Race-Switching Behavior and Regional Migration in Brazil

by

G. Branham Culpepper

(Under the Direction of Ian Schmutte)

Abstract

Brazilian labor market data shows that some workers change reported race when changing employers. I use this variation in race to separate its effect from time-invariant factors, showing that reclassification is systematically associated with wages. I further show that race switching behavior is associated with migration. After controlling for other factors, workers who move out of a region are more likely to reclassify toward the majority race of the origin region, compared to those who remain within it, indicating that racial reclassification is likely driven by employer perceptions.

Index words: Race, Brazil, Labor Markets, Discrimination, Wages, Passing, Migration

Race-Switching Behavior and Regional Migration in Brazil

by

G. Branham Culpepper

A Thesis Submitted to the Graduate Faculty of the
University of Georgia in Partial Fulfillment of the Requirements for the Degree

Master of Arts

Athens, Georgia

2021

Race-Switching Behavior and Regional Migration in Brazil

by

G. Branham Culpepper

Major Professor: Ian Schmutte

Committee: Christopher Cornwell
Gregorio Caetano

Electronic Version Approved:

Ron Walcott
Dean of the Graduate School
The University of Georgia
May 2021

Acknowledgments

I would like to thank my thesis committee chair, Dr. Ian Schmutte, for introducing me to this topic and guiding my thinking. I would also like to thank the other members of my committee, Dr. Christopher Cornwell and Dr. Gregorio Caetano, for their unique perspectives and invaluable advice.

Contents

List of Tables

Chapter 1

Introduction

1.1 Overview

Estimating the effect of race on wages is greatly complicated by the difficulty of separating the effects of the multifarious factors that determine wage, in addition to the fact that many of these factors, such as education, are often endogenous. This reality has given rise to a substantial literature on the subject and many techniques that seek to more closely pursue the effect, such as the use of various test scores as proxies for underlying ability (e.g., in Card (1993) or Blackburn and Neumark (1993)) or the Oaxaca-Blinder decomposition (Oaxaca (1973); Blinder (1973)). However, these techniques generally cannot overcome the challenge of separating the effect of race from other time-invariant unobserved factors, such as might be captured in an individual fixed effect.

While the economics literature generally models race as a fixed characteristic, this is not always the case. In certain situations, subjectivity or ambiguity in racial classification can result in changes to observed classifications over time. Furthermore, given the existence of racially discriminatory behavior and the ability of (at least some) people to actively manipulate perceived or reported race, these changes have economic implications and may in fact be endogenous. These phenomena are widespread in Brazil, the country on which I focus my analysis, as it has a history of both substantial racial discrimination and manipulable racial classification (Telles 2004).

Similarly to Cornwell, Rivera, and Schmutte (2017), I use variation in racial classification of workers in Brazil's *Relação Anual de Informações Sociais* (RAIS, Annual Social Information Survey) to estimate the effect of race on wages for workers who change jobs. I verify that some of their main results in the 2010 data hold true in the most recent data from 2017. However, that work did not consider geographic heterogeneity in racial distributions or the potential effects of migration, which could impact both employers' perceptions of workers' race and workers' incentives to reclassify. Tables A.1, A.3, and A.4 show that Brazil's demography varies widely between regions and that racial churn is fairly common. In particular, about 20% of job changers also change reported race.

My results indicate that the racial wage premia identified by Cornwell, Rivera, and Schmutte (2017) may vary across regions in ways consistent with qualitative knowledge about regional race relations and labor markets, and that race-switching behavior is related to cross-regional migration. In particular, workers who move across region boundaries are more likely to reclassify toward the majority race of their destination region, but after conditioning on observables, the opposite is true. This implies that passing may be related to movement across observable categories, but that race churn within those groups is likely driven more by employers' perceptions than by active manipulation on the part of workers.

1.2 Relation to Existing Literature

While fields such as sociology have considered questions of subjective identity for some time,[1] the economics literature on this subject is relatively recent, albeit quickly growing. The consideration of racial "passing" (that is, representing oneself to others as a member of a given racial group for the sake of certain benefits) in this paper is motivated by the theoretical work of Akerlof and Kranton (2000), which examine the implications of allowing identity to affect the payoffs for behavior. They conclude that "choice of identity may be the most important 'economic' decision people make. Individuals may––more or less consciously—choose who they want to be ... Previous economic analyses of, for example, poverty, labor supply, and schooling have not considered these possibilities."

The passing narrative is supported by evidence from other contexts. Empirical work by Dahis, Nix, and Qian (2020) demonstrates the prevalence of white-passing behavior among African American men in the US using census data from 1880 to 1940. They show that, as in my analysis, racial reclassification was substantially associated with migration. They find that passing was more common for those moving out of the South or living in areas where passing carried greater benefits. A similar conclusion appears in Jia and Persson (2020). Examining data from China, they find that children of mixed marriages with minority mothers were more likely to be classified according to the mother's ethnicity in contravention of social norms in provinces with government benefit programs for minority groups.

Furthermore, evidence suggests that active choice of racial classification is common in Brazil as well. Francis-Tan and Tannuri-Pianto (2015) shows that Brazilians manipulate perceived race to gain university admissions, and De Micheli (2021) finds that recent expansion of education for lower-class individuals has led individuals to reclassify as black as part of the formation of a racialized political identity. However, despite all of the above, the passing narrative fails to adequately explain my results. The possibility of obtaining better labor market outcomes (in particular, higher wages) provides a clear incentive for workers to reclassify toward white. If racial churn is driven by workers, then changing toward white should dominate (in the absence of other motivations, such as the political phenomena theorized by De Micheli (2021)). This is not the case; it is fact slightly more common for workers to switch to nonwhite, particularly when we group by the second job's region (see Tables A.3 and A.4). Race change conditioned on observables seems to be more related to firms' choices, similar to the firm "preference for white workers" identified

[1]For example, Harris (1964) discusses construction of Brazilian racial identity and census categories; Rogler (1944) gives a linguistics-focused analysis of racial classification in Puerto Rico; Petersen (1969) discusses the complexities of classifying Hawaiian populations. More recent scholarship is of course also extensive, and these examples are chosen to illustrate the relatively large length of time that sociological research has explored these questions.

by Gerard et al. (2018). However, overall migration patterns among racial groups give some support for passing across observable categories.

Chapter 2

Race in Brazil

Brazil's history is defined by immigration from various locations and substantial racial mixing among these groups.[1] This circumstance has given rise to an understanding of race markedly different from that in the United States. Whereas in the US race is often considered in terms of the ethnic origin of a person's progenitors, the Brazilian formulation of identity is primarily defined in terms of skin tone, with some consideration given to other phenotypical features as well. This creates an opportunity for a complex and nuanced understanding of race, with a wide variety of racial identifiers in common use. In a survey intended to probe Brazilian's understanding of their race, the *Instituto Brasileiro Geográfico e Estatístico* (IBGE, Brazilian Geographic and Statistical Institute) received over 130 answers to the racial classification question (Racusen 2009).[2] In official sources, including the RAIS data employed by this work, race is generally given by five categories: *branco* (white), *pardo* (brown/mixed), *preto* (black), *amarelo* (Asian, literally "yellow"), and *indígena* (indigenous). However, there is a cultural understanding that "you are what you say you are" (Telles 2004), and Brazilians have been known to modify the way they discuss their own or others' races to match social contexts (Racusen 2012). Racial classification is also associated with socioeconomic status; that is, "money whitens." Telles (2004) shows that surveyors record subjects as white more often when they are of higher socioeconomic status, even when those subjects self-report as nonwhite.

