The Effect of Racial and Ethnic Attitudes on Hispanic Identity in the U.S

Hussain Hadah [*]

September 3, 2023

Abstract

I study the determinants of the choice to identify as Hispanic among those who could—those whose parents, grandparents, or selves were born in a Spanish-speaking country. I find that individuals with Hispanic ancestry are significantly less likely to self-identify as Hispanic if they live in states with high levels of implicit ethnic bias. A one standard deviation increase in bias decreases self-reported Hispanic identity by 7 and 13 percentage points for first and second-generation Hispanics, respectively. These effects are more prominent among second-generation immigrants with both parents born in a Spanish-speaking country than among children of inter-ethnic parents. These findings have implications for the interpretation of research on ethnic gaps in economic outcomes and the correct counting of the population.

Keywords: Economics of Minorities, Race, and Immigrants; Discrimination and Prejudice

JEL Classification: I310, J15, J71

1

1 INTRODUCTION

The self-reported Hispanic population in the United States has tripled over the last three decades as Hispanics overtook Blacks as the largest minority group in America.[1] An extensive literature on Hispanics–non-Hispanic White gaps has emerged, finding differences in health (Antecol and Bedard 2006; Antman, Duncan, and Trejo 2016), wages (Trejo 1997), and schooling (Antman, Duncan, and Trejo 2022). However, defining and measuring these racial and ethnic groups is not straightforward, especially when considering self-reported identity. To the extent that self-report Hispanic identity is negatively selected, these gaps could be biased upward.

Various factors, including prejudice, can influence the manner in which individuals select their ethnic identity. In this paper, I explore the determinants of Hispanic identity and how Hispanics self-select into Hispanic and White identities. In particular, I study how bias against minorities influences their decisions to identify, or not, as a member of the ethnic minority. This is important as it affects our interpretations of a variety of findings. First, if individuals react to prejudice by choosing not to identify with their targeted group, standard analyses attempting to identify components of ethnic gaps in outcomes could be overestimated in the most biased states. Second, how individuals identify may impact measured changes in labor market outcomes among groups differentiated by race and ethnicity. As a result, Mexican immigrants' assimilation rates could appear slower than other groups.

To explore how individual characteristics and social attitudes toward Hispanics affect self-reported Hispanic identity. I use identity and ancestry information from the Current Population Survey (CPS) along with a proxy for state-level bias from Harvard's Project Implicit Association Test (IAT). [2] I motivate my analysis with a simple model in the vein of Akerlof and Kranton (2000). This makes explicit a path through which actions affect individuals' utility via their identity and introduces an externality where the actions of others—or prejudice—have different effects on a person's well-being and identity. Therefore, if a person can choose their identity credibly and this choice is affected by the prejudice of others, then they will choose it to maximize their outcomes.

Measuring identity choices outside of a laboratory is challenging because it requires objective and self-reported identity measures. I use data from a person's birth-

---

1. The 2020 Census counted more than 62 million Hispanics—19 percent of the population—triple the number of Hispanics counted three decades earlier (Flood et al. 2021a). The Hispanic population numbers are based on the author's calculations from the Current Population Survey and US Census data.

2. The IAT data is retrieved from Harvard's Project Implicit (Greenwald, McGhee, and Schwartz 1998). The implicit bias toward minorities, as measured by IAT, is widely used by psychologists and is growing in use among economists. IAT scores were shown to be correlated with economic outcomes (Chetty et al. 2020; Glover, Pallais, and Pariente 2017), voting behavior (Friese, Bluemke, and Wänke 2007), and health (Leitner et al. 2016).

place and ancestry to construct an ostensibly objective measure of identity. I find self-reported identity to be negatively correlated with individual and parental characteristics, i.e., parental education. I also find that they are negatively associated with discrimination and ethnic attitudes that reflect the social environment.

Among individuals of Hispanic ancestry, I find that higher state-level bias against people with darker skin is correlated with a lower self-reported Hispanic identity among Hispanic immigrants. I find that a one standard deviation increase in bias correlates with a 7 percentage points decrease in the self-reported Hispanic identity among first-generation immigrants and a 13 percentage points decrease among second-generation immigrants. Additionally, a one standard deviation increase in bias is correlated with a 15 percentage points drop in self-reported Hispanic identity among second-generation Hispanic children with both parents born in a Spanish-speaking country. Consequently, as the more economically successful Hispanic immigrants—educated and wealthy immigrants—do not self-report Hispanic identity, economic research using subjective ethnic measures will overestimate White-Hispanic gaps in the most biased states.

This paper fits in the economics of immigration and assimilation. Abramitzky, Boustan, and Eriksson (2016) measured the speed at which immigrants from Europe, Asia, and Latin America assimilate in the United States. They find that assimilation increases over time.[3] Fouka, Mazumder, and Tabellini (2022) investigated the effect of the inflow of Black Americans migrating from the South to the North on the assimilation of European immigrants. The authors found that immigrants in places that received more Black migrants assimilated faster. Meng and Gregory (2005) studied the effect of intermarriage on assimilation and found that immigrants who intermarry earn significantly more than those in an endogamous marriage. Antman, Duncan, and Trejo (2016) show that among immigrants from Mexico, the least economically successful self-identify as being of Mexican origin, while the most successful do not.

Other researchers studied the assimilation of Hispanic immigrants. Antecol and Bedard (2006) documented an interesting puzzle where non-native-born Hispanics have better health outcomes than native-born Hispanics, and Trejo (1997) showed that Mexican men earn substantially less than Whites.[4] Smith (2003) offered a more optimistic view of the assimilation of Hispanic immigrants. The longer Hispanic and Latino immigrants reside in the US, the more they can close the educational gap with White men. Moreover, some of the poor showings of how well Hispanic immigrants assimilate in the United States could be explained by ethnic attrition and the use of

---

3. For more on immigrant assimilation, see Abramitzky, Boustan, and Connor (2020), Abramitzky et al. (2019), Abramitzky et al. (2020), and Abramitzky, Boustan, and Eriksson (2014)

4. The Hispanic health paradox has led many researchers to try to explain it (Giuntella 2016; Giuntella and Stella 2017; Giuntella et al. 2018; Giuntella 2017; Antman, Duncan, and Trejo 2016; Antman, Duncan, and Trejo 2020).

self-reported Hispanic identity to study Hispanics (Duncan and Trejo 2017; Duncan and Trejo 2011b; Meng and Gregory 2005; Duncan and Trejo 2018a, 2018b; Antman, Duncan, and Trejo 2016; Antman, Duncan, and Trejo 2020). The ethnic attrition was driven by the children of inter-ethnic marriages (Meng and Gregory 2005; Duncan and Trejo 2005). Once the attrition was accounted for, Hispanic immigrants would appear healthier and thus more assimilated than previously thought (Antman, Duncan, and Trejo 2016; Antman, Duncan, and Trejo 2020). This research, however, does not explore whether the self-reported identity itself is a function of other factors like prejudice.

This paper is most closely related to Antman, Duncan, and Trejo (2016) and Antman and Duncan (2015, 2021) where the authors studied the ethnic attrition of Hispanic immigrants and how minorities change their self-reported identity to changes in policies.[5] Taking into consideration the ethnic attrition that Antman, Duncan, and Trejo (2016) document, I investigate the determinants of what drives a person to self-report, or not, their Hispanic identity. I aim to decompose some of the complexity associated with endogenous identity by exploring some of the personal and environmental determinants of identity. The empirical analysis in this paper documents how some observable, i.e., personal characteristics and societal attitudes, affect the self-reported identity of Hispanics.

The rest of this paper is structured as follows. First, I will discuss the conceptual framework in section (2). Second, I will describe the data I use in section (3). Third, I will introduce an empirical model and the results in section (4). Fourth, I will discuss robustness checks and discuss the results in section (5). Finally, I conclude in section (6).

2 CONCEPTUAL FRAMEWORK

I discuss a conceptual framework of identity in the spirit of Akerlof and Kranton (2000). A person belongs to some ethnic group, and their actions either affirm or deny their ethnic identity. Actions that deviate from what is proscribed of the ethnic identity are costly.

