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

Hussain Hadah †

January 27, 2025

Click Here to Get the Most Updated Version

Abstract

I study the determinants of the choice to identify as Asian among those who could—those whose parents, grandparents, or selves were born in an Asian country. Using a multiple proxy regression approach, I construct a bias measure based on the Implicit Association Test (IAT), the American National Election Studies (ANES), and hate crimes against Asians. I find that individuals with Asian ancestry are significantly less likely to self-identify as Asian if they live in states with high levels of bias. A one standard deviation increase in bias decreases self-reported Asian identity by 9 percentage points for all immigrants. A one standard deviation increase in bias decreases self-reported Asian identity by 8 and 9 percentage points for second- and third-generation Asians respectively. These findings have implications for the interpretation of research on racial and ethnic gaps in economic outcomes and the correct counting of the population. **JEL**: I310, J15, J71, Z13

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

*I thank Patrick Button, Willa Friedman, Chinhui Juhn, Vikram Maheshri, and Yona Rubinstein for their support and advice. I also thank Aimee Chin, Steven Craig, German Cubas, Elaine Liu, Fan Wang, and the participants of the Applied Microeconomics Workshop at the University of Houston, the European Society for Population Economics (ESPE), the anonymous referees, and the editor for their helpful feedback.

†Hussain Hadah: Department of Economics, Tulane University, 6823 St. Charles Ave., New Orleans, LA 70118, United States (e-mail: hhadah@tulane.edu, phone: +1-602-393-8077).

1 Introduction

The Asian population in the United States has almost doubled over the last two decades.[1] An extensive literature on Asian–White labor market gaps has emerged (Arabsheibani and Wang 2010; Chiswick 1983; Duleep and Sanders 2012; Hilger 2016). However, defining and measuring these racial and ethnic groups is not straightforward, especially when considering self-reported identity. To the extent that self-reporting Asian racial identity is positively selected, these gaps could be biased.

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

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

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

[1] The 2020 Census counted more than 20 million Asian Americans—6.4 percent of the population—almost double the number of Asians counted two decades earlier (Flood, King, et al., Integrated Public Use Microdata Series, Current Population Survey). The Asian population numbers are based on the author's calculations from the Current Population Survey and US Census data.

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

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

Among individuals of Asian ancestry, I find that higher bias—against Asians—is correlated with a lower self-reported Asian identity among Asian immigrants. I find that an increase of one standard deviation in bias correlates with a statistically significant 9 percentage point decrease in the self-reported Asian identity among first-generation immigrants, a statistically insignificant 5 percentage point decrease among second-generation immigrants, and a statistically significant 8 percentage point decrease in the self-reported Asian identity among third-generation immigrants. Additionally, a one standard deviation increase in bias is correlated with a 5 percentage points (insignificant) drop in self-reported Asian identity among second-generation Asian children with both parents born in a Asian country, a 15 percentage point decrease in self-reported Asian identity among children of Asian fathers-White mothers, and a 10 percentage point decrease in self-reported Asian identity among children of White fathers-Asian mothers. Consequently, as the more economically successful Asian immigrants—educated and wealthy immigrants—may self-report Asian identity, economic research using subjective ethnic measures will underestimate White-Asian gaps in the most biased states.

This paper most closely fits in the literature of stratification economics. The interplay between racial identity, economic status, and social outcomes forms a complex web that various scholars have sought to untangle. Darity (2022) and Darity, Mason, and Stewart (2006) provide a foundational understanding of the economics of identity and stratification, suggesting that both historical and contemporary economic factors contribute to the persistence of racial norms and inequality. This theme is extended in the context of labor and marital markets by Diette et al. (2015), Goldsmith, Hamilton, and Darity (2007), and Hamilton, Goldsmith, and Darity (2009) who explore how skin color influences economic and social prospects among African Americans. Similarly, Golash-Boza and Darity Jr (2008) reveal the nuanced racial self-identification processes among Latinos, affected by skin color and discrimination. The significant impact of political and national events on racial identity is also evident in Mason and Matella (2014) study of Arab and Islamic Americans post-September 11 and Mason (2017) examination of the 2008 Presidential Election's effect on African American racial identity. These studies collectively underscore the multidimensional nature of racial identity and its profound implications for economic and social stratification. I contribute to the literature of stratification economics by providing evidence that Asian identity formation is influenced by societal factors, i.e. discrimination and prejudice.

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

This paper is most closely related to Antman and Duncan (2015, 2021), Antman, Duncan, and Trejo (2016), and Hadah (forthcoming) where the authors studied the ethnic attrition of Hispanic immigrants and how minorities change their self-reported identity to changes in policies.[4] Taking into consideration the ethnic attrition that Antman, Duncan, and Trejo (2016) document, I investigate the determinants of what drives a person to self-report, or not, their Asian identity. Hadah (forthcoming) finds that bias and self-reported Hispanic identity are negatively associates among a sample of objectively Hispanic immigrants. I aim to decompose some of the complexity associated with endogenous identity by exploring certain personal and environmental determinants of identity. The empirical analysis in this paper documents how certain observable factors, namely personal characteristics and societal attitudes, affect the self-reported identity of Asians.

The rest of this paper is structured as follows. First, I will discuss the conceptual framework in section (2). Second, I will describe the data I use in section (3). Third, I will introduce an empirical model and the results in sections (4) and (5). Fourth, I will discuss robustness checks and discuss the results in section (6). Finally, I conclude in section (7).

2 Conceptual Framework

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

Formally, a person i belongs to racial group $e_i \in \{A, W\}$, where A is Asian and W is White. Agent i's utility depends on their actions and the extent to which their actions affirm their identity I_i:

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

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

[4] Ethnic attrition is when a person with an ethnic minority ancestry fails to self-identify with the group.

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

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

Where $\boldsymbol{a_i}$ is the actions of person i. $\boldsymbol{a_{-i}}$ is the actions of others that would affect i's identity, i.e., societal bias. I_i is the identity function. Each group has an associated set of behaviors that society proscribes them to conform to, which I denote as $\mathbf{B}_{e_i}$.[5]

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

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

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

That is i will change identities when the benefits of doing so $\tilde{U_i}^\star - U_i^\star$ exceed the costs c. These net benefits are non-zero only if $\frac{dI_i}{da_i} \neq 0$ and $\frac{\partial U_i}{\partial I_i} \neq 0$. This suggests that an empirical analysis of the determinants of identity choice should focus on: (1) individual characteristics that would lead to different a_i under different identities, (2) contextual characteristics that would lead to different a_{-i}—bias—under different identities, (3) the analysis should focus on a sample of the population with small c, and (4) the sub-sample with a utility that is greatly affected by their identity—i.e., $\frac{\partial U_i}{\partial I_i} \neq 0$). From the empirical analysis, I could investigate the characteristics that would affect i's actions to take different identities from point (1). These characteristics could be the generation immigrants belong to, whether their parents are interracial or endogamous, etc. I also investigate how different 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-Asian Whites with non-Asian ancestry.