2.1 History of Race and Racial Ideas in Brazil

Brazil's first non-indigenous settlers arrived from Portugal, followed by numerous waves of immigrants from other locations. Enslaved Africans were brought to work on sugar plantations, and became concentrated in the Northeast region after the abolition of slavery in 1888. During the period after the abolition of slavery and into the early 20th century, a policy of *branqueamento*, "whitening," encouraged immigration

[1] See Telles (2004) for a more thorough discussion.

[2] Most of these were used very rarely; the most common 6 terms (*branco*, *moreno*, *pardo*, *moreno-claro*, *preto* and *negro*; respectively "white," "brown/dark/brunette," "brown/mixed," "light brown," "black," and "black") accounted for 95% of the sample.

from Europe, and thus Brazil received many arrivals from Germany and northern Italy, as well as Iberia. Many of these new immigrants settled in the South and Southeast of the country, creating a pronounced racial divide between regions. Brazil also experienced a large influx of immigrants from Japan, and is home to the largest Japanese population outside of that country.

While Brazil did not witness the same level of legally enforced racial hierarchy that occurred in the United States, racial divides and discrimination have played a sizable role in Brazil's history. The Eurocentric attitude espoused by the advocates of *branqueamento* gave way to an ostensibly more positive attitude toward miscegenation, particularly following the 1933 publication of Gilberto Freyre's *Casa-Grande e Senzala*[3] and the rise of the idea of Brazil as a "racial democracy," where any discrimination was minor or even unintentional. However, later scholarship criticized this idea, pointing out sizable disparities between racial groups and clear evidence of discrimination (Hanchard 1994). To give one example, the outlawing of explicit racial hiring practices was followed by the use of euphemistic and coded language in job postings, such as "good appearance" (Telles 2004).

[3]Rendered in English as "The Masters and the Slaves"; see Freyre, Putnam, and Maybury-Lewis (1986)

Chapter 3

Empirical Setting

The data employed in this work derive from the 2017 edition of RAIS, the most recent available. RAIS is conducted annually by the IBGE and contains linked employer-employee data on Brazil's formal sector, collected at the plant level. Completion of the survey is legally required, as it is used to administrate mandatory yearly bonus pay requirements (the *abono salarial* or *décimo terceiro salario*, "13th salary"); compliance is thus very high, as penalties for failure to produce the relevant data are large. In small firms, data may be prepared by owners, while large firms may have human resources staff or other dedicated employees to collect data.

3.1 Collection of RAIS data

Workers have a document known as a *Carteira de Trabalho e Previdência Social* (CTPS, "Work and Social Security Booklet"[1]) which contains demographic information, such as name, age, birthdate, gender, home address, and an identification number. Notably, the CTPS does not include race. In addition to the CTPS information, workers provide a photo and evidence of requisite education for entry into a *Livro de Registro dos Empleados* (LRE, "Employee Registration Book") which is maintained by employers for reporting requirements. Workers' entry of information into the LRE is reviewed by the individuals responsible for reporting compliance or hiring. Thus both employees and employers have some ability to manipulate reported race through this mechanism.

3.2 Sample Construction

The sampling method used here is largely identical to that employed by Cornwell, Rivera, and Schmutte (2017). The sample is drawn from those who held full-time jobs with at least 40 contracted hours per week prior to January 1, 2017 (those who Cornwell, Rivera, and Schmutte (2017) call "continuing workers.")

[1]The law governing the CTPS is available at http://www.planalto.gov.br/ccivil_03/LEIS/L8260.htm; additional information, including about the movement to digital CTPS, is available at https://www.gov.br/trabalho/pt-br/assuntos/trabalhador/carteira-de-trabalho (Both in Portuguese).

This is then narrowed to only those who start exactly one other full time job during the year (so that they have exactly one other entry in the RAIS database for 2017). Finally, the sample includes only those workers racially classified as *branco*, *pardo*, or *preto*, excluding those categorized as *amarelo* and *indígena*, as well as those whose race is not reported. The two excluded racial categories are both very small (around 2% of the population) and geographically localized, *amarelos* in the Southeast and *indígenas* in the North. Those with unreported race tend to work at plants which uniformly do not report race, and generally work in the public sector or military. From this point I will refer to members of the *branco* category as "white" and members of the *pardo* and *preto* groups as "nonwhite." Additionally, I construct race history indicators in the same manner as Cornwell, Rivera, and Schmutte (2017), which indicate whether a worker was reported as white on each job by a 1; thus $rc11$ indicates a worker recorded as white on both jobs, while $rc01$ indicates those who switch from nonwhite on the first job to white on the second, etc.

I also utilize several plant-level averages, including the mean log wage, the share of the plant's workers listed as white, the share listed as male, and the total number of workers at the plant. Origin plants' characteristics are based on beginning-of-year levels, and destination plants' on end-of-year. The separation rate is the number of positions at the plant that ended divided by the average of beginning-of-year and end-of-year employment. These variables are also created according to the methodology of Cornwell, Rivera, and Schmutte (2017).

3.3 Overview of 2017 Data and Descriptive Statistics

Table A.1 includes means of selected individual characteristics, including race, age, gender, and log wage, broken down by region and by origin and destination job. It is clear that there is little difference in the origin and destination traits for a given region, or between regions for most indicators. However, the difference in racial makeup is stark; fully 86% of the Southern subsample is white on the origin job, compared to only 20% for the North. The geographic variation here is obvious, and may play a role in race-switching behavior. The consistency in racial proportions across jobs indicates that churn happens in both directions; under the hypothesis that workers in a discriminatory labor market seek to reclassify to white to obtain better employment outcomes, this result would imply that workers do not exercise full control of racial categorization.

Table A.2 shows means for selected plant variables, again broken down by region and origin versus destination. The regional variation in racial proportions is once again on display. There is also some measure of variability across regions and jobs for employment. However, there is no clear pattern here; while the Southeast and North show movement towards smaller plants, the opposite is true for the other three regions.

To examine race-switching behavior further I present each race history's proportion of observations by region in Tables A.3 and A.4. Sorting by the second job's region produces some changes to the distribution. Churn now appears less symmetrical; for the North, Northeast, and South, the most three racially homogeneous regions, racial reclassification leans toward the majority race of the region. Only around 6% of the sample moves between regions, but migration patterns may impact race-switching behavior. These results may reflect differing constructions of race by employers in each region, or employees' attempts to match the social context of the region or the destination plant.

Just as in Cornwell, Rivera, and Schmutte (2017), I find that race changers move between plants where large majorities of the workforce match the racial classification of the worker; i.e., those in the $rc10$ category move from heavily white plants to heavily nonwhite plants, and vice versa for the $rc01$ group. These values are reported in Table A.5. This pattern does not occur for those who do not switch race; white shares remain flat across the job change, but for each job they are comparable in magnitude to those for race changers in the same racial category on that job. Thus, the first-job white shares for the $rc10$ and $rc11$ groups are similar, etc.

Chapter 4

Empirical Model and Identification Strategy

4.1 Differenced Wage Models

I estimate variants of the reduced form model given by Cornwell, Rivera, and Schmutte (2017). This comprises a first-differences model regressing the difference in log of nominal monthly wage[1] on a set of race history indicators (R_i), individual characteristics (X_i), plant characteristics (P_j), and job characteristics (T_i).[2]

$$\Delta log(wage_i) = \alpha + \beta R_i + \gamma X_i + \delta P_j + \zeta T_i + \epsilon_i \tag{4.1}$$

Note that since $R_i = [rc11 \;\; rc10 \;\; rc01]'$ and those classified as nonwhite on both jobs form the omitted category, the coefficients on these indicators should be understood as relative to the wage trajectory of that group. If the wage effect of being white for a given job is constant, then the differencing should eliminate the race effect for those in the $rc11$ category and its coefficient should be 0. X_i includes indicators for several levels of education, a quadratic in age, and gender. P_i includes the differences in the plants' mean log wage, share white, share male, total employment, and separation rate, as well as indicators for establishment size, industry, and state for both plants. T_i includes the differences in tenure and contracted total weekly hours, as well as indicators for occupation, contract type, and hire and separation type (whether hired from unemployment or directly from another job, etc.; and whether the first job ended by termination, resignation, retirement, etc.). I assume that the idiosyncratic error term ϵ_i is uncorrelated with the explanatory variables, and that the absence of misclassification affirmed by Cornwell, Rivera, and Schmutte (2017) in the 2010 RAIS data generalizes to the 2017 data. Under these assumptions the wage effect of race-switching is identified, and ordinary least squares provides consistent estimates of the model

[1] The data derive from a single year; I assume that inflation will not seriously impact the estimation.

