

# Essays on Applied Microeconomics

by  
Hussain Hadah

A dissertation submitted to the Department of Economics,  
College of Liberal Arts and Social Sciences  
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy  
in Economics

Chair of Committee: Chinhui Juhn  
Co-Chair of Committee: Vikram Maheshri  
Committee Member: Willa Friedman  
Committee Member: Yona Rubinstein

University of Houston  
May 2023

Copyright 2023, Hussain Hadah

---

Chinhui Juhn

---

Vikram Maheshri

---

Willa Friedman

---

Yona Rubinstein

## **ACKNOWLEDGMENTS**

I am grateful to the many individuals who have generously contributed their time and support to my research. My sincere appreciation goes to my committee members, Chinhui Juhn, Vikram Maheshri, Willa Friedman, and Yona Rubinstein. I have also benefited greatly from the insights and guidance of the distinguished faculty members at the University of Houston, including Aimee Chin, Steven Craig, German Cubas, Elaine Liu, and Fan Wang. My sincere thanks also goes to the participants of the Applied Microeconomics Workshop at the University of Houston for their invaluable feedback over the years. I would also like to extend my appreciation to my friends Gael Compta, Mehdi Oni, Ali Alsaffouri, and Faisal Alhussari for their unwavering support and guidance during this journey. I am deeply grateful to my mother, Racha Alelewi, and my siblings, Sedra and Mera Hadah, for their constant encouragement and support. Finally, I want to express my heartfelt thanks to Kirsten Gilliland, MD for your love and constant encouragement and assistance throughout this process.

## ABSTRACT

In this dissertation, I explore two topics in applied microeconomics. In chapter one, I study discrimination against Hispanics in the labor market. I compare the children of inter-ethnic marriages to study the impact of having a Hispanic last name. While males born to Hispanic father-White mothers earn less than those born to White father-Hispanic mothers, the gap could be completely explained by educational differences. I also study the effect of identifying as Hispanic on earnings. I find that men that identify as Hispanic earn significantly less than those that do not, even after controlling for educational differences. In chapter two, study the determinants of the choice to identify as Hispanic among those who could—those whose parents, grandparents, or selves were born in a Spanishspeaking 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 seven and 13 percentage points for first and second-generation Hispanics, respectively. These effects are more prominent among second-generation immigrants whose mothers and fathers were born in a Spanish-speaking country than among children of inter-ethnic parents. These findings have implications for the interpretation of economic research on racial and ethnic gaps in a variety of contexts, including labor market outcomes, redistributive policies, and political representation.

# TABLE OF CONTENTS

<b>ACKNOWLEDGMENTS</b>	<b>iv</b>
<b>ABSTRACT</b>	<b>v</b>
<b>LIST OF FIGURES</b>	<b>viii</b>
<b>LIST OF TABLES</b>	<b>ix</b>
<b>1 THE IMPACT OF HISPANIC LAST NAMES AND IDENTITY ON LABOR MARKET OUTCOMES</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.2 Data . . . . .	5
1.3 Children of the four parental types . . . . .	6
1.3.1 Synthetic parents . . . . .	10
1.4 Empirical Strategy . . . . .	14
1.4.1 Estimating the effect of having a Hispanic last name . . . . .	15
1.4.2 Estimating the effect of identifying as Hispanic . . . . .	15
1.5 Results and Discussion . . . . .	16
1.5.1 The effect of having a Hispanic last name on labor market outcomes .	17
1.5.2 The effect of identifying as Hispanic on labor market outcomes . . .	18
1.5.3 Discussion . . . . .	23
1.6 Conclusion . . . . .	24
<b>2 THE EFFECT OF RACIAL AND ETHNIC ATTITUDES ON HISPANIC IDENTITY IN THE U.S</b>	<b>25</b>
2.1 Introduction . . . . .	25
2.2 Theoretical Model . . . . .	30
2.3 Data . . . . .	32
2.3.1 Measuring Hispanic Identity . . . . .	32
2.3.2 Measuring Prejudice . . . . .	41
2.4 Empirical Strategy . . . . .	44
2.4.1 The Determinants of Hispanic Identity . . . . .	44
2.4.2 Estimating the Relationship Between Bias, Interethnic Marriages, and Migration . . . . .	49
2.5 Results and Discussion . . . . .	52
2.5.1 Relationship between Attitudes and Identity . . . . .	52
2.5.2 The Relationship Between Bias, Interethnic Marriages, and Migration	64
2.6 Discussion . . . . .	68
2.7 Conclusion . . . . .	72
<b>A Figures</b>	<b>74</b>

<b>B Data</b>	<b>88</b>
---------------	-----------

<b>C Tables</b>	<b>89</b>
-----------------	-----------

## LIST OF FIGURES

1	Diagram of the Three Different Generations of Hispanic Immigrants. . . . .	37
2	Bias and Self-reported Hispanic Identity in the Least and Most Biased Places.	39
3	Maps of State-level Implicit Association Test Bias Over Time Measure with Census Division Regional Boundaries. . . . .	40
4	Maps of State-level Implicit Association Test Bias 2004-2021 Measure with Census Division Regional Boundaries. . . . .	41
5	Relationship Between Self-Reported Hispanic Identity and Bias: By Generation.	48
6	Relationship Between Self-Reported Hispanic Identity and Bias: By Parental Types. . . . .	49
7	Relationship Between Bias and Self-reported Identity on Second-Generation Hispanics: Interaction. . . . .	59
8	Relationship Between Bias and Self-reported Identity on third-generation Hispanics: Interaction. . . . .	61
9	Proportion of Non-Whites in the United States from 1995 to 2019. . . . .	74
10	Hispanics as a Percentage of the Total Population in the United States from 1995 to 2019. . . . .	75
11	Inter-racial Marriages as a Percent of All Marriages in the United States from 1995 to 2020. . . . .	76
12	Distribution of the Four Types of Children. . . . .	77
13	Chart Explaining Which Synthetic Parents and Children and When We Observe Them. . . . .	78
14	Relationship Between Bias and Self-reported Identity on third-generation Hispanics: Interaction . . . . .	79
15	Scatter Plot of Proportion Subjectively Hispanic on Bias . . . . .	80
16	Scatter Plot of Proportion Second-Generation and Both Parents Born in a Spanish-Speaking Country on Bias . . . . .	81
17	Hispanic Identification Among Hispanic Immigrants: By Generation . . . . .	82
18	Correlation Between Implicit and Explicit Biases: IAT Explicit Bias Question	84
19	Correlation Between Implicit and Explicit Biases: GSS Close to Black People Question . . . . .	85
20	Correlation Between Implicit and Explicit Biases: GSS Vote for A Black President Question . . . . .	86
21	Scatter Plot of Non-Hispanic Second-Generation Hispanic-Hispanic Parents on Bias . . . . .	87
22	Examples of an Implicit Association Test . . . . .	88

## LIST OF TABLES

1	Number of Children from the Different Types of Parents . . . . .	8
2	Couples' Type . . . . .	9
3	Summary statistics of outcomes using parent's place of birth . . . . .	12
4	Summary statistics of outcomes using parent's place of birth only for those that self-identify as Hispanic . . . . .	13
5	Effect of Having Hispanic Last Name . . . . .	20
6	Effect of Having Hispanic Last Name . . . . .	21
7	Hsiapnic–White Earnings Gap Using Hispanic Variable Only . . . . .	22
8	CPS Summary Statistics . . . . .	35
9	Hispanic Self-identification by Generation . . . . .	38
10	Skin Tone Implicit Association Test (IAT) Scores and Current Population Survey (CPS) Samples . . . . .	43
11	Relationship Between Bias and Self-Reported Hispanic Identity: By Generation . . . . .	54
12	Relationship Between Bias and Self-Reported Hispanic identity Among Second-Generation Hispanic Immigrants: By Parental Type . . . . .	58
13	Relationship Between Bias and Self-Reported Hispanic identity Among Third-Generation Hispanic Immigrants: By Grandparental Type . . . . .	60
14	Relationship Between Bias and Migration . . . . .	66
15	Relationship Between Bias and Interethnic Marriages . . . . .	68
16	Main Effect of Proxy on Second-Generation's Hispanic Self-identification . . . . .	70
17	Relationship Between Bias and Self-Reported Hispanic Identity: By Proxy Respondent . . . . .	89
18	Relationship Between Bias and Self-Reported Hispanic Identity . . . . .	90
19	Self-Reported Hispanic Identity and Charles and Guryan (2008a) Prejudice Measure: By Generation . . . . .	91
20	Relationship Between Self-Reported Hispanic Identity and Charles and Guryan (2008a) Prejudice Measure Among Second Generation Hispanic Immigrants: By Parental Type . . . . .	92
21	Relationship Between Self-Reported Hispanic identity and Charles and Guryan (2008a) Prejudice Measure Among Third-Generation Hispanic Immigrants: By Grandparental Type . . . . .	93

# 1 The Impact of Hispanic Last Names and Identity on Labor Market Outcomes

## 1.1 Introduction

There is a plethora of evidence of ethnic and racial discrimination in the labor market (Bayer and Charles, 2018; Charles and Guryan, 2008b; Card and Krueger, 1992; Fryer Jr and Levitt, 2004; Rubinstein and Brenner, 2014; Bertrand and Mullainathan, 2004a; Juhn et al., 1991). Hispanics constitute a large and growing portion of the population in the United States. As the number of Hispanics keeps increasing in the labor force, determining whether ethnic discrimination affects their labor market outcomes becomes increasingly crucial (Chetty et al., 2014b, 2016, 2017; Abramitzky et al., 2020a, 2014, 2016; Chetty et al., 2014a), how identity could shift public opinions toward trade (Grossman and Helpman, 2021), and how racial and gender attitudes affect the racial and gender earnings gaps (Charles and Guryan, 2008a; Charles et al., 2018). Thus, it is important to understand whether a person's ethnicity affects their labor market outcomes. Assimilation and mobility are crucial because they reflect how well Hispanics can integrate into society and move up the socioeconomic ladder.

In this paper, I answer the following questions. Does having a Hispanic last name affect labor market outcomes? What is the effect of identifying as Hispanic on earnings? Moreover, I aim to show that comparing Hispanic Whites to non-Hispanic Whites might create an artificially higher earnings gap since the two groups differ on many observable characteristics.<sup>1</sup>

<sup>2</sup> Others have attempted to compare how native-born White Hispanics fare to non-Hispanic Whites and foreign-born Hispanics. In Antman et al. (2020a, 2016a,b, 2020b), the authors

---

<sup>1</sup>By identifying as Hispanic, I refer to individuals who self-report their ethnicity as Hispanic on surveys or other data collection instruments.

<sup>2</sup>Observable characteristics refer to factors that can be measured and quantified, such as education level, work experience, and immigration status.

compare the health and educational outcomes of Hispanic Whites to non-Hispanic Whites and native-born Hispanics to foreign-born Hispanics. They find gaps in education and health between Hispanics and Whites. They also find that native-born Hispanics are more likely than their foreign-born counterparts to report poor health. Dávila and Mora (2008), the authors documents many gaps in labor market outcomes between Hispanics and Whites. They attribute a big part of this gap to differences in education, experience, immigration status, and regional differences.

The US population is growing in diversity. The proportion of non-Whites has increased by more than 10 percentage points from 1995 to 2019 (see figure 9), and the number of Hispanics has grown by 9 percentage points over the same period (see figure 10). Native-born White Hispanic men earn 21% less than White men (Duncan and Trejo, 2018a). However, a substantial portion of the earnings gap is due to educational differences between Hispanics and Whites (Duncan et al., 2006; Duncan and Trejo, 2018c). Therefore, understanding discrimination against Hispanics and their children is an important task. Discrimination against Hispanics has a couple of negative consequences, for example discrimination against Hispanics can lead to reduced job opportunities, lower wages, and hindered assimilation. First, if Hispanics face a substantial labor market penalty for their ethnic identity, it could hinder assimilation and integration. Second, a significant labor market penalty contributes to the growing wage gap between Whites and other minority groups. In this paper, I examine the role of having a Hispanic last name and identifying as Hispanic on labor market outcomes.

Identifying discrimination is difficult because of factors that affect labor market outcomes that are unobservable to economists—like innate ability. Also, I should be able to disentangle skills differentials between Hispanics and non-Hispanic Whites.<sup>3</sup> To causally identify ethnic discrimination against Hispanics and its effects on labor market outcomes, a researcher should

---

<sup>3</sup>Skills differentials refer to differences in education, work experience, and other factors that can affect a person's job performance and wages.

compare two identical people that only differ in their last name. Since this approach is unfeasible, researchers took different approaches. Bertrand and Mullainathan (2004a) and Fryer Jr and Levitt (2004) conducted audit studies to examine the effect of Black sounding names on employment discrimination. This approach, however, has its drawbacks. Audit studies only observe callbacks, not wages.

This study utilizes a method developed by Rubinstein and Brenner (2014) in this research. I will compare children from inter-ethnic marriages. More precisely, I will reach children of Hispanic fathers and White mothers (henceforth HW) to children of White fathers and Hispanic mothers (henceforth WH). This approach stems from the fact that marriages are not random. Couples match on several observable characteristics like income, schooling, socio-economic background, etc. (Averett et al., 2008; Averett and Korenman, 1996; Becker, 1973, 1974, 1993; Browning et al., 2006; Chiappori et al., 2012). Children of HW and WH marriages have more similar observable characteristics than children of homogeneous marriages— i.e., White fathers-White mothers and Hispanic fathers-Hispanic mothers. Moreover, children from a Hispanic father and White mother household will have a Hispanic last name from their fathers, enabling the investigation of how ethnic signals, such as having a Hispanic last name, affect annual log earnings.

Using the Current Population Survey (CPS) from 1994 to 2019 and the 1970-1990 US censuses, I will estimate the effect of having a Hispanic last name on the labor market outcomes. In other words, I will approximate the variation in labor market perception of ethnic signals, in this case having a Hispanic last name, on annual log earnings. I also study the effect of identifying as Hispanic on labor market outcomes.

To retrieve the effect of having a Hispanic last name, I compare the children of inter-ethnic marriages between Whites and Hispanics. I find that while men born to White father-Hispanic mothers earn less than men born to Hispanic father-White mothers, this gap is

entirely explained by educational differences. Thus, I do not find a significant effect of having a Hispanic last name. Finally, I find that men identifying as Hispanic earn significantly less than those who do not, even when I control ancestry and education. Inter-ethnic children with a White father, and thus having a non-Hispanic last name, earn 14% less than children of endogamous White parents. Inter-ethnic children with a Hispanic father, and thus having a Hispanic last name, earn 20% less than children of endogamous White parents. Endogamous children with Hispanic parents earn 24% less than children of endogamous White parents. A big part of the gaps could be explained by education. Inter-ethnic children with a White father and identify as Hispanic, and thus having a non-Hispanic last name, earn 14% less than children of endogamous White parents. Inter-ethnic children with a Hispanic father and identify as Hispanic, and thus having a Hispanic last name, earn 18% less than children of endogamous White parents. Endogamous children with Hispanic parents and identify as Hispanic earn 10% less than children of endogamous White parents.

A long list of economic papers investigated taste-based discrimination, statistical discrimination, and Black-White and gender earnings gaps that date back to the seminal work of Becker (1957). Bayer and Charles (2018) investigated the Black-White wage gap's narrowing and then divergence. Card and Krueger (1992) argued that the Black-White wage gap narrowed from 1960 to 1980 due to improvement in the quality of Black schools, and Juhn et al. (1991) showed that skills play a big part in the Black-White wage gap. Moreover, Charles and Guryan (2008b) provided evidence that 25% of the Black-White wage gap is due to prejudice. Black (1995) introduced a search model showing that with a taste for discrimination, minorities earn less than the majority group even when hired by an unprejudiced firm. In comparison, Altonji and Pierret (2001) found that firms do not statistically discriminate based on race and ethnicity. Little has been done, however, about discrimination against Hispanics, which offers an opportunity to exploit the observable racial similarities between

Hispanics and Whites.

Previous studies have also used names as a proxy for race and ethnicity (Fryer Jr and Levitt, 2004; Rubinstein and Brenner, 2014; Bertrand and Mullainathan, 2004a). Fryer Jr and Levitt (2004) point out that names can be a predictor of a person's race. Specifically, they provide a rising pattern among Blacks having different names than Whites. They did, however, find that having a Black name, after controlling for the home environment at birth, does not affect their labor market outcomes. Rubinstein and Brenner (2014) compared the children of mixed marriages between Sephardic and Ashkenazis in Israel.<sup>4</sup> They found that workers with Sephardic last names earn substantially less than those with an Ashkenazi last name. Bertrand and Mullainathan (2004a) conducted an audit study by sending employers identical resumes that differ in the ethnic and racial signal of a name (Black sounding name versus a White sounding one). They found that resumes with Black-sounding names received substantially fewer callbacks than their White counterparts.

The rest of the paper is organized as follows. Section 1.2 provides an overview of the data and summary statistics of the sample I am using. Section 1.4 introduces a formal empirical model. I go over the results in section 1.5. Finally, I offer a brief conclusion in section 1.6.

## 1.2 Data

For this paper, I will use two data sets. The Integrated Public Use Microdata Series (IPUMS) Current Population Survey (CPS) Annual Social and Economic (ASEC) (Flood et al., 2020) and the 1970 to 1990 US censuses (Ruggles et al., 2020).

I use the CPS' rich data set to study the effect of having a Hispanic last name on a person's labor market outcomes. I take advantage of the fact that the CPS asks for parents' place of birth, ethnicity, and race. The data spans the period between 1994—the earliest

---

<sup>4</sup>Ashkenazi and Sephardic are two distinct Jewish ethnic groups.

sample to ask about a parent's place of birth— to 2019. Moreover, since the CPS does not provide data on parents' characteristics, essential to determine the family background, I use the census to construct synthetic parents. The census offers a larger sample of potential parents. Similar to the CPS, the 1970 to 1990 censuses ask about the person's place of birth and the individual's race and ethnicity. I employ this information to construct "synthetic parents" using a method developed by Rubinstein and Brenner (2014). I construct the synthetic parents by linking husbands and wives in the census data to each other. I assume that parents have children between the ages of 25 and 40, I then link these synthetic parents using the year they were surveyed to the year of birth of the "children" in the CPS sample.

### 1.3 Children of the four parental types

I will use the CPS for my primary analysis of the effect of having a Hispanic last name on earnings. I restrict my sample to Whites, United States-born citizens aged 25 to 40 and born between 1970 to 1990. Taking advantage of data on parents' place of birth, I divide the sample into four groups depending on their parent's ethnicity. Mothers or fathers are Hispanic if they were born in a Spanish-speaking country and Puerto Rico, and White if they were born in the United States.<sup>5</sup> Therefore, an observation can be the product of four types of parents:

1. White father and White mother (hereafter WW)
2. White father and Hispanic mother (hereafter WH)
3. Hispanic father and White mother (hereafter HW)
4. Hispanic father and Hispanic mother (hereafter HH).

---

<sup>5</sup>The list of Spanish Speaking countries include: Argentina, Bolivia, Chile, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, El Salvador, Guatemala, Equatorial Guinea, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, Spain, Uruguay, Venezuela.

The distribution of the four types of children is presented in figure 12. The majority of the sample, 95.14%, is WW children. The second biggest group is HH, which constitutes 3.36% of the sample. Inter-ethnic children, WH and HW, make up 1.78% of the sample with 60,147 observations. Even though WH and HW are only 1.78% of the sample, I have plenty of observations to carry out an analysis. The summary statistics for the children of the four types of marriages are presented in table 3. Children of WW marriages (column i) do better on every measure while children of HH parents (column iv) do worse than other children on every measure.

On the one hand, WW children are more educated, with approximately 14 years of education for men and women. man and woman WW children have an employment rate of 95%. They have log hourly earnings of 2.557 for men and 2.435 for women. Also, their annual log earnings are 10.529 for men and 10.297 for women.

On the other hand, children of HH marriages have an average of 12.898 and 13.204 years of education for men and women, respectively. HH men and women have an employment rate of 93%. Their annual log earnings are 10.365 and 10.175 for men and women, respectively.

Inter-ethnic children, WH and HW, have lower educational attainment and income than WW children, but they are better than HH children. Also, WH and HW are comparable, unlike WW and HH children. These two groups are important to the analysis. Children of HW are going to be the treatment group. HW's children have a Hispanic father, who will carry his Hispanic last name. At the same time, the WH children have White fathers and will take his non-Hispanic, or White, last name. A first glance at the summary statistics of these two groups, table 3 columns (ii) and (iii) can show that they are different from the children of non-inter-ethnic marriages.

TABLE 1: Number of Children from the Different Types of Parents

		Parent's Type			
Parents' Type	White Father	White Father	Hispanic Father	Hispanic Father	
	White Mother	Hispanic Mother	White Mother	Hispanic Mother	
<b>Observations</b>	3,828,024 (0.95)	24,306 (0.01)	35,841 (0.01)	135,300 (0.03)	

<sup>1</sup> The data is restricted to people interviewed between 1994 and 2019, also White, married, and between the ages of 18 and 65. I identify the ethnicity of a person's parents through the parent's place of birth. A parent is Hispanic if both her parents were born in a Spanish-speaking country. A parent is White if born parents were born in the United States.

TABLE 2: Couples' Type

	Couples' Type			
	White Husband	White Husband	Hispanic Husband	Hispanic Husband
	White Wife	Hispanic Wife	White Wife	Hispanic Wife
<b>Observations</b>	1,286,731	7,178	7,606	20,911
	(0.97)	(0.01)	(0.01)	(0.02)

<sup>1</sup> The data is restricted to people interviewed in 1970 and 1960 and also White and married. I identify the ethnicity of a person through their place of birth. A parent is Hispanic if they were born in a Spanish-speaking country. A parent is White if they were born in the United States.

<sup>2</sup> The table includes information on the proportion of the four types of synthetic parents that I have constructed.

WH children have an average of 13.366 and 13.741 years of education for men and women. The employment rate for men is 94.5% and for women 95.0%. HW man and woman children obtained 13.072 and 13.270 years of schooling. Men's employment rate is 92.7% and women's 93.1%.

In panel D of table 3, I present the rates at which members of each group self-identify as Hispanic <sup>6</sup>. The vast majority of HH children, 97%, identify as Hispanic. There is considerable attrition in Hispanic identification of inter-ethnic children. Among WH children, 82.1% of men and 85.0% of women identified as Hispanic. Among HW children, 87.9% of men and 85.9% of women identified as Hispanic. These results align with the findings of a substantial body of literature that noted this attrition among children of immigrants, especially those coming from inter-ethnic marriages (Duncan and Trejo, 2017a, 2018a; Duncan

---

<sup>6</sup>Hispanic self-identification is observed when an individual notes that they are Hispanic on the CPS questionnaire

et al., 2020; Antman et al., 2020a). Finally, a small portion of WW children, 6%, identified as Hispanic. These 6% of WW children are most likely third generation—or more—immigrants. For my analysis, I dropped the WW observations identified as Hispanic to compare with non-Hispanic Whites.