3 Data

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

[5] Akerlof and Kranton (2000) refer to B_{e_i} as proscription.

measure of bias using the implicit association test, the American National Election Studies, and hate crimes against Asians. My composite measure is created using the method of Lubotsky and Wittenberg (2006) to reduce attenuation bias.

3.1 Measuring Asian Identity

I measure Asian identity using the Current Population Survey (CPS), which allows me to construct an objective measure of the Asian identity of minors living with their parents. I will use the information on the place of birth, parent's place of birth, and place of birth of grandparents to construct an objective Asian measure.[6] Thus, I could perfectly identify and construct a dataset of first-, second-, and third-generation Asian immigrants (see Figure 2 for a visual representation). This will consequently allow me to build an objective measure of the Asian identity of minors under the age 17 living with their parents.

The objective measure of identity—unlike the self-reported measure where respondents answer affirmatively when asked if they are Asian—depends on the birthplaces of the individual, their two parents, and four grandparents. Thus, the three identifiable generations are: 1) first-generation immigrants that are born in an Asian country with both parents also being born in an Asian country, 2) second-generation immigrants are native-born citizens to at least one parent that was born in an Asian country, 3) third-generation immigrants are native-born citizens to two native-born parents and at least one grandparent that was born in an Asian country.[7] I restrict the sample to Asian, first-, second-, and third-generation immigrants who are 17 year old and younger and still live with their parents between 2004 and 2021. I present a summary of the sample statistics in the Table (1).

While the CPS relies on household respondents (parents or caregivers) to report children's ethnic identity, this proxy reporting is likely to accurately reflect children's true identity since parents play a crucial role in identity formation. Duncan and Trejo (2011) support this view, showing no variation in Asian identification based on the household respondent. The data confirms this—Asian identity reporting is consistently high regardless of whether the mother (72%), father (72%), or child/other caregiver (87%) is the respondent, see Table 6.[8] Additionally, since my analysis compares high and low bias states, the estimates remain valid as long as reporting tendencies do not systematically differ between these states.

The overall sample is 49% female, and 65% of the sample self-report their identity as

[6] Following the works of Antman, Duncan, and Trejo (2016) and Antman, Duncan, and Trejo (2020).

[7] 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.

[8] 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.

Asian—answered yes to the question "what is your race". The average age is 8.4-year-old. Almost 52% of mothers have a college degree, and 52% of fathers have a college degree. I provide the rest of the summary statistics for the overall sample and for both the overall sample and each generation in Table (1).

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, I can divide parents of second-generation children into three objective types:

1. Objectively Asian-father-Asian-mother (AA)
2. Objectively Asian-father-White-mother (AW)
3. Objectively White-father-Asian-mother (WA)

Similarly, using the place of birth of grandparents, I can divide grandparents of third-generation children into 15 objective types: (1) objectively Asian paternal grandfather-Asian paternal grandmother-Asian maternal grandfather-Asian maternal grandmother (AAAA); (2) objectively White paternal grandfather-Asian paternal grandmother-Asian maternal grandfather-Asian maternal grandmother (WAAA); (3) objectively Asian paternal grandfather-White paternal grandmother-Asian maternal grandfather-Asian maternal grandmother (AWAA), etc...

My analysis uses a sub-sample of the US population; I show in Table (2) that I have enough observations in each generation. Consistent with the literature on ethnic attrition among Asians, I find significant attrition among third-generation Asian immigrants.[9] These results are displayed in Table (2): most first- and second-generation Asian immigrants self-report their identity as Asian. Of the first-generation Asian immigrants, 96% self-report their identity as Asian. 73% of the second-generation Asian immigrants self-report their identity as Asian, and 31% of third-generation Asian immigrants identified as Asian. The attrition among second- and third-generation Asian immigrants is primarily driven by children of interracially married parents.

3.2 Measuring Prejudice

To construct a measure of prejudice, I use the implicit association test, the American National Election Studies, and hate crimes against Asians. 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 (A.1) shows a few examples of what a test taker would see on a skin tone implicit association test by Harvard's Project Implicit.

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

I use Asian implicit association test data to construct a proxy of prejudice (Greenwald, McGhee, and Schwartz 1998). This measure has been used in the social sciences, especially in psychology. Previous work has shown that IAT test scores are hard to manipulate (Egloff and Schmukle 2002).

The IAT aims to measure the direction and magnitude of bias in people. It also aims to measure unconscious biases in people or biases that they are unwilling to report. On the one hand, in a meta-analysis of more than 122 papers that used IAT, Greenwald, McGhee, and Schwartz (1998) find that IAT measures had significantly higher predictive validity than self-report measures. On the other hand, some research disputes the claims of the IAT's predictive validity.[10] The Implicit Association Test (IAT) may not reliably measure or predict implicit prejudice or biased behaviors. Some research shows that implicit biases undergo minor and temporary changes through interventions. Additionally, implicit bias does not predict dictator game giving or being influenced by social pressure, highlighting the distinction between implicit bias and biased actions (Arkes and Tetlock 2004; Forscher et al. 2019; Lee 2018). Therefore, I supplement the IAT with a measure of explicit bias from the American National Election Studies (ANES) to construct a composite measure of bias.

I construct another proxy measure of racial animus using the ANES survey American National Election Studies (2021) to measure animus, or discrimination, against Black Americans. ANES is a survey that has been conducted since 1948 and is widely used in political science. The survey asks respondents about their attitudes toward different racial groups, voting intentions, and other political questions. I use several questions from the ANES surveys conducted between 2004 and 2020 to measure racial animus. The racial animus index is constructed by taking the average of the responses to several questions measuring racial animus. [11]

Lastly, I incorporate data from the Uniform Crime Reports (UCR) to quantify hate crimes against Asians (Bureau of Justice Statistics 2023). Hate crime data offers a tangible measure of racially motivated aggression and discrimination, which, when combined with implicit and explicit bias measures, allows for a fuller understanding of the landscape of prejudice across states. This combination of implicit and explicit bias measures, along with hate crime statistics, offers a multidimensional approach to understanding the nature and prevalence of racial prejudice.