[2] i indexes individuals, j plants. Since continuous variables are differenced, but indicators may be present within a given vector for both periods, I omit subscripts for origin vs. destination job in the model equation.

parameters. I estimate nested variants of this model for the full sample and each region, starting with a restricted model that imposes $\gamma = \delta = \zeta = 0$, then introducing additional controls.

4.2 Race History Models

As is shown in the following chapter, the inclusion of additional control variables substantially alters the results of the differenced wage models in a way that varies across regions, implying that the relationship between race-switching behavior and personal or job characteristics depends on location. To examine this relationship I also estimate linear probability models of race history of the form:

$$rc = \eta + \xi X_i + \pi P_i + \tau T_i + e_i \tag{4.2}$$

where rc can be any race history, and destination state has been replaced by destination region. These models gauge the relationship between race-switching behavior and migration, among other variables, and permit comparison of race switchers to the baseline $rc11$ group.

Chapter 5

Results and Analysis

5.1 Cross Sectional Wage Gap

I begin by estimating the cross-sectional wage gap on the origin job for each region. The dependent variable is the log of nominal monthly average earnings[1], and the primary regressor is an indicator for being white. Table A.6 reports the uncontrolled cross sectional wage gap for each region. Estimates range from about 0.082 in the North to 0.215 in the Northeast. Table A.7 adds controls for individual, plant, and job type characteristics. Introducing these controls reduces the estimated gap substantially, explaining at least one third of the initial gap in every region, and reaching up to around two thirds for the Northeast and the sample as a whole. In the South, Southeast, and Center-West, the additional controls remove around half the original wage gap.

5.2 Wage Effects of Race Switching

Table A.8 displays results from a regression of the difference in log wages for each job on race indicators, for each region and the full sample. These results show little evidence of a consistent relationship between race history and change in wages. While several parameter estimates are statistically significant, some magnitudes are so small that the effects are economically negligible, such as those for the full sample (0.005 and 0.004 for $rc10$ and $rc01$ respectively). Furthermore, there is no consistent pattern across in the sign of race-switching effect estimates. Of the three subsamples with significant estimates for both $rc10$ and $rc01$, the full sample and Center-West both show a positive wage effect from switching race, regardless of direction (0.012 and 0.014 in the Center-West), whereas the estimates for the North subsample show a negative effect of switching (-0.013 and -0.014). These results show that, without considering other factors, the wage trajectories for the four race history groups do vary between regions; however, this variation is not systematic, and does not fit with the expectations established by theory. Since many different types of

[1] This is not quite the same as wages; while wage is often specified on a monthly basis in Brazil, the monthly average earnings here are adjusted for those who leave a job midmonth to be comparable to measures based on a full month's work.

workers are grouped together in this specification, there is enough variation in individual wage trajectories that a clear effect of race cannot be identified.

As I have already shown, however, reported race is closely tied to race distributions in the plant; thus the results discussed above are clearly biased. Consider, now, the changes that result from adding controls for individual and plant characteristics, as reported in Table A.9. There is now a consistent pattern across regions of wage decrease for those switching to nonwhite and an increase for those who switch to white. Notably, the loss to wages of those in the former category outweighs the gain to wages of those in the latter.[2] Effect sizes and symmetries also vary greatly across regions. While the wage effects in the Northeast are very symmetric (the asymmetry measure is statistically indistinguishable from 0), all other regions report a significant asymmetry. The largest of these is in the North, which reports no statistically significant effect of switching to white and a relatively large effect of -0.053 for switching to nonwhite. The other subsamples fall in between these, with symmetry gaps from -0.015 to -.021. If these results were assumed valid, it would imply that employers punish switching to nonwhite more than they reward switching to white. This could be the case if, for example, those who switch to nonwhite are moving into occupations or industries where wages are lower. It is notable that the two heavily nonwhite regions, the North and Northeast, are substantially different from the other regions, but also from each other. Again assuming these results to be valid, this would indicate substantial differences in discriminatory behavior in these areas; a more complete accounting of this will be left to a later moment.

I now add job type controls to the model and reestimate in Table A.11.[3] This once again changes the estimates substantially and demonstrates that the omission of job type controls (in particular, hire and separation type) produces a significant downward bias in wage premium estimates, particularly in the North. This implies that these variables must be substantially correlated with race history. There is an interesting relationship between the North and Northeast, in that the Northeast's premium for switching to white (0.032) closely matches the North's penalty for switching to nonwhite (-0.035). In the other regions, effect magnitudes in both directions hover around 0.02, with $r01$ coefficients slightly outweighing those for $rc10$. Since observations are sorted by region of origin, these results may be consistent with a model in which race effects interact with urban/rural migration patterns. If the $rc10$ race history in the North is associated with movement into lower-paying rural work (i.e., if these positions are considered "nonwhite" jobs) and likewise the $rc01$ history in the Northeast is associated with movement into a "white" job (perhaps in one of the Northeast's large cities, such as Fortaleza, Recife, or Salvador), then we may see these exaggerated wage effects, as part of the "money whitening" paradigm. Most of the $rc11$ estimates are insignificant, and the ones that are significant are nonetheless small. This is consistent with a model of log wages in which race has a constant additive effect on both jobs, and thus for those coded as white on both jobs, this effect is differenced out.

To examine this further I regroup observations by the destination job's region. These results appear in Table A.13.[4] This resorting of subsamples produces more symmetric wage effects; only the full sample has a

[2]Formal examinations of the symmetry gap in effect sizes (i.e., the sum of the coefficients for $rc10$ and $rc01$) appear in Table A.10

[3]Symmetry tests in Table A.12

[4]Symmetry tests in Table A.14

statistically significant asymmetry, and the magnitude of the latter is very small, at 0.007. $rc01$ coefficients continue to outweigh the $rc10$ values except in the North. The exaggerated treatment effect sizes in the North and Northeast for both directions of switching relative to the other regions is likely a reflection of the fact that, in comparison to the more highly-developed South and Southeast, where labor markets are more robust, the North and Northeast retain some elements of a "caste society" with distinct racial hierarchies (Freyre, Putnam, and Maybury-Lewis 1986).

5.3 Predicting Race Switching Behavior and Migration

Tables A.15 through A.17 report two-way frequencies for origin and destination job's region in the full sample (as a baseline) as well as for the $rc01$ and $rc10$ categories. There is some symmetry in the movement patterns for the latter two subsamples, as those who change race seem more likely to move to regions where their second reported race is more common. For example, in the $rc01$ group 4948 individuals move from the Northeast to the Southeast, and 1356 in the opposite direction. However, for the $rc10$ group this pattern is reversed; 1762 people move from Northeast to Southeast, while 4228 move the other way. Recall that Table A.5 shows that race changers generally move between plants where their reported race on the respective job forms a large majority of the workforce.