Formally, a person i belongs to ethnic group $e_i \in \{H, NH\}$, where H is Hispanic and NH is non-Hispanic. Agent i's utility depends on their actions and the extent to which their actions affirm their identity I_i:

$$U_i = U_i(\boldsymbol{a_i}, \boldsymbol{a_{-i}}, I_i) \tag{1}$$

A person's identity, I_i, is influenced by their own actions, the actions of others, and the behavior proscribed by their ethnicity. I write this as:

---

5. Ethnic attrition is when a person with Hispanic ancestry fails to self-identify as Hispanic.

4

$$I_i = I_i(\mathbf{a_i}, \mathbf{a_{-i}}; \mathbf{B}_{e_i}) \qquad (2)$$

Where $\mathbf{a_i}$ is the actions of person i. $\mathbf{a_{-i}}$ is the actions of others that would affect i's identity, i.e., societal bias. I_i is the identity function. Each group has an associated set of behaviors that society proscribes them to conform to, which I denote as $\mathbf{B}_{e_i}$.[6]

A person i chooses action a_i that maximizes their utility function given ethnic group e_i, proscribed appropriate behavior $\mathbf{B}_{e_i}$, and the actions of others $\mathbf{a_{-i}}$. This implies the following first-order condition (F.O.C.):

$$\frac{\partial U_i}{\partial a_i} + \frac{\partial U_i}{\partial I_i} \cdot \frac{dI_i}{da_i} = 0 \qquad (3)$$

Whose solution $a_i^\star$ yields utility $U_i^\star$. Now, suppose a person can choose their ethnic identity at a cost of c. They will do so if $\tilde{U}_i^\star \geqslant U_i^\star + c$. Where $\tilde{U}_i^\star$ is the utility obtained from optimal actions $\tilde{a}_i^\star$ under the counterfactual ethnicity.

That is i will change identities when the benefits of doing so $\tilde{U}_i^\star - U_i^\star$ exceed the costs c. These net benefits are non-zero only if $\frac{dI_i}{da_i} \neq 0$ and $\frac{\partial U_i}{\partial I_i} \neq 0$. This suggests that an empirical analysis of the determinants of identity choice should focus on: (1) individual characteristics that would lead to different a_i under different identities, (2) contextual characteristics that would lead to different a_{-i}—bias—under different identities, (3) the analysis should focus on a sample of the population with small c, and (4) the sub-sample with a utility that is greatly affected by their identity—i.e., $\frac{\partial U_i}{\partial I_i} \neq 0$). From the empirical analysis, I could investigate the characteristics that would affect i's actions to take different identities from point (1). These characteristics could be the generation immigrants belong to, whether their parents are interethnic or endogamous, etc. I could also investigate how different state-level biases could affect identity. Finally, restricting the sample to people with a small cost of changing identity c guarantees that I do not include populations that would never change identities otherwise—for example, non-Hispanic Whites with non-Hispanic ancestry.

3 DATA

In this section, I describe the datasets I use. To study the association between social attitudes and self-reported Hispanic identity, I must measure subjective and objective Hispanic identities to select a subgroup of Hispanic immigrants for analysis. Thus, I use the Integrated Public Use Microdata Series (IPUMS) Current Population Survey (CPS) (Flood et al. 2021a) and use information on ancestry to construct an objective identity measure. I also use implicit association test (IAT) as a measure of bias, in

---

6. Akerlof and Kranton (2000) refer to B_{e_i} as proscription.

which bias will shift the costs of identifying as a particular ethnic identity (Hispanic).

3.1 MEASURING HISPANIC IDENTITY

I measure Hispanic identity using the Current Population Survey (CPS), which allows me to construct an objective measure of the Hispanic identity of minors living with their parents. I will use the information on the place of birth, parents' place of birth, and place of birth of grandparents to construct an objective Hispanic measure.[7] Thus, I could perfectly identify, and construct a dataset of first-, second-, and third-generation Hispanic immigrants (see figure 1 for a visual representation). This will consequently allow me to build an objective measure of the Hispanic identity of minors under the age 17 living with their parents.

The objective measure of identity—unlike the self-reported measure where respondents answer affirmatively when asked if they are Hispanic or Latino—depends on the birthplaces of the individual, their two parents, and four grandparents. Thus, the three identifiable generations are: 1) first-generation immigrants that were born in a Spanish-speaking country with both parents also being born in a Spanish-speaking country, 2) second-generation immigrants are native-born citizens to at least one parent that was born in a Spanish-speaking country, 3) third-generation immigrants are native-born citizens to two native-born parents and at least one grandparent that was born in a Spanish-speaking country.[8] I restrict the sample to Hispanic Whites, first-, second-, and third-generation immigrants who are 17 year old and younger and still live with their parents between 2004 and 2021. I present a summary of the sample statistics in the table (1).

The overall sample is 49% female, and 91% of the sample self-reportedly identifies as Hispanic—answered yes to the question "are you Latino/Hispanic?". The average age is 8.6 year old. Almost 14% of mothers have a college degree, and 14% of fathers have a college degree. I provide the rest of the summary statistics for the overall sample and broken down by generation in table (1).

---

7. Following the works of Antman, Duncan, and Trejo (2016) and Antman, Duncan, and Trejo (2020).

8. I restrict first-generation immigrants whose parents were born in a Spanish country to avoid including naturally born US citizens that were born abroad to US parents.

Table 1: CPS Summary Statistics with Skin IAT Scores

Characteristic	Overall	By Generation		
	All Sample N = 318,404	**First** N=40,033	**Second** N=199,294	**Third** N=79,077
Female	0.49	0.54	0.49	0.49
Asian	0.65	0.95	0.72	0.33
Age	8.7 (5.1)	11.2 (4.5)	8.5 (5.1)	7.7 (5.0)
College Graduate: Father	0.50	0.56	0.49	0.48
College Graduate: Mother	0.48	0.55	0.47	0.48
Total Family Income (1999 dollars)	78,763 (80,171)	69,732 (72,672)	78,909 (81,395)	82,811 (80,254)

[1] The samples include children ages 17 and below who live in intact families. First-generation Asian immigrant children that were born in a Asian country. Native-born second-generation Asian immigrant children with at least one parent born in a Asian country. Finally, native-born third generation Asian immigrant children with native-born parents and at least one grand parent born in a Asian country.

[2] Data source is the 2004-2021 Current Population Survey.

Moreover, using the place of birth of parents and grandparents, I can objectively identify their ethnic ancestry. Consequently, I can identify different types of parents and grandparents. Using the place of birth of parents data, I can divide parents of second-generation children into three objective types:

1. Objectively Hispanic-father-Hispanic-mother (HH)

2. Objectively Hispanic-father-White-mother (HW)

3. Objectively White-father-Hispanic-mother (WH)

Similarly, using the place of birth of grandparents, I can divide grandparents of third-generation children into 15 objective types: (1) objectively Hispanic paternal

grandfather-Hispanic paternal grandmother-Hispanic maternal grandfather-Hispanic maternal grandmother (HHHH); (2) objectively White paternal grandfather-Hispanic paternal grandmother-Hispanic maternal grandfather-Hispanic maternal grandmother (WHHH); (3) objectively Hispanic paternal grandfather-White paternal grandmother-Hispanic maternal grandfather-Hispanic maternal grandmother (HWHH), etc...

Figure 1: Diagram of the Three Different Generations of Hispanic Immigrants.
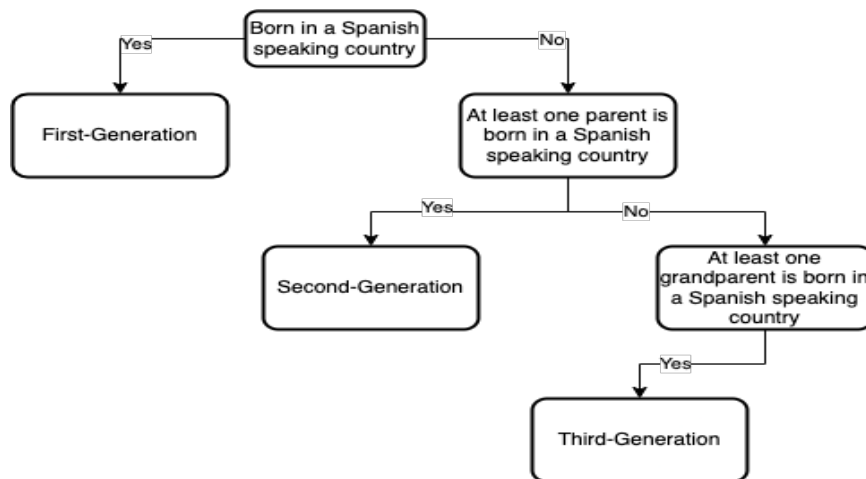
My analysis depends on a sub sample of the US population, I show in table (2) that I have enough observations in each generation. Consistent with the literature on ethnic attrition among Hispanics, I find significant attrition among third-generation Hispanic immigrants.[9] These results are displayed in table (2): most first- and second-generation Hispanic immigrants self-reportedly identified as Hispanic. Of the first-generation Hispanic immigrants, 96% self-reportedly identified as Hispanic. Similarly, 95% of the second-generation Hispanic immigrants self-reportedly identified as Hispanic, and 85% of third-generation Hispanic immigrants identifying as Hispanic. That is a more than threefold increase in attrition rates. Most of the attrition among third-generation Hispanics is driven by attrition among the children of inter-ethnic marriages.