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

[11]The questions used are similar to those used by Charles and Guryan (2008). The questions are: (1) "Conditions Make it Difficult for Blacks to Succeed", (2) "Blacks Should Not Have Special Favors to Succeed", (3) "Blacks Must Try Harder to Succeed", (4) "Blacks Gotten Less than They Deserve Over the Past Few Years", and (5) "Feeling Thermometer Toward Asians."

Moreover, to reduce attenuation bias and measurement error, I follow Lubotsky and Wittenberg (2006) in constructing a composite bias measure using the IAT, the ANES racial animus measure and hare crimes against Asians.[12] Figure (1a) shows a graphical representation of the bias measure over time in the most and least biased locations. Figure (1b) shows a graphical representation of self-reported Asian identity in the two most and least biased locations. A lower score implies less bias, whereas a higher score implies higher racial animus. A one standard deviation increase in bias is equivalent to moving from Washington, DC, or Vermont to North Dakota in 2020. I also show the average bias over time in the maps in Figure (3) and the overall average from 2004 to 2021 in Figure (4).

4 Estimation and Results

To understand the association between Asian self-identity and bias, I estimate regressions of the following form for each generation g:

$$\begin{aligned} A^{g}_{ist} = {} & \beta^{g}_{1}\mathit{Bias}_{st} + \beta^{g}_{2}\mathit{DadCollegeGrad}_{ist} + \beta^{g}_{3}\mathit{MomCollegeGrad}_{ist} \\ & + \beta^{g}_{4}\mathit{Women}_{ist} + X^{g}_{ist}\pi + \gamma_{rt} + \varepsilon_{ist}; \text{where } g \in \{1,2,3\} \end{aligned} \tag{4}$$

Where A^{g}_{ist} be the self-reported Asian identity of person i in state s at the time of interview t, let Bias_{st} be the average bias in state s at time t, $\mathit{DadCollegeGrad}_{ist}$, and $\mathit{MomCollegeGrad}_{ist}$ are indicator variables that are equal to one if the father or mother graduated from college, Women_{ist} is an indicator variable for sex, and X_{ist} is a vector of controls.[13] Additionally, γ_{rt} is region-time fixed effects that controls for region $\times$ year specific shocks.[14] The region $\times$ year also controls for systematic differences between regions in the overall Asian population and bias toward Asians, even if they vary over time. Throughout the analysis, I cluster the standard errors at the state level to account for correlation in the error term ε_{ist} within a state, overtime.

Since the specification includes region $\times$ year, γ_{rt}, the β^{g}_{1} coefficient summarizes individual's i responsiveness to bias changes in the state which they live. In other words, β^{g}_{1} captures the association between self-reported Asian identity and bias across states within a Census division region. Additionally, the γ_{rt} fixed effects account for any regional and national trends in bias over time. Consequently, β^{g}_{1} provide the correlation between self-reported Asian identity and bias above and beyond the national and regional trends in bias. If individuals in states within a region responded similarly to changes in bias, then β^{g}_{1} will be equal to zero.

[12]More on the method in the Data Online Appendix, see Section C.1.

[13]The controls include quartic age, fraction of population that is Asian in state s, type of parents (WA, AW, or AA), type of grandparents (AAAA, AAAW, etc.), and dummy variables the generation to which person i belong.

[14]I do not include state fixed effects because of lack of with-in state variation in bias.

5 Results

The results from the regression framework described above provide consistent evidence aligning with the following findings. First, bias is negatively associated with self-reported Asian identity. Second, first- and second-generation Asian immigrant children of endogamous marriages' self-reported Asian identity are more negatively associated with bias.[15]

I report the main results of estimating equation (4) in Figure (5).[16] I present the results of estimating the main specification for second-generation immigrants in panel (A) and for sub-samples of AA, AW, and WA children in panels (B), (C), and (D), respectively. I find that bias and self-reported Asian identity are negatively associated. A one standard deviation increase in bias is associated with a 9 percentage points decrease in self-reported Asian identity. Among first- and second-generation Asian immigrants, a one standard deviation increase in bias is associated with 5 and 8 percentage points decrease in self-reported Asian identity. The coefficient is not statistically significant for first-generation, but the confidence interval is mostly in the negative territory. Finally, among third-generation Asian immigrants, a one standard deviation increase in bias is associated with a 8 percentage points decrease in self-reported Asian identity.

I report the results of the same regression but on sub-samples of second-generation immigrants by type of parents—interracial and endogamous parents—in Figure (6). I present the results of estimating the main specification on second-generation immigrants in panel (A) and on sub-samples of AA, AW, and WA children in panels (B), (C), and (D), respectively. I find that children of interracial parents marriages are more influenced by bias. I find that a one standard deviation increase in bias is associated with a 5 percentage points decrease in self-reported Asian identity among children of endogamous parents—the estimate is statistically insignificant. However, a one standard deviation increase in bias is associated with a 15 percentage points decrease in self-reported Asian identity among children of Asian fathers-White mothers, and a 10 percentage points decrease in self-reported Asian identity among children of White fathers-Asian mothers.

I also report the results of the regression on sub-samples of third-generation immigrants by the number of Asian grandparents in Table (3). The overall effect of bias on the different type of Asian children is negative, however, they are also mostly statistically insignificant. I find that a one standard deviation increase in bias is associated with a 69 percentage points decrease in self-reported Asian identity among Asian children that have three grandparents that are born in an Asian country.

[15] I show the results with different Fixed Effects in Table (A.1).

[16] You can find the table results in Tables A.2 and A.3.

6 Robustness Checks and Discussions

In this section, I explore the empirical relationship between bias and interracial marriages. I also analyze migration patterns of second-generation Asian immigrants and the effect of proxy response on my results as robustness checks to my main analysis. I examine the impact of biases on the likelihood of interracial marriages, focusing on interracial couples, and the migration decisions of Asian individuals within the United States.

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

$$\mathtt{interracial}^2_{ist} = \beta^2_1 \mathtt{Bias}_{st} + X^2_{ist}\pi + \gamma_{rt} + \varepsilon_{ist} \tag{5}$$

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

I present the results of estimating equation (5) in Table (4). I find that a one standard deviation increase in bias increases the probability of having interracial parents by 4 percentage points. Moreover, I break down the analysis by the ethnicity of the couples. A one standard deviation increase in bias is associated with 1 percentage points decrease in the chances of a Asian husband marrying a White wife. A one standard deviation increase in bias is associated a 3 percentage points increase in the chances of a Asian wife having a White husband. The fact that bias and interracial marriage are positively correlated could be a result of the fact that Asian immigrants in states with high bias might aim to decrease the likelihood that their children will display Asian ethnicity signals. For example, Asian women in high bias states might marry a non-Asian White husband, so their children will have a non-Asian last name.[17] This positive correlation between bias and interracial marriage provides a robustness check on my main results, as it suggests an additional mechanism through which bias affects identity choices beyond direct self-reporting decisions.

