Working Papers

FEDERAL RESERVE BANK OF BOSTON

No. 24-13

The Impact of Market Factors on Racial Identity: Evidence from Multiracial Survey Respondents

Osborne Jackson

Abstract:
This paper examines the reported race of multiracial persons in the US Current Population Survey (CPS) before 2003, when limited response options exogenously constrained respondents to identify as a single race. Using this survey attribute and the 16-month longitudinal design of the basic monthly CPS, I explore whether market factors help causally determine racial identity. I find that pre-2003 race responds to state-level (1) racial composition, due largely to household composition, and (2) unemployment rates and wages by race. Although these findings suggest potential endogeneity of race, estimation of how race affects individual-level labor market outcomes indicates minimal bias.

JEL Classifications: J15, R23, C18
Keywords: Racial identity, multiracial, bias, Current Population Survey

Osborne Jackson (osborne.jackson@bos.frb.org) is a senior economist in the Federal Reserve Bank of Boston Research Department.

The author appreciates the helpful feedback he received from Jeff Thompson as well as the beneficial edits by Brad Arndt and Lawrence Bean. Katie Cox and Matt Sexton provided outstanding research assistance.

This version: October 2024 **https://doi.org/10.29412/res.wp.2024.13**

1 Introduction

The share of the US population that identifies as two or more races has grown considerably in recent years. According to the US Census Bureau, the multiracial population in 2010 represented 2.9 percent of the total population (9 million people), and by 2020, it had grown to 10.2 percent of the total population (33.8 million people), reflecting a 276 percent increase. This observed change could have been due to various factors, including demographic changes. Accordingly, the nation's multiracial population was the youngest of any racial group in 2020, with the lowest median age (29.5 years old), a large share under age 18 (32.5 percent), and a modal age of 12. However, the census notes that the large rise in the multiracial population from 2010 to 2020 was likely due primarily to improvements in survey design, data processing, and coding, allowing for a more complete and accurate measurement of how people prefer to be identified.[1] In addition to being substantively important for capturing demographic trends, the evolving survey measurement of the US multiracial population may provide an opportunity to broaden understanding around racial identity and its determinants for multiracial individuals and, perhaps, society at large.

This paper estimates the impact of market factors on racial identity, leveraging evidence from multiracial respondents to the US Current Population Survey (CPS) at the time of a change in that survey's design. In 2003, the survey allowed respondents to identify as multiple races rather than solely as a single race. I use this expansion of race-reporting options in the basic monthly CPS to study whether certain market factors help causally determine racial identity. The geographic market boundary that I opt for is a state, as a balance between more and less localized options (for instance, a town or census division) given traits that vary across individuals and can suggest contrasting market boundaries (such as labor market size varying by skill). I find that the pre-2003 reported race of multiracial persons responds

[1] See Nicholas Jones, Rachel Marks, Roberto Ramirez, and Merarys Rios-Vargas, "2020 Census Illuminates Racial and Ethnic Composition of the Country," US Census Bureau, August 12, 2021. See also Brittany Rico, Paul Jacobs, and Alli Coritz, "2020 Census Shows Increase in Multiracial Population in All Age Categories," US Census Bureau, June 1, 2023.

to state-level market factors—namely, racial composition (market demographics) as well as unemployment rates and wages, both by race (labor markets).

The effect of state-level racial composition on racial identity operates primarily through the racial composition of persons living in the same household rather than the racial composition of external peers who reside in the same state but outside one's household. State-level unemployment rates also seem to influence racial identity but only at a one-month lag. This suggests that time may be needed for respondents to learn about labor market conditions. Effects of lagged state-level wages on the reported race are also contingent on being defined for a subgroup (for instance, women, consistent with lagged state-level wages by race and gender) and pertaining to respondents from that same subgroup (for instance, multiracial women). This finding might indicate that narrower reference frames are necessary for wages to affect racial identity. While these results suggest that race is potentially endogenous, estimation of how race affects individual-level labor market outcomes indicates that any bias due to the omission of market factors from such analysis is likely minimal.

This study contributes to a large body of work on the economics of identity (for instance, reflected by discussions on identity and the economics of education or organizations from Akerlof and Kranton 2002 and 2005, respectively). More specifically, this paper adds to an established literature, largely by sociologists, on the factors causing or related to the racial identity of multiracial persons, with examined factors including gender, religion, age, physical appearance, parental race, and parental education (Davenport 2016; Doyle and Kao 2007; Harris and Sim 2002; Hitlin, Elder, Jr., and Brown 2006; Reece 2019; Sims 2016; Xie and Goyette 1997). Harris and Sim (2002) analyze how multiracial adolescents self-identify when prompted to select a single race by a supplemental question in the National Longitudinal Study of Adolescent Health. In binary logit models of the single-race response, one of the regressors the authors include as a market factor is local racial composition, as measured by the population share in a census tract that is non-Hispanic White. Among the differences with Harris and Sim (2002), this paper's estimation sample focuses primarily on adults whose

responsiveness to labor market factors could be larger if more potential work experience increases awareness of market conditions. Also, I use a multinomial choice model (well suited to a decision involving multiple races) and analyze multiple market factors, including market-level unemployment (distinct from studies that examine how individual-level unemployment is associated with racial identity). Davenport (2016) analyzes socioeconomic status as a market factor using neighborhood median income. By contrast, this paper examines market wages rather than earnings plus non-labor income; in addition, the wage measure is specific to each race within an area. Such specificity in market factors can help distinguish whether changes in an aggregate market factor (namely one that is not race-specific) cause individuals to identify as a particular race due to that race having a larger race-specific marginal effect or, rather, due to greater representation of that race in the market. This paper also contributes to the literature on sources of racial fluidity over time and the related implications for the determinants of racial identity (Dahis, Nix, and Qian 2020; Liebler et al. 2017; Rademakers and van Hoorn 2021; Saperstein and Gullickson 2013; Saperstein and Penner 2012). Lastly, the findings of this study help further an understanding of the extent to which race is endogenous and may result in biased estimation in some contexts.

The remainder of the paper is organized as follows. Section 2 describes the Current Population Survey and how its measurement of race has changed over time, including respondents' ability to identify as multiracial. Section 3 outlines the empirical strategy for estimating the impact of market factors on racial identity, while section 4 explores descriptive patterns. Section 5 presents the main causal findings, and section 6 discusses additional findings. Section 7 concludes.

2 Current Population Survey

2.1 Description and 2003 Change in Racial Categories

The Current Population Survey was first administered in 1940 to measure national unemployment. It is the main source of labor force statistics for the United States and is sponsored jointly by the US Census Bureau and the US Bureau of Labor Statistics (BLS). The basic monthly survey (BMS) component of the CPS relies on a rotating sample of 60,000 households whose responses on numerous topics refer to activities during the preceding week that includes the 12th of the month.[2] Households are in the CPS for four consecutive months and out for eight months. They then return for four months before leaving the sample permanently (US Census Bureau 2006). With this 4-8-4 design, the BMS has the scope to be used as a longitudinal survey, although it is typically used as a pooled cross-section. The Minnesota Population Center provides CPS data as part of its online Integrated Public Use Microdata Series (IPUMS) (Flood et al. 2020). The center's website and linking methods facilitate use of the BMS for longitudinal research (Drew, Flood, and Warren 2014).

Additional earnings information is collected in only the fourth and eighth months-in-sample (outgoing rotation groups) from civilians who are aged 15 and older, currently employed as wage/salary workers, and not self-employed. Collectively, these data are known as the "earner study" (Flood et al. 2020). I use information from the earner study on usual earnings per week at the current job, combined with information from the BMS on usual hours worked per week at the main job, to calculate an hourly wage measure.[3]

I leverage the fact that the BMS question on race, applicable to all persons since 1976, did not provide options for racial categories that identified multiple races until January 2003.

[2]The CPS sample size has grown over time and was most recently expanded from 50,0000 households to 60,000 households in 2001. See Ryan T. Helwig, Randy E. Ilg, and Sandra L. Mason, "Expansion of the Current Population Survey Sample Effective July 2001," US Bureau of Labor Statistics, August 2001.

[3]Hourly wages available directly from the earner study are limited to workers paid hourly. Also, to measure hours, I do not opt for usual hours per week from outgoing rotation groups in the earner study or usual hours per week at all jobs from the BMS. The former measure is limited to workers paid hourly, and the latter measure does not appear to align as closely with the earnings measure, which instead references a singular "job" rather than multiple [all] "jobs."

Before that change, response options for race identified only specific single races.[4] Due to the longitudinal design of the BMS and the 2003 changes in racial categories available, there are individuals who identified as multiracial in January 2003 or later but had been constrained to identify as a single race in December 2002 or earlier. This likely exogenous occurrence provides a unique opportunity to analyze the causal determinants of racial identity, specifically the impact of pre-2003 factors on the single-race survey choice for persons who identify as multiracial in the CPS from 2003 onward. Rather than concentrating on individual-level factors, I focus on market-level factors (pertaining to racial composition and labor market outcomes) to further aid the identification of causal effects on racial identity. Compared with other surveys, especially panel surveys, the relatively large sample size of the CPS and the focus of the BMS on labor market information also helps support such market-related measures. Nevertheless, the 16-month calendar span of the BMS limits the number of persons who bridge the January 2003 expansion of racial categories, and the intervening eight months out-of-sample for each worker complicates data construction and analysis.

2.2 Sample Restrictions and Selection

Using various sample restrictions, I try to drop from the sample persons with inconsistent or otherwise erroneous responses. Such restrictions include ensuring that, for each person, race does not vary over time in the response period before 2003 or the response period on or after January 2003.[5] For a given person, race is allowed to vary only between the last

[4]From 1976 through 1988, only the categories "white," "black," and "other" (single race, not classified elsewhere) were available in the BMS. Incorporating a 1988 change from the CPS Annual Social and Economic Supplement, the BMS added categories for "American Indian/Eskimo/Aleut" and "Asian or Pacific Islander" in 1989. The BMS dropped the "other" category in 1996, leaving four racial categories until 2003: "white," "black," "Asian or Pacific Islander," or "American Indian/Eskimo/Aleut." In this paper, I refer to the American Indian/Eskimo/Aleut category as "Native American." Starting in January 2003, the Asian or Pacific Islander group was replaced with two groups—"Asian" and "Hawaiian/Pacific Islander," and, crucially for this paper, respondents could specifically identify as more than one race. Those changes gave survey respondents a total of 21 racial categories from which they could select (although this total included two multiple-race categories—"two or three races, unspecified" and "four or five races, unspecified"—that did not identify specific races). The number of categories ultimately expanded to 26 in 2013 (Flood et al. 2020).