---

9. In Duncan and Trejo (2018a, 2018b), Antman, Duncan, and Trejo (2016), and Antman, Duncan, and Trejo (2020), the authors find substantial attrition among Hispanics.

Table 2: Asian Self-identification by Generation

	Self-identify as Asian	Self-identify as non-Asian	% Self-identify as Asian	% Self-identify as non-Asian
1st Gen.	38,198	1,835	0.95	0.05
2nd Gen.	143,066	56,228	0.72	0.28
Asian on:				
Both Sides	116,593	4,272	0.96	0.04
One Side	26,473	51,956	0.34	0.66
3rd Gen.	25,797	53,280	0.33	0.67
Asian on:				
Both Sides	11,204	547	0.95	0.05
One Side	6,188	19,225	0.24	0.76

[1] The samples include children ages 17 and below who live in intact families. First-generation Asian immigrant children that were born in a Asian country. Native-born second-generation Asian immigrant children with at least one parent born in a Asian country. Finally, native-born third-generation Asian immigrant children with native-born parents and at least one grandparent born in a Asian country.

[2] Data source is the 2004-2021 Current Population Survey.

Figure 2: Bias and Self-reported Hispanic Identity in the Least and Most Biased Places

(a) Skin Tone Implicit Association Bias Over Time
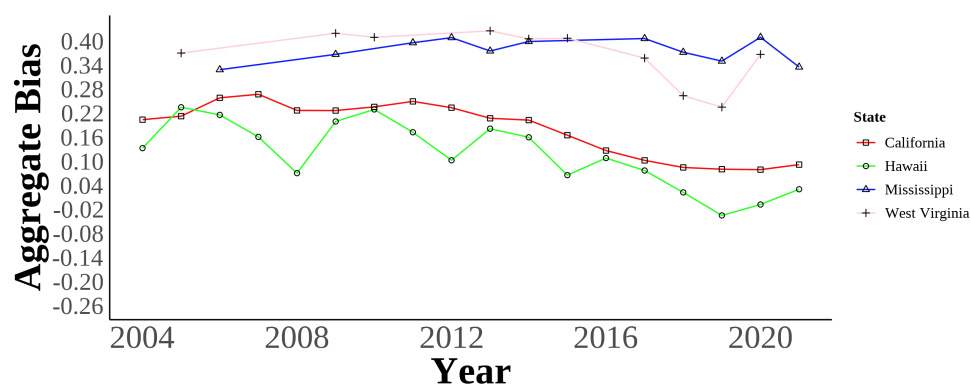
(b) Self-reported Hispanic Identity Over Time
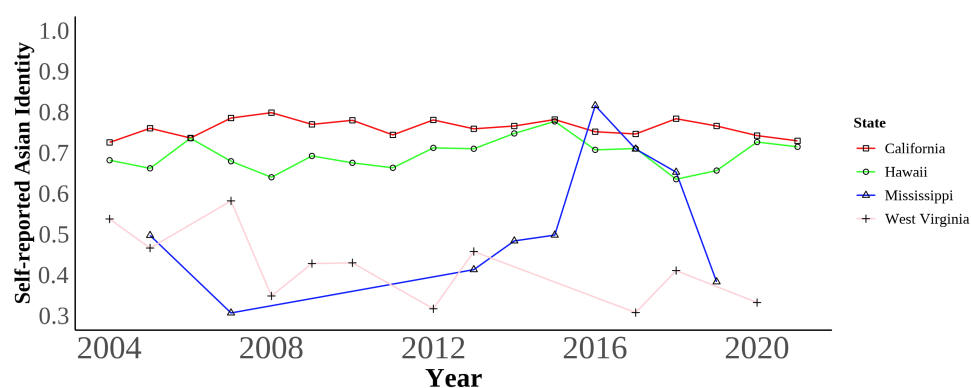
These two panels show the trends in implicit bias (panel a) and self-reported Hispanic identity among Hispanic immigrants (panel b) of the least and most biased places in the data. The District of Colombia is the least biased geographical area, and North Dakota is the most biased. The bias units are in standard deviations. Self-reported Hispanic identity is among first, second, and third-generation Hispanic immigrants aged 17 and younger still living in intact families.

Bias data is from the 2004-2021 Harvard's Project Implicit Association Test scores. Identity data is from the 2004-2021 Current Population Survey (CPS).

Figure 3: Maps of State-level Implicit Association Test Bias Over Time Measure with Census Division Regional Boundaries
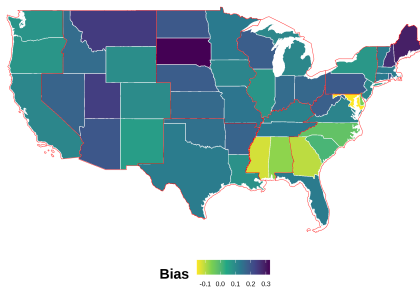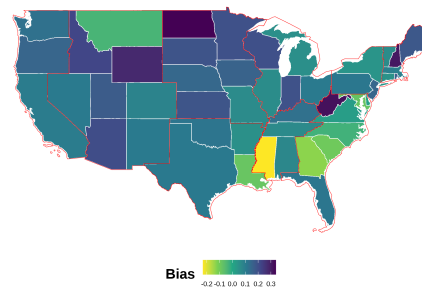
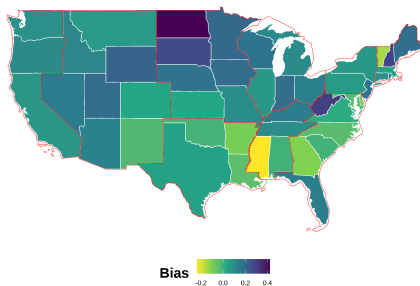(a) State-level Bias in 2004
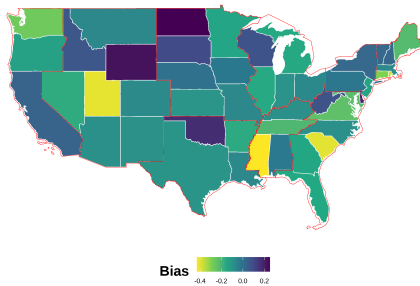
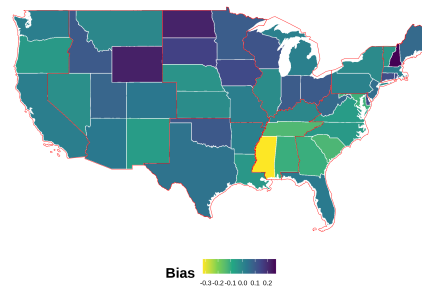(b) State-level Bias in 2006
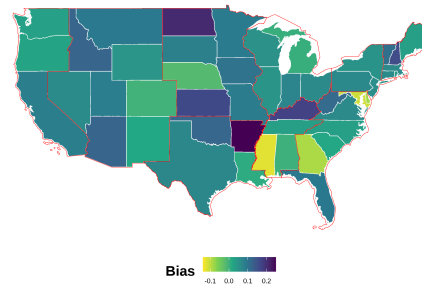
(c) State-level Bias in 2008
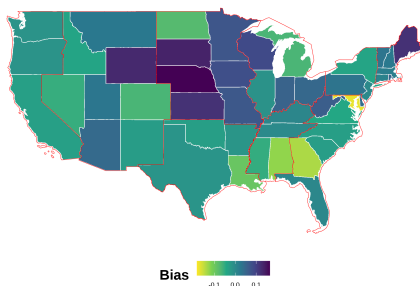
(d) State-level Bias in 2010
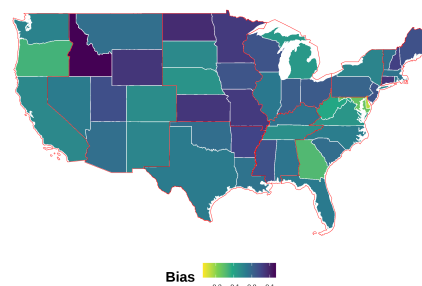
(e) State-level Bias in 2012

(f) State-level Bias in 2014

(g) State-level Bias in 2016

(h) State-level Bias in 2018

In this figure, I show the state-level implicit bias in different years in the sample. Each panel presents state-level bias during a certain year. The boundaries in red represent the different Census divisions in the United States. Notice how there is a variation across states with-in a region.