In columns (v) and (vi), I calculate the differences in means between HH-WW (column v) and HW-WH (column vi). Overall, HH children do worse than WW children. HH has one less year of education compared to WW, 2.1 percentage points less likely to be employed, HH men earn 18% less than WW men, and HH women make 12.3% less than WW women. Therefore, these two groups, HH and WW, are different in many observable.

The differentials between WH and HW are not as severe as the ones between HH and WW. HW children are still worse off than WH but are way better than HH. HW men attain 0.294 fewer years of education than WH men. HW women achieve 0.471 fewer years of schooling than WH women. HW children are 1.8-1.9 percentage points less likely to be employed. HW men earn the same annual earnings as WH men, and HW women earn 10.3% less. These differences show that HW and WH children are different from children of homogeneous marriages and, thus, are more comparable.

Moreover, in table 4, I present the summary statistics for the four types and limit the sample to those that only identify as Hispanic<sup>7</sup>. The same pattern discussed above between HH-WW and HW-WH persists. The one difference is that those who identify as Hispanic do slightly worse.

### 1.3.1 Synthetic parents

Using the 1970 and 1990 censuses, I constructed a panel of synthetic parents. The sample includes married White men and women. Even though the census asks a person whether

---

<sup>7</sup>Limiting the sample to those that identify as Hispanic only affects WH, HW, and HH. Members of the WW group are not affected by this restriction.

they are Hispanic or not, I took advantage of the questions on place of birth to create a proxy for ethnicity. I consider a Hispanic person as White and born in a Spanish-speaking country. Consequently, Whites in the sample are people who are White and native-born. Using the information provided in the census, I can link husbands and wives with each other. I assume that parents have children between the ages of 25 and 45. Therefore, my sample consists of married White men and women that were born in the 1930 to 1965 cohorts<sup>8</sup>.

I show the distribution of the four types of couples in table 1. White husbands and White wives (WW) make up the majority of couples in the sample, 98.18% (681,206 couples). Hispanic husbands and wives (HH) are the second largest group making up 1.04% (7,193 couples) of all couples. White husbands and Hispanic wives (WH) couples are 0.40% (2,750 couples) of the sample while Hispanic husbands and White wives (HW) are 0.39% (2,710 couples).

I present the summary statistics of the four types of synthetic parents in table 3's parent's panel. WW couples have higher education, 11.9 years for husbands and 11.8 for wives. As a household, WW couples have 25.194 years of schooling. Husbands in HH marriages have 8.5 years of education, while women have 8.2 years of schooling. As a household, HH couples have 17.539 years of education.

WH husbands have 10.3 years of education, while wives have an average of 9.6 years, the lowest of any group. WH household attained a total of 22.308 years of schooling in total. HW husbands and wives obtained 9.7 and 9.8 years of education, respectively. HW household attained a total of 21.473 years of schooling in total.

---

<sup>8</sup>The construction of "synthetic parents" follows the method used by Rubinstein and Brenner (2014).

TABLE 3: Summary statistics of outcomes using parent's place of birth

Variables	Father's and Mother's Ethnicities				Differences	
	White Father	White Father	Hispanic Father	Hispanic Father		
	White Mother	Hispanic Mother	White Mother	Hispanic Mother	HH - WW	HW - WH
	(WW)	(WH)	(HW)	(HH)	(v)	(vi)
<b>Panel A: Parent's</b>						
Husband's education (Total Years)	13.05 (2.44)	12.32 (3.33)	10.65 (4.39)	8.93 (4.41)	-4.11** (0.02)	-1.67** (0.04)
Wife's education (Total Years)	12.74 (2.12)	11.03 (3.92)	11.54 (3.12)	8.6 (4.13)	-4.13** (0.02)	0.51** (0.04)
Total Household seduction (Total Years)	25.78 (4.08)	23.35 (6.51)	22.19 (6.69)	17.54 (7.83)	-8.25** (0.03)	-1.16* (0.07)
<b>Panel B: Education</b>						
Men's education (Total Years)	13.91 (2.39)	13.58 (2.35)	13.21 (2.32)	12.91 (2.26)	-1.00*** (0.01)	-0.36** (0.03)
Women's education (Total Years)	14.29 (2.41)	13.87 (2.47)	13.42 (2.35)	13.27 (2.37)	-1.01*** (0.01)	-0.46** (0.03)
<b>Panel C: Employment and Earnings</b>						
Men's Employment Rate	0.95 (0.22)	0.94 (0.23)	0.92 (0.26)	0.93 (0.26)	-0.02*** (0.00)	-0.02*** (0.00)
Women's Employment Rate	0.96 (0.2)	0.95 (0.22)	0.93 (0.25)	0.94 (0.24)	-0.02*** (0.00)	-0.02*** (0.00)
Men's Log Hourly Earnings	2.48 (0.45)	2.41 (0.45)	2.4 (0.44)	2.41 (0.43)	-0.07*** (0.01)	-0.00** (0.02)
Women's Hourly Earnings	2.33 (0.49)	2.33 (0.46)	2.27 (0.45)	2.3 (0.41)	-0.02*** (0.01)	-0.06** (0.02)
Men's Log Annual Earnings	10.25 (1.01)	10.08 (1.05)	10.04 (1.06)	10.01 (1.04)	-0.25** (0.01)	-0.04** (0.03)
Women's Hourly Earnings	9.46 (1.78)	9.54 (1.64)	9.47 (1.57)	9.53 (1.52)	0.07** (0.02)	-0.07* (0.06)
<b>Panel D: Hispanic Identity</b>						
Men	0.05	0.81	0.88	0.97		
Women	0.06	0.85	0.87	0.97		

<sup>1</sup> The data is restricted to native-born United States citizens between 1994 and 2019 who are also White and between the ages of 25 and 40. I identify the ethnicity of a person's parents through the parent's place of birth. A parent is Hispanic if they were born in a Spanish-speaking country. A parent is White if they were born in the United States.

<sup>2</sup> In each column, I present the average statistics of the different types of people based on the ethnicities of their parents. In column one, I show the summary statistics of children of White fathers and White mothers. In column two, I present the summary statistics of children of White fathers and Hispanic mothers. In column three, I show the summary statistics of children of Hispanic fathers and White mothers. In column four, I present the summary statistics of children of Hispanic fathers and mothers.

<sup>3</sup> Columns five and six have data on the HH-WW gaps (column five) and the HW-WH gaps (column six).

TABLE 4: Summary statistics of outcomes using parent's place of birth only for those that self-identify as Hispanic

Variables	Father's and Mother's Ethnicities				Differences	
	White Father	White Father	Hispanic Father	Hispanic Father	HH - WW	HW - WH
	White Mother	Hispanic Mother	White Mother	Hispanic Mother	(v)	(vi)
	(WW)	(WH)	(HW)	(HH)	(v)	(vi)
<b>Panel A: Parent's</b>						
Husband's education (Total Years)	13.05 (2.44)	12.32 (3.33)	10.65 (4.39)	8.93 (4.41)	-4.11** (0.02)	-1.67** (0.04)
Wife's education (Total Years)	12.74 (2.12)	11.03 (3.92)	11.54 (3.12)	8.6 (4.13)	-4.13** (0.02)	0.51** (0.04)
Total Household seducation (Total Years)	25.78 (4.08)	23.35 (6.51)	22.19 (6.69)	17.54 (7.83)	-8.25** (0.03)	-1.16* (0.07)
<b>Panel B: Education</b>						
Men's education (Total Years)	12.97 (2.15)	13.45 (2.37)	13.13 (2.27)	12.89 (2.25)	-0.08** (0.01)	-0.32** (0.03)
Women's education (Total Years)	13.23 (2.25)	13.75 (2.41)	13.32 (2.34)	13.26 (2.37)	0.03** (0.01)	-0.43** (0.03)
<b>Panel C: Employment and Earnings</b>						
Men's Employment Rate	0.93 (0.26)	0.94 (0.23)	0.92 (0.27)	0.93 (0.26)	0.00*** (0.00)	-0.02*** (0.00)
Women's Employment Rate	0.94 (0.25)	0.94 (0.23)	0.93 (0.26)	0.94 (0.24)	0.00*** (0.00)	-0.02*** (0.00)
Men's Log Hourly Earnings	2.4 (0.44)	2.41 (0.45)	2.4 (0.44)	2.41 (0.43)	0.01*** (0.01)	-0.01** (0.02)
Women's Hourly Earnings	2.26 (0.43)	2.32 (0.45)	2.27 (0.45)	2.3 (0.41)	0.04*** (0.01)	-0.05** (0.02)
Men's Log Annual Earnings	10.02 (1.02)	10.06 (1.06)	10.03 (0.99)	10 (1.04)	-0.02** (0.01)	-0.03** (0.04)
Women's Hourly Earnings	9.44 (1.59)	9.55 (1.59)	9.47 (1.55)	9.52 (1.52)	0.08** (0.02)	-0.08* (0.05)

<sup>1</sup> The data is restricted to native-born United States citizens between 1994 and 2019 who are also White and between the ages of 25 and 40. I identify the ethnicity of a person's parents through the parent's place of birth. A parent is Hispanic if they were born in a Spanish-speaking country. A parent is White if they were born in the United States.

<sup>2</sup> In each column, I present the average statistics of the different types of people based on the ethnicities of their parents. In column one, I show the summary statistics of children of White fathers and White mothers. In column two, I present the summary statistics of children of White fathers and Hispanic mothers. In column three, I show the summary statistics of children of Hispanic fathers and White mothers. In column four, I present the summary statistics of children of Hispanic fathers and mothers.

<sup>3</sup> Columns five and six have data on the HH-WW gaps (column five) and the HW-WH gaps (column six).

In columns (v) and (vi) of table 3, I calculated the differences in means between HH-WW households and HW-WH households. An HH Hispanic husband attained 3.779 fewer years of education than a WW husband. While an HH wife acquired 3.876 fewer years of schooling than a White WW wife. The total educational difference between HH and WW households equals 7.655 years of education. The educational differences between HW and WH are less severe than those between HH and WW. A Hispanic husband in intermarriage attained 1.560 fewer years of schooling than a White husband in intermarriage. A White wife in intermarriage attained 0.726 years of education, more than a Hispanic wife in intermarriage. The difference in total years of education between HW and WH households is 0.835, which is an 89% decrease from the HH–WW gap. These differences show that couples that intermarry are better and different than those that do not.

## 1.4 Empirical Strategy

In this section, I will present two empirical strategies. The first empirical strategy will estimate the effect of having a Hispanic last name on annual log earnings. The second empirical strategy will estimate the effect of identifying as Hispanic on log yearly earnings.

The difference in means between Hispanics and non-Hispanic Whites could result from discrimination. It can also be caused by differences in innate abilities, skills, and parental investments during infancy. To avoid these threats to identification, I compare children of inter-ethnic marriages, HW and WH. WH and HW children are equally Hispanic physically but provide employers/labor market with different signals.<sup>9</sup> WH children will have a non-Hispanic last name, while HW children will have a Hispanic last name. This is a method developed by Rubinstein and Brenner (2014).

---

<sup>9</sup>WH and HW children are both half White, half Hispanic.

### 1.4.1 Estimating the effect of having a Hispanic last name

Let  $Y_{ist}$  be the annual log earnings of person  $i$  in state  $s$  at time  $t$ .  $WH_{ist}$ ,  $HW_{ist}$  and  $HH_{ist}$  are indicator variables for the type of parents person  $i$  has.  $X_{ist}$  is a vector of controls,  $\gamma_t$  are year fixed effects,  $\lambda_s$  are state fixed effects, and  $\phi_{ist}$  represents the error term. The equation for this strategy is written as follows:

$$Y_{ist} = \beta_1 WH_{ist} + \beta_2 HW_{ist} + \beta_3 HH_{ist} + X_{ist}\pi + \gamma_t + \lambda_s + \phi_{ist} \quad (1)$$

$\beta_1$ ,  $\beta_2$  and  $\beta_3$  are the coefficients of interest in this specification. They estimate the earnings gaps between the three groups and children of White-White parents. The difference between  $\beta_1$  and  $\beta_2$  captures the effect of having a Hispanic last name. This comes from the fact that the WH and HW children are similar. On the one hand, WH children will have a White father and carry his White last name. On the other hand, HW children will have a Hispanic father and carry his Hispanic last name. Therefore, *ceteris paribus*, the difference will capture the effect of having a Hispanic last name.

### 1.4.2 Estimating the effect of identifying as Hispanic

In this section, I will present an estimation strategy that would allow me to capture the effect of identifying as a Hispanic. The empiricism in the section follows the theoretical identity model developed by Akerlof and Kranton (2000a).

Let  $Y_{ist}$  be the annual log earnings of person  $i$  in state  $s$  at time  $t$ .  $WH_{ist}$ ,  $HW_{ist}$  and  $WW_{ist}$  are indicator variables for the type of parents person  $i$  has.  $X_{ist}$  is a vector of controls,  $\gamma_t$  are year fixed effects,  $\lambda_s$  are state fixed effects, and  $\phi_{ist}$  represents the error term. The addition to the previous model is the  $Hispanic_{ist}$  indicator variable.  $Hispanic_{ist}$  is equal to one if a person identifies as Hispanic and zero otherwise. I interact this term with the

$WH_{ist}$ ,  $HW_{ist}$  and  $HH_{ist}$ . The regression is presented in the following equation.

$$Y_{ist} = \beta_1 HW_{ist} + \beta_2 WH_{ist} + \beta_3 HH_{ist} + \beta_4 WH_{ist} \cdot Hispanic_{ist} + \beta_5 HW_{ist} \cdot Hispanic_{ist} + \beta_6 HH_{ist} \cdot Hispanic_{ist} + X_{ist}\pi + \gamma_t + \lambda_s + \phi_{ist} \quad (2)$$

The coefficients of interest from equation 2 are  $\beta_4$ ,  $\beta_5$  and  $\beta_6$ . The interaction terms capture the earnings gaps between members of  $WH_{ist}$ ,  $HW_{ist}$ , and  $WW_{ist}$  that identify as Hispanic and White WW children.

## 1.5 Results and Discussion

In this section, I present the results from estimating the two specifications in equations 1 and 2. The results are shown in tables 5 and 6. All estimations are done on a sample of White, native-born men between the ages of 25 and 40 employed for full-time full-year (FTFY) and are waged and salaried workers.

I find that the earnings gaps between WH and WW are 14%, 18% between HW and WW, and 24% between HH and WW. Moreover, the difference in earnings between HW and WH children, that captures the effect of having a Hispanic last name, equals a statistically significant 6 percentage points. Meaning by comparing inter-ethnic children, a person with a Hispanic last name earns 6 percentage points less than someone with a White last name. However, more than half of the earnings gaps could be explained by educational differences between WW children and WH, HW, and HH. When I control for education, the gap between the three groups and WW is cut between by 50% to 64.0%. Furthermore, the last name effect decreases to a statistically insignificant 2 percentage points earnings gap.

I also find a significant earnings gap between those that identify as Hispanic and WW children. The gaps between WH, HW, and HH that identify as Hispanic and WW are, respectively, 14%, 18%, and 10%. The difference between WH and HW that identify as Hispanic—captures the last name effect of those that identify as Hispanic—is equal to 3 percentage points and is statistically insignificant. Controlling for education explains a big part of the gap. The gaps between WH, HW, and HH that identify as Hispanic and WW after controlling for education are 8%, 12%, and 6%. The WH–WW and HH–WW gaps for those identifying as Hispanic become statistically insignificant after controlling for education. Finally, the difference between WH and HW that identify as Hispanic that identifies as Hispanic remains statistically insignificant but increases to 4 percentage points.

### **1.5.1 The effect of having a Hispanic last name on labor market outcomes**

I provide the results to the estimation of equation 1 in table 5. The sample in this analysis includes full-time, full-year, and waged and salaried men, and I control for hours worked and age and include year fixed effects (FE). Column one in table 5 is the results without controlling for education. Overall, there is big Hispanic–White earnings gaps. Children of White father-Hispanic mothers (WH) earn 14% less than children of White fathers-White mothers (WW) families. The gap between children of Hispanic father-White mothers (HW) and WW children is 20%. The gap between children of Hispanic father-Hispanic mothers (HH) and WW children is 24%. Therefore, Hispanic children, inter-ethnic or children of homogeneous HH marriages, face a substantial earnings gap.

The difference between the estimates of WH and HW would capture the effect of having a Hispanic last name. A WH child will have a White last name since their father is White, while an HW child will have a Hispanic last name like their Hispanic father. When I take the difference between HW and WH, I find a statistically significant gap of 6 percentage points,

which means that an inter-ethnic person with a Hispanic last name earns 6 percentage points less than a similar inter-ethnic person but with a White last name.

The literature on the earnings gaps shows that educational differences could explain a big part of these gaps (Duncan et al., 2006; Duncan and Trejo, 2017a, 2018a; Duncan et al., 2020). Therefore, I run the equation 1, but this time I control for education. The regression results are in column two of table 5. The sample includes full-time, full-year, and waged and salaried men, and I control for hours worked and age and include year fixed effects (FE). The gap between WH and WW is 9%, a reduction of 35% from the baseline after controlling for education. The earning gaps between HW and HH, and WW were respectively reduced by 45% and 50%. After controlling for education, an HW man earns 11% less than a WW man, and an HH man earns 12% less. The effect of having a Hispanic last name on earnings decreased from 6 to 2 percentage points and became statistically insignificant after controlling for education.

Therefore, Hispanics face big earnings gaps compared to non-Hispanic Whites. By comparing the children of inter-ethnic marriages, I could study the impact of having a Hispanic last name on earnings. Having a Hispanic last name carries a cost in the labor market. A person with a Hispanic last name earns 6 percentage points less than someone with a White name. However, educational differences between Hispanics and Whites could explain more than 50% of the earning gaps and all of the Hispanic last name costs.

### **1.5.2 The effect of identifying as Hispanic on labor market outcomes**

In this section, I present the results of equation 2 that estimates the effect of identifying as Hispanic on annual log earnings. I show the results in table 6. The sample is restricted to full-time full-year (FTFY) salaried and waged men. I control for age and hours worked and include year fixed effects.

I run the estimating equation without controlling for education in table 6 column 1. I find that those that identify as Hispanic face a big earnings gap compared to those that do not. A WH man that identifies as Hispanic earns 14% less than a WW man. An HW man who identifies as Hispanic makes 18% less than a WW man, and an HH man who is determined as Hispanic earns 10% less than a WW man. Comparing the coefficients of the interaction terms to those of the  $WH_{ist}$ ,  $HW_{ist}$ , and  $HH_{ist}$  shows that members of the three groups that do not identify as Hispanic earn about the same as WW. Moreover, the difference between  $HW_{ist} \cdot Hispanic_{ist}$  and  $WH_{ist} \cdot Hispanic_{ist}$  captures the effect of having a Hispanic last name and identifying as Hispanic compared to individuals who identify as Hispanic but have a White last name. A person with a Hispanic last name that identifies as Hispanic earns 3 percentage points less than a person that identifies as Hispanic but has a White last name. The difference between the two interaction terms is statistically insignificant.

In table 6 column 2, I show the results of equation 2 while controlling for education. Education explains all the identity costs for members of WH and HH. Also, controlling for education decreases the earnings gap for identifying as Hispanic and being a member of HW by 33%. An HW member that identifies as Hispanic earns 12% less than a WW man. Furthermore, the difference between  $HW_{ist} \cdot Hispanic_{ist}$  and  $WH_{ist} \cdot Hispanic_{ist}$  increases to 4 percentage points but remains statistically insignificant.

TABLE 5: Effect of Having Hispanic Last Name

	(1)	(2)
	Log annual earnings	Log annual earnings
$WH_i$	-0.14*** (0.03)	-0.09*** (0.02)
$HW_i$	-0.20*** (0.02)	-0.11*** (0.01)
$HH_i$	-0.24*** (0.02)	-0.12*** (0.01)
$HW_i - WH_i$	-0.06** (0.03)	-0.02 (0.02)
Observations	129 359	129 359
State FE	X	X
Year FE	X	X
<i>Controlling for:</i>		
Hours Worked	X	X
Age	X	X
Education		X

\* p &lt; 0.1, \*\* p &lt; 0.05, \*\*\* p &lt; 0.01

<sup>1</sup> This table includes the estimation results of equation (1).

TABLE 6: Effect of Having Hispanic Last Name

	(1)	(2)
	Log annual earnings	Log annual earnings
$WH_i \times Hispanic_i$	-0.14*** (0.05)	-0.08 (0.05)
$HW_i \times Hispanic_i$	-0.18*** (0.07)	-0.12* (0.06)
$HH_i \times Hispanic_i$	-0.10* (0.06)	-0.06 (0.05)
$HW_i \times Hispanic_i - WH_i \times Hispanic_i$	-0.03 (0.07)	-0.04 (0.06)
Observations	129 359	129 359
State FE	X	X
Year FE	X	X
<i>Controlling for:</i>		
Hours Worked	X	X
Age	X	X
Education		X

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

<sup>1</sup> This table includes the estimation results of equation (2).

<sup>2</sup> The four groups stand for White Husband White Wife (WW), White Husband Hispanic Wife (WH), Hispanic Husband White (HW), and Hispanic Husband Hispanic Wife (HH). *Hispanic<sub>i</sub>* is a dummy variable equal to one if a person identifies as Hispanic and zero otherwise.

<sup>3</sup> The sample is restricted to men working full-time full-year and are waged and salaried workers.

<sup>4</sup> Column one has the regression results when controlling for hours worked, age, and years of fixed effects. Column two has the results after controlling for education.

<sup>5</sup> Standard errors are clustered on the state level.

TABLE 7: Hsiapnic–White Earnings Gap Using Hispanic Variable Only

	(1)	(2)
	Log annual earnings	Log annual earnings
$Hispanic_i$	-0.17*** (0.00)	-0.07*** (0.00)
Observations	137977	137977
<i>Controlling for:</i>		
Hours Worked	X	X
Age	X	X
Year FE	X	X
Education		X

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

<sup>1</sup> This table includes the estimation results of equation (3).

<sup>2</sup>  $Hispanic_i$  is a dummy variable equal to one if a person identifies as Hispanic and zero otherwise.

<sup>3</sup> The sample is restricted to men working full time full-year and are waged and salaried workers.