[5]Although the stable, within-period reported race may also be causally affected by market factors, I cannot credibly identify such potential effects.

available response of 2002 and the first available response of 2003. These sample restrictions for data quality and analysis are listed in Appendix Table A1. As expected, some restrictions are minimally binding, if at all, but are imposed for assurance purposes. For instance, the restriction to "Drop if state varies across time" is expectedly non-binding since the CPS does not follow respondents who move. After incorporating any relevant information for each person from their last month-in-sample (MIS; this ends up being MIS8 and in 2003, although these were not explicit restrictions), I focus on cross-sectional data in 2002 from MIS4 for each person, thus allowing hourly wages to be included in the analysis.

Additionally, given numerous sample restrictions, I create descriptive weights for all sample individuals. These weights incorporate both a CPS sample design weight and a post-stratification weight, with the latter intended to capture inadvertent sample selection along various dimensions, including sex, education, and area (see Appendix). With the weights applied, sample statistics for share measures reasonably reflect the national population of interest.[6] Regarding unweighted counts, the estimation sample contains 899 multiracial persons, including 61 persons who do not specify any particular multiple-race categories (again, based on 2003 survey responses). Much of the analysis, to be explained later, also incorporates as a control group "single-race-fluid" persons, whose reported single race changes from 2002 to 2003 but is consistent within each year. The estimation sample includes 1,439 such single-race-fluid individuals. Lastly, limited analysis, which also will be discussed later, further includes "single-race-consistent" persons, whose reported single race remains unchanged from 2002 through 2003. The estimation sample includes 73,128 such single-race-consistent individuals.

Figure 1 illustrates the extent to which the racial composition of the weighted estimation sample, either in total or when solely comprising multiracial persons, differs from that of the US population aged 16 years and older.[7] Racial composition for 1996 through 2002 displays

[6]In validity checks, the estimation sample with weights applied closely replicates targeted population statistics such as the share that is female (0.52 in 2002 according to census estimates and, likewise, estimated as 0.52 in the estimation sample with weights applied).

[7]To generate statistics that are representative of the US population, the first two sets of figure bars use

expected patterns: The majority of the population is White (83.3 percent), followed by the Black share (11.8 percent), the Asian or Pacific Islander share (4.0 percent), and the Native American share (0.9 percent). Restricting the year analyzed to 2002 does not have much effect on racial composition, nor does additionally imposing estimation-sample restrictions.[8] The latter result further supports the validity of the descriptive weights. However, focusing solely on multiracial persons and the single race they reported in 2002 (that is, dropping the single-race persons in 2002 who remained single-race in 2003) results in a distinct racial composition in Figure 1, with a notably smaller White share (62.6 percent) and distinctly larger shares for the Asian or Pacific Islander and Native American racial groups (12.3 percent and 13.1 percent, respectively). The Black share (12.0 percent) remains comparable to its other estimates in the figure.

3 Empirical Strategy

3.1 Multinomial Choice Model

To determine the impact of market factors on racial identity, I use a multinomial choice model estimated by maximum likelihood. Market factors and related interaction terms vary across each race alternative and are thus "alternative-specific," while the multiracial indicator and control measures vary by individual but not by alternatives and are thus "case-specific." As outlined by Cameron and Trivedi (2005), due to the presence of at least some alternative-specific measures, a conditional logit model can be applied (which could alternatively be referred to as a mixed logit model here, given the inclusion of some case-specific regressors).

Using the framework of an additive random-utility model, for individual i and race al-

person-level weights provided by IPUMS-CPS ($WTFINL$; see Appendix), while the second two sets of figure bars use the descriptive weights outlined earlier.

[8]The second set of figure bars reflects the following shares: White, 82.6 percent; Black, 11.8 percent; Asian or Pacific Islander, 4.5 percent; and Native American, 1.1 percent. The third set of bars reflects the following shares: White, 83.5 percent; Black, 11.3 percent; Asian or Pacific Islander, 4.2 percent; and Native American, 1.1 percent (which sums to 100.1 percent due to rounding error).

ternative j, where $i = 1, ..., N$ and $j = 1, ..., 4$ (White, Black, Asian or Pacific Islander, Native American), one can assume that utility, U_{ij}, is the sum of a deterministic component dependent on regressors and parameters, $V_{ij} = \mathbf{X}'_{ij}\beta + \mathbf{Z}'_i\gamma_j$, and an unobserved random component, ε_{ij}:

$$U_{ij} = \mathbf{X}'_{ij}\beta + \mathbf{Z}'_i\gamma_j + \varepsilon_{ij}.$$

Note that $\mathbf{X}_{ij}$ are alternative-specific regressors (market factors and their interaction terms), and $\mathbf{Z}_i$ are case-specific regressors (multiracial indicator and controls). The outcome $Y_i = j$ is observed if race alternative j has the highest utility among all the race alternatives, k.[9] This implies:

$$\begin{aligned}\Pr(Y_i = j|\mathbf{X}_{ij}, \mathbf{Z}_i) &= \Pr(U_{ij} > U_{ik}), \forall k \\ &= \Pr(U_{ik} - U_{ij} \leq 0), \forall k \\ = \Pr(\varepsilon_{ik} - \varepsilon_{ij} \leq [\mathbf{X}'_{ij}\beta + \mathbf{Z}'_i\gamma_j] - [\mathbf{X}'_{ik}\beta + \mathbf{Z}'_i\gamma_k]), &\forall k.\end{aligned}$$

If one further assumes that ε_{ij} follows a Type 1 extreme value distribution, then it is possible to specify the probability that outcome $Y_i = j$ is observed as:

$$\Pr(Y_i = j|\mathbf{X}_{ij}, \mathbf{Z}_i) = \frac{\exp(\mathbf{X}'_{ij}\beta + \mathbf{Z}'_i\gamma_j)}{\Sigma_{k=1}^{4}\exp(\mathbf{X}'_{ik}\beta + \mathbf{Z}'_i\gamma_k)}, \quad (1)$$

where, as shown, the probability in equation (1) is constrained to lie in the [0,1] interval, and, for identification, one of the γ_j terms is set to zero as a "base" (in this paper, the White race alternative). Focusing on alternative-specific regressors—since the paper centers on market factors—the derived "own" and "cross" partial derivatives of interest, respectively, are as follows for some market-factor regressor, X_r, with coefficient β_r and X_{rik} reflecting the value

[9] Studies like Akerlof and Kranton (2000), Bodenhorn and Ruebeck (2003), and Darity Jr., Mason, and Stewart (2006) provide alternative frameworks for incorporating identity into utility maximization.

of X_r for individual i and race alternative k:

$$\frac{\partial \Pr(Y_i = j|\mathbf{X}_{ij}, \mathbf{Z}_i)}{\partial X_{rik}} = \Pr(Y_i = j|\mathbf{X}_{ij}, \mathbf{Z}_i)[1 - \Pr(Y_i = j|\mathbf{X}_{ij}, \mathbf{Z}_i)]\beta_r, \quad \forall j = k, \tag{2}$$

and

$$\frac{\partial \Pr(Y_i = j|\mathbf{X}_{ij}, \mathbf{Z}_i)}{\partial X_{rik}} = -\Pr(Y_i = j|\mathbf{X}_{ij}, \mathbf{Z}_i)\Pr(Y_i = k|\mathbf{X}_{ij}, \mathbf{Z}_i)\beta_r, \quad \forall j \neq k. \tag{3}$$

The first partial derivative (own-effect) in equation (2) reflects the change in the probability of choosing race j if the value of market factor X_r increases by one unit for option j. This partial derivative has the same sign as the coefficient, β_r. The second partial derivative (cross-effect) in equation (3) has the opposite sign of the first partial derivative (and β_r) and reflects the change in the probability of choosing race j if the value of market factor X_r increases by one unit for option k not equal to j. When reporting conditional logit results in this paper, for brevity, I focus on own-effects averaged across individuals (average marginal effects).

3.2 Market Factors, Controls, and Control Group

For market factors, $\mathbf{X}_{ij}$, I focus on measures that have reasonably clear sign predictions and that align with data available in the CPS. These criteria result in three race-specific market factors related to demographics and the labor market: the population share (reflecting all persons of a given race and market, unless noted otherwise), the unemployment rate, and the log average hourly wage (with both labor market measures reflecting persons in the labor force who are aged 16 and older of a given race and market, unless noted otherwise).[10] As noted, the geographic size of a market in this paper is a state, although for limited indicated analysis, I also allow for alternative geographic boundaries (for instance, a region, equal to a census division). Assuming that increased exposure to a race (weakly) increases related

[10]Wages are converted to constant 1999 US dollars using the BLS Consumer Price Index for All Urban Consumers (CPI-U) and are further adjusted by adding $1 to levels before averaging and calculating log values.

affinity through effects that are internal (household, likely reflecting family) and/or external (state, likely reflecting peers), I anticipate $\beta_r \geq 0$. By contrast, since unemployed workers are seeking employment by definition, I anticipate a higher race-specific unemployment rate to be perceived (weakly) negatively, consistent with $\beta_r \leq 0$. Lastly, all else being equal, I expect higher race-specific hourly wages to be viewed (weakly) positively given the increased consumption afforded (including leisure if labor supply bends backward), aligning with $\beta_r \geq 0$. As mentioned, (average) own-effects have those same sign predictions.[11]

Given the known CPS-based source of the timing and reasoning for multiracial persons identifying as a single race before 2003, combined with market-level regressors being potentially uncorrelated with any omitted individual-level measures, there may be no reason for concern about biased estimation of causal effects. Still, conservatively, I incorporate control variables and a control group, respectively, into some estimation. For control variables, $\mathbf{Z}_i$, I first try to account for life events that might have coincided with when the respondent was first able to identify as multiracial, thus possibly serving as the source of the change in reported race. When included in estimation, these life-event controls reflect a change in age, marital status, citizenship, family size, educational attainment, employer, employment or labor force status, or health (as captured by information about reasons for job-search inactivity or activity while not in the labor force). Additionally, conditional on inclusion of a control group (discussed later), I sometimes incorporate additional control variables that help control for either the person's contribution to market factors (namely an indicator for being employed and the log hourly wage in constant 1999 US dollars) or observed significant differences between the treatment group (multiracial) and the control group. (Single-race-fluid persons, as defined later. The controls are indicators for [1] education: some college [including associate degree] and college or more; [2] region: Middle Atlantic, West North Central, South Atlantic, East South Central, and West South Central; [3] industry: pub-

[11] If individuals distinguish their racial identity at work from their racial identity at home or among peers, then labor market factors may affect the former while demographic factors affect the latter. In such a case, the identity reflected by the reported race in the CPS would have effects aligning with the predictions, while the other identity would have attenuated effects due to measurement error.

lic administration; [4] occupation: operators, fabricators, and laborers; and [5] at least one parent in household.)