Figure 4: Maps of State-level Implicit Association Test Bias 2004-2021 Measure with Census Division Regional Boundaries
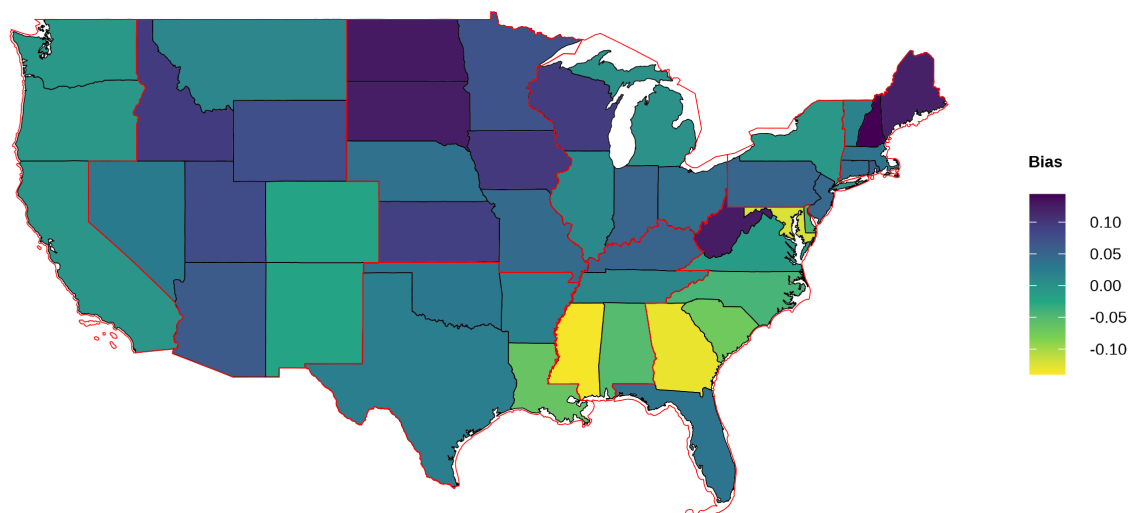
In this figure, I show the state-level implicit bias in the sample of IAT tests from 2004 to 2021. The boundaries in red represent the different Census divisions in the United States. Notice how there is a variation across states with-in a region.

3.2 MEASURING PREJUDICE

The implicit association test measures how people associate concepts—for example, Black and dark-skinned people—and evaluations—good, bad. Respondents are asked to quickly match words into categories shown on a screen. Figure (9) shows a few examples of what a test taker would see on a skin tone implicit association test by Harvard's Project Implicit.

I use skin tone implicit association test data to construct a measure of state-level prejudice (Greenwald, McGhee, and Schwartz 1998). This measure has been used in the social sciences, especially in psychology. Previous work has shown that IAT test scores are hard to manipulate (Egloff and Schmukle 2002). [10] Figure (2a) shows a

---

10. Research showed that the IAT tests are correlated with economic outcomes (Chetty et al. 2020; Glover, Pallais, and Pariente 2017), voting behavior (Friese, Bluemke, and Wänke 2007), and health (Leitner et al. 2016). Participation in the IAT, an online test, is voluntary. Therefore, the samples are not random and might suffer from selection bias in who decides to take the exam. However, bias reflected by Implicit Association Test (IAT) scores as a way to measure prejudiced attitudes has been used as a proxy for prejudiced attitudes in an area Chetty et al. (2020).

graphical representation of the bias measure over time in the most and least biased places. A lower score implies less light-skin bias, whereas a higher score implies more discrimination against dark-skinned people. One-half of a standard deviation increase in bias is equivalent to the moving from Washington DC to North Dakota in 2012. I also show the state-level average bias over time in the maps in figure (3) and the overall average from 2004 to 2021 in figure (4).

4 ESTIMATION AND RESULTS

Let H_{ist}^g be the self-reported Hispanic identity of person i in state s at the time of interview t, let $Bias_{st}$ be the average state-level bias in state s at time t, $DadCollegeGrad_{ist}$, and $MomCollegeGrad_{ist}$ are indicator variables that are equal to one if the father or mother graduated from college, $Women_{ist}$ is an indicator variable for sex, and X_{ist} is a vector of controls.[11] Additionally, γ_{rt} is region-time fixed effects that controls for region $\times$ year specific shocks. The region $\times$ year also controls for systematic differences between regions in the overall Hispanic population, or bias toward Hispanics, even if they vary over time.

I first estimate if state-level bias is correlated with self-reported Hispanic identity. I estimate different specifications for each generation g. The basic equation I estimate is:

$$
\begin{aligned}
H_{ist}^g = {} & \beta_1^g Bias_{st} + \beta_2^g DadCollegeGrad_{ist} + \beta_3^g MomCollegeGrad_{ist} \\
& + \beta_4^g Women_{ist} + X_{ist}^g \pi + \gamma_{rt} + \varepsilon_{ist}; \text{where } g \in \{1, 2, 3\} \quad (4)
\end{aligned}
$$

I report the main results of estimating equation (4) in figure (5). I present the results of estimating the main specification on all three generations in panel (A) and on sub-samples of first-, second-, and third-generation immigrants in panels (B), (C), and (D) respectively. I find that bias and self-reported Hispanic identity are negatively associated, while parental education and self-reported Hispanic identity are positively associated. A one standard deviation increase in state-level bias is associated with an 11 percentage points decrease in self-reported Hispanic identity. Among first- and second-generation Hispanic immigrants, a one standard deviation increase in state-level bias is associated with 7 and 13 percentage points decrease in self-reported Hispanic identity.

I report the results of the same regression but on sub-samples of second-generation

---

11. The controls include quartic age, fraction of population that is Hispanic in state s, type of parents (WH, HW, or HH), type of grandparents (HHHH, HHHW, etc.), and dummy variables the generation to which person i belong.

immigrants by type of parents—interethnic and endogamous parents—in figure (6). I present the results of estimating the main specification on second-generation immigrants in panel (A) and on sub-samples of HH, HW, and WH children in panels (B), (C), and (D), respectively. I find that children of endogamous Hispanic marriages are more influenced by state-level bias than those of interethnic marriages. I find that a one standard deviation increase in state-level bias is associated with 15 percentage points decrease in self-reported Hispanic identity among children of HH parents.
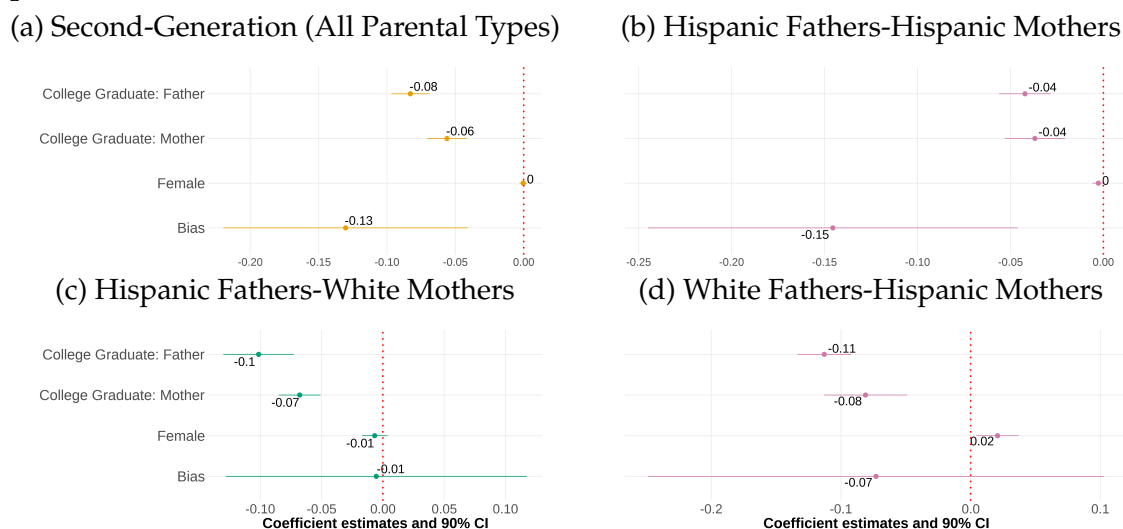
I also report the results of the regression on sub-samples of third-generation immigrants by the number of Hispanic grandparents in table (3). I find that state-level bias is not significantly associated with the self-reported Hispanic identity of children of inter-ethnic grandparents. However, it is negatively associated with lower self-reported Hispanic identity of children of Hispanic grandparents. I find that a one standard deviation increase in state-level bias is associated with a 14 percentage points decrease in self-reported Hispanic identity.