<sup>4</sup> Column one has the regression results when controlling for hours worked, age, and fixed effects. Column two has the results after controlling for education.  
22

### 1.5.3 Discussion

The results presented in tables 5 and 6 have an implication that affects how Hispanic-White earnings gaps should be calculated. First, A big portion of the earnings gaps could be explained by education. However, more unobservable characteristics, like innate ability, could affect earnings outcomes. Therefore, comparing Hispanics to Whites could yield inflated gaps since the two groups are different. The reduced form regression of this approach could be written as:

$$Y_{ist} = \beta_1 Hispanic_{ist} + X'_{ist} + \gamma_t + \lambda_s + \phi_{ist} \quad (3)$$

Where  $Y_{ist}$  in equation 3 is the annual log earnings of person  $i$  at time  $t$ .  $Hispanic_{ist}$  is a dummy variable equal to one if person  $i$  is Hispanic and zero otherwise.  $X_{ist}$  is a vector of controls,  $\gamma_t$  is time fixed effects and  $\phi_{ist}$  is the error term.

From equation 3,  $\beta_1$  estimates the earnings gap between Hispanics and Whites. When calculating the equation, before controlling for education, I find an earnings gap that is equal to 15.6%. This gap is close to the gap between HH and WW. After controlling for education, the gap decreases by more than half and becomes 6%, which is also similar to the gap between HH and WW after controlling for education. Since the differences between Whites and Hispanics, both observable and unobservable, are large, estimating the gap HW-WH could provide us with a bound on how much of the discrimination against Hispanics is taste-based.

The point estimate of the difference between HW and WH from equation 1 is equal to 1.3 percentage points. The upper bound of the confidence interval is equivalent to 4.5 percentage points. Therefore, I can put a potential upper bound on the portion of the Hispanic-White earnings gap that could be explained by taste based discrimination. In this case, the upper

bound is equal to 4.5 percentage points.

## 1.6 Conclusion

As the Hispanic population keeps growing in the United States, studying discrimination against this group is also important. In this paper, I examine discrimination against Hispanics in the labor market. More specifically, I examine the impact of Hispanic last names and Hispanic identification on annual log earnings.

I compare the children of inter-ethnic marriages to study the labor market impact of having a Hispanic last name. The earning gap between WH and WW is 14%, 20% between HW and WW, and 24% between HH and WW. When I compare the earnings of HW and WH children, which captures the effect of having a Hispanic last name, HW children earn 6% less than WH children. Thus, by comparing inter-ethnic children, a person with a Hispanic last name makes 6% less than someone with a White last name. Controlling for education, however, explains more than half of the earnings gaps between WW children and WH, HW, and HH. When I control for education, the gap between the three groups and WW is cut between 50-64.0%. Furthermore, the last name effect decreases to a statistically insignificant 2% earnings gap.

I also find that men who identify as Hispanic earn significantly less than those who do not, even when I control ancestry and education. The gaps between WH, HW, and HH that identify as Hispanic and WW are, respectively, 14%, 18%, and 10%. Controlling for education explains a big part of the gap. The gaps between WH, HW, and HH that identify as Hispanic and WW after controlling for education are 8%, 12%, and 6%.

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

### 2.1 Introduction

How should we count the sizes of various groups defined by race, ethnicity, and gender? Recent censuses have allowed individuals to check off multiple categories regarding race and ethnicity, reflecting a US population that has become increasingly multi-racial and multi-ethnic.<sup>10</sup> How individuals self-identify into various categories has significant political consequences on policy and economic research focusing on outcome gaps.

In this paper, I examine the determinants of self-identification as “Hispanic” among immigrants from Spanish-speaking countries. I focus on how much prejudice in a person’s place of residence impacts their identity decisions. Akerlof and Kranton (2000b) posit that actions have benefits and costs that are identity dependent. For example, an individual can engage in seemingly self-destructive behavior such as self-starvation if being thin accords with the proscription of being a woman. Adopting a specific identity then shapes people’s decisions, investments, and well-being in potentially profound ways.

Understanding the determinants of self-identification is vital for at least three reasons. First, the rates of self-identification by the group may have consequences for political representation and the distribution of resources. Second, how individuals identify may impact measured changes in labor market outcomes among groups differentiated by race and ethnicity. Antman et al. (2016a) show that among immigrants from Mexico, the least economically successful self-identify as being of Mexican origin, while the most successful do not. As a result, Mexican immigrants’ assimilation rates could appear slower than other groups. Third,

---

<sup>10</sup>For a discussion of the implications of multiple ethnic and racial categories and the future of the multi-racial and multi-ethnic demography of the United States see Bratter (2018); Alba (2020).

to the extent that individuals react to prejudice by not identifying with the group targeted by bias; the standard analyses that attempt to attribute some component of an ethnic gap in outcomes will be biased.

Most research on race and ethnicity relies on self-reported race and ethnic identity measures. Given the increasing importance of identity in contemporary American society and politics, as well as the growing number of individuals of Hispanic ancestry in the United States, questions related to assimilation, and mobility (Chetty et al., 2014b; Abramitzky et al., 2014), how identity could shift public opinions toward trade (Grossman and Helpman, 2021), and how racial and gender attitudes affect the racial and gender earnings gaps (Charles and Guryan, 2008a; Charles et al., 2018) are increasingly important.<sup>11</sup> <sup>12</sup>

To explore how individual characteristics and social attitudes toward Hispanics affect the self-reported identity of persons with Hispanic ancestry in the United States, 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. I motivate my analysis with a simple model in the vein of Akerlof and Kranton (2000b). The model makes explicit a path through which actions affect individuals’ utility via their identity. This model also introduces an externality where the actions of others have different effects on a person’s well-being and identity. Therefore, if a person has the ability to choose their identity credibly, then they will choose it in order to maximize their outcomes.

Measuring identity choices outside of a laboratory is challenging because I must observe and construct objective and self-reported measures of identity. I use data from the birthplace of a person and their ancestry to construct an objective measure of identity. I find the self-reported measure of identity to be correlated to individual and parental characteristics. I

---

<sup>11</sup>The number of Hispanics in the United States almost doubled between 1994 and 2021. Based on data from the Current Population Survey, the number of Hispanics increased from 9% in 1994 to 17% in 2021.

<sup>12</sup>For more on mobility Chetty et al. (2016, 2017, 2014a) and Abramitzky et al. (2020a, 2016); Chetty et al. (2014a) for more on assimilation.

also find that they are associated with variables on discrimination and racial 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 to lower self-reported Hispanic identity among Hispanic immigrants. I use data on the Implicit Association Test (IAT), which measures the implicit biases of participants, as a proxy for prejudice. The IAT data is retrieved from Harvard’s Project Implicit (Greenwald et al., 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 et al., 2017), voting behavior (Friese et al., 2007), and health (Leitner et al., 2016). I find that a one standard deviation increase in bias is associated with a seven percentage points decrease in the self-reported Hispanic identity among first-generation immigrants. Additionally, a one standard deviation increase in bias is associated with 13 percentage points decrease in the self-reported Hispanic identity among second-generation immigrants. Further, 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 parents born in a Spanish speaking country. Among third-generation Hispanic children with grandparents that were all born in a Spanish speaking country, a one standard deviation increase in bias is associated with 14 percentage points decrease in self-reported Hispanic identity. These results align with the findings of Charles et al. (2018) that sexism experienced by women at the place and time of birth (background sexism) and sexism experienced by women where they currently live (residential sexism) is correlated with lower wages and labor force participation. I also find that state-level bias is uncorrelated with migration decisions and is correlated with a higher probability of interethnic marriage formations.

This paper contributes to three different strands of literature. First, there is a growing

theoretical and empirical economic literature on racial and ethnic identities. Kranton et al. (2020) used an experimental lab design to deconstruct group biases and found significant preferences toward in-group reported identities. Bonomi et al. (2021) posit a model with class and cultural identities. They showed how cultural or class salient issues could increase polarization and affects distributive policies. Fryer and Torelli (2010) introduced a theory where black pupils would punish their peers when they strive for economic success by “Acting White.” There is an abundance of empirical studies investigating the effect of identity on many outcomes. Giuliano et al. (2009) found that non-White managers hire more Whites and fewer Black employees. In Giuliano et al. (2011), they found that employees of a certain race experience better outcomes when their manager shares the same race. Bagues et al. (2017) showed that having more women present on hiring committees for academic jobs in Italy and Spain did not decrease the gender gap. Åslund et al. (2014) found that immigrant managers hired more immigrant employees than native managers. Others have looked at the effect of distinct Black names and Hispanic last names on labor market outcomes (Bertrand and Mullainathan, 2004b; Fryer and Levitt, 2004; Hadah, 2020). Gershenson et al. (2016) found that non-Black teachers had lower expectations of Black students, and Dee (2005) showed that teachers had different evaluations when teacher and student did not share the same race.<sup>13</sup>

The second strand of the literature this paper fits in is the economics of immigration and

---

<sup>13</sup>Other works include Jardina (2019) reported that in the United States, people that identify more with a ‘White’ identity were more likely to support protectionist policies. Using a large experiment, Alesina et al. (2018) showed that by simply invoking the identity of native *versus* immigrant, voters supported less redistribution. Waters (2001) finds that West Indian immigrants in the US, especially second-generation, that keep their identity and resist assimilation into American culture (Americanization) were more likely to succeed economically. Grossman and Helpman (2021) constructed a theoretical model to explain what led to the recent de-liberalization of trade. More specifically, they investigated whether voters’ opinion of globalization is a function of their own self-interest and the self-interest of people that share their identity. The model showed that under some circumstances, as in changes in political and cultural conditions, more people would identify with the ‘broad nation.’ Their support for trade policies would hinge on the welfare of other members of the ‘broad nation’ and not their self-interest. Consequently, this increased the support for protectionist policies.

assimilation. Abramitzky et al. (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.<sup>14</sup> Fouka et al. (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 that intermarry earn significantly more than those in an endogamous marriage.

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.<sup>15</sup> 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 could 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 self-reported Hispanic identity to study Hispanics (Duncan and Trejo, 2017b, 2011b; Meng and Gregory, 2005; Duncan and Trejo, 2018b,d; Antman et al., 2016a, 2020b). 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 et al., 2016a, 2020b).

The third strand this paper fits in is the research on attitudes and discrimination toward minorities. Charles and Guryan (2008a); Charles et al. (2018) found that attitudes toward

---

<sup>14</sup>For more on immigrant assimilation, see Abramitzky et al. (2020c, 2019, 2020b, 2014)

<sup>15</sup>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 et al., 2016a, 2020b).

Black people (prejudice) and women (sexism) are correlated with the racial and gender gaps. Chetty et al. (2020) found that higher implicit bias toward Black People, captured by Implicit Association Test (IAT) scores, is associated with lower income ranks and, thus, lower intergenerational mobility. Glover et al. (2017) also found that bias, measured by Implicit Association Test (IAT) scores, of grocery stores' managers negatively affected minority job performance. Bursztyn et al. (2022) showed that the exposure to Arab-Muslims decreased the non-Arab-Muslim Whites' implicit bias toward Arab-Muslims.

This paper is most closely related to Antman et al. (2016a); 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.<sup>16</sup> Taking into consideration the ethnic attrition that Antman et al. (2016a) 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 introduce a theoretical model in section (2.2). Second, I will describe the data I use in section (2.3). Third, I will introduce an empirical model in section (2.4). Fourth, I will discuss the results in section (2.5). Finally, I conclude in section (2.7).

## 2.2 Theoretical Model

I introduce a model of identity in the spirit of Akerlof and Kranton (2000b). A person belongs to some ethnic group, and their actions either affirm or deny their ethnic identity.

---

<sup>16</sup>Ethnic attrition is when a person with Hispanic ancestry fails to self-identify as Hispanic.

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  $I_i$ :

$$U_i = U_i(\mathbf{a}_i, \mathbf{a}_{-i}, I_i) \quad (4)$$

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:

$$I_i = I_i(\mathbf{a}_i, \mathbf{a}_{-i}; \mathbf{B}_{e_i}) \quad (5)$$

Where  $\mathbf{a}_i$  is the actions of person  $i$ ,  $\mathbf{a}_{-i}$  is the actions of others, and  $I_i$  is the identity function. Each group has an associated set of behaviors that society proscribes them to conform to, which I denote as  $\mathbf{B}_{e_i}$ .<sup>17</sup>

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.):

$$\frac{\partial U_i}{\partial a_i} + \frac{\partial U_i}{\partial I_i} \cdot \frac{dI_i}{da_i} = 0 \quad (6)$$

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

That is  $i$  will change identities when the benefits of doing so  $\tilde{U}_i^* - U_i^*$  exceed the costs

---

<sup>17</sup>Akerlof and Kranton (2000b) refer to  $B_{e_i}$  as proscription.

c. These net benefits are non-zero only if  $\frac{dI_i}{da_i} \neq 0$ . 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  $a_{-i}$  under different identities, and (3) the analysis should focus on a sample of the population with small  $c$ . 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 no Hispanic ancestry.

## 2.3 Data

### 2.3.1 Measuring Hispanic Identity

I use multiple datasets to measure Hispanic identity and prejudice. The first dataset I use to measure Hispanic identity is the Integrated Public Use Microdata Series (IPUMS) Current Population Survey (CPS) (Flood et al., 2021a). I use the Integrated Public Use Microdata Series (IPUMS) Current Population Survey (CPS) dataset to measure Hispanic identity and prejudice in the US population. The CPS is a comprehensive dataset collected monthly by the United States Census Bureau and the Bureau of Labor Statistics, with the purpose of measuring unemployment rates in the country. The richness of the CPS dataset allows me to link household members with each other and construct an objective measure of the Hispanic identity of minors under the age of 17 living with their parents. The CPS is a comprehensive dataset collected monthly by the United States Census Bureau and the Bureau of Labor Statistics.

The survey is of US households to measure unemployment, and the monthly survey is of more than 65,000 households.

The richness of the CPS allows me to link household members with each other. Coupled with the fact that the CPS started asking respondents for the birthplace of their parents beginning in 1994, I could perfectly identify, and construct, a dataset of first-, second-, and third-generation Hispanic immigrants. This will consequently allow me to build an objective measure of the Hispanic identity of minors under the age 17 living with their parents. This 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.<sup>18</sup> 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 (8).

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. Finally, the average family income in the sample was \$39,882. I provide the rest of the summary statistics for the overall sample and broken down by generation in table (8).

As previously stated, data on the objective Hispanic measure is not readily available.

---

<sup>18</sup>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.

Consequently, it is hard to identify the different generations of Hispanics in the United States using public-use data. Therefore, much of the literature studying Hispanics relied on a self-reported measure of Hispanic identity. A person is identified as Hispanic if they answer affirmatively whether they are Hispanic, Spanish, or Latino. Using self-reported identity to check the assimilation and intergenerational mobility of Hispanics in the United States ignores that some immigrants might decide not to identify as Hispanic.

TABLE 8: CPS Summary Statistics

Characteristic	Overall		By Generation	
	All Sample	First	Second	Third
	N = 1,131,828	N=119,778	N=761,450	N=254,699
Female	0.49	0.48	0.49	0.49
Hispanic	0.91	0.96	0.94	0.82
Age	8.6 (5.1)	11.5 (4.3)	8.3 (5.0)	7.9 (5.0)
College Graduate: Father	0.14	0.15	0.11	0.23
College Graduate: Mother	0.14	0.15	0.11	0.22
Total Family Income (1999 dollars)	39,882 (48,692)	31,927 (38,804)	36,726 (45,353)	53,000 (58,984)

\* Mean (SD)

<sup>1</sup> The sample includes children ages 17 and below who live in intact families. First-generation Hispanic immigrant children that were born in a Spanish speaking country. 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 grand parent born in a Spanish speaking country.

<sup>2</sup> Data source is the 2004-2021 Current Population Survey.

To avoid the problem introduced above, 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.<sup>19</sup> The advantage is that the Current Population Survey (CPS) asks respondents about their parent's place of birth as part of the CPS survey; I can link children that are 17-year-old and younger that still live with their parents to their parents and identify four generations of Hispanic immigrants.

Furthermore, I can identify three generations of Hispanic immigrants. Figure (1) offers a visual representation of the algorithm I used to determine the different generations. The different generations are defined according to the following:

1. A person is a first-generation if they are born in a Spanish-speaking country with non-US-born parents
2. A person is a second-generation Hispanic if they are born in the US and at least one parent is born in a Spanish speaking country
3. A person is a third-generation Hispanic if they are born in the US, both parents are the US born, and they have at least one grandparent that is born in a Spanish speaking country

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:

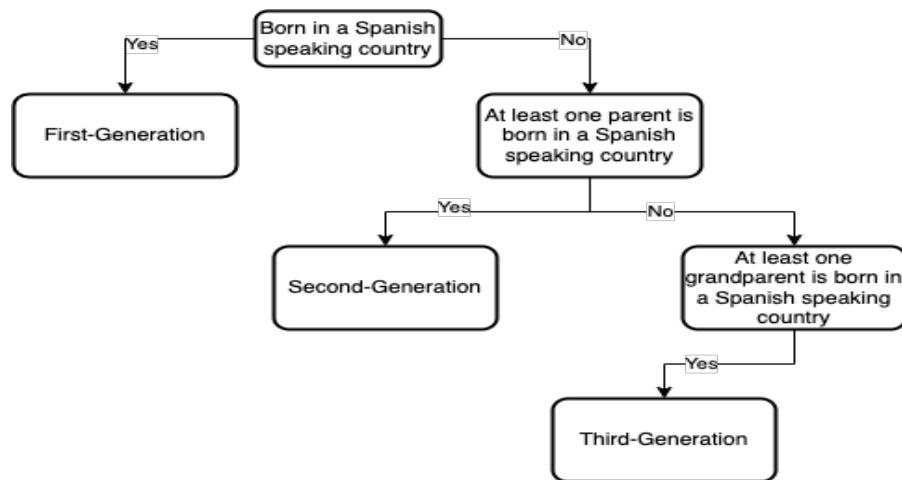
1. Objectively Hispanic-father-Hispanic-mother (HH)
2. Objectively Hispanic-father-White-mother (HW)
3. Objectively White-father-Hispanic-mother (WH)

---

<sup>19</sup>Following the works of Antman et al. (2016a, 2020b).

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 (9) 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.<sup>20</sup> These results are displayed in table (9): most first- and second-generation Hispanic immigrants self-reportedly identified as Hispanic. Among first-generation Hispanic immigrants, 96% of the group self-reportedly identified as Hispanic. Among second-generation Hispanic immigrants, 95% of the group self-reportedly identified as Hispanic. Attrition rates increase drastically after the second-generation, with 85% of third-generation Hispanic immigrants identifying as Hispanic. That is more than three folds increase in attrition rates. Most of the attrition among third-generation Hispanics is driven by attrition among the children of inter-ethnic marriages.

TABLE 9: Hispanic Self-identification by Generation

	Self-identify as Hispanic	Self-identify as non-Hispanic	% Self-identify as Hispanic	% Self-identify as non-Hispanic
<b>1st Gen.</b>	114657	5121	0.96	0.04
<b>2nd Gen.</b>	712916	48534	0.94	0.06
<b>Hispanic on:</b>				
<b>Both Sides</b>	516551	19318	0.96	0.04
<b>One Side</b>	196365	29216	0.87	0.13
<b>3rd Gen.</b>	209206	45493	0.82	0.18
<b>Hispanic on:</b>				
<b>Both Sides</b>	55401	2245	0.96	0.04
<b>One Side</b>	52879	17371	0.75	0.25

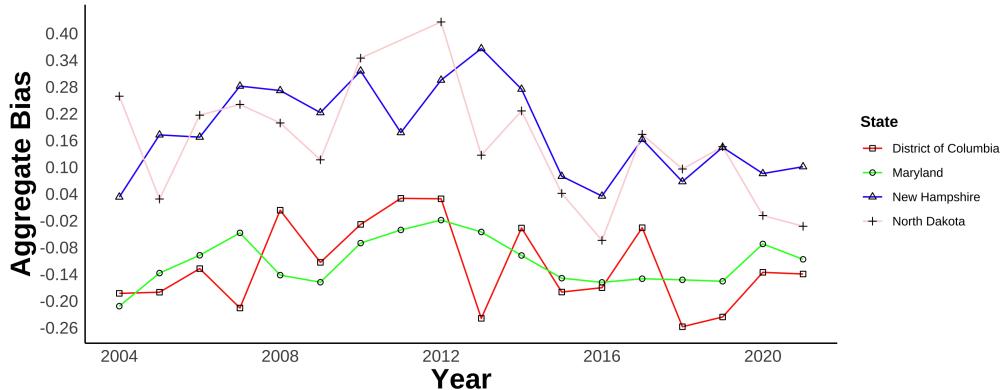
<sup>1</sup> The samples include children ages 17 and below who live in intact families. First-generation Hispanic immigrant children that were born in a Spanish speaking county. Native born second-generation Hispanic immigrant children with at least one parent born in a Spanish speaking country. Finally, native born third-generation Hispanic immigrant children with native born parents and at least one grandparent born in a Spanish speaking country.

<sup>2</sup> Data source is the 2004-2021 Current Population Survey.

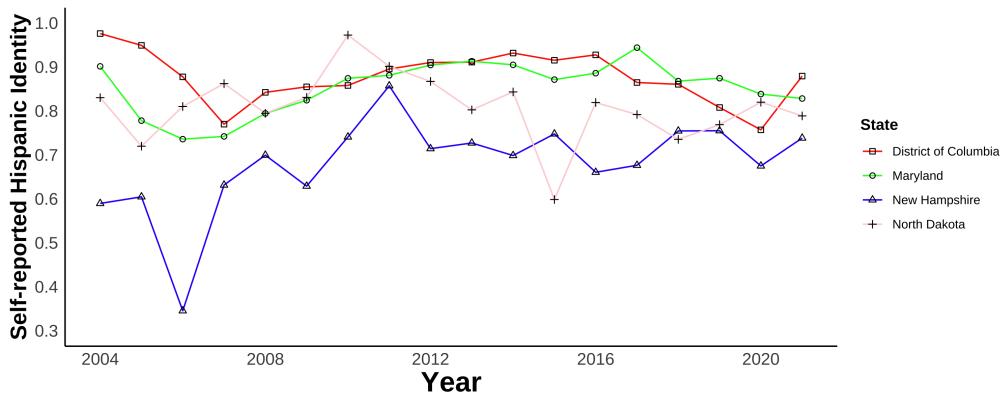
<sup>20</sup>In Duncan and Trejo (2018b,d); Antman et al. (2016a, 2020b), the authors find substantial attrition among Hispanics.