Although control variables can help account for possible observable sources of bias, some unobservable sources of bias might stem from unknown phenomena that coincide with the timing of the CPS change in race options. Such unobservable phenomena are captured by individuals who change from one single race to another single race when the expanded survey race options first become available to them in 2003. Specifically, such single-race-fluid persons report race changes from 2002 to 2003 but consistently report race within each year, as defined earlier.[12] They were not constrained by the available survey options and could have chosen the selected alternative race at any time, which seemingly indicates a change in some underlying phenomena that aligns with the timing of the survey change. I use these individuals as a control group in some specifications, thereby focusing on market factors interacted with the multiracial indicator as the key regressors.

3.3 Assessment of Omitted Variable Bias

A potential implication of the analysis in this paper is that race, if determined partly by market factors, may be endogenous in some analysis, resulting in biased estimation. To examine this possibility, I run ordinary least squares (OLS) estimation of individual-level labor market outcomes—namely unemployment status and log average hourly wages—on race indicators, with and without controls. I then examine whether the coefficients on the race indicators change substantially when omitted race-specific market factors are included. Such a result would suggest the presence of omitted variable bias when the market factors are excluded from estimation. For this analysis, in addition to multiracial and single-race-fluid persons, I incorporate "single-race-consistent" persons who, as defined earlier, have a reported race that remains unchanged from 2002 through 2003. Inclusion of such persons allows the sample to align more closely with samples typically used to estimate such racial

[12] Allowance of only one race change, in addition to linking methods by IPUMS-CPS, help support that the same person is measured over time (Drew, Flood, and Warren 2014).

disparities in labor market outcomes at the individual level.

4 Descriptive Patterns

Given this paper's focus on multiracial persons, it is helpful to assess how their characteristics differ from those of single-race persons, as Figure 1 shows to an extent. Table 1 examines the traits of single-race persons compared with multiracial persons in the estimation sample and indicates differences across several dimensions. Notably, multiracial individuals in the sample are younger, less likely to complete high school or college, more likely to live in the Pacific region or West South Central region, and more likely to have at least one parent in the household.

Another key interpretation issue for the paper's analysis is whether the single-race responses that multiracial persons gave in 2002 were informative of their preferred racial identity in 2003. If not, then any impact of market factors on such single-race responses does not provide insight into the determinants of racial identity.[13] Table 2 compares the reported race of multiracial persons in 2002 and 2003. In 2002, each person reports only one of the four racial groups, while in 2003, each person reports two or more of the racial groups.[14] Thus, a cell value in the table indicates the percentage of multiracial persons of a given race in 2002 (row) whose reported races in 2003 include the given race (column). Accordingly, cell values in a row do not sum to 100.0 percent. I also drop the 61 multiracial persons who do not specify race in 2003.[15]

The diagonal elements of the table show a significant persistence in race across years. This indicates that the race each multiracial person reported in 2002 is likely to be one

[13]For reasons noted earlier, I assume that the same person is measured over time, even in the absence of a strong relationship between reported race in 2002 compared with 2003.

[14]The exception is that "Asian or Pacific Islander" is listed in the CPS as two distinct racial groups in 2003 ("Asian" and "Hawaiian/Pacific Islander") and thus, on its own, may reflect multiracial identification in that year.

[15]Asian or Pacific Islander respondents account for a disproportionate 49 of the 61 multiracial persons with race unspecified in 2003 (80.3 percent), compared with 10 White respondents (16.4 percent) and one Black and one Native American respondent (each 1.6 percent), which sums to 99.9 percent due to rounding error.

of the races they reported in 2003. Specifically, the probabilities are as follows, by race: White, 98.5 percent; Black, 98.8 percent; Asian or Pacific Islander, 95.8 percent; and Native American, 100.0 percent. Thus, although responses in 2003 suggest that the constrained, single-race responses by multiracial persons in 2002 are likely not preferred, those 2002 responses nevertheless appear informative, as they are highly likely to reflect one of the races that multiracial respondents identify as in 2003.

Using scatter plots, Figure 2 depicts the association between the share of multiracial persons in a region (census division) reporting each single-race category in 2002 and values of each of the three market factors (averaged for unemployment rates and log wages), likewise by race and region.[16] Results in the figure generally support the relationships predicted earlier between market factors and the race of multiracial persons in 2002. In the plot for the first market factor, population shares by race and region, the fitted slope for every race is positive, as predicted. For instance, a 1 percentage point increase in a region's population share of Black persons is associated with a 0.7 percentage point increase in the share of multiracial persons in that region who identify as Black in 2002. Patterns in the chart for the second market factor, (average) unemployment rates by race and region, mostly align with predictions as well, with three of the four races displaying a negative relationship. For example, a 1 percentage point increase in the regional unemployment rate for Asian or Pacific Islander individuals is associated with a 1.8 percentage point decrease in the share of multiracial persons in the region who identify as Asian or Pacific Islander in 2002. Lastly, turning to the plot for the third market factor, (average) log average hourly wages (1999 US dollars), patterns for two of the four races display the predicted positive relationship. For instance, a one unit increase in the log average hourly wage for White individuals in a region is associated with a 0.5 percentage point increase in the share of multiracial persons in the region who identify as White in 2002.

[16] Although the causal analysis focuses primarily on states as the geographic boundary of a market, it is helpful to examine, in Figure 2, patterns at a higher level of aggregation to reduce the number of data points (similarly to the result of binned scatter plots).

5 Main Results

Turning to causal analysis, Table 3 examines differences in characteristics between the treatment group (multiracial persons) and the control group (single-race-fluid persons) that is sometimes incorporated into the analysis. Observed patterns are similar to those in Table 1 regarding traits that are significantly different between the two groups. As noted, in addition to life-event controls and controls for the person's contribution to market factors (again, an indicator for employment and the log hourly wage in constant 1999 US dollars), I also allow for controls for the measures that show significant differences between the treatment group and the control group in Table 3. This approach is in lieu of adding all variables listed in the table due to estimation challenges from the resulting loss in degrees of freedom.

Table 4 displays results from maximum likelihood estimation of a conditional logit model of the impact of state-level market factors on racial identity. As mentioned, for simplicity and brevity, I report own-effects (as average marginal effects) for the interaction between market factors and the multiracial indicator. Column (1), which does not weight estimation, reveals significantly positive effects on the reported race in 2002 for multiracial persons, stemming from increases in the race-specific market population shares and log wages. For instance, increasing the Black population share by 1.2 percentage points, or 10 percent of the mean Black share for multiracial persons, raises the probability that the reported race is Black by $(0.270 \times 1.2)/100 = 0.003$, or 0.3 percentage point. Alternatively, raising the Asian or Pacific Islander log wage by 0.15 units, or 10 percent of the mean log wage for multiracial persons, increases the probability that the reported race is Asian or Pacific Islander by $(0.020 \times 0.15)/1 = 0.003$, or 0.3 percentage point once again.

Results remain broadly similar once descriptive weights are incorporated in column (2), although there is no longer a significant own-effect of increases in the log wage on the reported race.[17] Adding life-event controls in column (3) does not noticeably change the

[17] I incorporate weights to adjust the contribution of each individual i to the likelihood function. All estimation also implements heteroskedasticity robust standard errors. However, since unaddressed heteroskedasticity in maximum likelihood estimation causes inconsistent estimates of coefficients as well as standard

average marginal effects, indicating minimal bias from not accounting for these observable life changes. However, incorporating a control group of single-race-fluid persons in column (4) notably reduces the size of the (differential) market-factor own-effects, although they remain highly significant in the case of the population share. Lastly, in the preferred specification reflected by column (5), adding additional controls (to account for the person's contribution to market factors and significant differences between the treatment group and control group) increases the magnitude of the own-effects for the population-share and unemployment-rate market factors. However, only the former remains significant—and robustly so across all five specifications in Table 4, consistent with Figure 2.

6 Additional Findings

6.1 Robustness

Table 5 examines various deviations from Table 4, column (5), to determine the robustness of the estimated effects. Column (1) shows that enlarging the geographic boundary of the market factors to regions results in reduced estimate precision and no significant market-factor own-effects.[18] Previous research outlines differences between types of racial identity. For instances, an "internal racial identity" could reflect what a person believes about their own race, while an "external racial identity" could reflect what others believe about a person's race (Harris and Sim 2002). Earlier estimates may reflect a mix of both types of racial identity since the reference person is not always the respondent. Thus, to reflect racial self-identity, columns (2) and (3) restrict the sample to instances when individuals in either the treatment group (column 2) or both the treatment and control groups (column 3) are the survey respondents, not just the reference persons. In both cases, the results are fairly

errors, I acknowledge that heteroskedasticity robust standard errors do not address the former bias. Thus, these standard errors serve as a safeguard in the event of heteroskedasticity that does not severely bias the coefficient estimates when robust inference may still be worthwhile.

[18] I do not pursue such estimation at the level of a metropolitan area or city due to data limitations.

similar to those in column (5) of Table 4, with comparably sized, significant own-effects for the population-share market factor.

Because it may take time to learn market-factor information, column (4) explores how such factors, lagged by one month, affect the chosen racial group. While own-effects for the population share and log wage remain similar to the estimates in Table 4, column (5), there is now a significantly negative impact of race-specific market unemployment rates on the reported race. For instance, according to mean values of the employed and labor force indicators for multiracial persons in Table 3, the estimated unemployment rate for this group is about $1 - (0.628/0.681) \times 100 = 7.8$ percent. Increasing the Native American lagged unemployment rate by 0.8 percentage point, or about 10 percent of the mean unemployment rate for multiracial persons, reduces the probability that the reported race is Native American by $(-0.502 \times 0.8)/100 = -0.004$, or 0.4 percentage point. Thus, although racial-composition market information affects reported race whether such information is contemporaneous or lagged by one month, some labor market information (unemployment rates) affects reported race only with a lag. Additionally, across the examples highlighting own-effect size, the effect magnitudes are similar for all three market factors. Column (5) of Table 5 shows that averaging market-factor information over all four pre-2003 months available for a person leads to results that are similar to those in Table 4, column (5), suggesting that incorporating contemporaneous or less recent unemployment rate conditions mitigates the impact on reported race.