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

[17] I show the results of estimating the relationship between bias and interracial marriages as a logistic regression in Table (A.4).

$$\mathtt{BirthPlaceMigration}^2_{ist} = \beta^2_1 \mathtt{Bias}_{st} + X^2_{ist}\pi + \gamma_{rt} + \varepsilon_{ist} \tag{6}$$

$$\mathtt{BirthPlaceMigration}^2_{ilb} = \beta^2_1 \mathtt{Bias}_{lb} + X^2_{ilb}\pi + \gamma_{lb} + \varepsilon_{ilb} \tag{7}$$

Where $\mathtt{BirthPlaceMigration}^2_{ist}$ is an indicator variable equal to one if person i in state s at the interview t lives in a state that is different from his or her birth state and zero otherwise. $\mathtt{BirthPlaceMigration}^2_{ilb}$ is an indicator variable that is equal to one if person i in birthplace l does not currently live in the same state he or she lived in at the year of birth b and zero otherwise. The analysis, restricted to second-generation Asian immigrants with both parents born in a Asian country, uses equations (6) and (7).

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

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

$$Y_{ist} = \beta_0 + \beta^2_1 \mathtt{Asian}_{ist} + X^2_{ist}\pi + \varepsilon_{ist} \tag{8}$$

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

The results of estimating equations (6), (7), and (8) are shown in Table (5) in columns (1), (2), and (3) respectively. I find that among second-generation immigrants, there is no significant correlation between bias and migration decisions. Among second-generation Asian immigrant movers, those who self-report Asian identity live in states with 0.06 standard deviations more biased than the state where they were born. Even though this result shows that there is selection into more biased states among second-generation immigrants, it does not affect my main results showing a correlation between bias and self-reported Asian identity. Since those identifying as Asians are the movers, my assessments of the relationship between bias and self-reported Asian identity might underestimate the effect of bias.

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

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

Moreover, Duncan and Trejo (2011) show that reported Asian identification does not vary with who is the household respondent. Additionally, I present the main effect of self-reported Asian identity by the household respondent in Table (6). The main effect of the reported Asian identity of children is 72 percentage points when the mother is the proxy, 72 percentage points when the father is the proxy, and 87 percentage points when the child or another caregiver was the household respondent.[18]

A second concern is that the IAT is voluntary and not representative of the population. While I do not claim that the IAT as a proxy for bias will represent the population, Egloff and Schmukle (2002) show that they are hard to manipulate. Several studies have shown that IAT is correlated with economic outcomes (Chetty et al. 2020; Glover, Pallais, and Pariente 2017), voting behavior (Friese, Bluemke, and Wänke 2007), decision-making (Bertrand, Chugh, and Mullainathan 2005; Carlana 2019), and health (Leitner et al. 2016). Another concern could be that the IAT test takers' characteristics change over time and, thus, are not the same. I address this concern by including non-parametric region $\times$ 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. Most importantly, I use the ANES racial animus measure and hate crimes against Asians to construct a composite measure of bias that reduces measurement error using Lubotsky and Wittenberg (2006).

Another concern could be reverse causality between having more Asian or Black people in a state and bias. It could be the case that the number of Asian people in a state affects the bias on the residents of that state. For example, having more Asians in Florida or Black people in Louisiana could affect the bias of the residents of Florida and Louisiana. To show that this is not the case, I provide Figures (A.2) as evidence. Figure (A.2) plots the percent of self-reported Asians in a state at a specific year against the average bias in the same state during that year. I find no correlation between bias and the number of Asians in a state, thus, making the case of reverse causality unlikely.

[18]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.

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

7 Conclusion

As the United States becomes more multi-racial and multi-ethnic, self-reported identity will significantly impact representation, distributive politics, and government transfers. The determinants of endogenous identity are particularly important to researchers interested in the role of discrimination on earnings gaps. In this paper, I show how individual characteristics and social attitudes toward racial and ethnic minorities affect the self-reported Asian identity of individuals with Asian ancestry in the United States. I find that people of Asian ancestry are less likely to identify as Asian in states with more significant bias. The relationship between self-reported Asian identity and bias among first-generation immigrants, where a one standard deviation increase in bias correlated with a 2 percentage points decrease in self-reported Asian identity; the results are not statistically significant. The relationship between self-reported Asian identity and bias is more prominent among second-generation immigrants, where a one standard deviation increase in bias correlated with a 4 percentage points decrease in self-reported Asian identity.

Additionally, bias has a more substantial effect among second-generation immigrant children with Asian fathers and Asian mothers. A one standard deviation increase in bias correlates with a 5 percentage points decrease in self-reported Asian identity among second-generation Asian immigrant children of objectively Asian parents. I also find that bias positively correlates with interracial marriage and not with migration decisions.

The results are important because of the consequences on the correct counting of Asians and minorities, 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 bias is negatively correlated with self-reported Asian identity, the characteristics of those who do not self-report Asian identity could have important consequences. For example, if the people whose identities are most likely affected by bias are the most educated. In this case, the racial and ethnic gaps will be overestimated in the most biased states. Furthermore, identity decisions are likely to affect people's choices, investments, and well-being profoundly.

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

References

Abramitzky, Ran, Leah Boustan, Katherine Eriksson, and Stephanie Hao. 2020. "Discrimination and the Returns to Cultural Assimilation in the Age of Mass Migration." *AEA Papers and Proceedings* 110 (May 1, 2020): 340–346. https://doi.org/10.1257/pandp.20201090.

Abramitzky, Ran, Leah Platt Boustan, and Dylan Connor. 2020. *Leaving the Enclave: Historical Evidence on Immigrant Mobility from the Industrial Removal Office.* w27372. Cambridge, MA: National Bureau of Economic Research, June. https://doi.org/10.3386/w27372.

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. 2014. "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration." *Journal of Political Economy* 122, no. 3 (June): 467–506. https://doi.org/10.1086/675805.

———. 2016. "Cultural Assimilation during the Age of Mass Migration." Prepublished, July. Working Paper. https://doi.org/10.3386/w22381. National Bureau of Economic Research: 22381.

Abramitzky, Ran, Leah Platt Boustan, Elisa Jácome, and Santiago Pérez. 2019. *Intergenerational Mobility of Immigrants in the US over Two Centuries.* w26408. Cambridge, MA: National Bureau of Economic Research, October. https://doi.org/10.3386/w26408.