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

(a) Skin Tone Implicit Association Bias Over Time



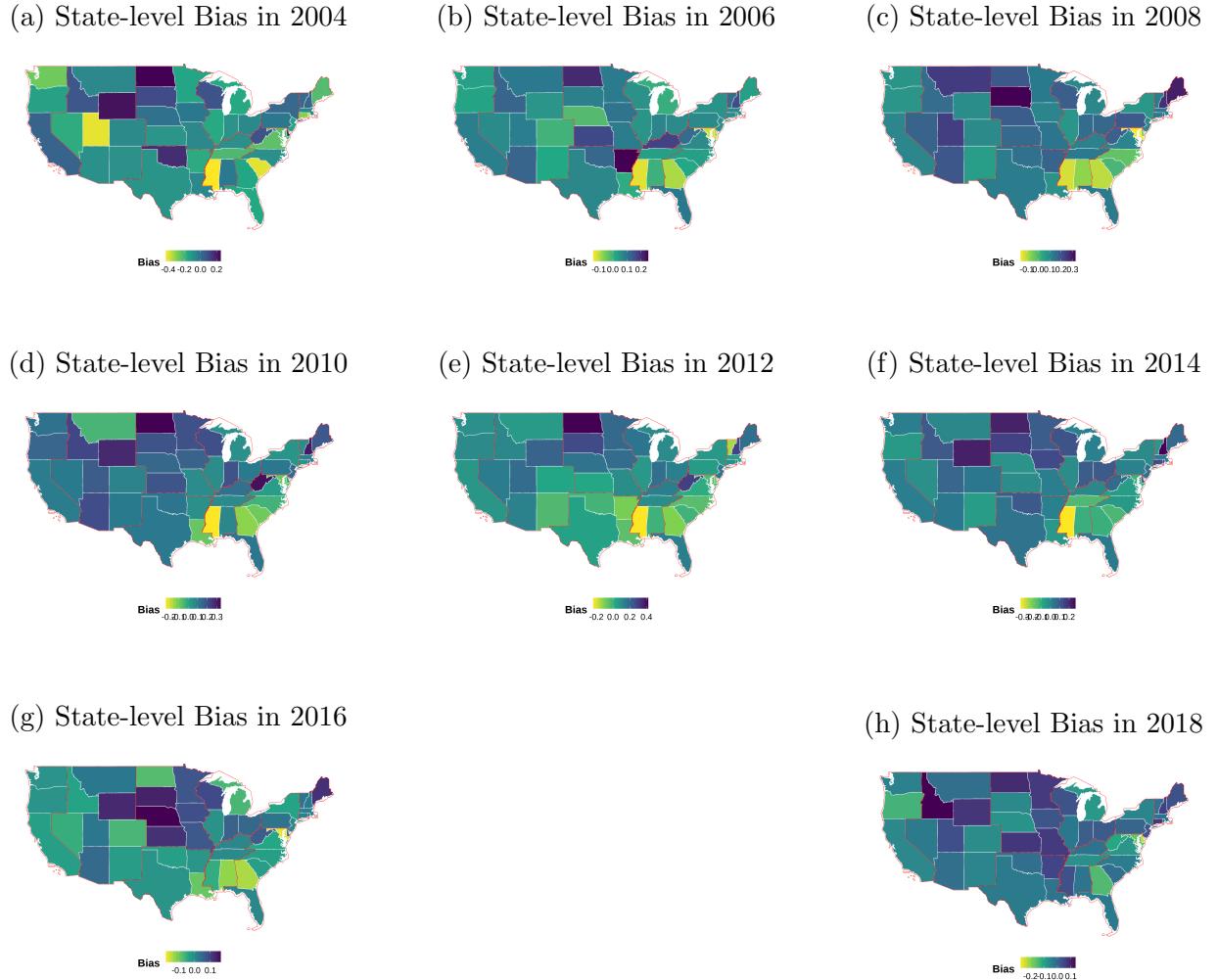
(b) Self-reported Hispanic Identity Over Time



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.

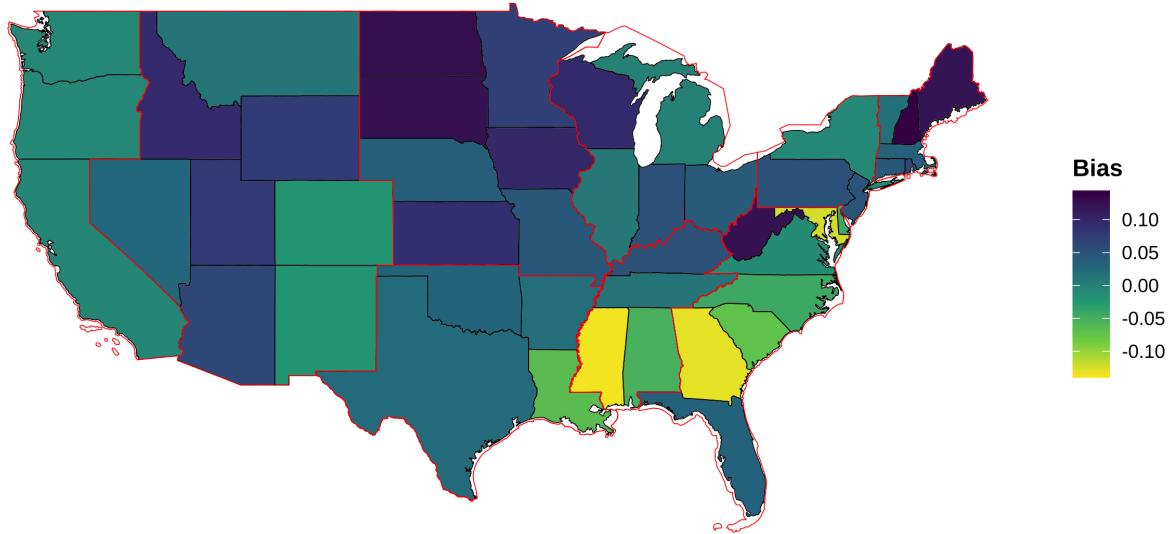
Bias data is from the 2004-2021 Harvard's Project Implicit Association Test scores. Identity data is from the 2004-2021 Current Population Survey (CPS).

Figure 3: Maps of State-level Implicit Association Test Bias Over Time Measure with Census Division Regional Boundaries.



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.

### 2.3.2 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 (22) 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 state-level prejudice (Greenwald et al., 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) and are correlated to economic outcomes (Chetty et al., 2020; Glover et al., 2017), voting behavior (Friese et al., 2007), and health (Leitner et al., 2016). Moreover, Bursztyn et al. (2022) show that exposure to Arab and Muslim immigrants is predictive of lower IAT scores. Figure (2a) shows a 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).

Participating 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). A summary of the sample statistics is presented in the table (10). The average age of an Implicit Association Test (IAT) test taker is 28 years. Of the test takers, 68% are women, and 62% are Whites or 56% non-Hispanic Whites. The test takers are highly educated. I also present the demographics of the representative Current Population Survey (CPS) sample over the same period in table (10). The CPS and IAT samples are different demographically, with the IAT sample being more female, less White, and with higher levels of education.

TABLE 10: Skin Tone Implicit Association Test (IAT) Scores and Current Population Survey (CPS) Samples

Characteristic	IAT	CPS
	N = 1,519,309	N = 29,981,618
<b>Age</b>	28	38
	(11)	(23)
<b>Female</b>	0.68	0.51
<b>White</b>	0.62	0.81
<b>Non-Hispanic White</b>	0.56	0.68
<b>Hispanic</b>	0.14	0.13
<b>Education Levels</b>		
Bachelor's degree	0.17	0.14
High school dropout	0.10	0.33
High school graduate	0.09	0.23
Master's degree	0.12	0.06
Other	0.48	0.21
Professional degree	0.05	0.02
<b>Bias</b>	0.30	
	(0.42)	

\* Mean (SD)

<sup>a</sup> Data source is the 2004-2021 Harvard's Project Implicit Association Test scores and 2004-2020 Current Population Survey (CPS).

## 2.4 Empirical Strategy

### 2.4.1 The Determinants of Hispanic Identity

I first estimate if state-level bias is correlated with self-reported Hispanic identity. Let  $H^g$  be the self-reported Hispanic identity variable for a person from the  $g^{th}$  generation, where  $g \in \{1, 2, 3\}$ , and  $X$  be a vector of controls. I estimate different specifications for each generation  $g$ . The regression I will estimate is as follows:

$$H_{ist}^g = \beta_1^g Bias_{st} + \beta_2^g DadCollegeGrad_{ist} + \beta_3^g MomCollegeGrad_{ist} + \beta_4^g Woman_{ist} \\ + X_{ist}^g \pi + \gamma_{rt} + \varepsilon_{ist}; \text{ where } g \in \{1, 2, 3\} \quad (7)$$

Where  $H_{ist}^g$  is the self-reported identity of person  $i$  (whether the answer to the question “Are you of Hispanic or Latino origin?” was yes or no) that lives in state  $s$  and was interviewed in year  $t$ ,  $Bias_{st}$  is the average bias in state  $s$  at year  $t$ ,  $DadCollegeGrad_{ist}$  and  $MomCollegeGrad_{ist}$  are indicator variables equal to 1 if the parents of person  $i$  are college graduates and zero otherwise, and  $Woman_{ist}$  is dummy variable that are equal to one if person  $i$  is a woman.  $X_{ist}^g$  is a vector of controls. Additionally,  $\gamma_{rt}$  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.

The coefficient of interest in equation (7),  $\beta_1^g$ , captures the relationship between state-level bias and self-reported Hispanic identity. If  $\beta_1^g > 0$ , then individuals living in a state with higher bias would be more likely to self-report Hispanic identity. The coefficient  $\beta_1^g$  captures the with-in region and across state variation in bias which could be seen in figures (4), (2a), and (2b). In other words, I am estimating the relationship between bias and self-reported

Hispanic identity by comparing states within a region to each other.

It could be the case that the desire to self-identify as Hispanic differs depending on the particular type of parents or grandparents. Therefore, I will further explore the relationship between bias and self-reported Hispanic identity by estimating the same equation by parents' and grandparental types. I present another specification where I estimate the heterogeneous effect of bias on the different kinds of second- and third-generation immigrants by type of parents and grandparents, equation (7). For example, I could estimate a regression on a sample of second-generation immigrants by parental type, where parental types depend on parental place of birth. I first estimate a model on a sample of second-generation immigrants by parental types, then on third-generation immigrants by grandparental types:

$$H_{ist}^2 = \beta_1^2 Bias_{st} + \sum_n \delta_n^2 I_{\{ParentType_{ist}=n\}} \quad (8)$$

$$+ \sum_m \alpha_m^2 Bias_{st} \times I_{\{ParentType_{ist}=m\}}$$

$$+ \beta_3^2 DadCollegeGrad_{ist} + \beta_4^2 MomCollegeGrad_{ist}$$

$$+ \beta_5^2 Woman_{ist} + X_{ist}^2 \pi + \gamma_{rt} + \varepsilon_{ist}$$

$$H_{ist}^3 = \beta_1^3 Bias_{st} + \sum_n \delta_n^3 I_{\{GrandParentType_{ist}=n\}} \quad (9)$$

$$+ \sum_m \alpha_m^3 Bias_{st} \times I_{\{GrandParentType_{ist}=m\}}$$

$$+ \beta_3^2 DadCollegeGrad_{ist} + \beta_4^2 MomCollegeGrad_{ist}$$

$$+ \beta_5^2 Woman_{ist} + X_{ist}^2 \pi + \gamma_{rt} + \varepsilon_{ist}$$

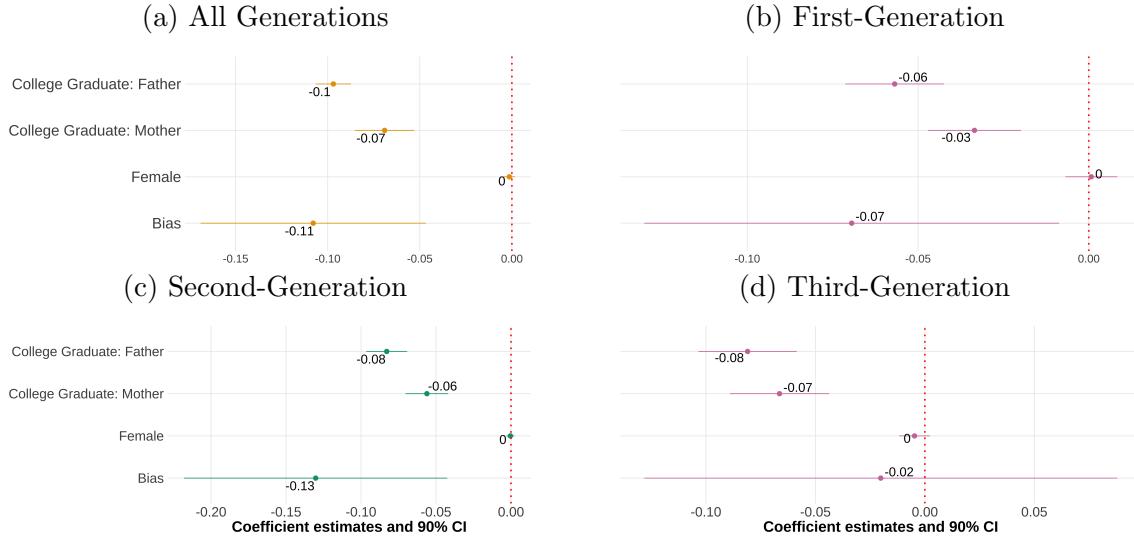
Where  $H_{ist}^2$  and  $H_{ist}^3$  are the self-reported identity of second and third-generation person  $i$  that lives in state  $s$  at time  $t$ . Regressions (8) and (9) include the state-year average bias, indicator variables  $I_{\{ParentType_{ist}=n\}}$  and  $I_{\{GrandParentType_{ist}=n\}}$ , and the interaction between the bias and parental and grandparent type. The variable  $I_{\{ParentType_{ist}=n\}}$  includes three categories of parents: 1) father born in a Spanish Speaking country and mother born in a Spanish-speaking country, or Hispanic-Hispanic (HH) and I will use as the reference category, 2) father born in a Spanish Speaking country and mother born in the US, or Hispanic-White (HW), 3) father born in the US and mother born in a Spanish speaking country, or White-Hispanic (WH). Thus, indicator variable  $I_{\{ParentType_{ist}=WH\}}$  is equal to one if the person  $i$  has an objectively White father-Hispanic mother, and zero otherwise. The variable  $I_{\{GrandParentType_{ist}=n\}}$  includes 15 possible combinations of the types of grandparents: 1) paternal grandfather born in a Spanish-speaking country, paternal grandmother born in a Spanish-speaking country, maternal grandfather born in a Spanish speaking country, and maternal grandmother born in a Spanish speaking country (HHHH) and I will use as the reference category, 2) paternal grandfather born in a Spanish speaking country, paternal grandmother born in a Spanish speaking country, maternal grandfather born in a Spanish speaking country, and maternal grandmother born in the US (HHHW), 3) paternal grandfather born in a Spanish speaking country, paternal grandmother born in a Spanish speaking country, maternal grandfather born in the US, and maternal grandmother born in the US (HWW), etc.

The coefficients of interest in equation (8),  $\alpha_n^2$ , capture the relationship between state-level bias and self-reported Hispanic identity by parents' type of second-generation immigrants. In other words,  $\alpha_n^2$  estimates the difference in self-reported Hispanic identity between children of Hispanic-White (HW) and White-Hispanic (WH) and the reference group Hispanic-Hispanic (HH) and its relationship with bias. For example, if  $\alpha_1^2$ , the coefficient of the indicator

variable  $I_{\{ParentType_{ist}=\text{White-Hispanic(WH)}\}}$ , is  $\alpha_1^2 > 0$ , then the difference between a child of an objectively White father and Hispanic mother (WH) and objectively Hispanic father and Hispanic mother (HH) is positively correlated with state-level bias. Consequently, an increase in bias is associated with higher self-reported Hispanic identity among objectively White father and Hispanic mother (WH) children compared to objectively Hispanic father and Hispanic mother (HH) children.

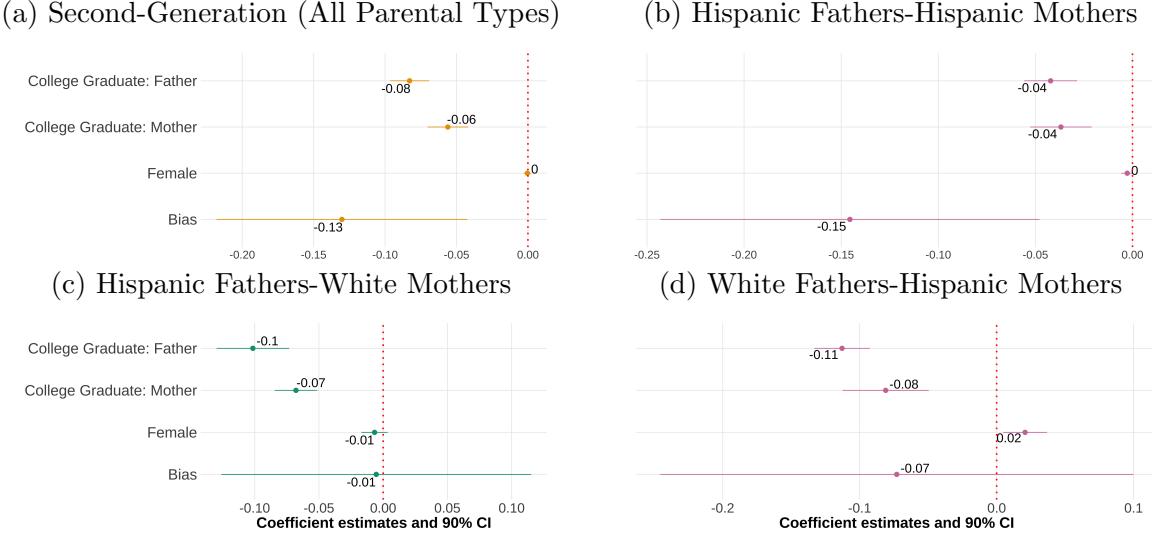
Furthermore, the coefficients of interest in equation (9),  $\alpha_m^3$ , captures the relationships between state-level bias and self-reported Hispanic identity by grandparents' type of third-generation immigrants. For example, if  $\alpha_1^3$ , the coefficient of the indicator variable  $I_{\{GrandParentType_{ist}=\text{Hispanic-Hispanic-White-White(HHWW)}\}}$ , is  $\alpha_1^3 > 0$ . In this case, the difference between the child of objectively Hispanic-Hispanic-White-White grandparents (HHWW) and objectively Hispanic-Hispanic-Hispanic-Hispanic (HHHH) grandparents is positively correlated with state-level bias. Consequently, an increase in bias is correlated with higher self-reported Hispanic identity among objectively Hispanic-Hispanic-White-White (HHWW) children compared to objectively Hispanic-Hispanic-Hispanic-Hispanic (HHHH).

Figure 5: Relationship Between Self-Reported Hispanic Identity and Bias: By Generation.



I show four panels of estimating equation (7). I include region  $\times$  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 country. 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.

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



I show four panels of estimating equation (7). I include region  $\times$  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.

#### 2.4.2 Estimating the Relationship Between Bias, Interethnic Marriages, and Migration

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

$$\begin{aligned}
 HH_{ist}^2 &= \beta_1^2 Bias_{st} + X_{ist}^2 \pi + \gamma_{rt} + \varepsilon_{ist} \\
 HW_{ist}^2 &= \beta_1^2 Bias_{st} + X_{ist}^2 \pi + \gamma_{rt} + \varepsilon_{ist} \\
 WH_{ist}^2 &= \beta_1^2 Bias_{st} + X_{ist}^2 \pi + \gamma_{rt} + \varepsilon_{ist}
 \end{aligned} \tag{10}$$

Where  $HH_{ist}^2$ ,  $HW_{ist}^2$ , and  $WH_{ist}^2$  are variables representing the type of parents of a second-generation Hispanic immigrant  $i$  in state  $s$  at time  $t$ .  $HH_{ist}^2$  represents Hispanic-husband-Hispanic-wife couples,  $HW_{ist}^2$  represents Hispanic-husband-White-wife couples, and  $WH_{ist}^2$  represents White-husband-Hispanic-wife couples.  $Bias_{st}$  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 year of immigration to the United States.

There are two ways to estimate the effect of bias on the probability of an interethnic marriage forming. One way is by estimating equations (10) in three separate linear probability models. Another way is by estimating an ordered multinomial logistic regression. The dependent variable in the logistic regression will be an ordinal variable with a value of (1) zero corresponding to an endogenous marriage; (2) one corresponding to an interethnic Hispanic-husband-White-wife marriage; (3) two corresponding to an interethnic White-husband-Hispanic-wife marriage. The coefficient of interest from (10) is  $\beta_1$ , which estimates the effect of bias on the probability of forming an interethnic marriage. If  $\beta_1 < 0$ , higher bias is correlated with lower log odds of an interethnic marriage forming.

The relationship between bias and interethnic marriage could also be estimated using the following regression.

$$Interethnic_{ist}^2 = \beta_1^2 Bias_{st} + X_{ist}^2 \pi + \gamma_{rt} + \varepsilon_{ist} \quad (11)$$

Where  $Interethnic_{ist}^2$  is an indicator variable that is equal to one if a couple are interethnic, i.e. Hispanic husband-White wife or White husband-Hispanic wife. I could estimate equation (11) using a linear probability model.

I am also interested in investigating the relationship between state-level bias and migration. For this purpose, I use the following models to estimate the relationship between state-level bias and migration:

$$BirthPlaceMigration_{ist}^2 = \beta_1^2 Bias_{st} + X_{ist}^2 \pi + \gamma_{rt} + \varepsilon_{ist} \quad (12)$$

$$BirthPlaceMigration_{ilb}^2 = \beta_1^2 Bias_{lb} + X_{ilb}^2 \pi + \gamma_{lb} + \varepsilon_{ilb} \quad (13)$$

Where  $BirthPlaceMigration_{ist}^2$  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 their state of birth and zero otherwise.  $BirthPlaceMigration_{ilb}^2$  is an indicator variable that is equal to one if person  $i$  in birthplace  $l$  does not currently live in the same state they lived in at the year of birth  $b$  and zero otherwise. The analysis of equations (12) and (13) is restricted to second-generation Hispanic immigrants with both parents born in a Spanish speaking country. The coefficient of interest from the regressions is  $\beta_1^2$ , which captures the effect of bias on migration.

Furthermore, to study the relationship between bias and the migration variables introduced above, I use two different ways to define the bias variable. In the first specification from equation (12), I estimate the relationship between the average bias at the time of the interview  $t$  in state  $s$  and  $BirthPlaceMigration_{ist}^2$ . In the Second specification from equation (13), I estimate the relationship between the average bias in state of birth  $l$  at the year of birth  $b$  on  $BirthPlaceMigration_{ilb}^2$ .