Tables A.18 through A.20 report the estimated coefficients of the destination region dummies in the race history models outlined in Chapter 4 for both directions of race change, as well as for those consistently reported white, as a baseline. These results were estimated for the full sample and for each origin region. The origin region of each subsample is the omitted level of the indicator, so that results are relative to those who do not change regions (for the full-sample estimates, I omit the Southeast). The estimates present clear patterns of geographic variation. The full-sample results (columns (1) through (3) in Table A.18) show an elevated chance of being white on both jobs relative to the Southeast for all destination regions except the Center-West. They also show a decreased chance of reclassifying toward the majority race in the South, North, and Northeast. This is notably inconsistent with the movement patterns shown in Tables A.16 and A.17. Furthermore, the regional subsamples display some important differences. In the heavily white South (columns (1) through (3) of Table A.20), cross-region migration was significantly associated with an increased probability of switching to white, regardless of the destination; the effect is strongest (0.063) for those moving to the largely nonwhite North. Likewise, the probability of switching to nonwhite is decreased for those moving out of the South, with the greatest effect occurring for those heading to the majority-nonwhite Northeast (-0.082); those moving to the Center-West are an exception (0.022). In contrast to the South, the results for the North (columns (4) through (6) of Table A.18) are largely reversed. Switching to nonwhite is more common for those headed to the South and Southeast (0.059 and 0.041, respectively), while there is no apparent effect of moving to the Northeast or Center-West. Those in the North who move to the Southeast are less likely to be recoded as white (-0.017); though statistically insignificant, the magnitude of the coefficient for the South is comparable at -0.019. Even though there are more total migrants switching toward the majority race of their destination region, after conditioning on observables, migration is associated with an increased chance of switching toward the majority race of the origin region. This implies that race-switching may be associated with these controlling factors.

Additional coefficient estimates are presented in tables A.21 through A.25. These results are largely consistent with those observed in the 2010 data by Cornwell, Rivera, and Schmutte (2017). Moving to a higher wage plant is associated with reclassification toward nonwhite; likewise moving to a more male plant is associated with switching to white. By far the best predictors are plants' shares of white workers, which is consistent with the workplace segregation we observed earlier. Interestingly, more highly educated people are significantly more likely to switch race in both directions in the full sample. They are more likely to switch to white in the Southeast and Northeast, and there are no statistically significant effects on switching to nonwhite in these regions. Education is not a significant predictor anywhere else. The other variables present such as age and individual gender seem to have little explanatory power in most subsamples; where they are significant, they have small effects.

Based on the above, a narrative of race switching behavior emerges. After controlling for observables, racial churn seems to mostly be a result of employer perceptions. This would explain why workers are more likely to reclassify toward their origin region's majority race. A possible mechanism is that geographic knowledge of a worker's origin directly affects others' perceptions of race (e.g., all else equal, employers are more likely to consider someone from the South to be white, someone from the North to be nonwhite, etc). Alternatively, construction of race with regard to skin tone likely differs between regions; if those from the South on average have lighter skin, and people in the North have a different standard of what skin tones are considered "white" versus "black" than those in the South, then these social differences could produce the trends observed here. However, simply examining inter-regional flows without conditioning on other factors reveals the opposite pattern: workers are more likely to reclassify toward the majority race of their destination region. This is consistent with the passing narrative. Migrants from the North or Northeast may attempt to pass for white in the South or Southeast to obtain better employment outcomes; those moving the other way may "reverse pass" to nonwhite. Passing behavior could allow a worker a greater chance at entering certain occupations or industries, or firms may be more likely to reclassify workers based on how they are hired. The relatively large premium earned by the $rc01$ group coming from the North (column (4) in Table A.11) is consistent with this narrative.

Chapter 6

Conclusion

I have shown that the results established for 2010 RAIS data by Cornwell, Rivera, and Schmutte (2017) largely persist in the 2017 data. Workers often change reported race on changing jobs, and this source of variation permits the separation of race's effect on wages from other, time-invariant characteristics. Even after accounting for these and observed factors, there remains an unexplained racial effect. Furthermore, average treatment effect sizes vary across regions. Grouping by destination job's region (due to the probability that race-switching is more heavily influenced by employers than workers, as detailed below), I find that wage effects of switching race are larger in the less-developed and majority-nonwhite North and Northeast, which I attribute to these regions' relatively less robust market institutions.

Race-switching behavior likewise exhibits regional variation. After controlling for observables and individual fixed effects, job changers who move to a different region show an elevated tendency to reclassify toward the majority race of their region of origin. Furthermore, while whites see a wage premium relative to nonwhites, moving to a higher-wage plant is associated with switching to nonwhite, and workers' reported race on a given job is very likely to correspond to a large majority group within the corresponding plant. Thus, within the groups defined by observable factors, race change seems to be driven by employers' perceptions and regional differences in the construction of race moreso than workers' attempts at passing for wage benefit. However, overall migration flows reveal a shift toward the majority race of the destination region. This may imply that passing behavior is in fact widespread, but that this manipulation of race is used to gain access to higher-paying or more prestigious industries or to different kinds of plants. This hypothesis is consistent with the cultural tendency to associate high socioeconomic status with whiteness. Furthermore, since wages and other benefits are highly linked to industry and other observables, this passing behavior would be rational.

There remains substantial room for additional study of race-switching phenomena in Brazil. A theoretical model that takes into account heterogeneity in employers' construction of race as well as the decision by workers to present as a member of another race would be valuable. Empirical improvements, such as more sophisticated probability modeling for race history categories or a more granular study of worker movement that also examines intra-regional migration, would benefit the understanding of this subject. Qualitative data, such as interviews or surveys, could be useful in building an understanding of the underlying economic decisions observed here.

Endogenous racial churn is observed outside the Brazilian context in countries including the US (Dahis, Nix, and Qian 2020) and in China (Jia and Persson 2020). Understanding race-switching phenomena in Brazil is likely to unlock new insights into the role of race and identity in labor markets across the world.

Appendix A

Tables

Table A.1: Means of Individual Characteristics by Region

	(1) Full Sample	(2) North	(3) Northeast	(4) Southeast	(5) South	(6) Center-West
White						
—Origin	0.57	0.20	0.24	0.61	0.86	0.38
—Destination	0.56	0.20	0.24	0.59	0.84	0.37
Pardo						
—Origin	0.37	0.76	0.72	0.33	0.10	0.57
—Destination	0.38	0.76	0.71	0.33	0.12	0.58
Preto						
—Origin	0.06	0.03	0.05	0.07	0.04	0.04
—Destination	0.06	0.03	0.05	0.07	0.04	0.05
Log Wage						
—Origin	7.55	7.48	7.33	7.62	7.52	7.49
—Destination	7.60	7.54	7.38	7.67	7.57	7.54
Weekly Hours						
—Origin	43.67	43.82	43.80	43.59	43.74	43.78
—Destination	43.62	43.64	43.75	43.55	43.69	43.75
Age						
—Origin	33.94	33.29	34.48	34.22	33.17	33.21
—Destination	33.93	33.29	34.48	34.21	33.17	33.21
Male						
—Origin	0.62	0.67	0.66	0.62	0.61	0.64
—Destination	0.62	0.67	0.66	0.62	0.61	0.64
Observations	2249181	85030	286179	1276009	424607	177356