Akerlof, George A., and Rachel E. Kranton. 2000. "Economics and Identity." *Quarterly Journal of Economics* 115, no. 3 (August): 715–753. https://doi.org/10.1162/003355300554881.

American National Election Studies. 2021. *ANES 2020 Time Series Study Full Release [Dataset and Documentation].* https://www.electionstudies.org.

Antman, Francisca, and Brian Duncan. 2015. "Incentives to Identify: Racial Identity in the Age of Affirmative Action." *Review of Economics & Statistics* 97, no. 3 (July): 710–713. https://doi.org/10.1162/REST_a_00527.

———. 2021. "American Indian Casinos and Native American Self-Identification" (August 29, 2021).

Antman, Francisca, Brian Duncan, and Stephen J. Trejo. 2016. "Ethnic Attrition and the Observed Health of Later-Generation Mexican Americans." *The American Economic Review* 106 (5): 467–471. http://www.jstor.org/stable/43861065.

Antman, Francisca M., Brian Duncan, and Stephen J. Trejo. 2020. "Ethnic Attrition, Assimilation, and the Measured Health Outcomes of Mexican Americans." *Journal of Population Economics* (Heidelberg, Netherlands) 33, no. 4 (October): 1499–1522. https://doi.org/https://doi-org.ezproxy.lib.uh.edu/10.1007/s00148-020-00772-8.

Arabsheibani, G Reza, and Jie Wang. 2010. "Asian-white male wage differentials in the United States." *Applied Economics Letters* 17 (1): 37–43.

Arkes, Hal R., and Philip E. Tetlock. 2004. "Attributions of Implicit Prejudice, or "Would Jesse Jackson 'Fail' the Implicit Association Test?"" *Psychological Inquiry* (US) 15 (4): 257–278. https://doi.org/10.1207/s15327965pli1504_01.

Bertrand, Marianne, Dolly Chugh, and Sendhil Mullainathan. 2005. "Implicit Discrimination." *American Economic Review* 95, no. 2 (May): 94–98. https://doi.org/10.1257/000282805774670365.

Bureau of Justice Statistics. 2023. *Uniform Crime Reporting Program Data Series [Dataset].* https://catalog.data.gov/dataset/uniform-crime-reporting-program-data-series-16edb.

Carlana, Michela. 2019. "Implicit Stereotypes: Evidence from Teachers' Gender Bias." *Quarterly Journal of Economics* 134, no. 3 (August): 1163–1224. https://doi.org/10.1093/qje/qjz008.

Charles, Kerwin Kofi, and Jonathan Guryan. 2008. "Prejudice and Wages: An Empirical Assessment of Becker's The Economics of Discrimination." *Journal of Political Economy* 116, no. 5 (October): 773–809. https://doi.org/10.1086/593073.

Chetty, Raj, Nathaniel Hendren, Maggie R Jones, and Sonya R Porter. 2020. "Race and Economic Opportunity in the United States: An Intergenerational Perspective*." *The Quarterly Journal of Economics* 135, no. 2 (May 1, 2020): 711–783. https://doi.org/10.1093/qje/qjz042.

Chiswick, Barry R. 1983. "An analysis of the earnings and employment of Asian-American men." *Journal of Labor Economics* 1 (2): 197–214.

Darity, William A. 2022. "Position and Possessions: Stratification Economics and Intergroup Inequality." *Journal of Economic Literature* 60, no. 2 (June 1, 2022): 400–426. https://doi.org/10.1257/jel.20211690.

Darity, William A., Patrick L. Mason, and James B. Stewart. 2006. "The Economics of Identity: The Origin and Persistence of Racial Identity Norms." *Journal of*

Economic Behavior & Organization 60, no. 3 (July 1, 2006): 283–305. https://doi.org/10.1016/j.jebo.2004.09.005.

Diette, Timothy M., Arthur H. Goldsmith, Darrick Hamilton, and William Darity. 2015. "Skin Shade Stratification and the Psychological Cost of Unemployment: Is There a Gradient for Black Females?" *The Review of Black Political Economy* 42, nos. 1-2 (January 1, 2015): 155–177. https://doi.org/10.1007/s12114-014-9192-z.

Duleep, Harriet Orcutt, and Seth G Sanders. 2012. "The economic status of Asian Americans before and after the Civil Rights Act."

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

Duncan, Brian, and Stephen J. Trejo. 2018a. "Identifying the Later-Generation Descendants of U.S. Immigrants: Issues Arising from Selective Ethnic Attrition." *The ANNALS of the American Academy of Political and Social Science* 677, no. 1 (May 1, 2018): 131–138. https://doi.org/10.1177/0002716218763293.

———. 2018b. *Socioeconomic Integration of U.S. Immigrant Groups over the Long Term: The Second Generation and Beyond.* Working Paper, Working Paper Series 24394. National Bureau of Economic Research, March. https://doi.org/10.3386/w24394.

Egloff, Boris, and Stefan C. Schmukle. 2002. "Predictive Validity of an Implicit Association Test for Assessing Anxiety." *Journal of Personality and Social Psychology* (US) 83:1441–1455. https://doi.org/10.1037/0022-3514.83.6.1441.

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, J. Robert Warren, and Michael Westberry. 2021. (Integrated Public Use Microdata Series, Current Population Survey: Version 9.0. V. 9.0). https://doi.org/10.18128/D030.V9.0.

Flood, Sarah, Goeken Ronald, Schouweiler Megan, and Sobek Matthew. 2021. (Integrated Public Use Microdata Series, USA: Version 12.0. V. 9.0). https://doi.org/10.18128/D030.V9.0.

Forscher, Patrick S., Calvin K. Lai, Jordan R. Axt, Charles R. Ebersole, Michelle Herman, Patricia G. Devine, and Brian A. Nosek. 2019. "A Meta-Analysis of Procedures to Change Implicit Measures." *Journal of Personality and Social Psychology* (US) 117 (3): 522–559. https://doi.org/10.1037/pspa0000160.

Fouka, Vasiliki, Soumyajit Mazumder, and Marco Tabellini. 2022. "From Immigrants to Americans: Race and Assimilation during the Great Migration." *The*

Review of Economic Studies 89, no. 2 (March 1, 2022): 811–842. https://doi.org/10.1093/restud/rdab038.

Friese, Malte, Matthias Bluemke, and Michaela Wänke. 2007. "Predicting Voting Behavior with Implicit Attitude Measures: The 2002 German Parliamentary Election." *Experimental Psychology* (Germany) 54 (4): 247–255. https://doi.org/10.1027/1618-3169.54.4.247.