I will also estimate the selection into states based on bias levels. In other words, I want to estimate whether there's a relationship between the difference in state-level bias between the state  $i$  is currently living in and their birthplace and self-reported Hispanic identity. The estimation equation for the relationship is:

$$Y_{ist} = \beta_0 + \beta_1^2 Hispanic_{ist} + X_{ist}^2 \pi + \varepsilon_{ist} \quad (14)$$

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 state of birth  $l$  at the birth year  $b$ . The analysis is restricted to second-generation Hispanic immigrants with both parents born in a Spanish speaking country ‘that migrated from the state they were born in  $b$  to another state  $s$ . The coefficient of interest is  $\beta_1^2$ , which captures whether a person that self-reports Hispanic identity self-selects and migrates into states that are more or less biased. For example, if  $\beta_1^2 > 0$ , then a person who self-reports Hispanic identity, compared to a person that does not, moves from states with less bias to a state with more bias.

## 2.5 Results and Discussion

In this section, I will present and discuss the results that I find to answer the following question. Is Hispanic self-identification associated with attitudes toward ethnic minorities in places individuals live in? Is the probability of an interethnic marriage associated with attitudes toward ethnic minorities? Are migration decisions associated with attitudes toward ethnic minorities in places they live in?

### 2.5.1 Relationship between Attitudes and Identity

I find that state-level bias is negatively correlated with self-reported Hispanic identity. I find that a one standard deviation increase in state-level bias is correlated with 7 percentage points decrease in self-reported Hispanic identity among first-generation Hispanic immigrants

and a decrease of 13 percentage points in self-reported Hispanic identity among second-generation Hispanic immigrants. I also find a negative correlation between bias and self-reported Hispanic identity in second-generation Hispanic immigrants children of parents born in a Spanish-speaking country and third-generation Hispanic immigrants children of four grandparents born in a Spanish-speaking country.

TABLE 11: Relationship Between Bias and Self-Reported Hispanic Identity: By Generation

	(1)	(2)	(3)	(4)
	All Gens	First Gen	Second Gen	Third Gen
	$H_{ist}$	$H_{ist}^1$	$H_{ist}^2$	$H_{ist}^3$
Bias	-0.10** (0.04)	-0.07* (0.04)	-0.13** (0.05)	-0.02 (0.07)
Female	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
College Graduate: Mother	-0.05*** (0.01)	-0.03*** (0.01)	-0.06*** (0.01)	-0.07*** (0.01)
College Graduate: Father	-0.06*** (0.01)	-0.06*** (0.01)	-0.08*** (0.01)	-0.08*** (0.01)
N	844 481	85 390	560 100	198 991
Mean	0.91	0.96	0.94	0.82
Year $\times$ Region FE	X	X	X	X

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

<sup>1</sup> Each column is an estimation of a heterogeneous effect of regression (7) by generation with region  $\times$  year fixed effects. I include controls for sex, quartic age, fraction of Hispanics in a state, and parental education. I also added parents' (HH, HW, and WH) and grandparents' (HHHH, HHHW, HHWH, etc.) type dummy variables to the regression on second and third generation immigrants, where H is objectively Hispanic (born in a Spanish Speaking Country) and W is objectively White (native born). Standard errors are clustered on the state level.

<sup>2</sup> 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.

<sup>3</sup> Data source is the 2004-2021 Current Population Survey.

I show the results of estimating equation (7) using the different specifications by generation in figure (5). Each panel evaluates a different specification with controls for parental education, sex, age, fraction of Hispanic people in a state, and region  $\times$  year fixed effect. Panel A of Figure (5) includes the results of estimating equation (7) on the overall sample—i.e. pooling first-, second-, and third-generations together. I find that state-level bias is associated with a ten percentage points decrease in self-reported Hispanic identity. Panel (B) of figure (5) includes the results of estimating equation (7) on a sample of first-generation Hispanic immigrants. I find a negative relationship between state-level bias and self-reported Hispanic identity. A one standard deviation increase in bias is correlated with a six percentage points decrease in self-reported Hispanic identity. Panel (C) of figure (5) includes the results of estimating equation (7) on a sample of second-generation Hispanic immigrants. I find a negative relationship between state-level bias and self-reported Hispanic identity among second-generation immigrants. A one standard deviation increase in bias is correlated with a ten percentage points decrease in self-reported Hispanic identity. Panel (D) of figure (5) includes the results of estimating equation (7) on a sample of third-generation Hispanic immigrants. I find no statistically significant relationship between bias and self-reported Hispanic identity among third-generation immigrants.

In table (12), I provide the results of estimating equation (7) on a sample of second-generation immigrants separately for each parent type—Spanish speaking country born father-Spanish speaking country born mother (HH), a native-born father-Spanish speaking country born mother (WH), and Spanish speaking country born father-native born mother (HW)—with region  $\times$  year fixed effects. I find a significant negative correlation between bias and endogenous Hispanic identity among second-generation Hispanic immigrant children of endogamously Hispanic parents—both the father and mother were born in Spanish speaking country. A one standard deviation increase in bias is correlated with a 15 percentage points

reduction in the self-reported Hispanic identity of second-generation Hispanic immigrant children of the Hispanic father-Hispanic mother (HH) (table 12 column 2). The result is statistically significant. A one standard deviation increase in bias is correlated with a one percentage points decrease in the Hispanic identity of second-generation Hispanic immigrant children of the Hispanic father-White mother (HW) (table 12 column 3). The result is statistically insignificant. A one standard deviation increase in bias is correlated with a seven percentage points decrease in the Hispanic identity of second-generation Hispanic immigrant children of the White father-Hispanic mother (WH) (table (12) column 4). The results are statistically insignificant.

The results above indicate that attitudes are not correlated with the self-reported Hispanic identity of third-generation immigrants. One reason for this finding is that first and second-generation immigrants aim to integrate and assimilate into their new community, while third-generation immigrants are already assimilated. The results are consistent with the findings of Abramitzky et al. (2020a) where they show that second-generation immigrants tend to change their names to adapt to the new culture, as suggested by Abramitzky et al. (2020a).

Furthermore, I estimate another specification of equation (7) by the number of grandparents that are objectively Hispanic—born in a Spanish-speaking country—and are third-generation Hispanic immigrants. For example, suppose a person has two grandparents born in a Spanish-speaking country. In that case, they have two ‘Hispanic’ grandparents, and I will estimate a specification for all of those that belong to this group. The result of evaluating the effect of bias on people with one objectively Hispanic grandparent is in table (13) column (1), with two objectively Hispanic grandparents is in table (13) column (2), with three objectively Hispanic grandparents is in table (13) column (3), and with four objectively Hispanic grandparents is in table (13) column (4). I find a statistically significant relationship between

and self-reported Hispanic identity among third-generation Hispanic immigrants with four objectively Hispanic grandparents. A one standard deviation increase in bias is associated with 14 percentage points decrease in the likelihood of a third-generation Hispanic immigrant with four objectively Hispanic grandparents to self-report Hispanic identity. Bias has no significant effect on children of interethnic grandparents.

TABLE 12: Relationship Between Bias and Self-Reported Hispanic identity Among Second-Generation Hispanic Immigrants: By Parental Type

Parents Type	All	Both Parents	Father	Mother
		from Spanish Speaking Country	from Spanish Speaking Country	from Spanish Speaking Country
		(HH)	(HW)	(WH)
		(1)	(2)	(3)
		$H_{ist}^2$	$H_{ist}^2$	$H_{ist}^2$
Bias		-0.13** (0.05)	-0.15** (0.06)	-0.01 (0.07)
Female		0.00 (0.00)	0.00 (0.00)	-0.01 (0.01)
College Graduate: Mother		-0.06*** (0.01)	-0.04*** (0.01)	-0.07*** (0.01)
College Graduate: Father		-0.08*** (0.01)	-0.04*** (0.01)	-0.10*** (0.02)
N		560 100	405 116	88 421
Year $\times$ Region FE		X	X	X
Mean		0.94	0.96	0.9
		0.83		

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

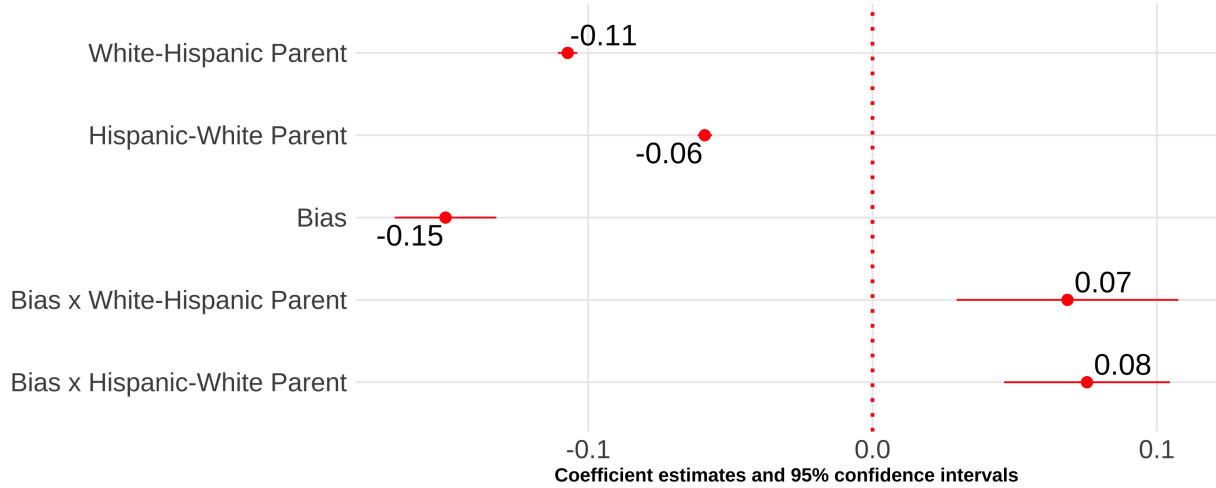
<sup>1</sup> Each column is an estimation of a heterogeneous effect of regression (7) by type of parents with region  $\times$  year fixed effects. I include controls for sex, quartic age, fraction of Hispanics in a state, and parental education. Standard errors are clustered on the state level.

<sup>2</sup> 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.

<sup>3</sup> Column (1) includes the results to regression (7) on all second-generation immigrants, column (2) includes the results to regression (7) on second-generation immigrants that who has a father and mother that were born in a Spanish speaking country (HH), column (3) includes the results to regression (7) on second-generation immigrants that who has a father that was born in a Spanish speaking country and a native born mother (HW), and column (4) includes the results to regression (7) on second-generation immigrants that who has a native born father and a mother that was born in a Spanish speaking country (WH).

<sup>4</sup> Data source is the 2004-2021 Current Population Survey.

Figure 7: Relationship Between Bias and Self-reported Identity on Second-Generation Hispanics: Interaction.



I show four panels of estimating equation (8). I include region  $\times$  year fixed effects with controls for sex, quartic age, fraction of Hispanics in a state, and parental education. Each panel is the results from the same regression but is presented with different combinations of grandparental types. Robust standard errors are reported. 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 13: 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 (0.11)	0.03 (0.09)	0.19 (0.26)	-0.14* (0.07)
Female	-0.01 (0.01)	0.00 (0.01)	-0.01 (0.01)	0.00 (0.01)
College Graduate: Mother	-0.11*** (0.03)	-0.07*** (0.02)	0.02 (0.02)	-0.02 (0.01)
College Graduate: Father	-0.11*** (0.03)	-0.08*** (0.01)	0.02 (0.01)	-0.03* (0.01)
N	55 051	74 100	12 194	57 646
Year × Region FE	X	X	X	X

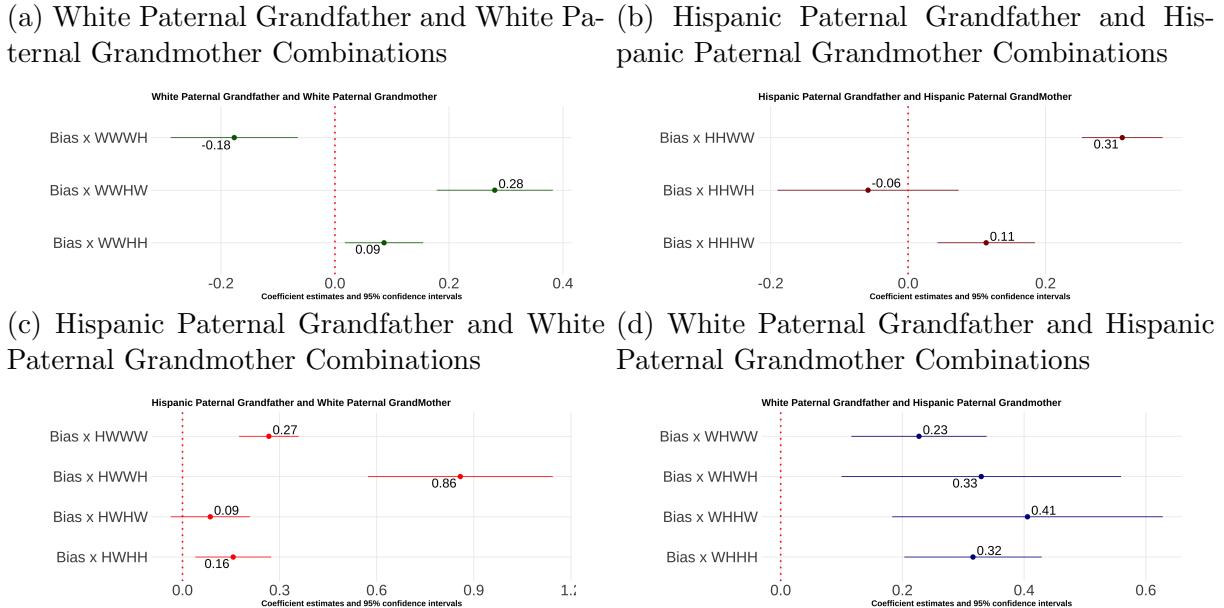
\* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01

<sup>1</sup> Each column is an estimation of equation (7) 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.

<sup>2</sup> 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.

<sup>3</sup> Data source is the 2004-2021 Current Population Survey.

Figure 8: Relationship Between Bias and Self-reported Identity on third-generation Hispanics: Interaction.



I show four panels of estimating equation (9). I include region  $\times$  year fixed effects with controls for sex, quartic age, and parental education. Each panel results from the same regression but is presented by different combinations of grandparental types. The syntax of the grandparents' type consists of four letters. The first letter represents the paternal grandfather, the second letter represents the paternal grandmother, the third letter represents the maternal grandfather, and the fourth letter represents the maternal grandmother. A grandparent would be of type  $H$  if they were born in a Spanish-speaking country and of type  $W$  if they were native-born. Robust standard errors are reported. The samples include third-generation Hispanic children ages 17 and below who live in intact families. Third-generation Hispanic immigrant children with native-born parents and at least one grandparent born in a Spanish-speaking country.

I estimate the heterogeneous relationship between bias and self-reported Hispanic identity by parental types of equation (8) and show the results in figure (7). This specification estimates the heterogeneous effect of bias on a sample of second-generation immigrants by parental type. Moreover, I find a statistically significant positive heterogeneous relationship

between bias and parents on self-reported Hispanic identity. I find that a one standard deviation increase in state-level bias correlates with a seven percentage points increase in Hispanic identity in children with an objectively Hispanic Father-White Mother (HW) compared to the reference group—children of an objectively Hispanic father-Hispanic mother (HH). I find that a one standard deviation increase in bias correlates with an eight percentage points increase in self-reported Hispanic identity in children with an objectively White Father-Hispanic Mother (WH) compared to children of an objectively Hispanic father-Hispanic mother.

I estimate the heterogeneous relationship between bias and self-reported Hispanic identity by grandparental types of equation(9) and show the results in figure (8). I mostly find statistically significant positive heterogeneous effect of bias on the different types of third-generation Hispanic immigrants compared to children of objectively Hispanic paternal grandfather-Hispanic paternal grandmother-Hispanic maternal grandfather-Hispanic maternal grandmother (HHHH).<sup>21</sup>

To simplify, I will discuss the results in four groups by the country of origin of the paternal grandparents with the combination of the maternal grandparents. Among third-generation immigrants with objectively White paternal grandfather-White paternal grandmother (figure 8a): (1) those with objectively Hispanic maternal grandparents are nine percentage points more likely than children with HHHH grandparents to self-identify as Hispanic; (2) those with objectively Hispanic maternal grandfather and White maternal grandmother are 28 percentage points more likely than children with HHHH grandparents to self-identify as Hispanic; (3) those with objectively White maternal grandfather and Hispanic maternal grandmother are 18 percentage points less likely than children with HHHH grandparents to

---

<sup>21</sup>The syntax of the grandparents' type consists of four letters. The first letter represents the paternal grandfather, the second letter represents the paternal grandmother, the third letter represents the maternal grandfather, and the fourth letter represents the maternal grandmother. A grandparent would be of type *H* if they were born in a Spanish-speaking country and of type *W* if they were native-born.

self-identify as Hispanic.

Among third-generation immigrants with objectively Hispanic paternal grandfather-Hispanic paternal grandmother (figure 8b): (1) those with objectively Hispanic maternal grandfather and White maternal grandfather are 11 percentage points more likely than children with objectively HHHH grandparents to self-identify as Hispanic; (2) those with objectively White maternal grandfather and Hispanic maternal grandmother are as likely as children with HHHH grandparents to self-identify as Hispanic; (3) those with objectively White maternal grandfather and White maternal grandmother are 31 percentage points more likely than children with HHHH grandparents to self-identify as Hispanic.

Among third-generation immigrants with objectively Hispanic paternal grandfather-White paternal grandmother (figure 8c): (1) those with objectively Hispanic maternal grandparents are 16 percentage points more likely than children with HHHH grandparents to self-identify as Hispanic; (2) those with objectively Hispanic maternal grandfather and White maternal grandmother are as likely as children with HHHH grandparents to self-identify as Hispanic; (3) those with objectively White maternal grandfather and Hispanic maternal grandmother are 86 percentage points more likely than children with HHHH grandparents to self-identify as Hispanic; (4) those with objectively White maternal grandfather and White maternal grandmother are 27 percentage points more likely than children with HHHH grandparents to self-identify as Hispanic.

Among third-generation immigrants with objectively White paternal grandfather-Hispanic paternal grandmother (figure 8d): (1) those with objectively Hispanic maternal grandparents are 32 percentage points more likely than children with HHHH grandparents to self-identify as Hispanic; (2) those with objectively Hispanic maternal grandfather and White maternal grandmother are 41 percentage points more likely than children with HHHH grandparents

to self-identify as Hispanic; (3) those with objectively White maternal grandfather and Hispanic maternal grandmother are 33 percentage points more likely than children with HHHH grandparents to self-identify as Hispanic; (4) those with objectively White maternal grandfather and White maternal grandmother are 23 percentage points more likely than children with HHHH grandparents to self-identify as Hispanic.

### 2.5.2 The Relationship Between Bias, Interethnic Marriages, and Migration

In this section, I will discuss the results of estimating the relationship between bias and inter-ethnic marriages (equation 10), and bias and migration (equations 12, 13, and 14). As the CPS does not report a person's state of birth, I use the 2004-2021 Censuses to construct a sample of second-generation Hispanic immigrants (Flood et al., 2021b). I construct a mover variable if they moved from their state of birth to another state. I find that there is no significant correlation between bias and migration decisions among second-generation immigrants. I find that second-generation immigrants who self-report as Hispanic tend to move from states with low bias to states with high bias. I find a statistically significant negative correlation between bias and the likelihood of inter-ethnic marriage formation.

I find no relationship between and the migration decisions objectively Hispanic parents of second-generation Hispanic children make. I show the results of estimating regressions 12, 13 , and 14 in table (14). The results of estimating regression 12 are in column (1) of table (14), where the dependent variable is an indicator variable that is equal to one if a person migrated from the state they were born in, and zero otherwise.  $Bias_{st}$  is the average bias in the state the interviewee is currently living  $s$  during the year of the survey  $t$ . I find that bias is not significantly correlated with the migration decisions. The results of estimating regression 13 are in column (2), where the dependent variable is an indicator variable that is equal to one if a person migrated from the state they were born in, and zero otherwise.

$Bias_{lb}$  is the average bias in the state the interviewee was born in  $l$  during their year of birth  $b$ . I find no statistically significant effect of bias on migration decisions taken by the parents of second-generation Hispanic immigrants.

By estimating equation (14), I find that among second generation Hispanic immigrants movers with objectively Hispanic parents, those that self-report Hispanic identity move from a state of birth with lower bias to a state with higher bias. The results to estimating (14) are shown in table (14) column (3). The dependent variable is the difference in bias between the state person  $i$  currently lives in at time  $t$  and bias in the state of birth  $l$  during birth year  $b$ , i.e.  $Bias_{ist} - Bias_{ilb}$ . Among second-generation Hispanic immigrant movers, those that self-report Hispanic identity live in states that are 0.02 standard deviations more biased than the state they were born. Even though this result shows that there is selection into more biassed states among second-generation immigrants, it does not affect my main results showing correlation between bias and self-reported Hispanic identity. Since those that identify as Hispanics are the movers, this implies that my assessments of the relationship between bias and self-reported Hispanic identity might underestimate the effect of bias.

TABLE 14: Relationship Between Bias and Migration

	(1)	(2)	(3)
Migrated from Birth Place	Migrated from Birth Place	$Bias_{ist} - Bias_{ilb}$	
$Bias_{st}$	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*** (0.01)	0.03*** (0.01)	0.00 (0.01)
College Graduate: Father	0.02** (0.01)	0.02*** (0.01)	-0.01 (0.01)
N	352 712	185 024	12 806
Mean	0.09	0.09	-0.02
Year ( $t$ ) $\times$ Region FE	X		
Birthyear ( $b$ ) $\times$ Birth Region FE		X	

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

<sup>1</sup> Each column is an estimation of equations (12) in column (1), (13) in column (2), and (14) in column (3).

<sup>2</sup> 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  $\times$  year fixed effects. Standard errors are clustered on the state level.

<sup>3</sup> 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.

<sup>4</sup> Data source is the 2004-2021 Census Data.

Moreover, I estimate equations (11) as a linear probability model and find that a one standard deviation increase in state-level bias increases the likelihood of interethnic marriages. In table (15) I present the regression results. I find that a one standard deviation increase in bias increases the probability of having inter-ethnic parents by five percentage points. Moreover, I break down the analysis by the ethnicity of the couples. Estimate the a regression on a sample of Hispanic men (table 15 column 1) and women (table 15 column 2) with a left hand side variable being a dummy variable that is equal to one if they married someone from a different ethnicity and zero otherwise, and the right hand side variable is state-level bias. A one standard deviation increase in state-level bias is associated six 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 seven 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 state with high bias would like to decrease the chances that their children would have Hispanic ethnicity signals. For example, Hispanic women in high bias states could marry a non-Hispanic White husband, and their children will have a non-Hispanic last name.