Table A.2: Means of Plant Characteristics by Region

	Full Sample	North	Northeast	Southeast	South	Center-West
Mean Log Wage						
—Origin	7.59	7.49	7.34	7.66	7.59	7.53
—Destination	7.62	7.53	7.39	7.69	7.62	7.56
Share Male						
—Origin	0.60	0.64	0.64	0.59	0.58	0.61
—Destination	0.61	0.65	0.64	0.60	0.59	0.62
Share White						
—Origin	0.58	0.20	0.24	0.61	0.86	0.39
—Destination	0.57	0.22	0.25	0.60	0.84	0.39
Employment						
—Origin	502.10	683.56	398.24	602.70	303.24	335.06
—Destination	389.14	318.85	448.15	408.84	309.99	375.21
Separation Rate						
—Origin	0.84	0.85	0.78	0.85	0.84	0.87
—Destination	0.89	0.90	1.01	0.87	0.87	0.88
Observations	2249181	85030	286179	1276009	424607	177356

Table A.3: Race History by Origin Region

	(1) Full Sample	(2) North	(3) Northeast	(4) Southeast	(5) South	(6) Center-West
White	0.46	0.12	0.14	0.49	0.79	0.24
Nonwhite	0.33	0.71	0.66	0.28	0.08	0.49
N. to W.	0.10	0.09	0.10	0.11	0.06	0.13
W. to N.	0.11	0.09	0.10	0.12	0.07	0.14
Observations	2249181	85030	286179	1276009	424607	177356

Columns correspond to the region of the worker's first employer

Table A.4: Race History by Destination Region

	(1) Full Sample	(2) North	(3) Northeast	(4) Southeast	(5) South	(6) Center-West
White	0.46	0.12	0.14	0.49	0.78	0.24
Nonwhite	0.33	0.71	0.67	0.28	0.08	0.49
N. to W.	0.10	0.07	0.09	0.11	0.07	0.12
W. to N.	0.11	0.10	0.11	0.12	0.06	0.16
Observations	2249181	82996	281431	1281677	424118	178959

Columns correspond to the region of the worker's second employer

Table A.5: White Share of Plants by Race History and Region

	(1) Full Sample	(2) North	(3) Northeast	(4) Southeast	(5) South	(6) Center-West
rc10						
—Origin	0.73	0.61	0.64	0.74	0.87	0.69
—Destination	0.36	0.12	0.13	0.40	0.61	0.23
Observations	248728	8252	31934	152977	27364	28201
rc01						
—Origin	0.34	0.13	0.14	0.37	0.52	0.23
—Destination	0.72	0.54	0.62	0.73	0.87	0.66
Observations	223274	5998	24005	143754	28383	21134

$rc10$ denotes workers reported as white by their first employer and nonwhite by their second;
$rc01$ denotes the opposite scenario.

Table A.6: Naive Cross-Sectional Wage Gap, 1st Job

	(1) Full Sample	(2) North	(3) Northeast	(4) Southeast	(5) South	(6) Center-West
White	0.204***	0.176***	0.180***	0.215***	0.082***	0.167***
	(0.0009)	(0.0064)	(0.0031)	(0.0012)	(0.0024)	(0.0034)
r2	0.0216	0.0113	0.0152	0.0210	0.0024	0.0148
N	2249181	85030	286179	1276009	424607	177356

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.025$; dependent variable is log earnings;
sole regressor is an indicator for being reported white.

Table A.7: Controlled Cross-Sectional Wage Gap, 1st Job

	(1) Full Sample	(2) North	(3) Northeast	(4) Southeast	(5) South	(6) Center-West
White	0.071***	0.089***	0.072***	0.072***	0.054***	0.079***
	(0.0008)	(0.0054)	(0.0028)	(0.0010)	(0.0021)	(0.0031)
r2	0.6599	0.6535	0.6287	0.6763	0.6050	0.6219
N	1945531	72639	245812	1104607	369468	153005

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.025$; dependent variable is log earnings; regressors include an indicator for being reported white, gender, education, a quadratic in age, tenure, occupation, and plant's share white, mean log wage, share male, industry, employment and separation rate.

Table A.8: Naive Differenced Models

	(1) Full Sample	(2) North	(3) Northeast	(4) Southeast	(5) South	(6) Center-West
rc11	0.004***	0.012***	0.001	0.001	0.001	0.008***
	(0.0006)	(0.0042)	(0.0021)	(0.0008)	(0.0020)	(0.0026)
rc10	0.005***	-0.013***	0.001	0.005***	-0.000	0.012***
	(0.0010)	(0.0056)	(0.0028)	(0.0013)	(0.0031)	(0.0032)
rc01	0.004***	-0.014***	0.024***	-0.001	0.005	0.014***
	(0.0011)	(0.0058)	(0.0028)	(0.0014)	(0.0033)	(0.0035)
r2	0.0000	0.0003	0.0003	0.0000	0.0000	0.0002
N	2249181	85030	286179	1276009	424607	177356

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.025$; columns correspond to first job's region; dependent variable is difference in log earnings; regressors are race history indicators.

Table A.9: Differenced Models with Individual and Plant Controls

	(1) Full Sample	(2) North	(3) Northeast	(4) Southeast	(5) South	(6) Center-West
rc11	0.001	0.023***	0.010***	-0.001	-0.006***	0.005
	(0.0008)	(0.0055)	(0.0025)	(0.0010)	(0.0025)	(0.0030)
rc10	-0.033***	-0.053***	-0.029***	-0.034***	-0.035***	-0.033***
	(0.0012)	(0.0071)	(0.0034)	(0.0015)	(0.0035)	(0.0039)
rc01	0.016***	0.009	0.023***	0.013***	0.016***	0.017***
	(0.0012)	(0.0070)	(0.0034)	(0.0016)	(0.0037)	(0.0040)
r2	0.1313	0.1626	0.1472	0.1235	0.1346	0.1688
N	1636198	59193	205684	931149	313430	126742

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.025$. Columns correspond to second job's region. Dependent variable is difference in log earnings. Regressors include race history indicators, gender, education, a quadratic in age, tenure on each job, and each plant's share white, mean log wage, share male, industry, employment and separation rate.

Table A.10: Symmetry Tests (rc10 + rc01) - Ind. and Plant Controls

Region	Point Estimate	P-Value	Confidence Interval
Full Sample	-0.016***	0.000	[-0.020, -0.013]
North	-0.044***	0.000	[-0.062, -0.026]
Northeast	-0.006	0.183	[-0.015, 0.003]
Southeast	-0.021***	0.000	[-0.025, -0.016]
South	-0.019***	0.002	[-0.030, -0.007]
Center-West	-0.015***	0.004	[-0.026, -0.005]

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.025$

Table A.11: Fully Controlled Differenced Models

	(1) Full Sample	(2) North	(3) Northeast	(4) Southeast	(5) South	(6) Center-West
rc11	0.001	0.005	0.001	-0.001	-0.001	-0.004
	(0.0008)	(0.0056)	(0.0026)	(0.0010)	(0.0025)	(0.0031)
rc10	-0.019***	-0.035***	-0.016***	-0.020***	-0.021***	-0.020***
	(0.0012)	(0.0071)	(0.0034)	(0.0015)	(0.0035)	(0.0039)
rc01	0.026***	0.022***	0.032***	0.022***	0.025***	0.023***
	(0.0012)	(0.0071)	(0.0034)	(0.0016)	(0.0037)	(0.0040)
r2	0.1521	0.2016	0.1713	0.1455	0.1612	0.1971
N	1636198	59193	205684	931149	313430	126742

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.025$. Columns correspond to first job's region. Dependent variable is difference in log earnings. Regressors include race history indicators; gender; education; a quadratic in age; tenure, occupation, hire type, separation type, and contract type on each job; and each plant's share white, mean log wage, share male, industry, employment and separation rate.