Figure 5: Relationship Between Self-Reported Hispanic Identity and Bias: By Generation
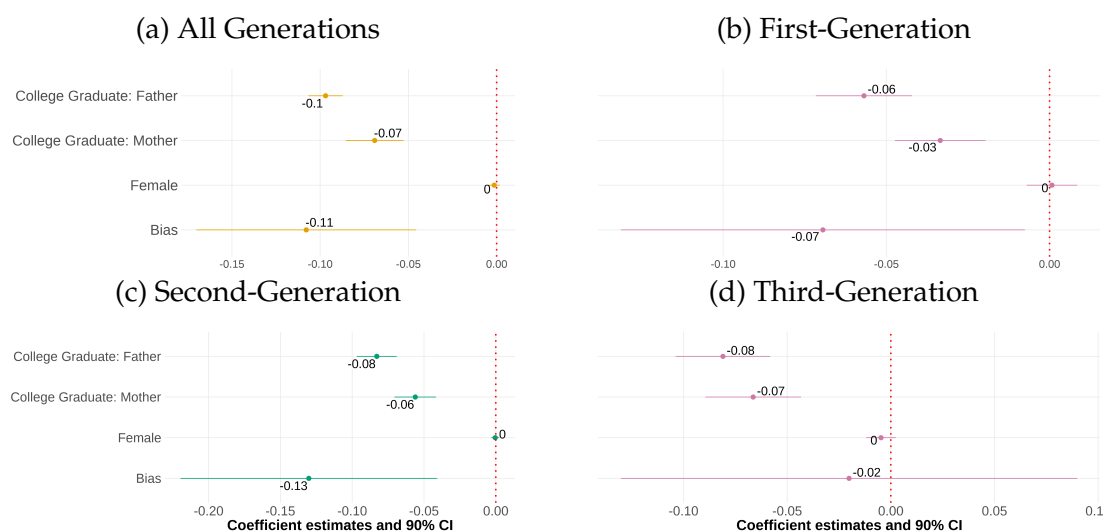
(a) All Generations

(b) First-Generation

(c) Second-Generation

(d) Third-Generation

I show four panels of estimating equation ([4]). I include region × year fixed effects with controls for sex, quartic age, and parental education. The dependent variable is self-reported Hispanic identity and the independent variable is state-level bias. Each panel is the results from the same regression but on different samples that are divided by generation. Standard errors are clustered on the state level. The samples include first-, second-, and third-generation Hispanic children ages 17 and below who live in intact families. First-generation Hispanic immigrants are children that were born in a Spanish-speaking county. Native-born second-generation Hispanic immigrants are children with at least one parent born in a Spanish-speaking country. Finally, native-born third-generation Hispanic immigrants are children with native-born parents and at least one grandparent born in a Spanish-speaking country.

15

Figure 6: Relationship Between Self-Reported Hispanic Identity and Bias: By Parental Types

(a) Second-Generation (All Parental Types)

(b) Hispanic Fathers-Hispanic Mothers

(c) Hispanic Fathers-White Mothers

(d) White Fathers-Hispanic Mothers

I show four panels of estimating equation (4). I include region × year fixed effects with controls for sex, quartic age, and parental education. The dependent variable is self-reported Hispanic identity and the independent variable is state-level bias. Each panel results from the same regression but on different samples divided by parental types. Standard errors are clustered on the state level. The samples include second-generation Hispanic children ages 17 and below who live in intact families. Native-born second-generation Hispanic immigrant children with at least one parent born in a Spanish-speaking country.

Table 3: Relationship Between Bias and Self-Reported Hispanic identity Among Third-Generation Hispanic Immigrants: By Grandparental Type

	Number of Hispanic Grandparents			
	(1)	(2)	(3)	(4)
	One	Two	Three	Four
Bias	-0.04	0.03	0.19	-0.14*
	(0.11)	(0.09)	(0.26)	(0.07)
Female	-0.01	0.00	-0.01	0.00
	(0.01)	(0.01)	(0.01)	(0.01)
College Graduate: Mother	-0.11***	-0.07***	0.02	-0.02
	(0.03)	(0.02)	(0.02)	(0.01)
College Graduate: Father	-0.11***	-0.08***	0.02	-0.03*
	(0.03)	(0.01)	(0.01)	(0.01)
Observations	55,051	74,100	12,194	57,646
Year × Region FE	X	X	X	X

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

[1] Each column is an estimation of equation (4) restricted to third-generation Hispanic immigrants by number of Hispanic grandparents with region × year fixed effects. I include controls for sex, quartic age, fraction of Hispanics in a state, and parental education. Standard errors are clustered on the state level.

[2] The samples include third-generation Hispanic children ages 17 and below who live in intact families. Native-born third-generation Hispanic immigrant children with at least one grandparent born in a Spanish-speaking country.

[3] Data source is the 2004-2021 Current Population Survey.

5 ROBUSTNESS CHECKS AND DISCUSSIONS

I will investigate the relationship between state-level bias and interethnic marriages. To this purpose, the regression specifications for the estimation will be as follows:

$$\text{Interethnic}_{ist}^2 = \beta_1^2 \text{Bias}_{st} + X_{ist}^2 \pi + \gamma_{rt} + \varepsilon_{ist} \tag{5}$$

Where $\text{Interethnic}_{ist}^2$ is an indicator variable that is equal to one if a couple is interethnic, i.e., a Hispanic husband-White wife or a White husband-Hispanic wife. Bias_{st} is the average bias in state s at time t, and X_{ist}^2 is a vector of partner-specific controls that would affect a marriage match that includes the wife's and husband's education, age, and years since immigrating to the United States.

I present the results of estimating equation (5) in table (4). I find that a one standard deviation increase in bias increases the probability of having inter-ethnic parents by a 5 percentage point. Moreover, I break down the analysis by the ethnicity of the couples. A one standard deviation increase in state-level bias is associated a 6 percentage points increase in the chances of a Hispanic husband having a White wife. A one standard deviation increase in state-level bias is associated a 7 percentage points increase in the chances of a Hispanic wife having a White husband. The fact that bias and interethnic marriage are positively correlated could be a result of the fact that Hispanic immigrants in states with high bias might aim to decrease the likelihood that their children will display Hispanic ethnicity signals. For example, Hispanic women in high bias states might marry a non-Hispanic White husband, so their children will have a non-Hispanic last name.

I am also interested in investigating the relationship between state-level bias and migration. As the CPS does not report a person's birth state, I use the 2004-2021 Censuses to construct a sample of second-generation Hispanic immigrants (Flood et al. 2021b). I construct a mover variable to indicate whether these second-generation Hispanic immigrants have moved from their birth state to another state. For this purpose, I use the following models to estimate the relationship between state-level bias and migration:

$$\text{BirthPlaceMigration}_{ist}^2 = \beta_1^2 \text{Bias}_{st} + X_{ist}^2 \pi + \gamma_{rt} + \varepsilon_{ist} \tag{6}$$

$$\text{BirthPlaceMigration}_{ilb}^2 = \beta_1^2 \text{Bias}_{lb} + X_{ilb}^2 \pi + \gamma_{lb} + \varepsilon_{ilb} \tag{7}$$

Where $\text{BirthPlaceMigration}_{ist}^2$ is an indicator variable equal to one if person i in state s at the interview t lives in a state that is different from his or her birth state and zero otherwise. $\text{BirthPlaceMigration}_{ilb}^2$ is an indicator variable that is equal to one if

person i in birthplace l does not currently live in the same state he or she lived in at the year of birth b and zero otherwise. The analysis, restricted to second-generation Hispanic immigrants with both parents born in a Spanish-speaking country, uses equations (6) and (7).

Furthermore, I use two ways to define the bias variable to study the relationship between bias and the migration variables introduced above. In the first specification from equation (6), I estimate the relationship between the average bias at the time of the interview t in state s and $BirthPlaceMigration^2_{ist}$. In the second specification from equation (7), I estimate the relationship between the average bias in birth state l at the year of birth b and $BirthPlaceMigration^2_{ilb}$.

I also estimate whether those who self-identify as Hispanic tend to move from high-bias to low-bias states. The estimation equation for the relationship is:

$$Y_{ist} = \beta_0 + \beta_1^2 Hispanic_{ist} + X^2_{ist}\pi + \varepsilon_{ist} \tag{8}$$

Where $Y_{ist} \equiv Bias_{ist} - Bias_{ilb}$, $Bias_{ist}$ is i's state-level bias in state s at the time of interview t, and $Bias_{ilb}$ is i's state-level bias in birth state l at the birth year b. The analysis is restricted to second-generation Hispanic immigrants with both parents born in a Spanish-speaking country who migrated from the state they were born in b to another state s.

The results of estimating equations (6), (7), and (8) are shown in table (5) in columns (1), (2), and (3) respectively. I find that among second-generation immigrants, there is no significant correlation between bias and migration decisions. Among second-generation Hispanic immigrant movers, those who self-report Hispanic identity live in states with 0.02 standard deviations more biased than the state where they were born. Even though this result shows that there is selection into more biased states among second-generation immigrants, it does not affect my main results showing a correlation between bias and self-reported Hispanic identity. Since those identifying as Hispanics are the movers, my assessments of the relationship between bias and self-reported Hispanic identity might underestimate the effect of bias.