TABLE 15: Relationship Between Bias and Interethnic Marriages

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

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

<sup>1</sup> This is the result to estimating (11) as a linear probability model.

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

<sup>3</sup> Data source is the 2004-2020 Current Population Survey Data.

## 2.6 Discussion

The results in this paper show that bias is negatively correlated with the self-reported Hispanic identity of Hispanic immigrants. In this paper, my aim is not to establish a causal effect of bias on self-reported Hispanic identity. I aim to establish a correlation between bias and self-reported identity to show that depending on whether the bias levels in a state, the racial and ethnic gaps that rely on self-reported identity might be overestimating or

underestimating the effect of discrimination.

A few questions could arise regarding the validity of the results. 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 a true 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 (16) and the results to estimation of equation (7) by the proxy respondent in table (17) on all the generation. The main effect of the reported Hispanic identity of children is 94% when the mother is the proxy, 92% when the father is the proxy, and 96% when the child themselves or another caregiver was the household respondent.<sup>22</sup> The estimation of equation (7) by the proxy respondent, table (17), mostly yields a negative effect of bias on self-reported Hispanic identity for all types of proxy respondents. The relationship between bias and self-reported Hispanic identity yields an increase of one percentage point when the mother proxy identifies as White, but the result is not statistically significant. The relationship between bias and self-reported Hispanic identity yields a reduction of six percentage points when the mother proxy identifies as Hispanic, but the result is not statistically significant. The relationship between bias and self-reported Hispanic identity is a reduction of 13 percentage points when the objectively White father is the proxy respondent, but the result is not statistically significant. The relationship between

---

<sup>22</sup>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.

bias and self-reported Hispanic identity is a reduction of seven percentage points when the objectively Hispanic father is the proxy respondent, and the result is statistically significant. The relationship between bias and self-reported Hispanic identity is a reduction of seven percentage points when the person themselves is the proxy respondent, but the result is not statistically significant. The relationship between bias and self-reported Hispanic identity is a reduction of 17 percentage points when another caregiver is the proxy respondent, and the result is statistically significant.

TABLE 16: Main Effect of Proxy on Second-Generation’s Hispanic Self-identification

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

A second concern is that the Implicit Association Test (IAT) is voluntary and not representative of the population. While I do not claim that the Implicit Association Test (IAT) as a proxy for bias will represent the population, Egloff and Schmukle (2002) show that they are hard to manipulate. Several studies had shown that Implicit Association Test (IAT) are correlated with economic outcomes (Chetty et al., 2020; Glover et al., 2017), voting behavior (Friese et al., 2007), decision-making (Bertrand et al., 2005; Carlana, 2019), and health (Leitner et al., 2016). Another concern could be that the Implicit Association Test (IAT) test takers’ characteristics is changing over time and thus, are not the same. I address this concern by including non-parametric region  $\times$  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.

Moreover, I use different specification and prejudice measures as a robustness check. First, I use the General Security Survey (GSS)—nationally representative survey—to construct a measure of racial prejudice following Charles and Guryan (2008a). A downside of using the GSS is that most of the questions used to build the prejudice measure in Charles and Guryan (2008a) were discontinued after 1996, making it difficult to merge the measure with the CPS data since the CPS only started to ask about the place of birth of parents in 1994. Thus, I use the GSS to construct a residual prejudice measure from 15-20 years before the CPS survey year, which means it reflects the level of prejudice at roughly the time of birth of the CPS sample. Using the GSS data, I mostly find negative effects of prejudice on self-reported Hispanic identity. However, these results are only significant among second-generation Hispanic immigrant children of native-born fathers and Spanish-speaking country-born mothers, third-generation Hispanic immigrants, and third-generation Hispanic immigrants with two grandparents who were born in a Spanish-speaking country (see tables 19, 20, and 21.) These results suggest that ‘residual’ prejudice mainly affects interethnic children.

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 (15) and (16) as evidence. Figure (15) plot the percent of self-reported Hispanics in a state at a specific year on the average implicit bias in the same state during that year. Figure (16) plots the percent of objectively second-generation Hispanic children of endogamous marriages in a state at a certain year on 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, estimating 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 in a section above that this is not the case (table 14). I find no evidence that there's a relationship between migration decisions and bias. Additionally, I find that those that report Hispanic identity move out of birthplaces that are less biased and live in states that are more biased at the time of the survey. Thus, my results might underestimate the relationship between bias.

## 2.7 Conclusion

As the United States becomes more multi-racial and multi-ethnic, self-reported identity will have significant effects on representation, distributional 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 seven and 13 percentage points decrease in self-reported Hispanic identity. Also, among second-generation immigrants children with Hispanic Fathers-Hispanic Mothers are affected more by state-level bias. A one standard deviation increase in bias is correlated with 15 percentage points decrease in self-reported Hispanic identity among second-generation Hispanic immigrants children of objectively Hispanic parents. I also find

that bias is positively correlated with interethnic marriage and is uncorrelated with migration decisions.

The results are important because of the consequences of 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 would have important consequences. For example, suppose the people whose identities are most likely to be 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 in profound way.

Moreover, this study could encourage further research into the relationship between bias and self-reported identities for other groups. The analysis of bias on self-reported Hispanic 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 that self-report to those that do not by using restricted administrative data.

## A Figures

Figure 9: Proportion of Non-Whites in the United States from 1995 to 2019.

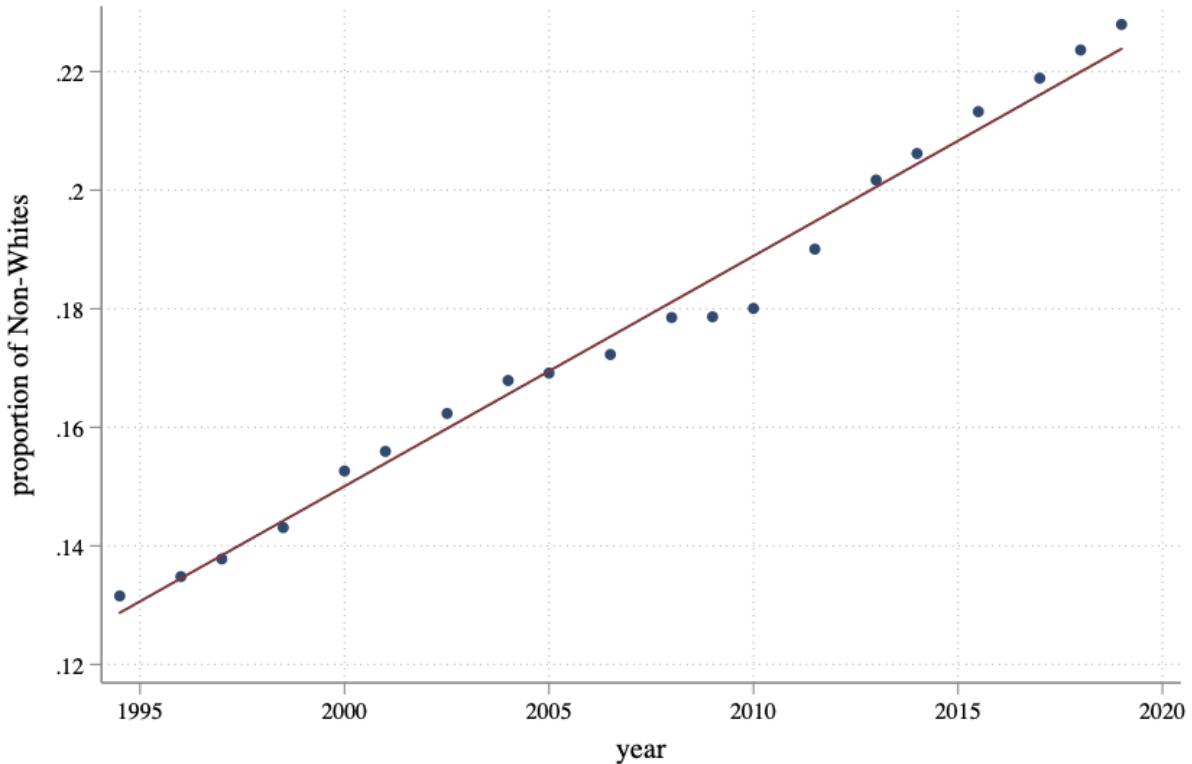


Figure 10: Hispanics as a Percentage of the Total Population in the United States from 1995 to 2019.

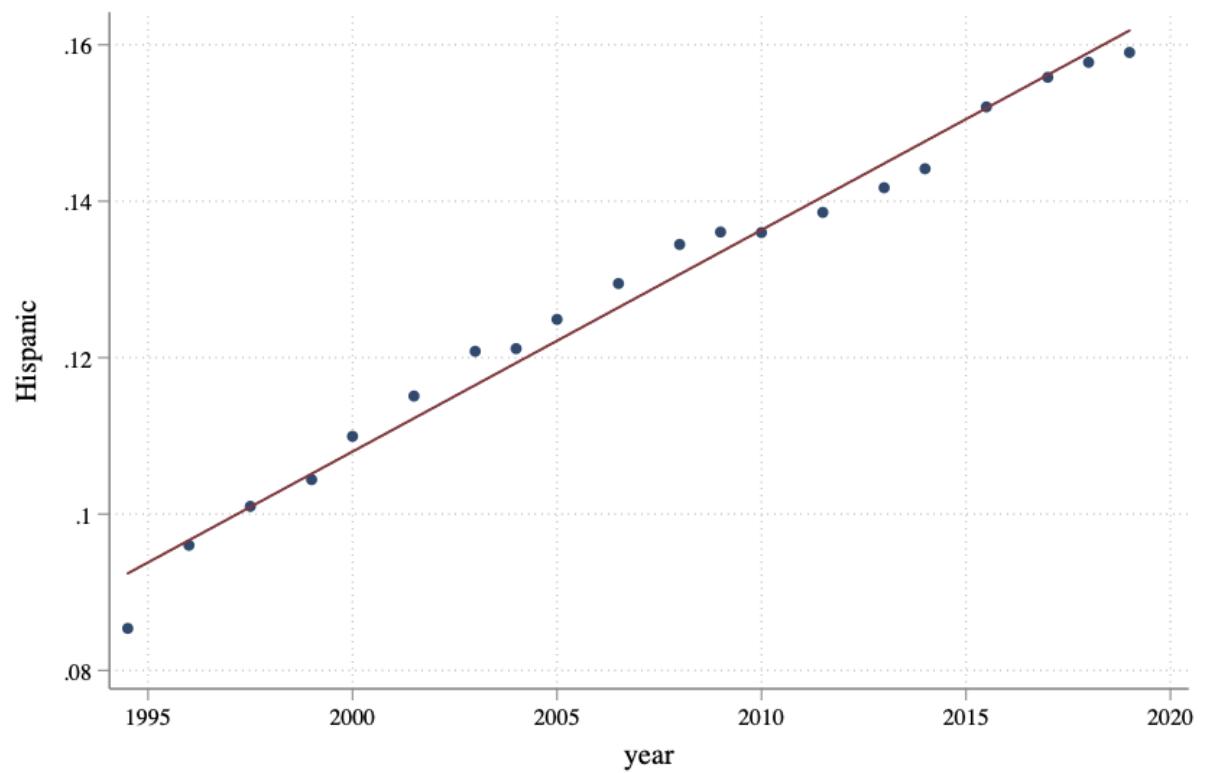


Figure 11: Inter-racial Marriages as a Percent of All Marriages in the United States from 1995 to 2020.

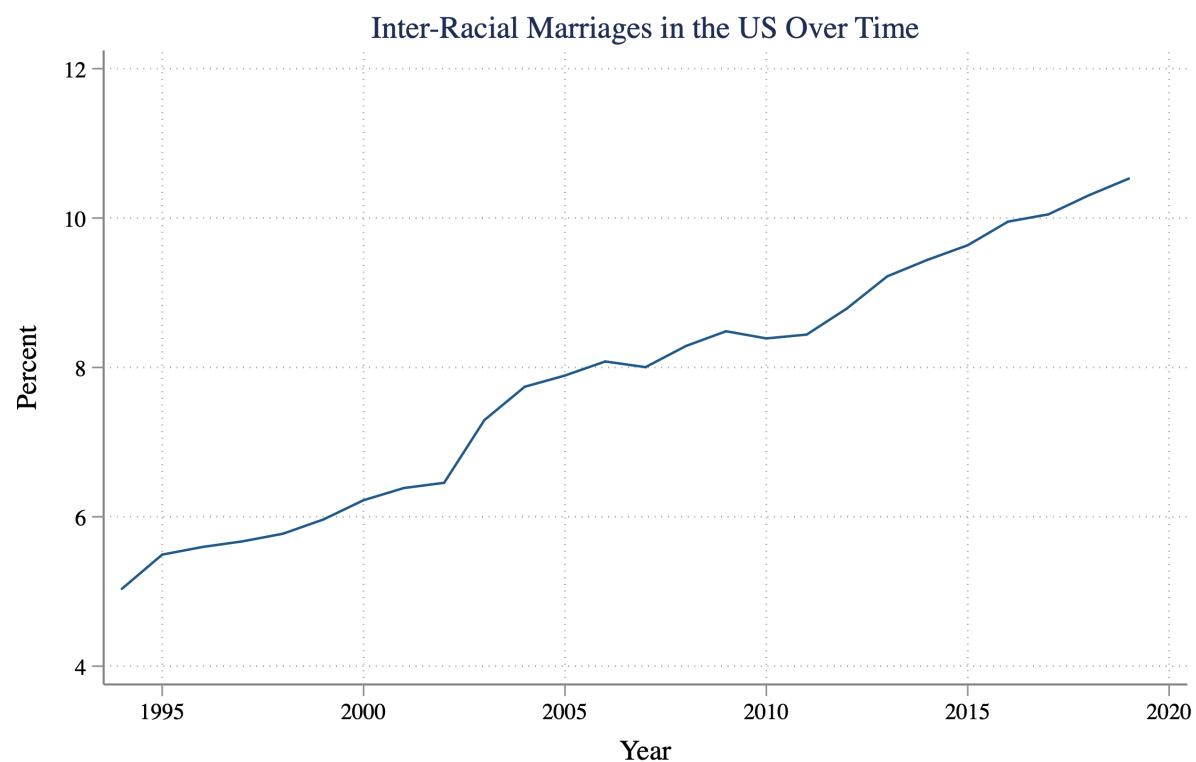


Figure 12: Distribution of the Four Types of Children.

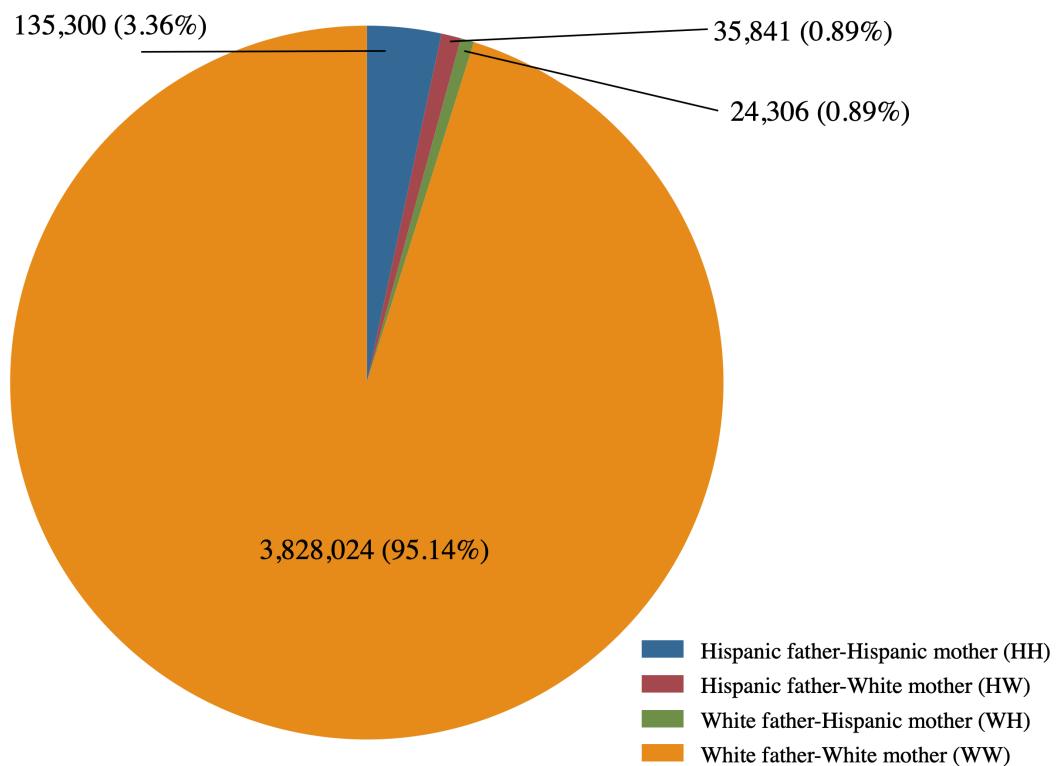


Figure 13: Chart Explaining Which Synthetic Parents and Children and When We Observe Them.

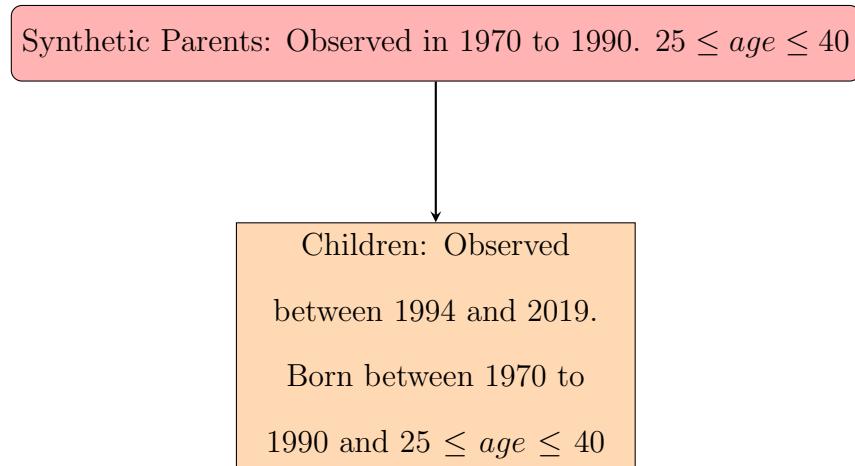
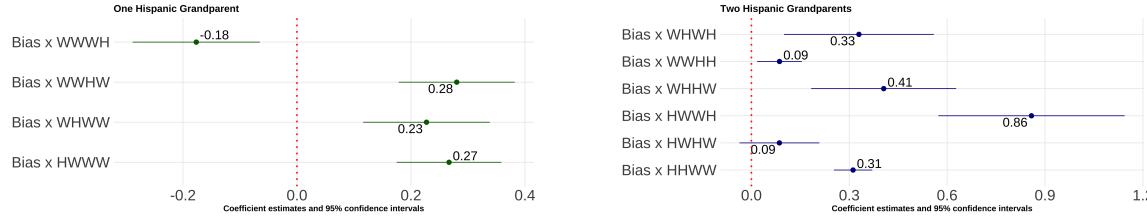
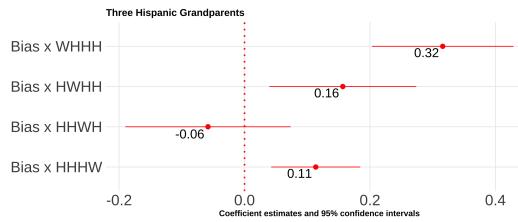


Figure 14: Relationship Between Bias and Self-reported Identity on third-generation Hispanics: Interaction

(a) Grandparents Combinations with One Objectively Hispanic Grandparent      (b) Grandparents Combinations with Two Objectively Hispanic Grandparents



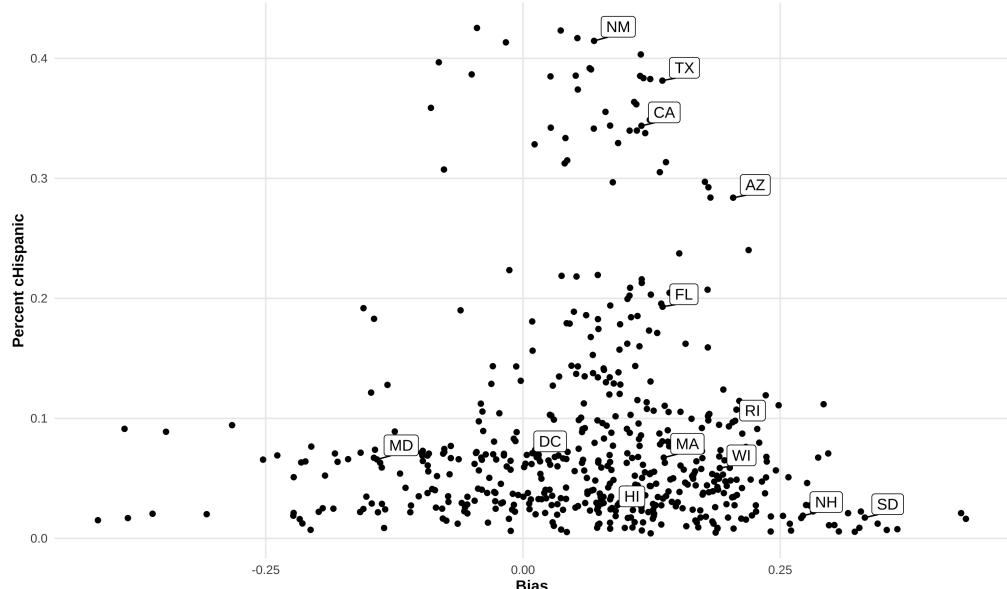
(c) Grandparents Combinations with Three Objectively Hispanic Grandparents



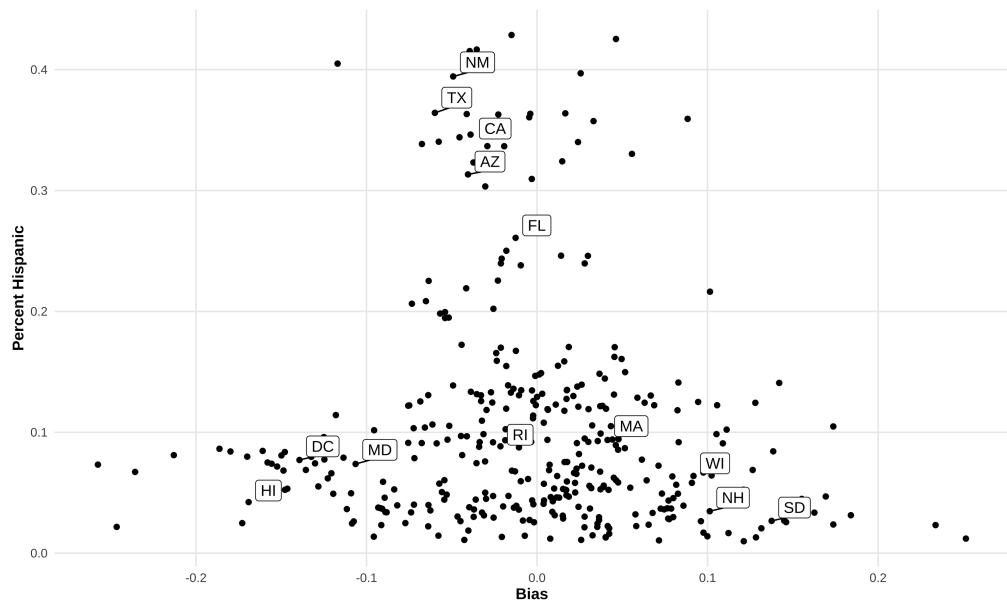
I show four panels of estimating equation (9). I include region  $\times$  year fixed effects with controls for sex, quartic age, and parental education. Each panel results from the same regression but of a different combination of grandparents types. Robust standard errors are reported. The samples include third-generation Hispanic children ages 17 and below who live in intact families. Third-generation Hispanic immigrant children with native-born parents and at least one grandparent born in a Spanish-speaking country.