Table A.12: Symmetry Tests (rc10 + rc01) - Full Controls

Region	Point Estimate	P-Value	Confidence Interval
Full Sample	0.007***	0.000	[0.004, 0.011]
North	-0.013	0.161	[-0.032, 0.005]
Northeast	0.016***	0.001	[0.007, 0.025]
Southeast	0.002	0.394	[-0.003, 0.007]
South	0.004	0.557	[-0.008, 0.015]
Center-West	0.003	0.618	[-0.008, 0.014]

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.025$

Table A.13: Differenced Models by Destination Region

	(1) Full Sample	(2) North	(3) Northeast	(4) Southeast	(5) South	(6) Center-West
rc11	0.001	0.008	0.011***	-0.002	-0.005**	0.001
	(0.0008)	(0.0057)	(0.0027)	(0.0010)	(0.0025)	(0.0030)
rc10	-0.019***	-0.032***	-0.025***	-0.019***	-0.023***	-0.023***
	(0.0012)	(0.0071)	(0.0034)	(0.0015)	(0.0036)	(0.0038)
rc01	0.026***	0.029***	0.027***	0.022***	0.021***	0.024***
	(0.0012)	(0.0076)	(0.0036)	(0.0016)	(0.0036)	(0.0041)
r2	0.1521	0.1975	0.1647	0.1454	0.1654	0.2032
N	1636198	57171	200329	938378	313105	127215

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.025$. Columns correspond to first job's region. Dependent variable is difference in log earnings. Regressors include race history indicators; gender; education; a quadratic in age; tenure, occupation, hire type, separation type, and contract type on each job; and each plant's share white, mean log wage, share male, industry, employment and separation rate.

Table A.14: Symmetry Tests (rc10 + rc01) - Destination

Region	Point Estimate	P-Value	Confidence Interval
Full Sample	0.007***	0.000	[0.004, 0.011]
North	-0.003	0.737	[-0.023, 0.016]
Northeast	0.002	0.637	[-0.007, 0.011]
Southeast	0.003	0.168	[-0.001, 0.008]
South	-0.002	0.772	[-0.013, 0.010]
Center-West	0.001	0.833	[-0.010, 0.012]

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.025$

Table A.15: Movement between Regions - Full Sample

	North	Northeast	Southeast	South	Center-West
North	74067	2404	4309	1320	2930
Northeast	2667	256962	19340	3345	3865
Southeast	3407	16021	1224163	19956	12462
South	708	3027	21841	395411	3620
Center-West	2147	3017	12024	4086	156082

Table reports frequency of moving between any two regions. Rows correspond to origin region, columns to destination region.

Table A.16: Movement between Regions - rc01

	North	Northeast	Southeast	South	Center-West
North	5345	291	1031	512	357
Northeast	186	21930	4948	1264	472
Southeast	250	1356	132754	3357	1176
South	34	151	2155	21938	230
Center-West	183	277	2866	1312	18899

Table reports frequency of moving between any two regions in the $rc01$ group.
Rows correspond to origin region, columns to destination region.

Table A.17: Movement between Regions - rc10

	North	Northeast	Southeast	South	Center-West
North	6563	189	325	66	409
Northeast	290	25972	1762	164	434
Southeast	773	4228	146123	1940	3105
South	268	992	3447	24854	1242
Center-West	358	553	1320	340	23011

Table reports frequency of moving between any two regions in the $rc10$ group.
Rows correspond to origin region, columns to destination region.

Table A.18: Race History Models, Migration - Full Sample and North

	Full Sample			North		
	(1)	(2)	(3)	(4)	(5)	(6)
	rc11	rc01	rc10	rc11	rc01	rc10
North	0.032***	0.003	-0.051***			
	(0.0027)	(0.0027)	(0.0027)			
Northeast	0.018***	0.004***	-0.033***	0.010	-0.004	0.000
	(0.0018)	(0.0017)	(0.0017)	(0.0060)	(0.0067)	(0.0060)
Southeast				-0.029***	-0.017***	0.041***
				(0.0059)	(0.0071)	(0.0050)
South	0.012***	-0.041***	0.007***	-0.061***	-0.019	0.059***
	(0.0019)	(0.0016)	(0.0015)	(0.0103)	(0.0134)	(0.0074)
Center-West	-0.015***	0.012***	0.006***	-0.011*	-0.017***	0.006
	(0.0021)	(0.0018)	(0.0019)	(0.0062)	(0.0062)	(0.0059)
r2	0.4461	0.2172	0.2249	0.2920	0.3099	0.3344
N	1636198	1636198	1636198	59193	59193	59193

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.025$. Multicolumn headings indicate subsample based on first job. Dependent variables are the race history indicators given above each column. Regressors include indicators for second job's region; gender; education; a quadratic in age; tenure, occupation, hire type, separation type, and contract type on each job; and each plant's share white, mean log wage, share male, industry, employment and separation rate.

Table A.19: Race History Models, Migration - Northeast and Southeast

	Northeast			Southeast		
	(1)	(2)	(3)	(4)	(5)	(6)
	rc11	rc01	rc10	rc11	rc01	rc10
North	0.012**	0.005	-0.011*	-0.010	0.045***	-0.042***
	(0.0056)	(0.0058)	(0.0061)	(0.0067)	(0.0053)	(0.0074)
Northeast				-0.029***	0.033***	0.004
				(0.0031)	(0.0025)	(0.0036)
Southeast	-0.027***	-0.012***	0.038***			
	(0.0029)	(0.0033)	(0.0025)			
South	-0.071***	0.012	0.068***	-0.004	-0.020***	0.041***
	(0.0063)	(0.0082)	(0.0043)	(0.0033)	(0.0027)	(0.0023)
Center-West	-0.012***	-0.020***	0.016***	-0.037***	0.043***	0.025***
	(0.0046)	(0.0053)	(0.0050)	(0.0038)	(0.0029)	(0.0040)
r2	0.3592	0.3200	0.3209	0.3509	0.2083	0.2156
N	205684	205684	205684	931149	931149	931149

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.025$. Multicolumn headings indicate subsample based on first job. Dependent variables are the race history indicators given above each column. Regressors include indicators for second job's region; gender; education; a quadratic in age; tenure, occupation, hire type, separation type, and contract type on each job; and each plant's share white, mean log wage, share male, industry, employment and separation rate.

Table A.20: Race History Models, Migration - South and Center-West

	South			Center-West		
	(1)	(2)	(3)	(4)	(5)	(6)
	rc11	rc01	rc10	rc11	rc01	rc10
North	-0.011	0.063***	-0.035**	0.013*	-0.008	-0.023***
	(0.0147)	(0.0094)	(0.0176)	(0.0069)	(0.0065)	(0.0077)
Northeast	0.015*	0.014***	-0.082***	0.003	0.011**	-0.022***
	(0.0077)	(0.0049)	(0.0084)	(0.0060)	(0.0055)	(0.0067)
Southeast	0.024***	0.029***	-0.028***	0.006	0.005	0.009***
	(0.0035)	(0.0023)	(0.0028)	(0.0042)	(0.0041)	(0.0034)
South				-0.004	0.000	0.030***
				(0.0071)	(0.0075)	(0.0050)
Center-West	-0.012	0.019***	0.022***			
	(0.0075)	(0.0044)	(0.0077)			
r2	0.3147	0.1936	0.2160	0.3354	0.2760	0.3012
N	313430	313430	313430	126742	126742	126742

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.025$. Multicolumn headings indicate subsample based on first job. Dependent variables are the race history indicators given above each column. Regressors include indicators for second job's region; gender; education; a quadratic in age; tenure, occupation, hire type, separation type, and contract type on each job; and each plant's share white, mean log wage, share male, industry, employment and separation rate.