Table 4: Relationship Between Bias and Interethnic Marriages

	(1) Interethnic	Hispanic Men (2) Interethnic	Hispanic Women (3) Interethnic
Bias	0.05***	0.06***	0.07***
	(0.01)	(0.01)	(0.01)
College Graduate: Wife	0.05***	0.04***	0.04***
	(0.00)	(0.00)	(0.00)
College Graduate: Husband	0.08***	0.04***	0.06***
	(0.00)	(0.00)	(0.00)
Observations	477,672	433,232	434,297
Year × Region FE	X	X	X

* p < 0.1, ** p < 0.05, *** p < 0.01

[1] This is the result to estimating (5) as a linear probability model.

[2] I include controls for partners' sex, age, education, and years since immigrating to the United States. Standard errors are clustered on the household level.

[3] Data source is the 2004-2020 Current Population Survey Data.

Table 5: Relationship Between Bias and Migration

	(1) Migrated from Birth Place	(2) Migrated from Birth Place	(3) $Bias_{ist} - Bias_{ilb}$
$Bias_{st}$	0.06		
	(0.12)		
$Bias_{lb}$		-0.09	
		(0.26)	
Hispanic			0.02***
			(0.01)
Female	0.00	0.00	0.00*
	(0.00)	(0.00)	(0.00)
College Graduate: Mother	0.03***	0.03***	0.00
	(0.01)	(0.01)	(0.01)
College Graduate: Father	0.02**	0.02***	-0.01
	(0.01)	(0.01)	(0.01)
Observations	352,712	185,024	12,806
Mean	0.09	0.09	-0.02
Year × Region FE	X		
Birthyear × Birth Region FE		X	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

[1] Each column is an estimation of equations (6) in column (1), (7) in column (2), and (8) in column (3).

[2] Column (1) is a regression where the left hand side variable is a dummy variable that is equal to one if a person migrated from the state were born in and the right hand side variable is bias the year of survey. Column (2) is a regression where the left hand side variable is a dummy variable that is equal to one if a person migrated from the state were born in and the right hand side variable is bias the year of birth in the state of birth. Column (3) is a regression where the left hand side variable is the difference between state-level bias during the year of the survey in the current state the respondent is living in, and state-level bias during the year of birth in the state of birth and the right hand side variable is self-reported Hispanic identity. This regression captures the selection of those that self-reported Hispanic identity into states with different levels of bias. I include controls for sex, quartic age, parental education, fraction of Hispanics in a state, and region × year fixed effects. Standard errors are clustered on the state level.

[3] The samples include children ages 17 and below who live in intact families. Native-born second-generation Hispanic immigrant children with both parents born in a Spanish-speaking country. The sample in the column (3) regression is further restricted to only those that migrated from their birth state.

[4] Data source is the 2004-2021 Census Data.

The findings presented in this paper indicate a negative correlation between bias

and the self-reported Hispanic identity of Hispanic immigrants. While my aim is not to establish a causal effect of bias on self-reported Hispanic identity, I intend to illustrate a correlation between bias and self-reported identity. This correlation suggests that depending on the levels of bias in a state, racial and ethnic gaps that rely on self-reported identity might either overestimate or underestimate the effect of discrimination.

There are a couple of concerns with this analysis. First, the self-reported identity in the Current Population Survey (CPS) is reported by a household respondent—parent or adult caregiver. Thus, the 'self-reported' ethnic identity might not reflect a child's true identity. I view the identity that a parent or a caregiver reports as an accurate representation of the child's identity since parents are essential in shaping their children's sense of self. Also, I compare states with a high and low bias for my analysis. The estimates will not be threatened as long as the likelihood of self-reporting does not differ between these states.

Moreover, Duncan and Trejo (2011a) show that reported Hispanic identification does not vary with who is the household respondent. Additionally, I present the main effect of self-reported Hispanic identity by the household respondent in table (6) and the results to the estimation of equation (4) by the proxy respondent in table (7) across all generations. The main effect of the reported Hispanic identity of children is 94 percentage points when the mother is the proxy, 92 percentage points when the father is the proxy, and 96 percentage points when the child or another caregiver was the household respondent.[12] The estimation of equation (4) by the proxy respondent, table (7), mostly yields a negative effect of bias on self-reported Hispanic identity for all types of proxy respondents.

Table 6: Main Effect of Proxy on Second-Generation's Hispanic Self-identification

Parents Type	All	Hispanic-Hispanic	Hispanic-White	White-Hispanic
Proxy:				
Mother	0.94	0.96	0.9	0.84
Father	0.92	0.96	0.86	0.8
Self	0.96	0.97	0.9	0.84
Others	0.96	0.97	0.92	0.9

A second concern is that the IAT is voluntary and not representative of the population. While I do not claim that the IAT as a proxy for bias will represent the population, Egloff and Schmukle (2002) show that they are hard to manipulate. Several

---

12. According to the Current Population Survey (CPS), a person can be the household respondent if they are at least 15 years old and have enough knowledge about the household. Thus, when the proxy is 'self,' the respondent is between the ages of 15 and 17.

studies have shown that IAT is correlated with economic outcomes (Chetty et al. 2020; Glover, Pallais, and Pariente 2017), voting behavior (Friese, Bluemke, and Wänke 2007), decision-making (Bertrand, Chugh, and Mullainathan 2005; Carlana 2019), and health (Leitner et al. 2016). Another concern could be that the IAT test takers' characteristics change over time and, thus, are not the same. I address this concern by including non-parametric region × year fixed effects that would control for the systematic difference in the characteristics of test takers between regions. These changes will be controlled for as long as the differences in the characteristics between test takers do not vary across states within a region.

Another concern could be reverse causality between having more Hispanic people in a state and implicit bias. It could be the case that the number of Hispanic people in a state affects the implicit bias on the residents of that state. For example, having more Hispanics in Florida could affect the implicit bias of the residents of Florida. To show that this is not the case, I provide figures (7) and (8) as evidence. Figure (7) plots the percent of self-reported Hispanics in a state at a specific year against the average implicit bias in the same state during that year. Figure (8) plots the percent of objectively second-generation Hispanic children of endogamous marriages in a state at a certain year against the average implicit bias in the same state during that year. I find no correlation between bias and the number of Hispanics in a state, thus, making the case of reverse causality unlikely.

Finally, the estimator of the relationship between bias (prejudice) and self-reported Hispanic identity could be biased if those that do not self-report Hispanic identity migrate to more prejudiced states. I have shown above that this is not the case (table 5). I find no evidence of a relationship between migration decisions and bias. Additionally, I find that those reporting Hispanic identity moved out of birthplaces with less bias and lived in more biased states at the time of the survey. Thus, my results might underestimate the relationship between bias and self-reported Hispanic identity.

6 CONCLUSION

As the United States becomes more multi-racial and multi-ethnic, self-reported identity will significantly affect on representation, distributive politics, and government transfers. The determinants of endogenous identity are particularly important to researchers interested in the role of discrimination on earnings gaps. In this paper, I show how individual characteristics and social attitudes toward racial and ethnic minorities affect the self-reported Hispanic identity of individuals with Hispanic ancestry in the United States. I find that people of Hispanic ancestry are less likely to identify as Hispanic in states with more significant bias. The relationship between

self-reported Hispanic identity and bias is more prominent among first- and second-generation immigrants, where a one standard deviation increase in bias correlated with a 7 and 13 percentage points decrease in self-reported Hispanic identity.

Additionally, state-level bias has a more substantial effect among second-generation immigrant children with Hispanic fathers and Hispanic mothers. A one standard deviation increase in bias correlates with a 15 percentage points decrease in self-reported Hispanic identity among second-generation Hispanic immigrant children of objectively Hispanic parents. I also find that bias positively correlates with interethnic marriage and not with migration decisions.

The results are important because of the consequences on the correct counting of Hispanics and minorities, and assimilation and mobility. They could indicate that bias could significantly affect how economists estimate the earnings gap. Most research concerning race and ethnicity relies on self-reported race and ethnic identity measures. Since state-level bias is negatively correlated with self-reported Hispanic identity, the characteristics of those who do not self-report Hispanic identity could have important consequences. For example, if the people whose identities are most likely affected by bias are the most educated. In this case, the racial and ethnic gaps will be overestimated in the most biased states. Furthermore, identity decisions are likely to affect people's choices, investments, and well-being profoundly.

Moreover, this study could encourage further research into the relationship between bias and self-reported identities for other groups. The analysis of the effect of bias on self-reported identity could be applied to other groups. For example, we could estimate the effect of bias on the identities of sexual minorities and other ethnic and racial minorities such as Asian American, Black, Native American, and Arab American populations in the United States. Researchers could also explore the differences in outcomes between the ethnic and racial minorities who self-report to those that do not by using restricted administrative data.

REFERENCES

Abramitzky, Ran, Leah Boustan, Katherine Eriksson, and Stephanie Hao. 2020. "Discrimination and the Returns to Cultural Assimilation in the Age of Mass Migration." *AEA Papers and Proceedings* 110 (May): 340–346.