Glover, Dylan, Amanda Pallais, and William Pariente. 2017. "Discrimination as a Self-Fulfilling Prophecy: Evidence from French Grocery Stores." *Quarterly Journal of Economics* 132, no. 3 (August): 1219–1260. https://doi.org/10.1093/qje/qjx006.

Golash-Boza, Tanya, and William Darity Jr. 2008. "Latino Racial Choices: The Effects of Skin Colour and Discrimination on Latinos' and Latinas' Racial Self-Identifications." *Ethnic and Racial Studies* 31, no. 5 (July 1, 2008): 899–934. https://doi.org/10.1080/01419870701568858.

Goldsmith, Arthur H., Darrick Hamilton, and William Darity. 2007. "From Dark to Light: Skin Color and Wages among African-Americans." *The Journal of Human Resources* 42 (4): 701–738. https://www.jstor.org/stable/40057327.

Greenwald, Anthony G., Debbie E. McGhee, and Jordan L. K. Schwartz. 1998. "Measuring Individual Differences in Implicit Cognition: The Implicit Association Test." *Journal of Personality and Social Psychology* (US) 74 (6): 1464–1480. https://doi.org/10.1037/0022-3514.74.6.1464.

Hadah, Hussain. Forthcoming. "The Effect of Racial and Ethnic Attitudes on Hispanic Identity in the U.S." *Southern Economic Journal.*

Hamilton, Darrick, Arthur H. Goldsmith, and William Darity. 2009. "Shedding "Light" on Marriage: The Influence of Skin Shade on Marriage for Black Females." *Journal of Economic Behavior & Organization* 72, no. 1 (October 1, 2009): 30–50. https://doi.org/10.1016/j.jebo.2009.05.024.

Hilger, Nathaniel. 2016. *Upward mobility and discrimination: The case of Asian Americans.* Technical report. National Bureau of Economic Research.

Lee, Daniel J. 2018. "Does Implicit Bias Predict Dictator Giving?" *Games* 9, no. 4 (4): 73. https://doi.org/10.3390/g9040073.

Leitner, Jordan B., Eric Hehman, Ozlem Ayduk, and Rodolfo Mendoza-Denton. 2016. "Racial Bias Is Associated with Ingroup Death Rate for Blacks and Whites: Insights from Project Implicit." *Social Science & Medicine* 170 (December 1, 2016): 220–227. https://doi.org/10.1016/j.socscimed.2016.10.007.

Lubotsky, Darren, and Martin Wittenberg. 2006. “Interpretation of Regressions with Multiple Proxies.” *The Review of Economics and Statistics* 88, no. 3 (August 1, 2006): 549–562. https://doi.org/10.1162/rest.88.3.549.

Mason, Patrick L. 2017. “Not Black-Alone: The 2008 Presidential Election and Racial Self-Identification among African Americans.” *The Review of Black Political Economy* 44, nos. 1-2 (January 1, 2017): 55–76. https://doi.org/10.1007/s12114-017-9247-z.

Mason, Patrick L., and Andrew Matella. 2014. “Stigmatization and Racial Selection after September 11, 2001: Self-Identity among Arab and Islamic Americans.” *IZA Journal of Migration* 3, no. 1 (October 23, 2014): 20. https://doi.org/10.1186/s40176-014-0020-9.

Meng, Xin, and Robert G. Gregory. 2005. “Intermarriage and the Economic Assimilation of Immigrants.” *Journal of Labor Economics* 23, no. 1 (January): 135–174. https://doi.org/10.1086/425436.

Table 1: CPS Summary Statistics

Characteristic	Overall	By Generation		
	All Sample N = 318,404	First N=40,033	Second N=199,294	Third N=79,077
Female	0.49	0.53	0.49	0.49
Asian	0.65	0.96	0.73	0.31
Age	8.4 (5.1)	10.9 (4.5)	8.3 (5.1)	7.7 (5.0)
College Graduate: Father	0.52	0.59	0.52	0.50
College Graduate: Mother	0.52	0.56	0.51	0.52
Total Family Income (1999 dollars)	87,031 (84,797)	75,815 (74,489)	88,295 (88,411)	89,436 (80,051)

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

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

Table 2: Asian Self-identification by Generation

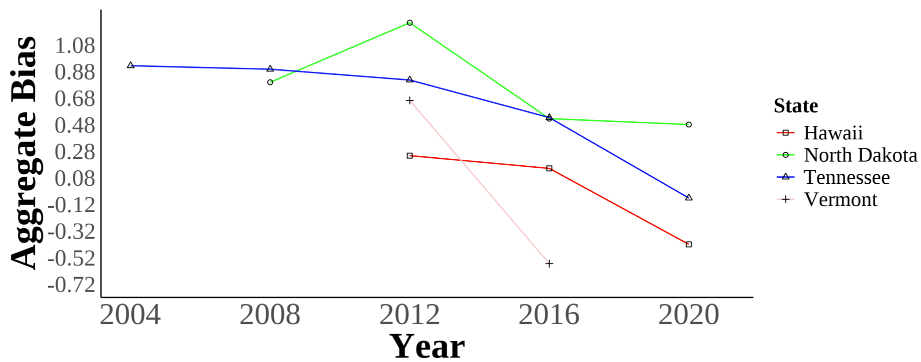
	Self-identify as Asian	Self-identify as non-Asian	% Self-identify as Asian	% Self-identify as non-Asian
1st Gen.	14,811	688	0.96	0.04
2nd Gen.	58,756	21,381	0.73	0.27
Asian on:				
Both Sides	49,118	1,717	0.97	0.03
One Side	9,638	19,664	0.33	0.67
3rd Gen.	10,394	23,048	0.31	0.69
Asian on:				
Both Sides	5,428	316	0.94	0.06
One Side	3,030	9,213	0.25	0.75

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

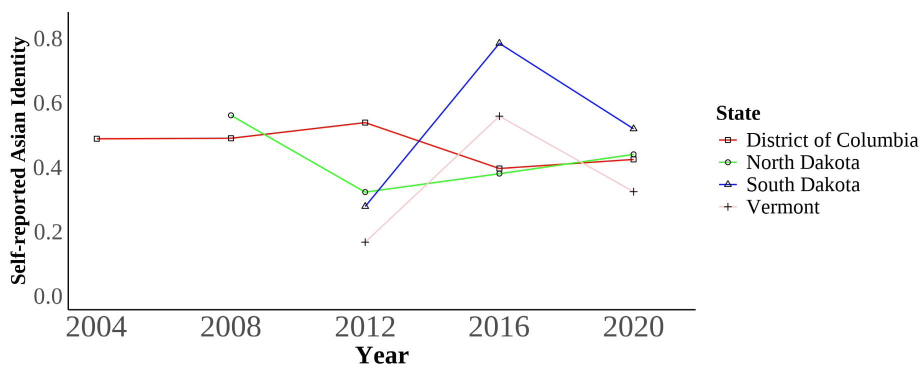
[2] Data source is the 2004-2021 Current Population Survey.