Figure 15: Scatter Plot of Proportion Subjectively Hispanic on Bias

(a) Year  $\in 2015$



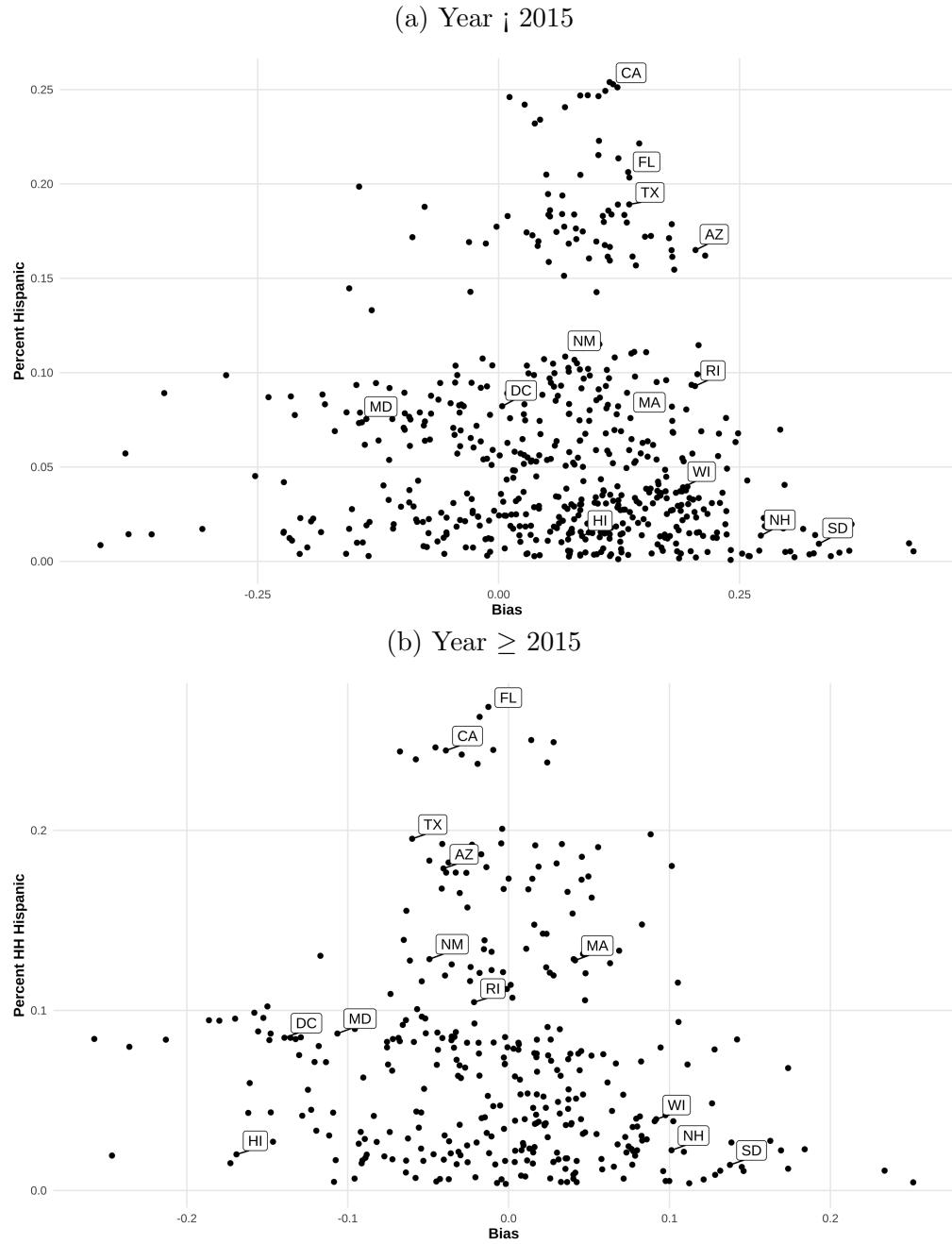
(b) Year  $\geq 2015$



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

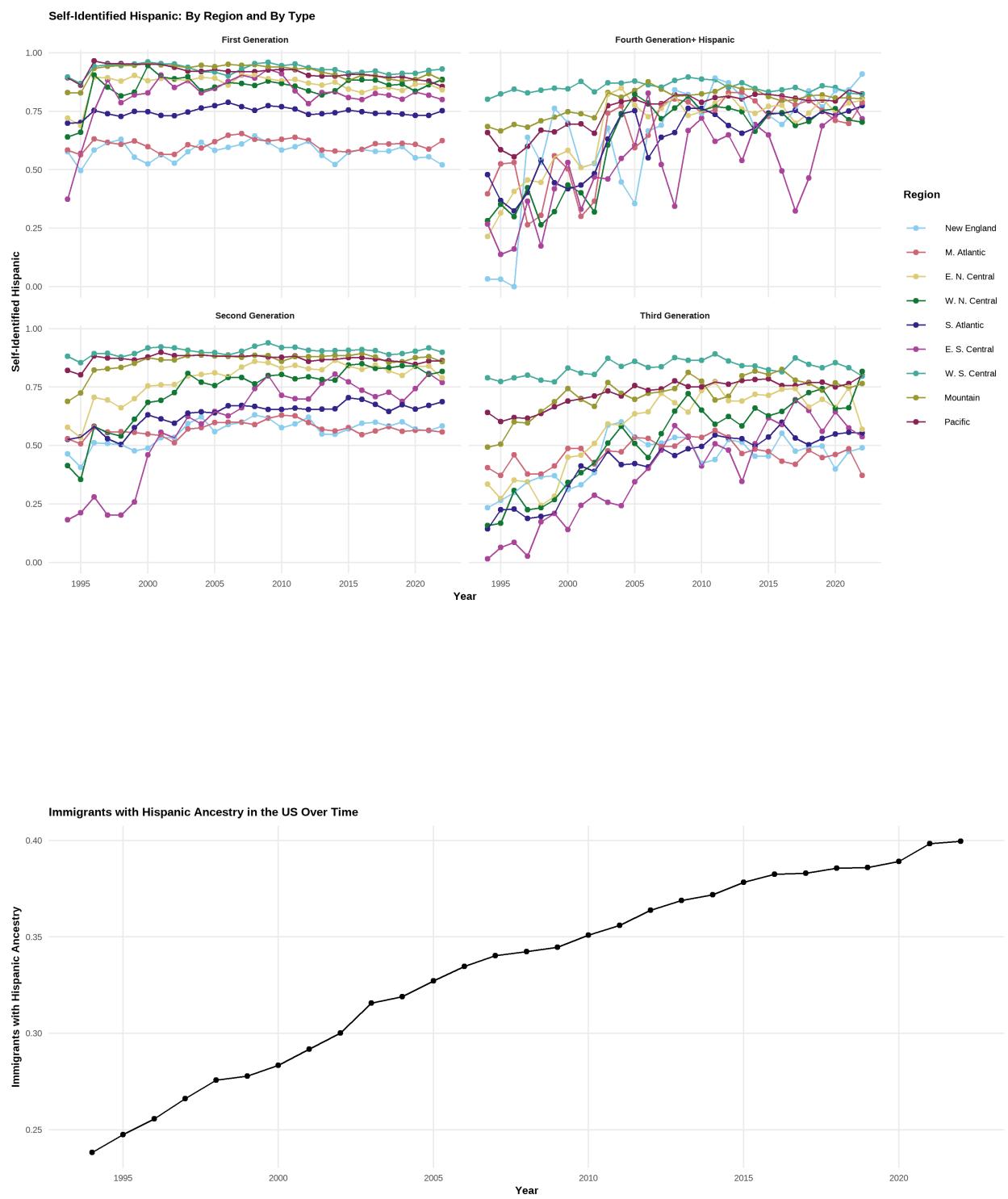


Here are two scatter plots of bias on subjective Hispanic population in a state. Each dot represents a state in a certain year.

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

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

Figure 17: Hispanic Identification Among Hispanic Immigrants: By Generation



Immigrants with Hispanic Ancestry in the US Over Time: By Region (Objective)

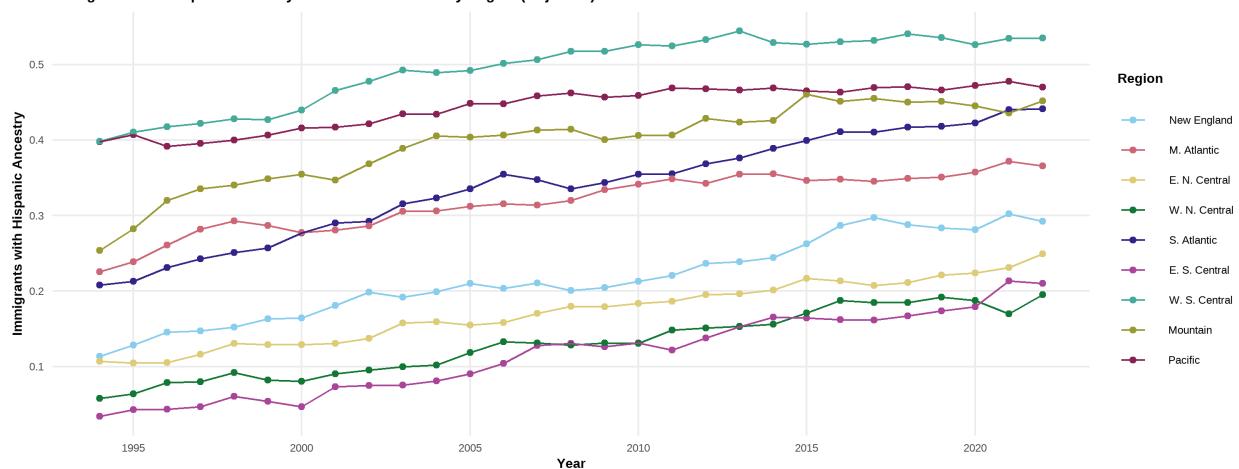


Figure 18: Correlation Between Implicit and Explicit Biases: IAT Explicit Bias Question

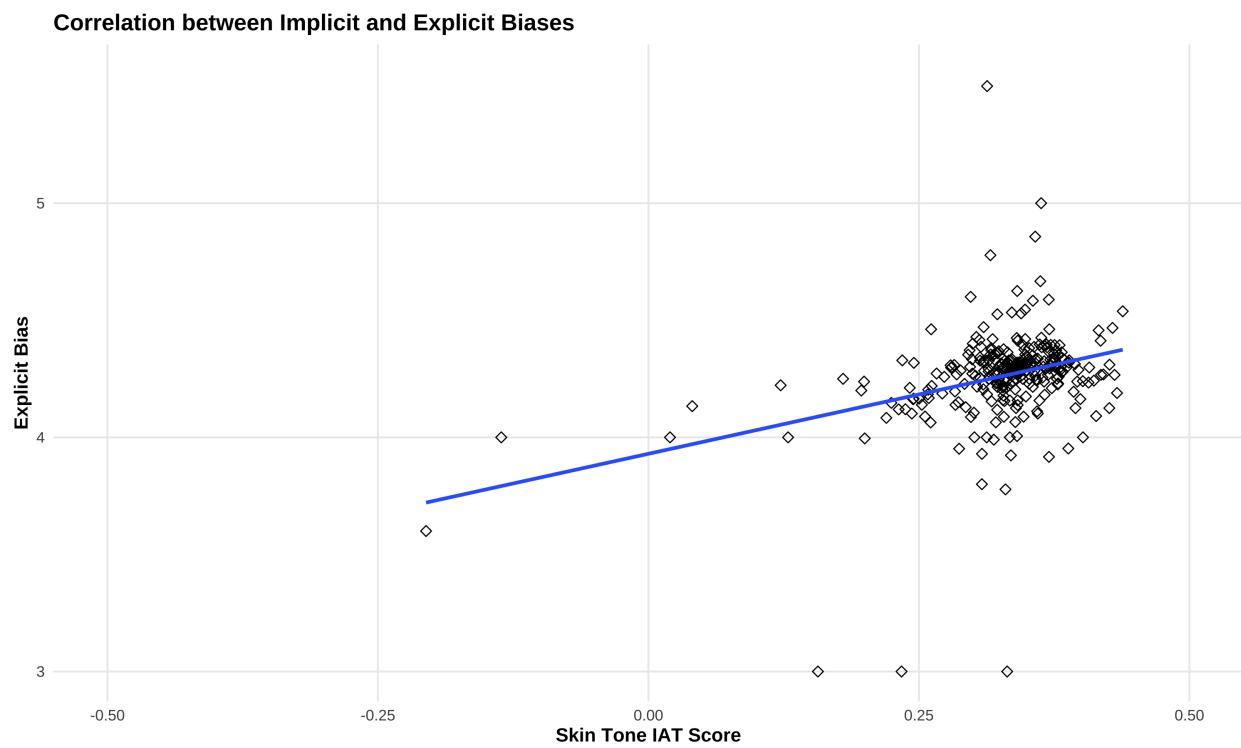


Figure 19: Correlation Between Implicit and Explicit Biases: GSS Close to Black People Question



Figure 20: Correlation Between Implicit and Explicit Biases: GSS Vote for A Black President Question

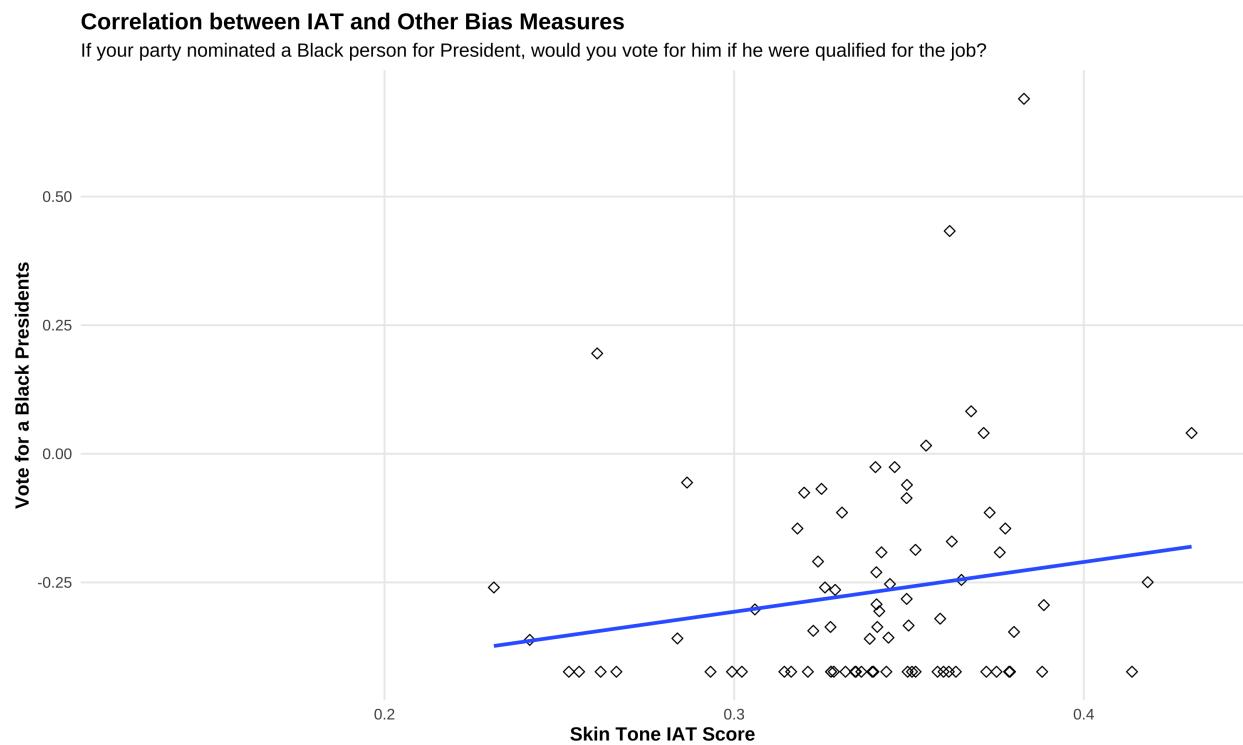
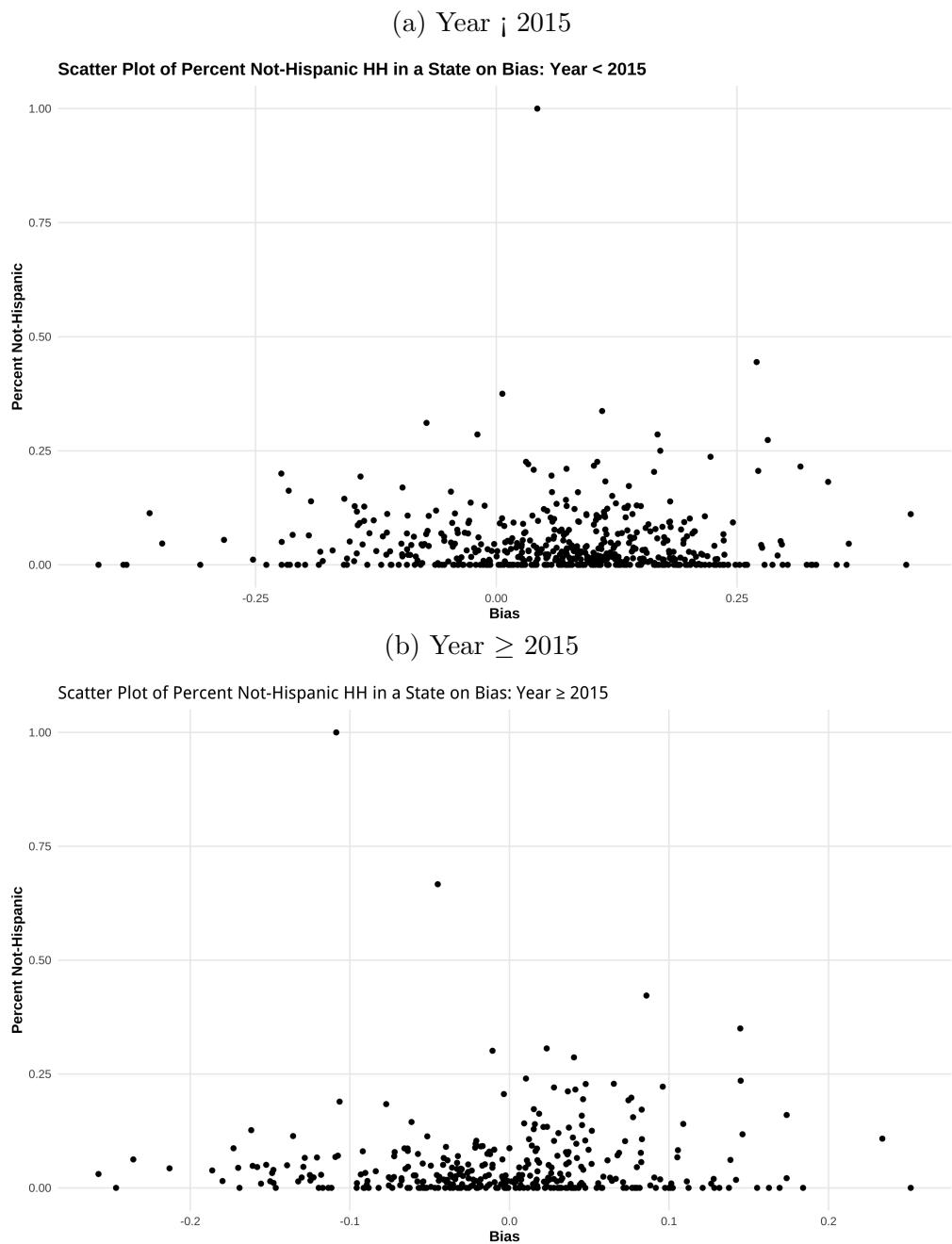


Figure 21: Scatter Plot of Non-Hispanic Second-Generation Hispanic-Hispanic Parents on Bias



## B Data

Figure 22: Examples of an Implicit Association Test

**Implicit Association Test**

Next, you will use the 'E' and 'I' computer keys to categorize items into groups as fast as you can. These are the four groups and the items that belong to each:

Category	Items
Good	Delight, Enjoy, Laughing, Excitement, Terrific, Lovely, Pleasure, Love
Bad	Humiliate, Annoy, Angry, Horrific, Despise, Ugly, Tragic, Evil

There are seven parts. The instructions change for each part. Pay attention!

**Continue**

Press "E" for Press "I" for

If you make a mistake, a red X will appear. Press the other key to continue.

**Part 1 of 7**

Put a left finger on the E key for items that belong to the category Light Skinned People.  
Put a right finger on the I key for items that belong to the category Dark Skinned People.  
Items will appear one at a time.

If you make a mistake, a red X will appear. Press the other key to continue.  
Go as fast as you can while being accurate.

Press the **space bar** when you are ready to start.

Press "E" for **Bad** Press "I" for **Good**

**Enjoy**

If you make a mistake, a red X will appear. Press the other key to continue.

Press "E" for **Bad** or

**Tragic**

If you make a mistake, a red X will appear. Press the other key to continue.