Lastly, the significant impact of market population shares on reported race may operate through external peer effects, internal household effects (that is, persons residing in the same home), or some combination of the two mechanisms, as also noted by Harris and Sim (2002). The potential contribution of household effects may help explain the existence of significant contemporaneous own-effects for only the population-share market factor, as such information may be known and not require time to learn, as external peer information might. To help separate the external and internal mechanisms, column (6) of Table

5 adds lagged household-level population shares by race and also interacts them with the multiracial indicator. The results in the table show that when these household measures are controlled for, the magnitude of the market-level own-effect is an order of magnitude smaller and less statistically significant. Thus, the effect of race-specific market population shares on reported race seems largely driven by household racial composition, suggesting that the internal mechanism may be relatively more salient than the external mechanism for determining racial identity.[19]

6.2 Heterogeneity

The analysis thus far assumes that all individuals of a given race and state define the reference frame for race-specific market factors to affect the reported race of multiracial persons. However, the reference group with which individuals compare themselves may be narrower, thus altering the relevant dimensions that market factors and respondents should reflect. Table 6 thus examines the heterogeneous impact of market factors on reported race. The listed subgroup for each specification indicates the applicable persons in the treatment and control groups, as well as the applicable persons used to construct the race-specific lagged market factors.

Compared with the estimates in Table 5, column (4), the gender-stratified estimates in columns (1) and (2) of Table 6 show relatively larger lagged population-share own-effects for women than men, in addition to showing that lagged unemployment-rate own-effects seem to be driven entirely by women. Additionally, the reported race of both multiracial men and multiracial women displays a significantly positive own-response to lagged race-specific market log wages by gender, aligning with predictions, whereas there is no significant response to log wages in Table 5. Thus, gender seems to be a relevant reference frame when

[19] A lagged household-level log wage cannot be constructed since the necessary earnings information is available only in MIS4, and, unlike at the state level, all persons at the household level have MIS3 information in the month before the MIS4 information used for analysis. Lagged household-level unemployment rates have limited identifying variation (since not every household member is in the labor force) and yield seemingly spurious positive effects on reported race when estimated.

estimating how (lagged) market log wages affect reported race, perhaps due to market wages for men compared with women.

Columns (3) and (4) of the table stratify own-effects by age. Lagged market factors generally do not have a significant impact on reported race for younger multiracial persons aged 16 through 24 years, with the sole exception of the own-effect for the Asian or Pacific Islander lagged population share. The absence of such effects seems at least partly due to a relatively small sample, with fewer than 200 persons for both the treatment and control groups and resultant decreases in estimate precision. By contrast, for multiracial persons aged 25 years and older, own-effects for the lagged population share and lagged unemployment rate are significant and have the expected signs.

Lastly, columns (5) and (6) of Table 6 stratify own-effects by educational attainment, distinguishing between high school (diploma or equivalent, including persons "not in universe" or with missing responses) or less compared with some college (including associate degree) or more, respectively. For both groups, there are significantly positive own-effects of the lagged population share on reported race, with generally larger effects for the some-college-plus specification. However, with only the high-school-or-less specification are significant own-effects with expected signs additionally observed for the lagged unemployment rate and lagged log wage.

6.3 Bias Implications

Given the estimated impact that lagged market factors have on racial identity, race may be endogenous in some analysis. Namely, in estimation of how race affects individual-level labor market outcomes for unemployment and log wages, the omission of race-specific market factors for population shares, unemployment rates, and log wages may lead to biased estimation. Such bias would occur if the omitted market factors are correlated with the included race indicators (which this paper's analysis suggests is true, at least for multiracial persons in 2002) and if the omitted market factors have an impact on the individual-level

labor market outcomes, conditional on other included regressors.

To examine this possibility, Tables 7 and 8 run the aforementioned regressions in 2002 for a multiracial cross-sectional sample and an overall cross-sectional sample. The regressions focus on the extent of racial disparities in outcomes for Black, Asian or Pacific Islander, and Native American persons, as compared with White persons (the omitted racial category), with and without control measures and, crucially in each case, with and without the market factors. Table 7 shows that across all specifications without lagged market factors, whether the outcome is unemployment or log wages and whether controls are included or excluded, none of the race indicators differs significantly from zero. The lack of significant disparities with the omitted White category seems due, at least in part, to a relatively small sample size, as reflected by imprecise estimates compared with Table 8. Thus, although including lagged market factors in estimation results in revised estimates of racial disparities in unemployment or wages, one cannot rule out the equivalence of such disparity estimates with and without lagged market factors. Hence, Table 7 suggests the presence of minimal, if any, omitted variable bias from the exclusion of the market factors.

Table 8 includes evidence of significant disparities in unemployment and wages for the included racial groups compared with the omitted White racial group. For instance, in column (3), where controls are included in estimation but market factors are not, being Black is associated with a 0.7 percentage point greater probability of being unemployed compared with being White, and in column (7), it is associated with a $(1 - \exp[-0.116] \times 100) = 11.0$ percent lower wage. However, for these racial disparities and the others in the table, with and without controls, the inclusion of lagged market factors in estimation typically changes the race coefficients of interest only slightly, if at all. Thus, once again, one cannot rule out the equivalence of disparity estimates with and without lagged market factors. Like Table 7, Table 8 indicates the existence of minimal, if any, omitted variable bias from the exclusion of the market factors.

7 Conclusion

This study estimates the impact of market factors on racial identity. Leveraging the 2003 introduction of multiple-race response options to the US Current Population Survey, I examine how the pre-2003, constrained, single-race responses of multiracial persons are affected by variation in state-level population shares by race, race-specific state-level average unemployment rates, and race-specific state-level average log wages. I find the most robust evidence that market racial composition, whether contemporaneous or lagged, affects racial identity, driven mostly by the racial composition of multiracial respondents' households rather than their market peers. However, I find that lagged race-specific market unemployment rates also affect racial identity, as do lagged race-specific market wages when pertaining to narrower subgroups defined by gender and education. These findings perhaps reflect that respondents need time to learn about labor market conditions. Additionally, although the results suggest the endogeneity of race and possible biased estimation when analyzing how race affects individual-level labor market outcomes such as unemployment or wages, I find negligible evidence of such bias.

These findings suggest that in individual-level estimation of racial disparities in labor market outcomes, there is reason to be encouraged that such inclusion of race measures as regressors is econometrically permissible. However, given the relatively small population and period of focus to identify the effects in this paper (that is, multiracial persons in 2002), some caution is warranted regarding the validity of the findings to broader populations and periods. Additionally, the results in this paper suggest that as surveys evolve and survey designs for eliciting race responses change, respondents may react strategically to constrained response options, perhaps in unanticipated ways. To the extent that such strategic behavior has implications for the accurate measurement of demographic trends and related public policies, it should be kept in mind, if not addressed through survey adjustments.

References

Akerlof, George A., and Rachel E. Kranton. 2000. "Economics and Identity." *Quarterly Journal of Economics* 115(3): 715–753.

Akerlof, George A., and Rachel E. Kranton. 2002. "Identity and Schooling: Some Lessons for the Economics of Education." *Journal of Economic Literature* 40: 1167–1201.

Akerlof, George A., and Rachel E. Kranton. 2005. "Identity and the Economics of Organizations." *Journal of Economic Perspectives* 19(1): 9–32.

Bodenhorn, Howard, and Christopher S. Ruebeck. 2003. "The Economics of Identity and the Endogeneity of Race." NBER Working Paper No. 9962. Cambridge, MA: National Bureau of Economic Research.

Cameron, A. Colin, and Pravin K. Trivedi. 2005. *Microeconometrics: Methods and Applications.* Cambridge, UK: Cambridge University Press.

Dahis, Ricardo, Emily Nix, and Nancy Qian. 2020. "Choosing Racial Identity in the United States, 1880–1940." Working Paper.

Darity Jr., William A., Patrick L. Mason, and James B. Stewart. 2006. "The Economics of Identity: The Origin and Persistence of Racial Identity Norms." *Journal of Economic Behavior and Organization* 60: 283–305.

Davenport, Lauren D. 2016. "The Role of Gender, Class, and Religion in Biracial Americans' Racial Labeling Decisions." *American Sociological Review* 81(1): 57–84.

Doyle, Jamie Mihoko, and Grace Kao. 2007. "Are Racial Identities of Multiracials Stable? Changing Self-identification among Single and Multiple Race Individuals." *Social Psychology Quarterly* 70(4): 405–423.

Drew, Julia A. Rivera, Sarah Flood, and John Robert Warren. 2014. "Making Full Use of the Longitudinal Design of the Current Population Survey: Methods for Linking Records across 16 Months." *Journal of Economic and Social Measurement* 39: 121–144.

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, and John Robert Warren. 2020. "Integrated Public Use Microdata Series, Current Population Survey: Version 8.0" [dataset]. Minneapolis, MN: IPUMS. https://doi.org/10.18128/D030.V8.0.

Harris, David R., and Jeremiah Joseph Sim. 2002. "Who Is Multiracial? Assessing the Complexity of Lived Race." *American Sociological Review* 67(4): 614–627.

Hitlin, Steven, Glen H. Elder, Jr., and J. Scott Brown. 2006. "Racial Self-categorization in Adolescence: Multiracial Development and Social Pathways." *Child Development* 77(5): 1298–1308.

Jackson, Osborne. 2021. "Job Displacement and Sectoral Mobility." Research Department Working Paper 21-19. Boston, MA: Federal Reserve Bank of Boston.

Liebler, Carolyn A., Sonya R. Porter, Leticia E. Fernandez, James M. Noon, and Sharon R. Ennis. 2017. "America's Churning Races: Race and Ethnicity Response Changes between Census 2000 and the 2010 Census." *Demography* 54: 259–284.

Rademakers, Robert, and Andre van Hoorn. 2021. "Ethnic Switching: Longitudinal Evidence on Prevalence, Correlates, and Implications for Measuring Ethnic Segregation." *Journal of Development Economics* 152: 1–16. https://doi.org/10.1016/j.jdeveco.2021.102694.

Reece, Robert L. 2019. "Coloring Racial Fluidity: How Skin Tone Shapes Multiracial Adolescents' Racial Identity Changes." *Race and Social Problems* 11: 290–298.

Saperstein, Aliya, and Aaron Gullickson. 2013. "A 'Mulatto Escape Hatch' in the United States? Examining Evidence of Racial and Social Mobility during the Jim Crow Era." *Demography* 50: 1921–1942.