Figure 1: Bias and Self-reported Asian Identity in the Least and Most Biased Places

(a) Bias Over Time



(b) Self-reported Asian Identity Over Time



These two panels show the trends in bias (panel a) and self-reported Asian identity among Asian 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 Asian identity is among first, second, and third-generation Asian immigrants aged 17 and younger still living in intact families.
Bias data is from the 2004-2021 Harvard's Project Implicit Association Test scores, American National Election Studies (ANES), and state-level hate crimes against Asians. Identity data is from the 2004-2021 Current Population Survey (CPS).

Figure 2: Diagram of the Three Different Generations of Asian Immigrants.

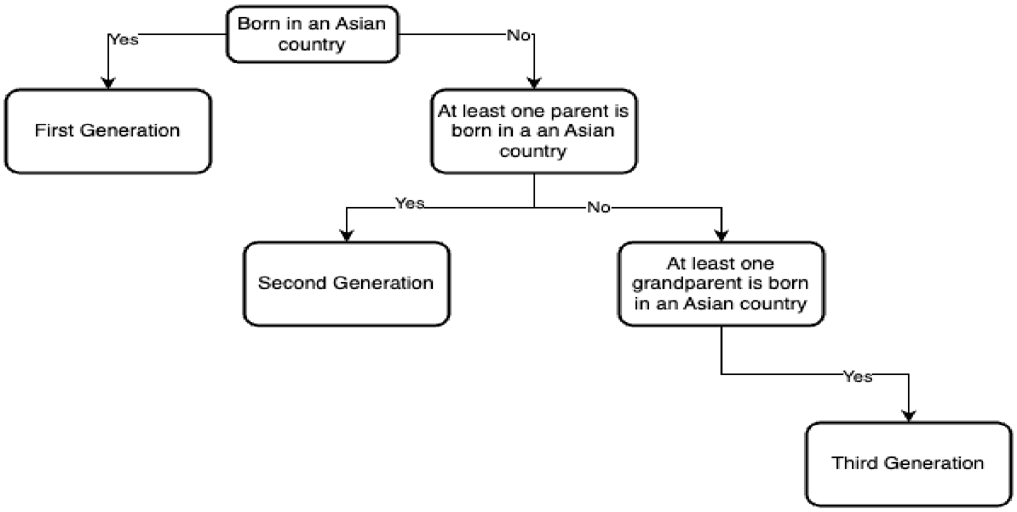
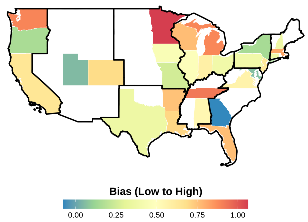
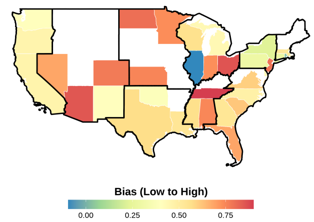


Figure 3: Maps of State-level Association Test Bias Over Time Measure with Census Division Regional Boundaries

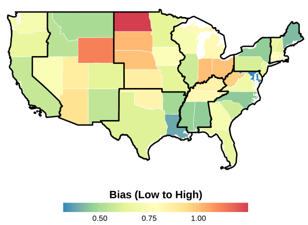
(a) State-level Bias in 2004



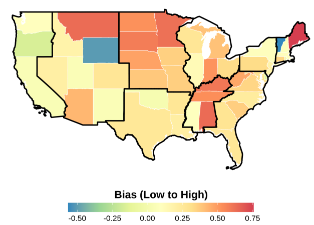
(b) State-level Bias in 2008



(c) State-level Bias in 2012

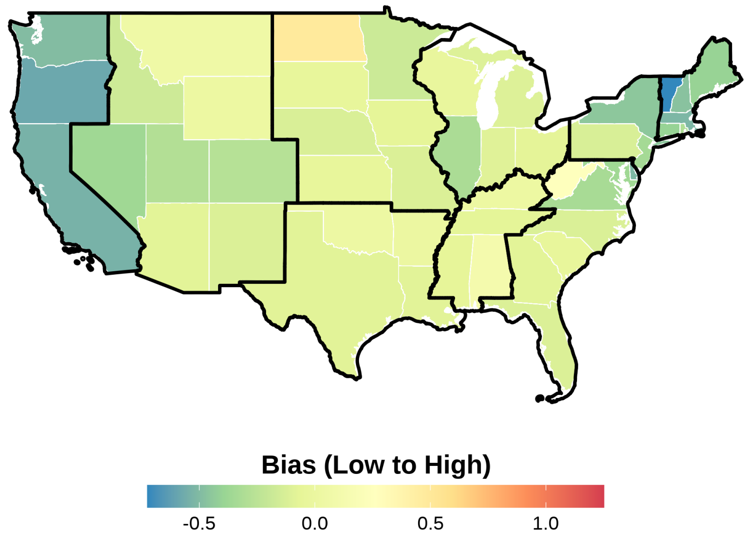


(d) State-level Bias in 2016



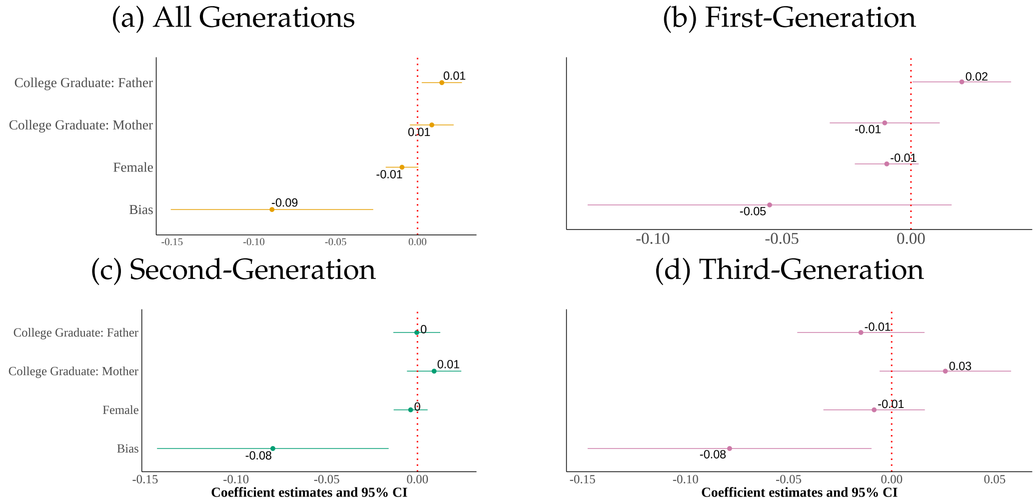
This figure shows the state-level bias index in different years in the sample. The bias units are in standard deviations and ranges from low to high bias. Bias index is constructed following Lubotsky and Wittenberg (2006). The data is from the 2004-2021 Harvard's Project Implicit Association Test scores, American National Election Studies (ANES), and state-level hate crimes against Asians. Each panel presents state-level bias during a certain year. The boundaries in black represent the different Census divisions in the United States. Notice how there is a variation across states within a region.