Abramitzky, Ran, Leah Platt Boustan, and Dylan Connor. 2020. *Leaving the Enclave: Historical Evidence on Immigrant Mobility from the Industrial Removal Office.* Technical report w27372. Cambridge, MA: National Bureau of Economic Research, June.

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. 2014. "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration." *Journal of Political Economy* 122, no. 3 (June): 467–506.

———. 2016. *Cultural Assimilation during the Age of Mass Migration.* Working Paper, 22381, July.

Abramitzky, Ran, Leah Platt Boustan, Elisa Jácome, and Santiago Pérez. 2019. *Intergenerational Mobility of Immigrants in the US over Two Centuries.* Technical report w26408. Cambridge, MA: National Bureau of Economic Research, October.

Akerlof, George A., and Rachel E. Kranton. 2000. "Economics and Identity." *Quarterly Journal of Economics* 115, no. 3 (August): 715–753.

Antecol, Heather, and Kelly Bedard. 2006. "Unhealthy Assimilation: Why Do Immigrants Converge to American Health Status Levels?" *Demography* 43, no. 2 (May): 337–360.

Antman, Francisca, and Brian Duncan. 2015. "Incentives to Identify: Racial Identity in the Age of Affirmative Action." *Review of Economics & Statistics* 97, no. 3 (July): 710–713.

———. 2021. "American Indian Casinos and Native American Self-Identification" (August).

Antman, Francisca, Brian Duncan, and Stephen J. Trejo. 2016. "Ethnic Attrition and the Observed Health of Later-Generation Mexican Americans." *The American Economic Review* 106 (5): 467–471.

Antman, Francisca M., Brian Duncan, and Stephen J. Trejo. 2020. "Ethnic Attrition, Assimilation, and the Measured Health Outcomes of Mexican Americans." *Journal of Population Economics* (Heidelberg, Netherlands) 33, no. 4 (October): 1499–1522.

Antman, Francisca M, Brian Duncan, and Stephen J Trejo. 2022. "Hispanic Americans in the Labor Market: Patterns over Time and across Generations."

Bertrand, Marianne, Dolly Chugh, and Sendhil Mullainathan. 2005. "Implicit Discrimination." *American Economic Review* 95, no. 2 (May): 94–98.

Carlana, Michela. 2019. "Implicit Stereotypes: Evidence from Teachers' Gender Bias." *Quarterly Journal of Economics* 134, no. 3 (August): 1163–1224.

Chetty, Raj, Nathaniel Hendren, Maggie R Jones, and Sonya R Porter. 2020. "Race and Economic Opportunity in the United States: An Intergenerational Perspective*." *The Quarterly Journal of Economics* 135, no. 2 (May): 711–783.

Duncan, B., and S. J. Trejo. 2011a. "Intermarriage and the Intergenerational Transmission of Ethnic Identity and Human Capital for Mexican Americans." *J Labor Econ* 29.

———. 2011b. "Who Remains Mexican? Selective Ethnic Attrition and the Intergenerational Progress of Mexican Americans." In *Latinos and the Economy: Integration and Impact in Schools, Labor Markets, and Beyond,* edited by D. L. Leal and S. J. Trejo. New York: Springer.

Duncan, Brian, and Stephen J. Trejo. 2005. *Ethnic Identification, Intermarriage, and Unmeasured Progress by Mexican Americans.* Working Paper, Working Paper Series 11423. National Bureau of Economic Research, June.

———. 2017. "The Complexity of Immigrant Generations: Implications for Assessing the Socioeconomic Integration of Hispanics and Asians." *ILR Review: The Journal of Work and Policy* 70 (5): 1146–1175.

———. 2018a. "Identifying the Later-Generation Descendants of U.S. Immigrants: Issues Arising from Selective Ethnic Attrition." *The ANNALS of the American Academy of Political and Social Science* 677, no. 1 (May): 131–138.

———. 2018b. *Socioeconomic Integration of U.S. Immigrant Groups over the Long Term: The Second Generation and Beyond.* Working Paper, Working Paper Series 24394. National Bureau of Economic Research, March.

Egloff, Boris, and Stefan C. Schmukle. 2002. "Predictive Validity of an Implicit Association Test for Assessing Anxiety." *Journal of Personality and Social Psychology* (US) 83:1441–1455.

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, J. Robert Warren, and Michael Westberry. 2021a. *Integrated Public Use Microdata Series, Current Population Survey: Version 9.0.*

Flood, Sarah, Goeken Ronald, Schouweiler Megan, and Sobek Matthew. 2021b. *Integrated Public Use Microdata Series, USA: Version 12.0.*

Fouka, Vasiliki, Soumyajit Mazumder, and Marco Tabellini. 2022. "From Immigrants to Americans: Race and Assimilation during the Great Migration." *The Review of Economic Studies* 89, no. 2 (March): 811–842.

Friese, Malte, Matthias Bluemke, and Michaela Wänke. 2007. "Predicting Voting Behavior with Implicit Attitude Measures: The 2002 German Parliamentary Election." *Experimental Psychology* (Germany) 54 (4): 247–255.

Giuntella, O. 2016. "Assimilation and Health: Evidence from Linked Birth Records of Second- and Third-Generation Hispanics." *Demography* 53.

Giuntella, O., Z. L. Kone, I. Ruiz, and C. Vargas-Silva. 2018. "Reason for Immigration and Immigrants' Health." *Public Health,* Special Issue on Migration: A Global Public Health Issue. 158 (May): 102–109.

Giuntella, Osea. 2017. "Why Does the Health of Mexican Immigrants Deteriorate? New Evidence from Linked Birth Records." *Journal of Health Economics* 54 (July): 1–16.

Giuntella, Osea, and Luca Stella. 2017. "The Acceleration of Immigrant Unhealthy Assimilation." *Health Economics* 26, no. 4 (April): 511–518.

Glover, Dylan, Amanda Pallais, and William Pariente. 2017. "Discrimination as a Self-Fulfilling Prophecy: Evidence from French Grocery Stores." *Quarterly Journal of Economics* 132, no. 3 (August): 1219–1260.

Greenwald, Anthony G., Debbie E. McGhee, and Jordan L. K. Schwartz. 1998. "Measuring Individual Differences in Implicit Cognition: The Implicit Association Test." *Journal of Personality and Social Psychology* (US) 74 (6): 1464–1480.

Leitner, Jordan B., Eric Hehman, Ozlem Ayduk, and Rodolfo Mendoza-Denton. 2016. "Racial Bias Is Associated with Ingroup Death Rate for Blacks and Whites: Insights from Project Implicit." *Social Science & Medicine* 170 (December): 220–227.

Meng, Xin, and Robert G. Gregory. 2005. "Intermarriage and the Economic Assimilation of Immigrants." *Journal of Labor Economics* 23, no. 1 (January): 135–174.

Smith, James P. 2003. "Assimilation across the Latino Generations." *The American Economic Review* (Nashville, United States) 93, no. 2 (May): 315.

Trejo, Stephen J. 1997. "Why Do Mexican Americans Earn Low Wages?" *Journal of Political Economy* 105, no. 6 (December): 1235.

A TABLES

Table 7: Relationship Between Bias and Self-Reported Hispanic Identity: By Proxy Respondent

	Proxy Respondent					
	White Mother	Hispanic Mother	White Father	Hispanic Father	Self	Other
	(1)	(2)	(3)	(4)	(5)	(6)
	H_{ist}	H_{ist}	H_{ist}	H_{ist}	H_{ist}	H_{ist}
Prejudice Measure	0.01	-0.06	-0.13	-0.07***	-0.07	-0.17**
	(0.11)	(0.03)	(0.11)	(0.02)	(0.05)	(0.07)
Female	-0.01	0.00	0.01	0.00***	0.00	0.00
	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)
College Graduate: Mother	-0.02	-0.03***	-0.04***	-0.03***	-0.03*	-0.03***
	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)
College Graduate: Father	-0.09***	-0.05***	-0.01	-0.03***	-0.10***	-0.06***
	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)
First Gen	-0.10*	0.07***	0.03	0.09***	0.13***	0.08***
	(0.06)	(0.01)	(0.05)	(0.02)	(0.02)	(0.02)
Second Gen	0.05*	0.04***	0.06***	0.05***	0.12***	0.07***
	(0.03)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)
Observations	64,052	380,625	47,233	267,690	10,472	69,301
Year × Region FE	X	X	X	X	X	X

* p < 0.1, ** p < 0.05, *** p < 0.01

[1] Each column is an estimation of a heterogeneous effect of regression (4) by the proxy household respondent with region × year fixed effects. I include controls for sex, quartic age, fraction of Hispanics in a state, parental education. Standard errors are clustered on the state level.