Saperstein, Aliya, and Andrew M. Penner. 2012. "Racial Fluidity and Inequality in the United States." *American Journal of Sociology* 113(3): 676–727.

Sims, Jennifer Patrice. 2016. "Reevaluation of the Influence of Appearance and Reflected Appraisals for Mixed-race Identity: The Role of Consistent Inconsistent Racial Perception." *Sociology of Race and Ethnicity* 2(4): 569–583.

US Census Bureau. 2006. "Current Population Survey: Design and Methodology, Technical Paper 66." https://www.census.gov/prod/2006pubs/tp-66.pdf (October).

Xie, Yu, and Kimberly Goyette. 1997. "The Racial Identification of Biracial Children with One Asian Parent: Evidence from the 1990 Census." *Social Forces* 76(2): 547–570.

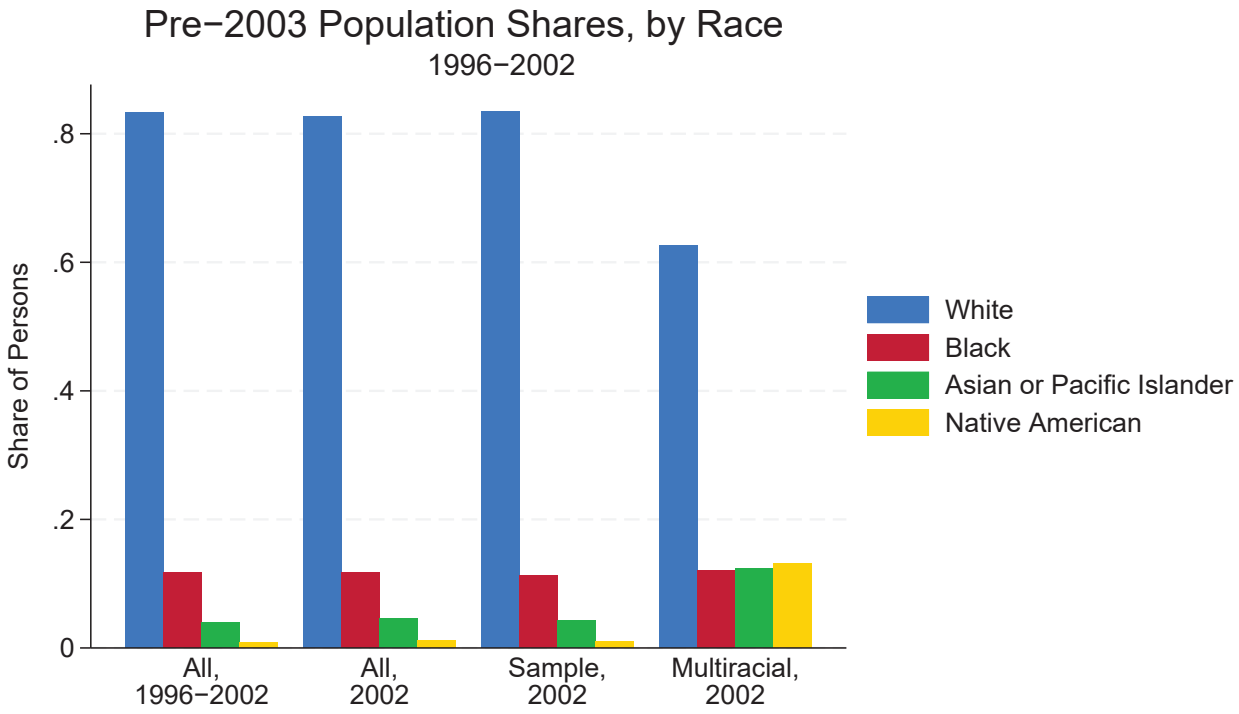

Figure 1: Sample Selection of Multiracial Persons, by Pre-2003 Race
Sources: 1996–2002 Current Population Survey data and author's calculations.

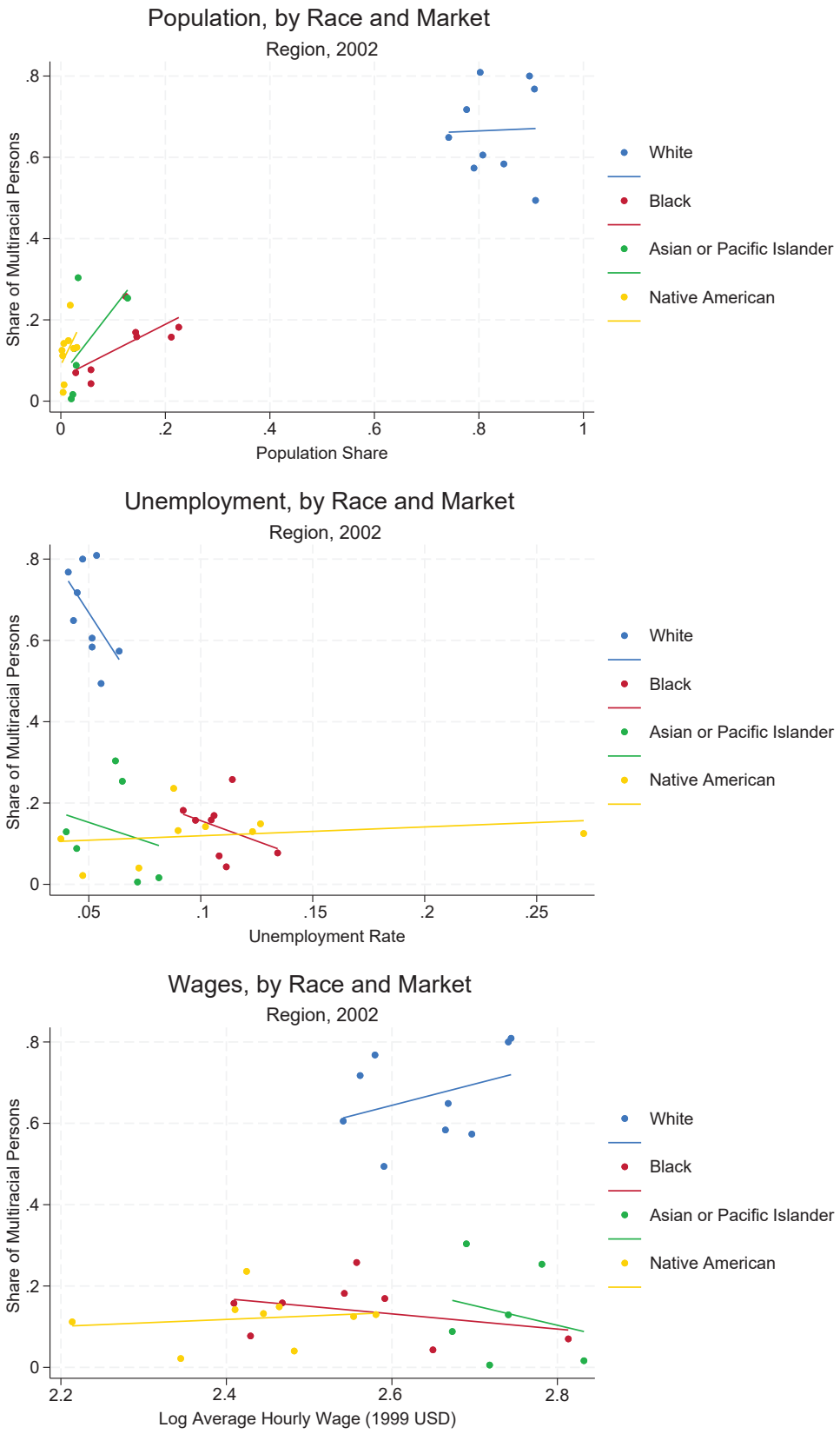

Figure 2: 2002 Race of Multiracial Persons versus Market Factors
Sources: 2002 Current Population Survey data and author's calculations.

Table 1: Characteristics of Single-Race versus Multiracial Persons, 2002

Indicator	Mean (Single Race)	Mean (Two or More Races)	Difference
Female	.520	.524	.004
Married, spouse present or absent	.594	.505	–.089***
Any own children in household	.438	.446	.008
Any own children less than age 5 in household	.128	.145	.018
Age 16-24	.147	.217	.070***
Age 25-54	.594	.595	.001
Age 55+	.259	.187	–.071***
Less than high school	.187	.226	.038**
High school	.321	.299	–.021
Some college	.254	.316	.061***
College	.237	.159	–.078***
White	.837	.626	–.211***
Black	.113	.120	.008
Asian or Pacific Islander	.041	.123	.082***
Native American	.009	.131	.122***
New England	.052	.029	–.023***
Middle Atlantic	.145	.061	–.084***
East North Central	.164	.105	–.058***
West North Central	.070	.068	–.002
South Atlantic	.184	.135	–.049***
East South Central	.059	.049	–.011
West South Central	.106	.168	.062***
Mountain	.064	.074	.010
Pacific	.155	.310	.155***
Missing industry	.288	.290	.002
Agriculture, forestry, and fishing	.012	.015	.004
Mining	.003	.006	.003
Construction	.041	.050	.009
Manufacturing	.111	.090	–.021*
Transportation, communication, and other utilities	.055	.051	–.004
Wholesale and retail trade	.144	.161	.017
Finance, insurance, and real estate	.045	.032	–.012*
Various services	.263	.246	–.018
Public administration	.038	.059	.021**
Missing occupation	.288	.290	.002
Managerial and professional specialty	.217	.178	–.039**
Technical, sales, and administrative support	.212	.214	.002
Service	.093	.116	.024*
Farming, forestry, and fishing	.013	.015	.002
Precision production, craft, and repair	.072	.090	.018
Operators, fabricators, and laborers	.105	.097	–.008
Reference person is household respondent	.504	.492	–.012
At least one parent in household	.139	.200	.061***
Labor force	.686	.681	–.005
Employed	.650	.628	–.022
Weekly earnings, constant 1999 USD (log)	4.023	3.802	–.221*
Hourly wage, constant 1999 USD (log)	1.640	1.505	–.136**
Usual hours worked per week at main job	23.678	22.076	–1.602*
Number of persons	74,567	899	.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Author's calculations using Current Population Survey data from 2002. Multiracial persons are identified using 2003 information, as pre-2003 information is single race only. Earnings and wages are adjusted by adding $1 (constant 1999) to levels before calculating log values. Means are weighted using descriptive weights.

