemporary Brazil that impact racial subjectivity and fluidity: *morenidade* versus *negritude* and racial democracy versus racial affirmation (Bailey 2009; Miranda 2015; Bailey and Fialho 2018). Regarding our demonstration of the salience of racial darkening, existing research on this period argues that this darkening dynamic results from the impact of both racial formation projects and widely enacted race-targeted affirmative action (Francis-Tan and Tannuri-Pianto 2015; Miranda 2015). Our finding on increasing racial response stability across time in Brazil, though novel due to our research design, echoes recent empirical work using repeated cross-sectional surveys (Bailey, Fialho, and Loveman 2018).

There is also evidence in the research literature supporting our findings of decreasing racial earnings inequality in the 2000s and the high impact of macroeconomic factors on such a decrease (Camargo et al. 2013; Paiva 2016). First, there was substantial economic growth in Brazil from 2003 to 2010. The Lula presidency’s (2003–2013) economic policy increased formal job creation and decreased unemployment rates (from 9.6% in 1999 to 6.3% in 2012). Second, the Lula administration increased the minimum wage, thereby contributing to an almost 49% real wage increase between 2004 and 2013. Third, redistributive social policies were reformed, reconfigured, and expanded during the Lula era. Its arguably most significant element was the Bolsa Família Program (BFP), a conditional cash transfer initiative created in 2004. The number of poor Brazilian families benefiting directly from the BFP grew from zero (at its initiation in 2004) to 8 million in 2005, 11 million in 2008, to 14 million in 2012, amounting to 50 million people or approximately one-fourth of the national population (Paiva 2016). Moreover, the BFP had disparate positive effects on the nonwhite population due to its overrepresentation among the poor and extremely poor (Camargo et al. 2013), helping to mitigate racial inequality through economic improvements (Bayer and Charles 2018; Manduca 2018; Paiva 2016). As a result of these progressive and intense policies of income redistribution,
not only racial inequality but also the Gini index for Brazil fell steeply throughout the
survey period (Paiva 2016).

While existing research corroborates many of our findings, our core result is quite
novel: an empirical demonstration of the insensitivity of estimates of racial earnings
inequality to racial response shift in Brazil. Moreover, we can isolate its mechanism:
bidirectional, countervailing racial reclassification flows of entering and exiting
individuals across racial subpopulations impacting relative population size and income
mass simultaneously (see Liebler et al. 2017). Even still, when using counterfactual
scenarios to simulate exclusive lightening and darkening flows – that is, removing the
possibility of countervailing force – the results further suggest the insensitivity of
estimates of racial income inequality to racial response shift.

In sum, our research suggests that racial response shift may not be among the
compositional factors that substantially impact estimates of racial earnings inequality, at
least not in our study’s context (Brazil’s six largest metropolitan areas) and period (2002–
2012). It remains an open empirical question, then, as to when racial response shift in a
given context might be significant enough to impact estimates of racial earnings
inequality. That noted, our analysis suggests that the size of racial response shift
documented by Liebler et al. (2017) in the United States between its 2000 and 2010
censuses is probably below the necessary threshold concerning its ability to impact
estimates of white-to-black income disparity.

8. Acknowledgments

An earlier version of this article was presented in the Department of Demography at the
University of California–Berkeley. We thank Mara Loveman, Joshua Goldstein, and
anonymous reviewers for their insightful comments. Muniz received financial support
from the National Council for Scientific and Technological Development (CNPq, process
304521/ 2019–01) and from the CAPES Foundation (process PVEX –
88881.170751/2018-01). Any views expressed in this article are those of the authors and
not those of the sponsoring institutions.
References


Muniz & Bailey: Does race response shift impact racial inequality?