Table A.21: Race History Models, Individual Characteristics - Full Sample and North

	Full Sample			North		
	(1)	(2)	(3)	(4)	(5)	(6)
	rc11	rc01	rc10	rc11	rc01	rc10
Age	-0.003***	-0.000***	0.001***	-0.002**	-0.002***	-0.000
	(0.0002)	(0.0001)	(0.0001)	(0.0008)	(0.0008)	(0.0007)
Age^2	0.000***	0.000*	-0.000***	0.000***	0.000***	0.000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Male						
—Origin	-0.014***	-0.002	0.004**	-0.008	-0.003	0.005
	(0.0022)	(0.0019)	(0.0020)	(0.0071)	(0.0092)	(0.0085)
—Destination	-0.015***	0.005***	0.000	-0.001	-0.007	-0.009
	(0.0022)	(0.0019)	(0.0020)	(0.0070)	(0.0091)	(0.0084)
Education						
—High School	-0.014	0.059***	0.031***	0.097***	0.060	0.018
	(0.0136)	(0.0100)	(0.0136)	(0.0232)	(0.0615)	(0.0420)
—Some College	0.012	0.058***	0.029**	0.099***	0.068	0.030
	(0.0137)	(0.0100)	(0.0136)	(0.0236)	(0.0616)	(0.0422)
—Bachelor's or Higher	0.028**	0.059***	0.027**	0.119***	0.068	0.023
	(0.0137)	(0.0100)	(0.0136)	(0.0235)	(0.0616)	(0.0421)
r2	0.4461	0.2172	0.2249	0.2920	0.3099	0.3344
N	1636198	1636198	1636198	59193	59193	59193

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.025$. Multicolumn headings indicate subsample based on first job. Dependent variables are the race history indicators given above each column. Regressors include indicators for second job's region; gender; education; a quadratic in age; tenure, occupation, hire type, separation type, and contract type on each job; and each plant's share white, mean log wage, share male, industry, employment and separation rate.

Table A.22: Race History Models, Plant Characteristics - Full Sample and North

	Full Sample			North		
	(1)	(2)	(3)	(4)	(5)	(6)
	rc11	rc01	rc10	rc11	rc01	rc10
Plant Mean Log Wage						
—Origin	-0.010***	0.033***	-0.049***	0.008	0.016***	-0.048***
	(0.0011)	(0.0008)	(0.0008)	(0.0047)	(0.0046)	(0.0046)
—Destination	-0.023***	-0.046***	0.042***	0.006	-0.052***	0.010**
	(0.0011)	(0.0008)	(0.0008)	(0.0050)	(0.0049)	(0.0043)
Plant Share Male						
—Origin	0.024***	-0.023***	0.026***	0.007	-0.013*	0.031***
	(0.0016)	(0.0013)	(0.0013)	(0.0065)	(0.0069)	(0.0067)
—Destination	0.024***	0.026***	-0.024***	0.009	0.036***	-0.020***
	(0.0016)	(0.0013)	(0.0013)	(0.0067)	(0.0069)	(0.0066)
Plant Share White						
—Origin	0.529***	-0.436***	0.451***	0.345***	-0.271***	0.634***
	(0.0012)	(0.0011)	(0.0011)	(0.0055)	(0.0043)	(0.0061)
—Destination	0.514***	0.418***	-0.459***	0.310***	0.592***	-0.273***
	(0.0012)	(0.0011)	(0.0011)	(0.0055)	(0.0065)	(0.0046)
Plant Employment						
—Origin	0.000*	-0.000	0.000***	0.000	0.000	-0.000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
—Destination	-0.000***	-0.000	0.000***	-0.000***	0.000	0.000**
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
r2	0.4461	0.2172	0.2249	0.2920	0.3099	0.3344
N	1636198	1636198	1636198	59193	59193	59193

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.025$. Multicolumn headings indicate subsample based on first job. Dependent variables are the race history indicators given above each column. Regressors include indicators for second job's region; gender; education; a quadratic in age; tenure, occupation, hire type, separation type, and contract type on each job; and each plant's share white, mean log wage, share male, industry, employment and separation rate.

Table A.23: Race History Models, Individual Characteristics - Northeast and Southeast

	Northeast			Southeast		
	(1)	(2)	(3)	(4)	(5)	(6)
	rc11	rc01	rc10	rc11	rc01	rc10
Age	-0.001***	-0.001	0.001	-0.003***	-0.001***	0.001***
	(0.0004)	(0.0004)	(0.0004)	(0.0003)	(0.0002)	(0.0002)
Age^2	0.000***	0.000*	-0.000	0.000***	0.000***	-0.000***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Male						
—Origin	-0.007	-0.003	-0.007	-0.013***	-0.004	0.008***
	(0.0042)	(0.0050)	(0.0048)	(0.0032)	(0.0027)	(0.0028)
—Destination	-0.011***	0.002	0.005	-0.020***	0.008***	-0.001
	(0.0042)	(0.0050)	(0.0048)	(0.0032)	(0.0027)	(0.0028)
Education						
—High School	-0.011	0.038***	0.016	0.073	0.073***	-0.026
	(0.0090)	(0.0120)	(0.0140)	(0.0647)	(0.0124)	(0.0444)
—Some College	-0.002	0.043***	0.027*	0.108*	0.070***	-0.031
	(0.0094)	(0.0123)	(0.0142)	(0.0648)	(0.0125)	(0.0444)
—Bachelor's or Higher	0.011	0.048***	0.033***	0.128**	0.070***	-0.034
	(0.0093)	(0.0122)	(0.0141)	(0.0648)	(0.0125)	(0.0444)
r2	0.3592	0.3200	0.3209	0.3509	0.2083	0.2156
N	205684	205684	205684	931149	931149	931149

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.025$. Multicolumn headings indicate subsample based on first job. Dependent variables are the race history indicators given above each column. Regressors include indicators for second job's region; gender; education; a quadratic in age; tenure, occupation, hire type, separation type, and contract type on each job; and each plant's share white, mean log wage, share male, industry, employment and separation rate.

Table A.24: Race History Models, Plant Characteristics - Northeast and Southeast

	Northeast			Southeast		
	(1)	(2)	(3)	(4)	(5)	(6)
	rc11	rc01	rc10	rc11	rc01	rc10
Plant Mean Log Wage						
—Origin	0.009***	0.012***	-0.055***	-0.017***	0.041***	-0.050***
	(0.0027)	(0.0026)	(0.0025)	(0.0014)	(0.0010)	(0.0011)
—Destination	-0.016***	-0.045***	0.028***	-0.028***	-0.048***	0.050***
	(0.0027)	(0.0027)	(0.0024)	(0.0015)	(0.0010)	(0.0011)
Plant Share Male						
—Origin	0.011***	-0.016***	0.028***	0.027***	-0.028***	0.028***
	(0.0036)	(0.0036)	(0.0035)	(0.0023)	(0.0017)	(0.0019)
—Destination	0.010***	0.038***	-0.005	0.031***	0.028***	-0.029***
	(0.0036)	(0.0037)	(0.0035)	(0.0024)	(0.0018)	(0.0018)
Plant Share White						
—Origin	0.389***	-0.316***	0.588***	0.564***	-0.466***	0.415***
	(0.0027)	(0.0022)	(0.0030)	(0.0016)	(0.0014)	(0.0013)
—Destination	0.350***	0.573***	-0.306***	0.550***	0.383***	-0.492***
	(0.0027)	(0.0031)	(0.0023)	(0.0016)	(0.0013)	(0.0015)
Plant Employment						
—Origin	0.000***	0.000	-0.000***	0.000***	-0.000***	-0.000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
—Destination	-0.000	-0.000***	0.000***	-0.000***	-0.000***	0.000***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
r2	0.3592	0.3200	0.3209	0.3509	0.2083	0.2156
N	205684	205684	205684	931149	931149	931149

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.025$. Multicolumn headings indicate subsample based on first job. Dependent variables are the race history indicators given above each column. Regressors include indicators for second job's region; gender; education; a quadratic in age; tenure, occupation, hire type, separation type, and contract type on each job; and each plant's share white, mean log wage, share male, industry, employment and separation rate.