Table 2: Comparing the Reported Race of Multiracial Persons in 2002 and 2003

	Race in 2003 (%)			
Measure	White	Black	Asian or P.I.	Native American
Race in 2002 (%)				
White	98.5	11.7	14.4	76.5
Black	45.2	98.8	2.4	66.7
Asian or P.I.	57.1	5.0	95.8	4.2
Native American	99.1	0.9	0.0	100.0

Notes: Author's calculations using Current Population Survey data from 2002–2003. Multiracial persons are identified using 2003 information, as pre-2003 information is single race only. Each cell displays the percentage of multiracial persons reporting the listed race in 2002 who subsequently reported the listed race in 2003. As indicated, racial groups are White, Black, Asian or Pacific Islander, and Native American. In 2002, each multiracial person reports one of the listed four racial groups. In 2003, each multiracial person reports at least two of the listed four racial groups, with the exception that "Asian or Pacific Islander" is listed as two distinct racial groups in 2003 ("Asian" and "Hawaiian/Pacific Islander") and therefore, on its own, may reflect multiracial identification in that year. Thus, cell values in a row do not sum to 100.0 percent. I exclude from the sample 61 persons who do not specify any multiple-race categories in 2003. Estimates are based on unweighted counts from a sample of 838 multiracial persons (899 persons – 61 persons).

Table 3: Estimation Sample Characteristics

Indicator	Overall Mean (All Persons)	Treatment Mean (Multi-racial)	Control Mean (Single-Race-Fluid)	Difference (Treatment – Control)
Female	.520	.524	.498	.025
Married, spouse present or absent	.593	.505	.543	–.038
Any own children in household	.438	.446	.486	–.040
Any own children less than age 5 in household	.128	.145	.172	–.027
Age 25-54	.594	.595	.625	–.029
Age 55+	.258	.187	.189	–.001
High school	.320	.299	.281	.018
Some college	.255	.316	.235	.081***
College	.237	.159	.220	–.060***
Black	.113	.120	.281	–.161***
Asian or Pacific Islander	.042	.123	.188	–.064***
Native American	.011	.131	.208	–.077***
Middle Atlantic	.144	.061	.167	–.106***
East North Central	.163	.105	.086	.019
West North Central	.070	.068	.032	.036***
South Atlantic	.184	.135	.198	–.063***
East South Central	.059	.049	.024	.024***
West South Central	.107	.168	.094	.074***
Mountain	.064	.074	.068	.006
Pacific	.157	.310	.286	.024
Agriculture, forestry, and fishing	.012	.015	.020	–.004
Mining	.003	.006	.001	.005
Construction	.042	.050	.043	.007
Manufacturing	.111	.090	.094	–.004
Transportation, communication, and other utilities	.055	.051	.058	–.007
Wholesale and retail trade	.144	.161	.165	–.005
Finance, insurance, and real estate	.045	.033	.037	–.004
Various services	.263	.246	.267	–.021
Public administration	.038	.059	.034	.025**
Managerial and professional specialty	.217	.178	.182	–.004
Technical, sales, and administrative support	.212	.214	.207	.007
Service	.093	.116	.114	.002
Farming, forestry, and fishing	.013	.015	.021	–.006
Precision production, craft, and repair	.072	.090	.071	.020
Operators, fabricators, and laborers	.105	.097	.125	–.027*
Reference person is household respondent	.504	.492	.460	.033
At least one parent in household	.139	.200	.159	.041**
Labor force	.686	.681	.692	–.012
Employed	.650	.628	.656	–.028
Weekly earnings, constant 1999 USD (log)	4.021	3.802	3.979	–.176
Hourly wage, constant 1999 USD (log)	1.639	1.505	1.607	–.102
Number of persons	75,466	899	1,439	.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Author's calculations using Current Population Survey data from 2002. Multiracial persons are identified using 2003 information, as pre-2003 information is single race only. The overall sample reflects multiracial persons (treatment group), single-race-fluid persons (control group), and single-race-consistent persons (excluded from estimation). Earnings and wages are adjusted by adding $1 (constant 1999) to levels before calculating log values. Omitted indicators are for age: 16–24; education: less than high school (diploma or equivalent, including not-in-universe or blank); race: White; region: New England; industry: agriculture, forestry, and fishing; and occupation: managerial and professional specialty. Means are weighted using descriptive weights.

Table 4: Impact of Market Factors on Racial Identity

	Unweighted	Weighted			
	(1)	(2)	(3)	(4)	(5)
'*Group*' *Pop. Share*					
× *Multiracial*					
White	0.586***	0.717***	0.706***	0.286***	0.346***
	(0.034)	(0.050)	(0.049)	(0.096)	(0.098)
Black	0.270***	0.359***	0.356***	0.225***	0.265***
	(0.031)	(0.045)	(0.043)	(0.076)	(0.076)
Asian or P.I.	0.282***	0.260***	0.260***	0.171***	0.201***
	(0.011)	(0.022)	(0.022)	(0.057)	(0.056)
Native American	0.327***	0.384***	0.374***	0.189***	0.217***
	(0.034)	(0.048)	(0.046)	(0.064)	(0.062)
'*Group*' *Unemp. Rate*					
× *Multiracial*					
White	0.101	–0.115	–0.112	0.080	0.182
	(0.127)	(0.192)	(0.193)	(0.242)	(0.261)
Black	0.046	–0.058	–0.057	0.063	0.139
	(0.058)	(0.096)	(0.097)	(0.191)	(0.200)
Asian or P.I.	0.048	–0.042	–0.041	0.048	0.106
	(0.062)	(0.069)	(0.071)	(0.145)	(0.152)
Native American	0.056	–0.062	–0.060	0.053	0.114
	(0.071)	(0.103)	(0.103)	(0.161)	(0.164)
'*Group*' *Log Wage*					
× *Multiracial*					
White	0.041*	0.052	0.049	0.032	0.031
	(0.022)	(0.037)	(0.037)	(0.043)	(0.046)
Black	0.019*	0.026	0.025	0.025	0.024
	(0.011)	(0.019)	(0.019)	(0.034)	(0.035)
Asian or P.I.	0.020*	0.019	0.018	0.019	0.018
	(0.011)	(0.014)	(0.014)	(0.026)	(0.027)
Native American	0.023*	0.028	0.026	0.021	0.019
	(0.013)	(0.020)	(0.020)	(0.029)	(0.029)
Life-event Controls	No	No	Yes	Yes	Yes
Other Controls	No	No	No	No	Yes
Control Group	No	No	No	Yes	Yes
McFadden's R-squared	0.193	0.118	0.133	0.077	0.109
Number of Observations	3,323	3,323	3,323	8,751	8,751
Number of Multiracial Persons	893	893	893	893	893
Number of Single-Race-Fluid Persons	.	.	.	1,432	1,432

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Author's calculations using Current Population Survey data from 2002. Multiracial persons are identified using 2003 information, as pre-2003 information is single race only. Each column displays a set of average marginal effects from estimation by maximum likelihood of a single conditional logit model of reported race in 2002, as noted. Each estimate shown thus reflects the average change in the probability of choosing the indicated race option from a one-unit increase in the corresponding "Group" market factor × a multiracial indicator. A Group is one of the following racial groups: White, Black, Asian or Pacific Islander, or Native American. A state-level market factor is a population share, an unemployment rate, or a log average hourly wage, with wages adjusted by adding $1 (constant 1999) to levels before averaging and calculating log values. Each observation reflects one outcome alternative for a person. In addition to the displayed "own-race" effects, "cross-race" effects are estimated as part of each specification but are omitted for brevity, as are controls. Estimation also includes distinct regressors for the multiracial indicator and each market factor, with the exception of models (1) through (3) in which those measures cannot be separately identified from the interaction terms of interest due to collinearity given a sample of only multiracial persons and no control group. In models (4) and (5), the control group is single-race-fluid persons whose reported race changes from 2002 to 2003 but is consistent within each year. When included, life-event controls reflect a change in age, marital status, citizenship, family size, educational attainment, employer, employment or labor force status, or health (as captured by information about reasons for job-search inactivity or activity while not in the labor force). When included, other controls reflect additional measures that help control for either the person's contribution to market factors (namely an indicator for employment and the log hourly wage in constant 1999 US dollars) or observed significant differences between the treatment and control groups (namely indicators for education: some college [including associate degree] and college or more; region: Middle Atlantic, West North Central, South Atlantic, East South Central, and West South Central; industry: public administration; occupation: operators, fabricators, and laborers; and at least one parent in household). Models (2) through (5) incorporate descriptive weights. Heteroskedasticity robust standard errors are in parentheses.

Table 5: Robustness of the Impact of Market Factors on Racial Identity

	(1) Region Market Factors	(2) Respondent Treatment	(3) Respondent Treatment and Control	(4) Lagged Market Factors	(5) Averaged Market Factors	(6) Household Pop. Share Added
'*Group*′ *Pop. Share* × *Multiracial*						
White	0.269	0.348***	0.380***	0.331***	0.341***	0.062*
	(0.248)	(0.114)	(0.135)	(0.100)	(0.100)	(0.033)
Black	0.200	0.285***	0.312***	0.251***	0.261***	0.027*
	(0.184)	(0.094)	(0.113)	(0.077)	(0.077)	(0.014)
Asian or P.I.	0.161	0.214***	0.201***	0.190***	0.198***	0.031*
	(0.149)	(0.069)	(0.071)	(0.057)	(0.057)	(0.017)
Native American	0.169	0.224***	0.229***	0.202***	0.222***	0.046*
	(0.156)	(0.074)	(0.083)	(0.062)	(0.066)	(0.025)
'*Group*′ *Unemp. Rate* × *Multiracial*						
White	0.424	0.283	0.086	–0.820***	–0.223	0.043
	(0.628)	(0.324)	(0.353)	(0.290)	(0.352)	(0.069)
Black	0.315	0.232	0.071	–0.622***	–0.171	0.019
	(0.467)	(0.265)	(0.291)	(0.221)	(0.269)	(0.030)
Asian or P.I.	0.254	0.174	0.045	–0.472***	–0.130	0.022
	(0.376)	(0.199)	(0.187)	(0.168)	(0.205)	(0.035)
Native American	0.266	0.183	0.052	–0.502***	–0.146	0.032
	(0.395)	(0.208)	(0.213)	(0.179)	(0.230)	(0.051)
'*Group*′ *Log Wage* × *Multiracial*						
White	0.126	0.028	–0.013	–0.001	–0.028	0.016
	(0.119)	(0.061)	(0.063)	(0.048)	(0.064)	(0.012)
Black	0.093	0.023	–0.011	–0.001	–0.022	0.007
	(0.088)	(0.050)	(0.052)	(0.037)	(0.049)	(0.005)
Asian or P.I.	0.075	0.017	–0.007	–0.001	–0.016	0.008
	(0.071)	(0.037)	(0.034)	(0.028)	(0.037)	(0.006)
Native American	0.079	0.018	–0.008	–0.001	–0.018	0.012
	(0.075)	(0.039)	(0.038)	(0.030)	(0.042)	(0.009)
McFadden's R-squared	0.092	0.086	0.122	0.113	0.113	0.889
No. of Observations	9,337	7,084	4,266	8,580	9,289	8,751
No. of Multiracial	898	449	449	894	899	893
No. of Single-Race-Fluid	1,439	1,432	693	1,393	1,439	1,432