[2] The samples include children ages 17 and below who live in intact families. First-generation Hispanic immigrant children that were born in a Spanish-speaking county. Native-born second-generation Hispanic immigrant children with at least one parent born in a Spanish-speaking country. Finally, native-born third-generation Hispanic immigrant children with native-born parents and at least one grandparent born in a Spanish speaking country.

[3] Data source is the 2004-2021 Current Population Survey.

B   Figures

Figure 7: Scatter Plot of Proportion Subjectively Hispanic on Bias

(a) Year < 2015
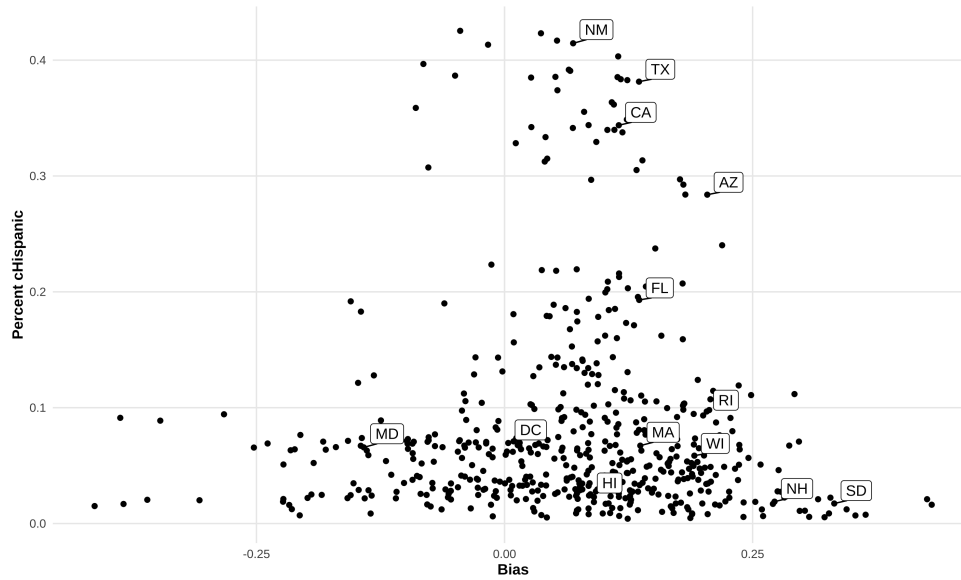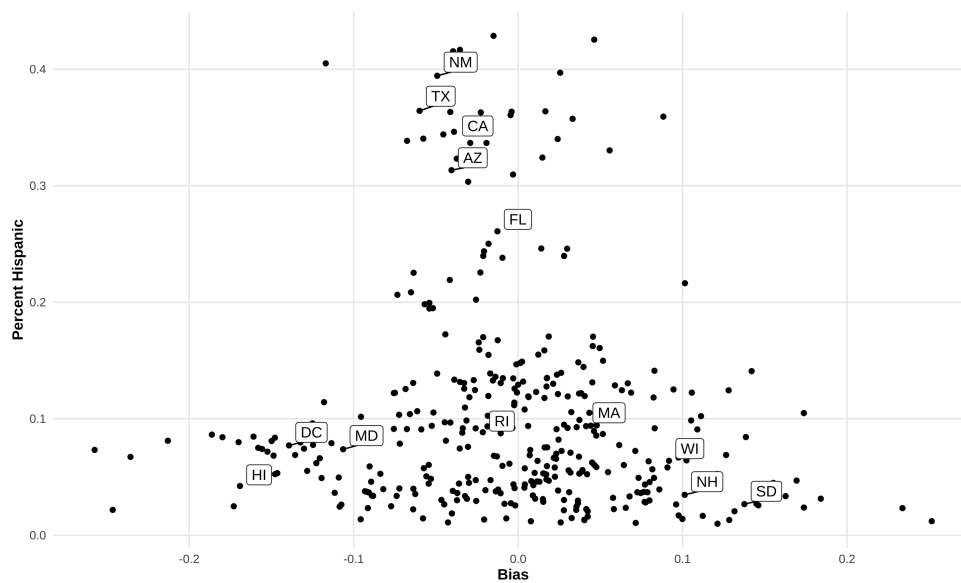
(b) Year ≥ 2015

Here are two scatter plots of bias on subjective Hispanic population in a state. Each dot represents a state in a certain year. Percent subjectively Hispanic = $\frac{\#\text{Hispanics}}{\text{Population}}$

Source. 2004-2021 Current Population Survey and 2004-2021 Implicit Association Test as a proxy for bias.

Figure 8: Scatter Plot of Proportion Second-Generation and Both Parents Born in a Spanish-Speaking Country on Bias

(a) Year < 2015
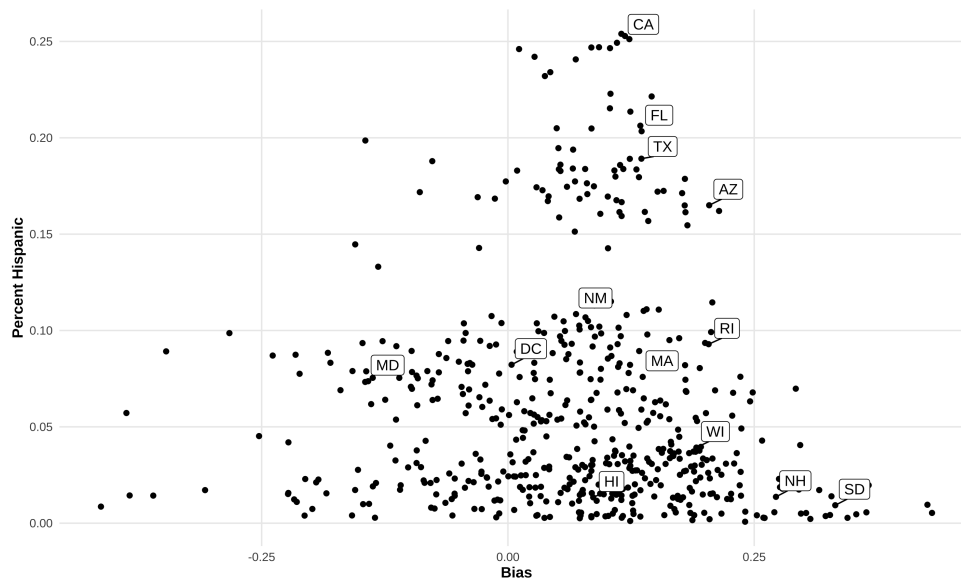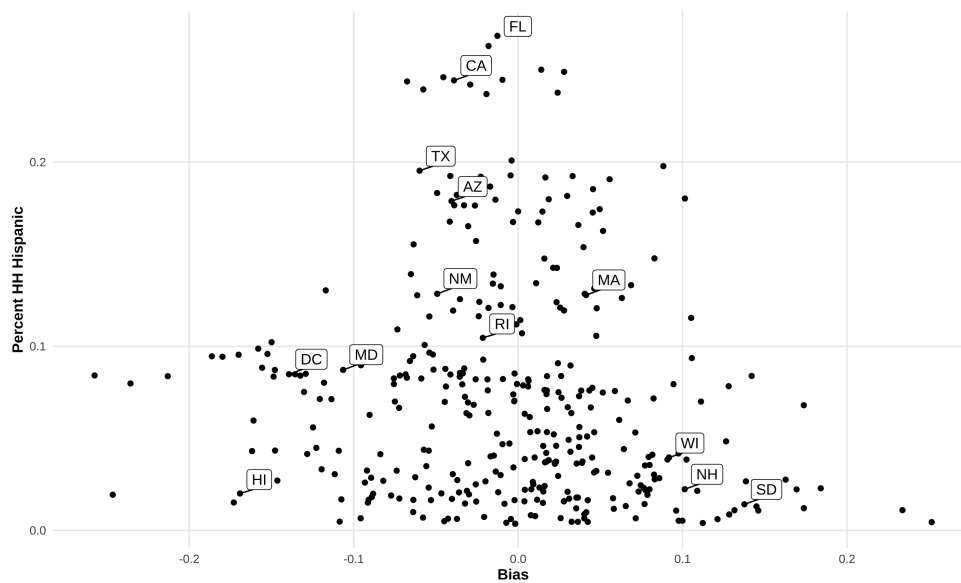
(b) Year ⩾ 2015

Here are two scatter plots of bias on subjective Hispanic population in a state. Each dot represents a state in a certain year.

Percent HH Hispanic = $\frac{\text{\#Hispanics with two parents born in a Spanish-speaking country}}{\text{Population}}$.

Source. 2004-2021 Current Population Survey and 2004-2021 Implicit Association Test as a proxy for bias.

C DATA

Figure 9: Examples of an Implicit Association Test

Here are a few examples of what a respondent would see on an implicit association test.