Press "I" for **Good** or

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

## C Tables

TABLE 17: 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 (0.11)	-0.06 (0.03)	-0.13 (0.11)	-0.07*** (0.02)	-0.07 (0.05)	-0.17** (0.07)
Female	-0.01 (0.01)	0.00 (0.00)	0.01 (0.01)	0.00*** (0.00)	0.00 (0.01)	0.00 (0.00)
College Graduate: Mother	-0.02 (0.02)	-0.03*** (0.01)	-0.04*** (0.02)	-0.03*** (0.01)	-0.03* (0.02)	-0.03*** (0.01)
College Graduate: Father	-0.09*** (0.02)	-0.05*** (0.01)	-0.01 (0.02)	-0.03*** (0.01)	-0.10*** (0.02)	-0.06*** (0.01)
First Gen	-0.10* (0.06)	0.07*** (0.01)	0.03 (0.05)	0.09*** (0.02)	0.13*** (0.02)	0.08*** (0.02)
Second Gen	0.05* (0.03)	0.04*** (0.01)	0.06*** (0.02)	0.05*** (0.01)	0.12*** (0.02)	0.07*** (0.02)
N	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

<sup>1</sup> Each column is an estimation of a heterogeneous effect of regression (7) by the proxy household respondent with region  $\times$  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.

<sup>2</sup> 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 country. 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.

<sup>3</sup> Data source is the 2004-2021 Current Population Survey.

TABLE 18: Relationship Between Bias and Self-Reported Hispanic Identity

	(1)	(2)	(3)	(4)	(5)
	$H_i$	$H_i$	$H_i$	$H_i$	$H_i$
Bias	-0.04 (0.03)	0.00 (0.05)	-0.03*** (0.01)	0.03 (0.03)	-0.10** (0.04)
Female	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
College Graduate: Mother	-0.05*** (0.01)	-0.06*** (0.01)	-0.05*** (0.01)	-0.05*** (0.01)	-0.05*** (0.01)
College Graduate: Father	-0.07*** (0.00)	-0.07*** (0.00)	-0.06*** (0.01)	-0.06*** (0.01)	-0.06*** (0.01)
First Gen	-1.18*** (0.03)	0.13*** (0.01)	0.36*** (0.03)	0.04*** (0.01)	0.26*** (0.03)
Second Gen	-1.19*** (0.02)	0.11*** (0.01)	0.34*** (0.02)	0.02*** (0.01)	0.24*** (0.03)
N	844 481	844 481	844 481	844 481	844 481
State FE			X	X	
Year FE		X		X	
Mean	0.91	0.91	0.91	0.91	0.91
Year $\times$ Region FE				X	

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

<sup>1</sup> Each column is an estimation of a regression the following regression (7) with different set of fixed effects. I include controls for sex, quartic age, parental education, fraction of Hispanics in a state, and generational, parents' and grandparents' type dummy variables. Standard errors are clustered on the state level.

<sup>2</sup> 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.

<sup>3</sup> Data source is the 2004-2021 Current Population Survey.

TABLE 19: Self-Reported Hispanic Identity and Charles and Guryan (2008a) Prejudice Measure: By Generation

	(1)	(2)	(3)	(4)
	All Gens	First Gen	Second Gen	Third Gen
	$H_{ist}$	$H_{ist}^1$	$H_{ist}^2$	$H_{ist}^3$
Bias	-0.03 (0.03)	-0.02 (0.02)	-0.01 (0.02)	-0.13** (0.06)
Female	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
College Graduate: Mother	-0.05*** (0.01)	-0.04*** (0.01)	-0.08*** (0.01)	-0.07*** (0.01)
College Graduate: Father	-0.07*** (0.01)	-0.04*** (0.01)	-0.10*** (0.01)	-0.10*** (0.02)
N	906 106	111 960	605 316	188 830
Mean	0.89	0.96	0.92	0.76
Year $\times$ Region FE	X	X	X	X

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

<sup>1</sup> Each column is an estimation of equation (7) by parents' type with region  $\times$  year fixed effects with Charles and Guryan (2008a) Prejudice measure. Charles and Guryan (2008a) use the General Security Survey of the most common racial questions between 1970-2000 for their measure of prejudice. To use the prejudice measure, I take the average of the GSS questions by year groups. The groups I use are as follows: (1) 1977 and 1982, (2) 1985 and 1988, and 1989, (3) 1990, 1991, and 1993, and (4) 1994 and 1996. I link these groups to the following grouped Current Population Survey (CPS) years: (1) 1994-1999, (2) 2000-2005, (3) 2006-2010, and (4) 2011-2016. In other words, I merge CPS data with the residual prejudice measure from 20 years before the survey. I include controls for sex, quartic age, parental education. Standard errors are clustered on the state level.

<sup>2</sup> 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 country. 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.

<sup>3</sup> Data source is the 2004-2021 Current Population Survey.

TABLE 20: Relationship Between Self-Reported Hispanic Identity and Charles and Guryan (2008a) Prejudice Measure Among Second Generation Hispanic Immigrants: By Parental Type

Parents Type	All	HH	HW	WH
	(1)	(2)	(3)	(4)
	$H_{ist}^2$	$H_{ist}^2$	$H_{ist}^2$	$H_{ist}^2$
Bias	-0.01 (0.02)	0.02 (0.02)	-0.01 (0.08)	-0.15** (0.07)
Female	0.00 (0.00)	0.00 (0.00)	-0.01 (0.01)	0.02 (0.01)
College Graduate: Mother	-0.08*** (0.01)	-0.04*** (0.01)	-0.10*** (0.01)	-0.11*** (0.02)
College Graduate: Father	-0.10*** (0.01)	-0.04*** (0.01)	-0.15*** (0.02)	-0.15*** (0.01)
N	605 316	433 135	97 699	74 482
Year $\times$ Region FE	X	X	X	X
Mean	0.92	0.96	0.83	0.77

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

<sup>1</sup> Each column is an estimation of equation (7) by parents' type with region  $\times$  year fixed effects with Charles and Guryan (2008a) Prejudice measure. Charles and Guryan (2008a) use the General Security Survey of the most common racial questions between 1970-2000 for their measure of prejudice. To use the prejudice measure, I take the average of the GSS questions by year groups. The groups I use are as follows: (1) 1977 and 1982, (2) 1985 and 1988, and 1989, (3) 1990, 1991, and 1993, and (4) 1994 and 1996. I link these groups to the following grouped Current Population Survey (CPS) years: (1) 1994-1999, (2) 2000-2005, (3) 2006-2010, and (4) 2011-2016. In other words, I merge CPS data with the residual prejudice measure from 20 years before the survey. I include controls for sex, quartic age, parental education. Standard errors are clustered on the state level

<sup>2</sup> 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.

<sup>3</sup> Data source is the 1994-2021 Current Population Survey.

TABLE 21: Relationship Between Self-Reported Hispanic identity and Charles and Guryan (2008a) Prejudice Measure Among Third-Generation Hispanic Immigrants: By Grandparental Type

	Numer of Hispanic Grandparents			
	(1)	(2)	(3)	(4)
	One	Two	Three	Four
Bias	-0.14 (0.10)	-0.19*** (0.07)	-0.13 (0.19)	-0.02 (0.03)
Female	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.00)
College Graduate: Mother	-0.10*** (0.02)	-0.08*** (0.02)	0.02 (0.01)	-0.04*** (0.02)
College Graduate: Father	-0.13*** (0.03)	-0.09*** (0.02)	0.00 (0.02)	-0.01 (0.02)
N	62 969	68 431	12 111	45 319
Year × Region FE	X	X	X	X

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

<sup>1</sup> Each column is an estimation of equation (7) by grandparents' type with region × year fixed effects with Charles and Guryan (2008a) prejudice measure. Charles and Guryan (2008a) use the General Security Survey of the most common racial questions between 1970-2000 for their measure of prejudice. To use the prejudice measure, I take the average of the GSS questions by year groups. The groups I use are as follows: (1) 1977 and 1982, (2) 1985 and 1988, and 1989, (3) 1990, 1991, and 1993, and (4) 1994 and 1996. I link these groups to the following grouped Current Population Survey (CPS) years: (1) 1994-1999, (2) 2000-2005, (3) 2006-2010, and (4) 2011-2016. In other words, I merge CPS data with the residual prejudice measure from 20 years before the survey. I include controls for sex, quartic age, parental education. Standard errors are clustered on the state level.

<sup>2</sup> 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.

<sup>3</sup> Data source is the 2004-2021 Current Population Survey.

## Bibliography

- Abramitzky, R., Boustan, L., and Eriksson, K. (2020a). Do Immigrants Assimilate More Slowly Today Than in the Past? *American Economic Review: Insights*, 2(1):125–141.
- Abramitzky, R., Boustan, L., Eriksson, K., and Hao, S. (2020b). Discrimination and the Returns to Cultural Assimilation in the Age of Mass Migration. *AEA Papers and Proceedings*, 110:340–346.
- Abramitzky, R., Boustan, L. P., and Connor, D. (2020c). Leaving the Enclave: Historical Evidence on Immigrant Mobility from the Industrial Removal Office. Technical Report w27372, National Bureau of Economic Research, Cambridge, MA.
- Abramitzky, R., Boustan, L. P., and Eriksson, K. (2014). A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration. *Journal of Political Economy*, 122(3):467–506.
- Abramitzky, R., Boustan, L. P., and Eriksson, K. (2016). Cultural Assimilation during the Age of Mass Migration.
- Abramitzky, R., Boustan, L. P., Jácome, E., and Pérez, S. (2019). Intergenerational Mobility of Immigrants in the US over Two Centuries. Technical Report w26408, National Bureau of Economic Research, Cambridge, MA.
- Akerlof, G. A. and Kranton, R. E. (2000a). Economics and identity. *The quarterly journal of economics*, 115(3):715–753.
- Akerlof, G. A. and Kranton, R. E. (2000b). Economics and Identity. *Quarterly Journal of Economics*, 115(3):715–753.

Alba, R. (2020). *The Great Demographic Illusion: Majority, Minority, and the Expanding American Mainstream*. Princeton University Press.

Alesina, A., Miano, A., and Stantcheva, S. (2018). Immigration and Redistribution.

Altonji, J. G. and Pierret, C. R. (2001). Employer learning and statistical discrimination. *The quarterly journal of economics*, 116(1):313–350.

Antecol, H. and Bedard, K. (2006). Unhealthy assimilation: Why do immigrants converge to American health status levels? *Demography*, 43(2):337–360.

Antman, F. and Duncan, B. (2015). Incentives to Identify: Racial Identity in the Age of Affirmative Action. *Review of Economics & Statistics*, 97(3):710–713.

Antman, F. and Duncan, B. (2021). American Indian Casinos and Native American Self-Identification.

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

Antman, F., Duncan, B., and Trejo, S. J. (2016b). Ethnic attrition and the observed health of later generation Mexican Americans. *Am Econ Rev*, 106.

Antman, F. M., Duncan, B., and Trejo, S. J. (2020a). Ethnic attrition, assimilation, and the measured health outcomes of mexican americans. *Journal of Population Economics*, pages 1–24.

Antman, F. M., Duncan, B., and Trejo, S. J. (2020b). Ethnic attrition, assimilation, and the measured health outcomes of Mexican Americans. *Journal of Population Economics*, 33(4):1499–1522.

Åslund, O., Hensvik, L., and Skans, O. N. (2014). Seeking Similarity: How Immigrants and Natives Manage in the Labor Market. *Journal of Labor Economics*, 32(3):405–441.

Averett, S. and Korenman, S. (1996). The Economic Reality of the Beauty Myth. *The Journal of Human Resources*, 31(2):304–330.

Averett, S. L., Sikora, A., and Argys, L. M. (2008). For better or worse: Relationship status and body mass index. *Economics and Human Biology*, 6(3):330–349.

Bagues, M., Sylos-Labini, M., and Zinovyeva, N. (2017). Does the Gender Composition of Scientific Committees Matter? *American Economic Review*, 107(4):1207–1238.

Bayer, P. and Charles, K. K. (2018). Divergent paths: A new perspective on earnings differences between black and white men since 1940. *The Quarterly Journal of Economics*, 133(3):1459–1501.

Becker, G. S. (1957). *The economics of discrimination*. University of Chicago press.

Becker, G. S. (1973). A Theory of Marriage: Part I. *Journal of Political Economy*, 81(4):813–846.

Becker, G. S. (1974). A Theory of Marriage: Part II. *Journal of Political Economy*, 82(2, Part 2):S11–S26.

Becker, G. S. (1993). *A Treatise on the Family*. Harvard University Press, Cambridge, UNITED STATES.

Bertrand, M., Chugh, D., and Mullainathan, S. (2005). Implicit Discrimination. *American Economic Review*, 95(2):94–98.

Bertrand, M. and Mullainathan, S. (2004a). Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination. *American economic review*, 94(4):991–1013.

Bertrand, M. and Mullainathan, S. (2004b). Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *American Economic Review*, 94(4):991–1013.

Black, D. A. (1995). Discrimination in an equilibrium search model. *Journal of labor Economics*, 13(2):309–334.

Bonomi, G., Gennaioli, N., and Tabellini, G. (2021). Identity, Beliefs, and Political Conflict\*. *The Quarterly Journal of Economics*, 136(4):2371–2411.

Bratter, J. L. (2018). Multiracial Identification and Racial Gaps: A Work in Progress. *The Annals of the American Academy of Political and Social Science*, 677:69–80.

Browning, M., Chiappori, P.-A., and Lechene, V. (2006). Collective and Unitary Models: A Clarification. *Review of economics of the household*, 4(1):5–14.

Bursztyn, L., Chaney, T., Hassan, T. A., and Rao, A. (2022). The Immigrant Next Door: Long-Term Contact, Generosity, and Prejudice.

Card, D. and Krueger, A. B. (1992). School quality and black-white relative earnings: A direct assessment. *The quarterly journal of Economics*, 107(1):151–200.

Carlana, M. (2019). Implicit Stereotypes: Evidence from Teachers' Gender Bias. *Quarterly Journal of Economics*, 134(3):1163–1224.

Charles, K. K. and Guryan, J. (2008a). Prejudice and Wages: An Empirical Assessment of Becker's The Economics of Discrimination. *Journal of Political Economy*, 116(5):773–809.

Charles, K. K. and Guryan, J. (2008b). Prejudice and wages: an empirical assessment of becker's the economics of discrimination. *Journal of political economy*, 116(5):773–809.

Charles, K. K., Guryan, J., and Pan, J. (2018). The Effects of Sexism on American Women: The Role of Norms vs. Discrimination.

Chetty, R., Grusky, D., Hell, M., Hendren, N., Manduca, R., and Narang, J. (2017). The fading American dream: Trends in absolute income mobility since 1940. *Science*, 356(6336):398–406.

Chetty, R., Hendren, N., Jones, M. R., and Porter, S. R. (2020). Race and Economic Opportunity in the United States: An Intergenerational Perspective\*. *The Quarterly Journal of Economics*, 135(2):711–783.

Chetty, R., Hendren, N., and Katz, L. F. (2016). The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment. *American Economic Review*, 106(4):855–902.

Chetty, R., Hendren, N., Kline, P., and Saez, E. (2014a). Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States \*. *The Quarterly Journal of Economics*, 129(4):1553–1623.

Chetty, R., Hendren, N., Kline, P., Saez, E., and Turner, N. (2014b). Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility. *American Economic Review*, 104(5):141–147.

Chiappori, P.-A., Orefice, S., and Quintana-Domeque, C. (2012). Fatter Attraction: Anthropometric and Socioeconomic Matching on the Marriage Market. *Journal of Political Economy*, 120(4):659–695.

Dávila, A. and Mora, M. T. (2008). Changes in the Relative Earnings Gap Between Natives and Immigrants Along the U.s.-Mexico Border\*. *Journal of Regional Science*, 48(3):525–545.

Dee, T. S. (2005). A Teacher Like Me: Does Race, Ethnicity, or Gender Matter? *The American Economic Review*, 95(2):158–165.

Duncan, B., Grogger, J., Leon, A. S., and Trejo, S. J. (2020). New evidence of generational progress for mexican americans. *Labour Economics*, 62:101771.

Duncan, B., Hotz, V. J., and Trejo, S. J. (2006). Hispanics in the us labor market. *Hispanics and the Future of America*, pages 228–290.

Duncan, B. and Trejo, S. J. (2005). Ethnic Identification, Intermarriage, and Unmeasured Progress by Mexican Americans. Working Paper 11423, National Bureau of Economic Research.

Duncan, B. and Trejo, S. J. (2011a). Intermarriage and the intergenerational transmission of ethnic identity and human Capital for Mexican Americans. *J Labor Econ*, 29.

Duncan, B. and Trejo, S. J. (2011b). Who remains Mexican? Selective ethnic attrition and the intergenerational Progress of Mexican Americans. In Leal, D. L. and Trejo, S. J., editors, *Latinos and the Economy: Integration and Impact in Schools, Labor Markets, and Beyond*. Springer, New York.

Duncan, B. and Trejo, S. J. (2017a). The complexity of immigrant generations: Implications for assessing the socioeconomic integration of hispanics and asians. *ILR Review*, 70(5):1146–1175.

Duncan, B. and Trejo, S. J. (2017b). 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.

Duncan, B. and Trejo, S. J. (2018a). Identifying the later-generation descendants of us immigrants: Issues arising from selective ethnic attrition. *The ANNALS of the American Academy of Political and Social Science*, 677(1):131–138.

Duncan, B. and Trejo, S. J. (2018b). 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(1):131–138.

Duncan, B. and Trejo, S. J. (2018c). Socioeconomic integration of us immigrant groups over the long term: the second generation and beyond. Technical report, National Bureau of Economic Research.

Duncan, B. and Trejo, S. J. (2018d). Socioeconomic Integration of U.S. Immigrant Groups over the Long Term: The Second Generation and Beyond. Working Paper 24394, National Bureau of Economic Research.

Egloff, B. and Schmukle, S. C. (2002). Predictive validity of an implicit association test for assessing anxiety. *Journal of Personality and Social Psychology*, 83:1441–1455.

Flood, S., King, M., Rodgers, R., Ruggles, S., and Warren, J. R. (2020). Integrated public use microdata series, current population survey: Version 7.0 [dataset]. Minneapolis, MN: IPUMS, 2020. <https://doi.org/10.18128/D030.V7.0>.

Flood, S., King, M., Rodgers, R., Ruggles, S., Warren, J. R., and Westberry, M. (2021a). Integrated Public Use Microdata Series, Current Population Survey: Version 9.0.

Flood, S., King, M., Rodgers, R., Ruggles, S., Warren, J. R., and Westberry, M. (2021b). Integrated Public Use Microdata Series, Current Population Survey: Version 9.0.

Fouka, V., Mazumder, S., and Tabellini, M. (2022). From Immigrants to Americans: Race and Assimilation during the Great Migration. *The Review of Economic Studies*, 89(2):811–842.

Friese, M., Bluemke, M., and Wänke, M. (2007). Predicting voting behavior with implicit attitude measures: The 2002 German parliamentary election. *Experimental Psychology*, 54(4):247–255.

Fryer, R. G. and Levitt, S. D. (2004). The Causes and Consequences of Distinctively Black Names. *The Quarterly Journal of Economics*, 119(3):767–805.

Fryer, R. G. and Torelli, P. (2010). An empirical analysis of ‘acting white’. *Journal of Public Economics*, 94(5):380–396.

Fryer Jr, R. G. and Levitt, S. D. (2004). The causes and consequences of distinctively black names. *The Quarterly Journal of Economics*, 119(3):767–805.

Gershenson, S., Holt, S. B., and Papageorge, N. W. (2016). Who believes in me? The effect of student–teacher demographic match on teacher expectations. *Economics of Education Review*, 52:209–224.

Giuliano, L., Levine, D. I., and Leonard, J. (2009). Manager Race and the Race of New Hires. *Journal of Labor Economics*, 27(4):589–631.

Giuliano, L., Levine, D. I., and Leonard, J. (2011). Racial Bias in the Manager-Employee Relationship: An Analysis of Quits, Dismissals, and Promotions at a Large Retail Firm. *Journal of Human Resources*, 46(1):26–52.

- Giuntella, O. (2016). Assimilation and health: Evidence from linked birth records of second- and third-generation Hispanics. *Demography*, 53.
- Giuntella, O. (2017). Why does the health of Mexican immigrants deteriorate? New evidence from linked birth records. *Journal of Health Economics*, 54:1–16.
- Giuntella, O., Kone, Z. L., Ruiz, I., and Vargas-Silva, C. (2018). Reason for immigration and immigrants' health. *Public Health*, 158:102–109.
- Giuntella, O. and Stella, L. (2017). The Acceleration of Immigrant Unhealthy Assimilation. *Health Economics*, 26(4):511–518.
- Glover, D., Pallais, A., and Pariente, W. (2017). Discrimination as a Self-Fulfilling Prophecy: Evidence from French Grocery Stores. *Quarterly Journal of Economics*, 132(3):1219–1260.
- Greenwald, A. G., McGhee, D. E., and Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6):1464–1480.
- Grossman, G. M. and Helpman, E. (2021). Identity Politics and Trade Policy. *The Review of Economic Studies*, 88(3):1101–1126.
- Hadah, H. (2020). The Impact of Hispanic Last Names and Identity on Labor Market Outcomes.
- Jardina, A. (2019). *White Identity Politics*. Cambridge Studies in Public Opinion and Political Psychology. Cambridge University Press, Cambridge.
- Juhn, C., Murphy, K. M., and Pierce, B. (1991). Accounting for the slowdown in black-white wage convergence. *Workers and their wages*, 107:143.

Kranton, R., Pease, M., Sanders, S., and Huettel, S. (2020). Deconstructing bias in social preferences reveals groupy and not-groupy behavior. *Proceedings of the National Academy of Sciences*, 117(35):21185–21193.

Leitner, J. B., Hehman, E., Ayduk, O., and Mendoza-Denton, R. (2016). Racial bias is associated with ingroup death rate for Blacks and Whites: Insights from Project Implicit. *Social Science & Medicine*, 170:220–227.

Meng, X. and Gregory, R. G. (2005). Intermarriage and the Economic Assimilation of Immigrants. *Journal of Labor Economics*, 23(1):135–174.

Rubinstein, Y. and Brenner, D. (2014). Pride and prejudice: Using ethnic-sounding names and inter-ethnic marriages to identify labour market discrimination. *Review of Economic Studies*, 81(1):389–425.

Ruggles, S., Flood, S., Goeken, R., Pacas, J., Schouweiler, M., and Sobek, M. (2020). Integrated public use microdata series, american community survey: Version 11.0 [dataset]. Minneapolis, MN: IPUMS, 2020. <https://doi.org/10.18128/D010.V11.0>.

Smith, J. P. (2003). Assimilation across the Latino generations. *The American Economic Review*, 93(2):315.

Trejo, S. J. (1997). Why do Mexican Americans earn low wages? *Journal of Political Economy*, 105(6):1235.

Waters, M. C. (2001). *Black Identities: West Indian Immigrant Dreams and American Realities*. Russell Sage Foundation Books at Harvard University Press. Russell Sage Foundation Books at Harvard University Press, Cambridge, MA.