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Author's calculations using Current Population Survey data from 2002. Multiracial persons are identified using 2003 information, as pre-2003 information is single race only. Each column displays a set of average marginal effects from estimation by maximum likelihood of a single conditional logit model of reported race in 2002, as noted. Each estimate shown thus reflects the average change in the probability of choosing the indicated race option from a one-unit increase in the corresponding "Group" market factor × a multiracial indicator. A Group is one of the following racial groups: White, Black, Asian or Pacific Islander, or Native American. A market factor is a population share, an unemployment rate, or a log average hourly wage, with the market at the state level unless otherwise noted. Wages are adjusted by adding $1 (constant 1999) to levels before averaging and calculating log values. Each observation reflects one outcome alternative for a person. In addition to the displayed "own-race" effects, "cross-race" effects are also estimated as part of each specification but are omitted for brevity, as are controls. All estimation includes a control group for single-race-fluid persons whose reported race changes from 2002 to 2003 but is consistent within each year. All estimation also includes: (1) distinct regressors for the multiracial indicator and each market factor; (2) life-event controls reflecting a change in age, marital status, citizenship, family size, educational attainment, employer, employment or labor force status, or health (as captured by information about reasons for job-search inactivity or activity while not in the labor force); and (3) other controls reflecting additional measures that help control for either the person's contribution to market factors (namely an indicator for being employed and the log hourly wage in constant 1999 US dollars) or observed significant differences between the treatment and control groups (namely indicators for education: some college [including associate degree] and college or more; region: Middle Atlantic, West North Central, South Atlantic, East South Central, and West South Central; industry: public administration; occupation: operators, fabricators, and laborers; and at least one parent in household). All estimation incorporates descriptive weights, and heteroskedasticity robust standard errors are in parentheses. Model (1) replaces state-level market factors with region-level market factors that reflect the nine census divisions. Models (2) and (3) restrict the treatment group (model 2) or both the treatment and control groups (model 3) to cases where the referenced person is the survey respondent. Model (4) lags market factors to reflect their value in the previous pre-2003 month for a person (month-in-sample 3). Model (5) averages market factors to reflect their mean value over all pre-2003 months for a person (months-in-sample 1 through 4). Model (6) adds race-specific population shares at the household level (individually and interacted with the multiracial indicator) to the existing race-specific population shares at the state level (included individually and interacted with the multiracial indicator).

Table 6: Heterogeneous Impact of Market Factors on Racial Identity

	(1) Male	(2) Female	(3) Age 16–24	(4) Age 25+	(5) HS or Less	(6) Some College+
'Group' Lagged Pop. Share × Multiracial						
White	0.243*	0.406***	0.385	0.355***	0.272**	0.336**
	(0.145)	(0.133)	(0.236)	(0.116)	(0.134)	(0.155)
Black	0.195*	0.293***	0.279	0.271***	0.192**	0.275**
	(0.117)	(0.098)	(0.173)	(0.089)	(0.096)	(0.128)
Asian or P.I.	0.126*	0.241***	0.223*	0.195***	0.110**	0.237**
	(0.075)	(0.078)	(0.134)	(0.063)	(0.054)	(0.107)
Native American	0.131*	0.221***	0.098	0.222***	0.166**	0.168**
	(0.078)	(0.073)	(0.060)	(0.073)	(0.083)	(0.078)
'Group' Lagged Unemp. Rate × Multiracial						
White	0.491	–1.405***	0.782	–0.686**	–0.601*	0.116
	(0.436)	(0.360)	(0.530)	(0.324)	(0.362)	(0.353)
Black	0.393	–1.013***	0.566	–0.525**	–0.424*	0.095
	(0.349)	(0.261)	(0.378)	(0.249)	(0.258)	(0.290)
Asian or P.I.	0.255	–0.835***	0.452	–0.377**	–0.243*	0.082
	(0.226)	(0.217)	(0.314)	(0.179)	(0.146)	(0.250)
Native American	0.264	–0.765***	0.199	–0.430**	–0.367*	0.058
	(0.236)	(0.200)	(0.134)	(0.205)	(0.221)	(0.177)
'Group' Lagged Log Wage × Multiracial						
White	0.127*	0.131*	–0.030	0.013	0.187**	0.069
	(0.072)	(0.077)	(0.105)	(0.054)	(0.087)	(0.050)
Black	0.102*	0.094*	–0.022	0.010	0.132**	0.057
	(0.057)	(0.055)	(0.076)	(0.042)	(0.062)	(0.041)
Asian or P.I.	0.066*	0.078*	–0.017	0.007	0.076**	0.049
	(0.038)	(0.046)	(0.060)	(0.030)	(0.036)	(0.035)
Native American	0.068*	0.071*	–0.008	0.008	0.114**	0.035
	(0.038)	(0.042)	(0.027)	(0.034)	(0.053)	(0.026)
McFadden's R-squared	0.159	0.115	0.196	0.119	0.146	0.099
No. of Observations	3,597	4,330	1,022	7,029	4,022	3,775
No. of Multiracial	390	484	159	715	447	433
No. of Single-Race-Fluid	622	738	181	1,167	718	634

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Author's calculations using Current Population Survey data from 2002. Multiracial persons are identified using 2003 information, as pre-2003 information is single race only. Each column displays a set of average marginal effects from estimation by maximum likelihood of a single conditional logit model of reported race in 2002, as noted. Each estimate shown thus reflects the average change in the probability of choosing the indicated race option from a one-unit increase in the corresponding "Group" lagged market factor × a multiracial indicator. A Group is one of the following racial groups: White, Black, Asian or Pacific Islander, or Native American. A state-level lagged market factor is a population share, an unemployment rate, or a log average hourly wage, with measures lagged to reflect their value in the previous pre-2003 month for a person (month-in-sample 3) and wages adjusted by adding $1 (constant 1999) to levels before averaging and calculating log values. Each observation reflects one outcome alternative for a person. In addition to the displayed "own-race" effects, "cross-race" effects are estimated as part of each specification but are omitted for brevity, as are controls. All estimation includes a control group for single-race-fluid persons whose reported race changes from 2002 to 2003 but is consistent within each year. All estimation also includes: (1) distinct regressors for the multiracial indicator and each market factor; (2) life-event controls reflecting a change in age, marital status, citizenship, family size, educational attainment, employer, employment or labor force status, or health (as captured by information about reasons for job-search inactivity or activity while not in the labor force); and (3) other controls reflecting additional measures that help control for either the person's contribution to market factors (namely an indicator for employment and the log hourly wage in constant 1999 US dollars) or observed significant differences between the treatment and control groups (namely indicators for education: some college [including associate degree] and college or more; region: Middle Atlantic, West North Central, South Atlantic, East South Central, and West South Central; industry: public administration; occupation: operators, fabricators, and laborers; and at least one parent in household). All estimation incorporates descriptive weights, and heteroskedasticity robust standard errors are in parentheses. State-level market factors are specific to indicated subgroups by gender (models 1 and 2), age (models 3 and 4), and education (models 5 and 6). Education (attainment) distinguishes between high school (diploma or equivalent, including persons "not in universe" or with missing responses) or less compared with some college (including associate degree) or more.

Table 7: Potential Bias from the Impact of Market Factors on Racial Identity, Multiracial

	(1) Unemp.	(2) Unemp.	(3) Unemp.	(4) Unemp.	(5) Log Wage	(6) Log Wage	(7) Log Wage	(8) Log Wage
Black	–0.018	–0.028	–0.012	–0.020	0.033	0.101	–0.037	–0.075
	(0.025)	(0.024)	(0.025)	(0.025)	(0.196)	(0.205)	(0.167)	(0.175)
Asian or P.I.	0.027	0.046	0.007	0.025	0.145	–0.091	0.244	0.103
	(0.037)	(0.059)	(0.041)	(0.063)	(0.171)	(0.261)	(0.161)	(0.231)
Native American	–0.004	–0.010	0.007	–0.003	0.268	0.291	0.169	0.175
	(0.030)	(0.029)	(0.029)	(0.027)	(0.190)	(0.196)	(0.179)	(0.183)
Black								
Lagged Pop. Share		0.324		0.412		–1.320		0.426
		(0.201)		(0.267)		(1.075)		(1.313)
Lagged Unemp.		0.344		0.290		–0.388		–0.078
		(0.229)		(0.216)		(1.252)		(1.108)
Lagged Log Wage		0.018		–0.005		0.099		0.168
		(0.020)		(0.020)		(0.185)		(0.161)
Asian or P.I.								
Lagged Pop. Share		–0.007		–0.005		0.540		0.494
		(0.078)		(0.081)		(0.388)		(0.342)
Lagged Unemp.		–0.085		0.025		2.326*		0.935
		(0.163)		(0.193)		(1.273)		(1.180)
Lagged Log Wage		–0.028		–0.035*		0.026		0.094
		(0.021)		(0.020)		(0.120)		(0.111)
Native American								
Lagged Pop. Share		0.179		0.312		–0.878		–0.717
		(0.316)		(0.296)		(1.724)		(1.672)
Lagged Unemp.		–0.011		–0.039		–0.271		–0.548
		(0.112)		(0.104)		(0.646)		(0.559)
Lagged Log Wage		0.022		0.026*		0.049		0.015
		(0.014)		(0.013)		(0.100)		(0.087)
Controls	No	No	Yes	Yes	No	No	Yes	Yes
Overall Y Mean	0.047	0.047	0.047	0.047	1.491	1.491	1.491	1.491
Multiracial Y Mean	0.047	0.047	0.047	0.047	1.491	1.491	1.491	1.491
R-squared	0.003	0.018	0.101	0.116	0.005	0.022	0.240	0.248
No. of Observations	677	677	677	677	677	677	677	677
No. of Multiracial	677	677	677	677	677	677	677	677

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Author's calculations using Current Population Survey data from 2002. Multiracial persons are identified using 2003 information, as pre-2003 information is single race only. Each column displays a set of coefficients from estimation by ordinary least squares of an individual-level labor market outcome in 2002, as noted. The outcomes are an indicator for unemployment or a log hourly wage, with the latter adjusted by adding $1 (constant 1999) to levels before calculating log values. Displayed regressors reflect indicators for Black, Asian or Pacific Islander, and Native American racial groups (the indicator for the White racial group is the omitted category). Corresponding state-level lagged market factors reflect the population share, unemployment rate, or a log average hourly wage, with measures lagged to reflect their value in the previous pre-2003 month for a person (month-in-sample 3) and wages adjusted by adding $1 (constant 1999) to levels before averaging and calculating log values. Each observation reflects a multiracial person. When included, controls are life-event measures reflecting a change in age, marital status, citizenship, family size, educational attainment, employer, employment or labor force status, or health (as captured by information on reasons for job-search inactivity or activity while not in the labor force); and other measures that control for observed significant differences between treatment and control groups in conditional logit models of reported race in 2002 (namely indicators for education: some college [including associate degree] and college or more; region: Middle Atlantic, West North Central, South Atlantic, East South Central, and West South Central; industry: public administration; occupation: operators, fabricators, and laborers; and at least one parent in household). All estimation incorporates descriptive weights, and heteroskedasticity robust standard errors are in parentheses.