Table A.25: Race History Models, Individual Characteristics - South and Center-West

	South			Center-West		
	(1)	(2)	(3)	(4)	(5)	(6)
	rc11	rc01	rc10	rc11	rc01	rc10
Age	-0.005***	0.001***	0.001***	-0.002***	-0.000	0.001*
	(0.0004)	(0.0003)	(0.0003)	(0.0006)	(0.0006)	(0.0006)
Age^2	0.000***	-0.000***	-0.000***	0.000***	0.000	-0.000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Male						
—Origin	-0.018***	0.007*	0.005	-0.016***	-0.007	-0.004
	(0.0052)	(0.0037)	(0.0039)	(0.0063)	(0.0066)	(0.0068)
—Destination	-0.011**	0.002	0.002	-0.005	0.004	-0.000
	(0.0052)	(0.0037)	(0.0039)	(0.0063)	(0.0066)	(0.0067)
Education						
—High School	0.117	-0.026	-0.049	0.057	0.032	-0.037
	(0.1059)	(0.0573)	(0.0847)	(0.1299)	(0.0388)	(0.1041)
—Some College	0.135	-0.033	-0.053	0.075	0.046	-0.030
	(0.1059)	(0.0573)	(0.0847)	(0.1299)	(0.0389)	(0.1042)
—Bachelor's or Higher	0.147	-0.034	-0.057	0.094	0.043	-0.035
	(0.1059)	(0.0573)	(0.0847)	(0.1299)	(0.0389)	(0.1042)
r2	0.3147	0.1936	0.2160	0.3354	0.2760	0.3012
N	313430	313430	313430	126742	126742	126742

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.025$. Multicolumn headings indicate subsample based on first job. Dependent variables are the race history indicators given above each column. Regressors include indicators for second job's region; gender; education; a quadratic in age; tenure, occupation, hire type, separation type, and contract type on each job; and each plant's share white, mean log wage, share male, industry, employment and separation rate.

Table A.26: Race History Models, Plant Characteristics - South and Center-West

	South			Center-West		
	(1)	(2)	(3)	(4)	(5)	(6)
	rc11	rc01	rc10	rc11	rc01	rc10
Plant Mean Log Wage						
—Origin	-0.017***	0.031***	-0.029***	0.005	0.013***	-0.054***
	(0.0025)	(0.0016)	(0.0017)	(0.0035)	(0.0031)	(0.0032)
—Destination	-0.026***	-0.025***	0.039***	-0.006*	-0.058***	0.028***
	(0.0025)	(0.0016)	(0.0018)	(0.0038)	(0.0033)	(0.0032)
Plant Share Male						
—Origin	0.032***	-0.024***	0.007***	0.019***	-0.012***	0.033***
	(0.0034)	(0.0021)	(0.0025)	(0.0054)	(0.0050)	(0.0050)
—Destination	0.024***	0.010***	-0.023***	0.015***	0.030***	-0.030***
	(0.0034)	(0.0022)	(0.0024)	(0.0055)	(0.0051)	(0.0050)
Plant Share White						
—Origin	0.724***	-0.598***	0.269***	0.420***	-0.352***	0.560***
	(0.0045)	(0.0042)	(0.0029)	(0.0033)	(0.0028)	(0.0033)
—Destination	0.721***	0.228***	-0.635***	0.404***	0.522***	-0.365***
	(0.0043)	(0.0026)	(0.0041)	(0.0034)	(0.0035)	(0.0030)
Plant Employment						
—Origin	0.000	-0.000***	0.000	0.000	-0.000**	0.000*
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
—Destination	-0.000***	0.000	0.000***	-0.000***	0.000	0.000*
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
r2	0.3147	0.1936	0.2160	0.3354	0.2760	0.3012
N	313430	313430	313430	126742	126742	126742

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.025$. Multicolumn headings indicate subsample based on first job. Dependent variables are the race history indicators given above each column. Regressors include indicators for second job's region; gender; education; a quadratic in age; tenure, occupation, hire type, separation type, and contract type on each job; and each plant's share white, mean log wage, share male, industry, employment and separation rate.

Bibliography

Akerlof, George A. and Rachel E. Kranton (Aug. 2000). "Economics and Identity". In: *The Quarterly Journal of Economics* 115.3, pp. 715–753.

Blackburn, McKinley L and David Neumark (1993). *Are OLS estimates of the return to schooling biased downward? Another look*. NBER Working Paper 4259. National Bureau of Economic Research.

Blinder, Alan S. (1973). "Wage Discrimination: Reduced Form and Structural Estimates". In: *The Journal of Human Resources* 8.4, pp. 436–455. ISSN: 0022166X.

Card, David (1993). *Using Geographic Variation in College Proximity to Estimate the Return to Schooling*. NBER working paper w4483.

Cornwell, Christopher, Jason Rivera, and Ian M Schmutte (2017). "Wage Discrimination When Identity Is Subjective: Evidence from Changes in Employer-Reported Race". In: *Journal of Human Resources* 52.3, pp. 719–755.

Dahis, Ricardo, Emily Nix, and Nancy Qian (2020). *Choosing Racial Identity in the United States, 1880-1940*. NBER Working Paper 26465.

De Micheli, David (2021). "Racial Reclassification and Political Identity Formation". In: *World Politics* 73.1, pp. 1–51.

Francis-Tan, Andrew and Maria Tannuri-Pianto (2015). "Inside the Black Box: Affirmative Action and the Social Construction of Race in Brazil". In: *Ethnic and Racial Studies* 38.15, pp. 2771–2790.

Freyre, G., S. Putnam, and D. Maybury-Lewis (1986). *The Masters and the Slaves*. Campus (Berkeley, Calif.) University of California Press. ISBN: 9780520056657.

Gerard, François et al. (2018). *Assortative Matching or Exclusionary Hiring? The Impact of Firm Policies on Racial Wage Differences in Brazil*. NBER Working Paper. National Bureau of Economic Research.

Hanchard, M.G. (1994). *Orpheus and Power: The Movimento Negro of Rio de Janeiro and São Paulo, Brazil, 1945-1988*. Political science: Latina American studies. Princeton University Press. ISBN: 9780691032924.

Harris, Marvin D. (1964). "Racial Identity in Brazil". In: *Luso-Brazilian Review* 1.2, pp. 21–28. ISSN: 00247413, 15489957.

Jia, Ruixue and Torsten Persson (June 2020). "Choosing Ethnicity: The Interplay between Individual and Social Motives". In: *Journal of the European Economic Association*.

Oaxaca, Ronald (1973). "Male-Female Wage Differentials in Urban Labor Markets". In: *International Economic Review* 14.3, pp. 693–709. ISSN: 00206598, 14682354.

Petersen, William (1969). "The Classification of Subnations in Hawaii: An Essay in the Sociology of Knowledge". In: *American Sociological Review* 34.6, pp. 863–877. ISSN: 00031224.

Racusen, Seth (2009). "Affirmative Action and Identity". In: *Brazil's new racial politics*, pp. 89–122.

— (2012). "The Grammar of Color Identity in Brazil". In: ed. by Bernd Reiter and Kimberly Eison Simmons, pp. 141–176.

Rogler, Charles C. (1944). "The Role of Semantics in the Study of Race Distance in Puerto Rico". In: *Social Forces* 22.4, pp. 448–453. ISSN: 00377732, 15347605.

Telles, Edward E. (2004). *Race in Another America: The Significance of Skin Color in Brazil*. Princeton University Press. ISBN: 9780691127927.