Table 8: Potential Bias from the Impact of Market Factors on Racial Identity, Overall

	(1) Unemp.	(2) Unemp.	(3) Unemp.	(4) Unemp.	(5) Log Wage	(6) Log Wage	(7) Log Wage	(8) Log Wage
Black	0.016***	0.018***	0.007*	0.007*	–0.245***	–0.232***	–0.116***	–0.123***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.024)	(0.024)	(0.021)	(0.022)
Asian or P.I.	0.008	0.006	0.006	0.006	0.003	0.007	–0.124***	–0.123***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.038)	(0.041)	(0.035)	(0.037)
Native American	0.016	0.016	0.008	0.009	–0.145**	–0.113*	–0.019	0.008
	(0.011)	(0.011)	(0.010)	(0.010)	(0.066)	(0.067)	(0.060)	(0.060)
Black								
Lagged Pop. Share		–0.029**		0.001		–0.460***		0.116
		(0.014)		(0.018)		(0.104)		(0.126)
Lagged Unemp.		0.014		0.010		–0.018		–0.138
		(0.020)		(0.020)		(0.150)		(0.137)
Lagged Log Wage		0.004		0.002		0.090***		0.051**
		(0.003)		(0.003)		(0.021)		(0.020)
Asian or P.I.								
Lagged Pop. Share		0.012		–0.004		–0.196*		–0.013
		(0.015)		(0.014)		(0.107)		(0.097)
Lagged Unemp.		0.010		0.004		–0.436***		–0.181
		(0.022)		(0.021)		(0.156)		(0.144)
Lagged Log Wage		0.001		–0.001		0.042***		0.026*
		(0.002)		(0.002)		(0.014)		(0.013)
Native American								
Lagged Pop. Share		–0.026		–0.032		–1.631***		–1.213***
		(0.041)		(0.040)		(0.304)		(0.286)
Lagged Unemp.		0.028***		0.031***		–0.144**		–0.156**
		(0.010)		(0.010)		(0.070)		(0.065)
Lagged Log Wage		–0.002		–0.002		0.003		–0.000
		(0.002)		(0.001)		(0.011)		(0.010)
Controls	No	No	Yes	Yes	No	No	Yes	Yes
Overall Y Mean	0.032	0.032	0.032	0.032	1.660	1.660	1.660	1.660
Multiracial Y Mean	0.047	0.047	0.047	0.047	1.491	1.491	1.491	1.491
S.-Race-Fl. Y Mean	0.036	0.036	0.036	0.036	1.560	1.560	1.560	1.560
S.-Race-Cons. Y Mean	0.032	0.032	0.032	0.032	1.664	1.664	1.664	1.664
R-squared	0.001	0.001	0.111	0.111	0.003	0.004	0.160	0.161
No. of Observations	52,527	52,527	52,527	52,527	52,527	52,527	52,527	52,527
No. of Multiracial	677	677	677	677	677	677	677	677
No. of S.-Race-Fl.	1,096	1,096	1,096	1,096	1,096	1,096	1,096	1,096
No. of S.-Race-Cons.	50,754	50,754	50,754	50,754	50,754	50,754	50,754	50,754

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Author's calculations using Current Population Survey data from 2002. Multiracial persons are identified using 2003 information, as pre-2003 information is single race only. Each column displays a set of coefficients from estimation by ordinary least squares of an individual-level labor market outcome in 2002, as noted. The outcomes are an indicator for unemployment or a log hourly wage, with the latter adjusted by adding $1 (constant 1999) to levels before calculating log values. Displayed regressors reflect indicators for Black, Asian or Pacific Islander, and Native American racial groups (the indicator for the White racial group is the omitted category). Corresponding state-level lagged market factors reflect the population share, unemployment rate, or a log average hourly wage, with measures lagged to reflect their value in the previous pre-2003 month for a person (month-in-sample 3) and wages adjusted by adding $1 (constant 1999) to levels before averaging and calculating log values. Each observation reflects a person who is multiracial, single-race-fluid, or single-race-consistent. Single-race-fluid persons are individuals whose reported race changes from 2002 to 2003 but is consistent within each year. Single-race-consistent persons are individuals whose reported race remains unchanged from 2002 through 2003. When included, controls are life-event measures reflecting a change in age, marital status, citizenship, family size, educational attainment, employer, employment or labor force status, or health (as captured by information about reasons for job-search inactivity or activity while not in the labor force); and other measures that control for observed significant differences between treatment and control groups in conditional logit models of reported race in 2002 (namely indicators for education: some college [including associate degree] and college or more; region: Middle Atlantic, West North Central, South Atlantic, East South Central, and West South Central; industry: public administration; occupation: operators, fabricators, and laborers; and at least one parent in household). All estimation incorporates descriptive weights, and heteroskedasticity robust standard errors are in parentheses.

A Appendix

A.1 Descriptive Weight Construction

As mentioned in the main text and similarly to the approach in Jackson (2021), given numerous sample restrictions for the data, I create a descriptive weight for all workers to reflect their nationally representative count. These weights incorporate both a "sample design" component and a "post-stratification" component. The two components are multiplied to generate the descriptive weight, $WTALL$, which is used for both descriptive and causal analysis, as noted in the main text and displays.

For the sample design weight component, I adjust the $WTFINL$ measure provided by IPUMS-CPS. $WTFINL$ is described as "the final person-level weight that should be used in analyses of basic monthly data" and "is based on the inverse probability of selection into the sample," with additional adjustments for various factors (Flood et al. 2020). I use the MIS4 value of $WTFINL$ for each person, as it reflects the focal, pre-2003 month-in-sample given availability of earnings information from an outgoing rotation group.

For the post-stratification weight component, the goal is to further adjust the descriptive weight for any differential sample selection across a set of key individual traits. Such selection is determined by comparing the baseline sample of individuals with person-level IDs noted in Appendix Table A1 (call this the "raw" sample) and the prospective analysis sample, also noted in Appendix Table A1, which further restricts to MIS4 and the last available MIS (which turns out to be MIS8; call this the "final" sample, although MIS8 observations are ultimately omitted from analysis once multiracial, single-race-fluid, and single-race-consistent persons are identified). Once again, I focus on the MIS4 value for each individual trait. Both the raw and final samples reflect the resulting 2002 calendar years spanned by workers in MIS4, and both samples are also restricted to persons aged 16 years and older since the analysis sample is constrained to such individuals. I focus on five categories of individual traits, with the corresponding number of values for each measure indicated in parentheses: sex (2), age (2), education (3), period (2), and area (9).[20] Every person is uniquely assigned to one bin among all 216 possible bin combinations from those five traits. For each bin and the corresponding workers assigned to those bins, the post-stratification weight, $WTPOST$, is the count of persons in the raw sample divided by the count of persons in the final sample.

The final descriptive weight for each worker in the analysis sample, $WTALL$, is thus $WTFINL \times WTPOST$. As noted in the main text, I run validity checks to compare various population statistics (shares) with those generated by the analysis sample with the $WTALL$ descriptive weight applied. In these validity checks, I closely replicate the chosen population statistics.

[20]Regarding category values: sex is male or female; age is 16 through 44 or 45 years and older; education is high school (diploma or equivalent, including persons "not in universe" or with missing responses) or less, some college (including associate degree), and college (bachelor's degree) or more; period is January 2002 through June 2002 or July 2002 through December 2002; and area (region) is New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain, and Pacific, reflecting census divisions and the associated states (Flood et al. 2020).

Table A1: Initial Sample Selection

Sample Restriction	Count			Percentage of Baseline		
	Household	Indiv.	Obs.	Household	Indiv.	Obs.
Baseline w/ CPSIDP (person ID)	517,557	1,353,517	6,672,148	100.00	100.00	100.00
Drop if CPSIDP only appears once	482,299	1,242,606	6,561,237	93.19	91.81	98.34
Drop if under age 16	482,263	941,471	4,995,447	93.18	69.56	74.87
Drop if not in BMS for at least 4 months-in-sample	168,366	310,152	2,445,974	32.53	22.91	36.66
Drop if MISH4 is not available	162,641	300,920	2,389,319	31.42	22.23	35.81
Drop if not in BMS for at least 4 months-in-sample before Jan 2003	102,885	190,296	1,511,273	19.88	14.06	22.65
Drop if not in BMS for at least 2 months-in-sample after Jan 2003 (inclusive)	56,328	104,088	828,250	10.88	7.69	12.41
Drop if race varies before Jan 2003 or after Jan 2003 (inclusive)	56,251	103,941	827,099	10.87	7.68	12.40
Drop if Hispanic varies across time	53,765	98,998	788,015	10.39	7.31	11.81
Drop if sex varies across time	53,410	98,366	783,623	10.32	7.27	11.74
Drop if age varies incorrectly across time	49,136	86,397	688,841	9.49	6.38	10.32
Drop if state unavailable for any month-in-sample	49,136	86,397	688,841	9.49	6.38	10.32
Drop if state varies across time	49,136	86,397	688,841	9.49	6.38	10.32
Drop if self-employed	46,008	75,721	603,766	8.89	5.59	9.05
Drop if not in MISH4 or last month-in-sample	46,008	75,721	151,442	8.89	5.59	2.27

Note: Author's calculations using Current Population Survey data from 2001–2004.