Figure 4: Maps of State-level Bias 2004-2021 Measure with Census Division Regional Boundaries



This figure shows the state-level bias index in the sample from 2004 to 2021. The bias units are in standard deviations and ranges from low to high bias. Bias index is constructed following Lubotsky and Wittenberg (2006). The data is from the 2004-2021 Harvard's Project Implicit Association Test scores, American National Election Studies (ANES), and state-level hate crimes against Asians. The boundaries in black represent the different Census divisions in the United States. Notice how there is a variation across states within a region.

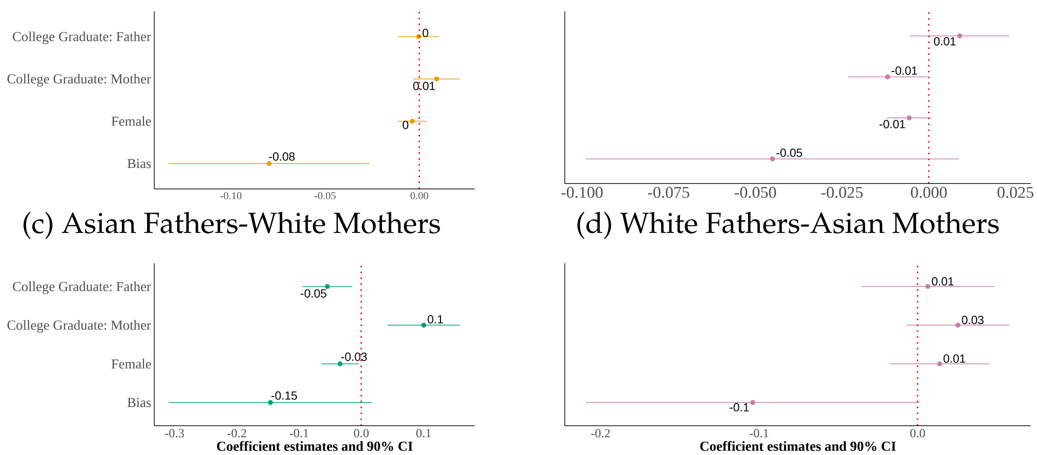
Figure 5: Relationship Between Self-Reported Asian Identity and Bias: By Generation



I show four panels of estimating equation (4). I include region $\times$ year fixed effects with controls for sex, quartic age, and parental education. The dependent variable is self-reported Asian 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 Asian children ages 17 and below who live in intact families. First-generation Asian immigrants are children that were born in a Asian country. Native-born second-generation Asian immigrants are children with at least one parent born in a Asian country. Finally, native-born third-generation Asian immigrants are children with native-born parents and at least one grandparent born in a Asian country.

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

(a) Second-Generation (All Parental Types) (b) Asian Fathers-Asian Mothers



I show four panels of estimating equation (4). I include region × year fixed effects with controls for sex, quartic age, and parental education. The dependent variable is self-reported Asian 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 Asian children ages 17 and below who live in intact families. Native-born second-generation Asian immigrant children with at least one parent born in a Spanish-speaking country.

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

	Number of Asian Grandparents			
	(1) One	(2) Two	(3) Three	(4) Four
Bias	-0.01	-0.09	-0.69**	-0.11
	(0.04)	(0.08)	(0.32)	(0.06)
Female	-0.01	-0.01	-0.04	-0.03**
	(0.01)	(0.02)	(0.06)	(0.01)
College Graduate: Mother	0.01	0.07**	0.08	0.00
	(0.01)	(0.03)	(0.09)	(0.03)
College Graduate: Father	-0.04***	0.00	-0.07	0.00
	(0.01)	(0.04)	(0.08)	(0.01)
Observations	14,453	12,678	567	5,744
Year $\times$ Region FE	X	X	X	X

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

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

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

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

Table 4: Relationship Between Bias and Interracial Marriages

		Asian Men	Asian Women
	(1)	(2)	(3)
	Interracial	Interracial	Interracial
Bias	0.04***	−0.01	0.03**
	(0.01)	(0.01)	(0.01)
College Graduate: Wife	0.04***	0.04***	0.05***
	(0.00)	(0.01)	(0.00)
College Graduate: Husband	−0.01*	−0.01	−0.02***
	(0.00)	(0.01)	(0.00)
Observations	69,800	52,103	60,214
Year × Region FE	X	X	X

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

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

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

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

Table 5: Relationship Between Bias and Migration

	(1) Migrated from Birth Place	(2) Migrated from Birth Place	(3) $\mathrm{Bias}_{ist} - \mathrm{Bias}_{ilb}$
Bias_{st}	0.13*		
	(0.07)		
Bias_{lb}		-0.03	
		(0.17)	
Asian			0.02
			(0.04)
Female	0.00	-0.01	0.00
	(0.00)	(0.00)	(0.02)
College Graduate: Mother	0.01***	0.00	-0.01
	(0.00)	(0.01)	(0.03)
College Graduate: Father	-0.03***	-0.03***	0.03
	(0.01)	(0.01)	(0.02)
Observations	73,563	41,641	2,075
Mean	0.15	0.15	-0.1
Year × Region FE	X		
Birthyear × Birth Region FE		X	

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

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

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

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

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

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

Parents Type	All	Asian-Asian	Asian-White	White-Asian
Proxy:				
Mother	0.72	0.97	0.37	0.3
Father	0.72	0.97	0.39	0.29
Self	0.87	0.97	0.23	0.31
Others	0.88	0.96	0.6	0.54

[1] The samples include children ages 17 and below who live in intact families. A proxy is the person that answered the Current Population Survey questionnaire.

[2] Data source is the 1994-2021 Current Population Survey.

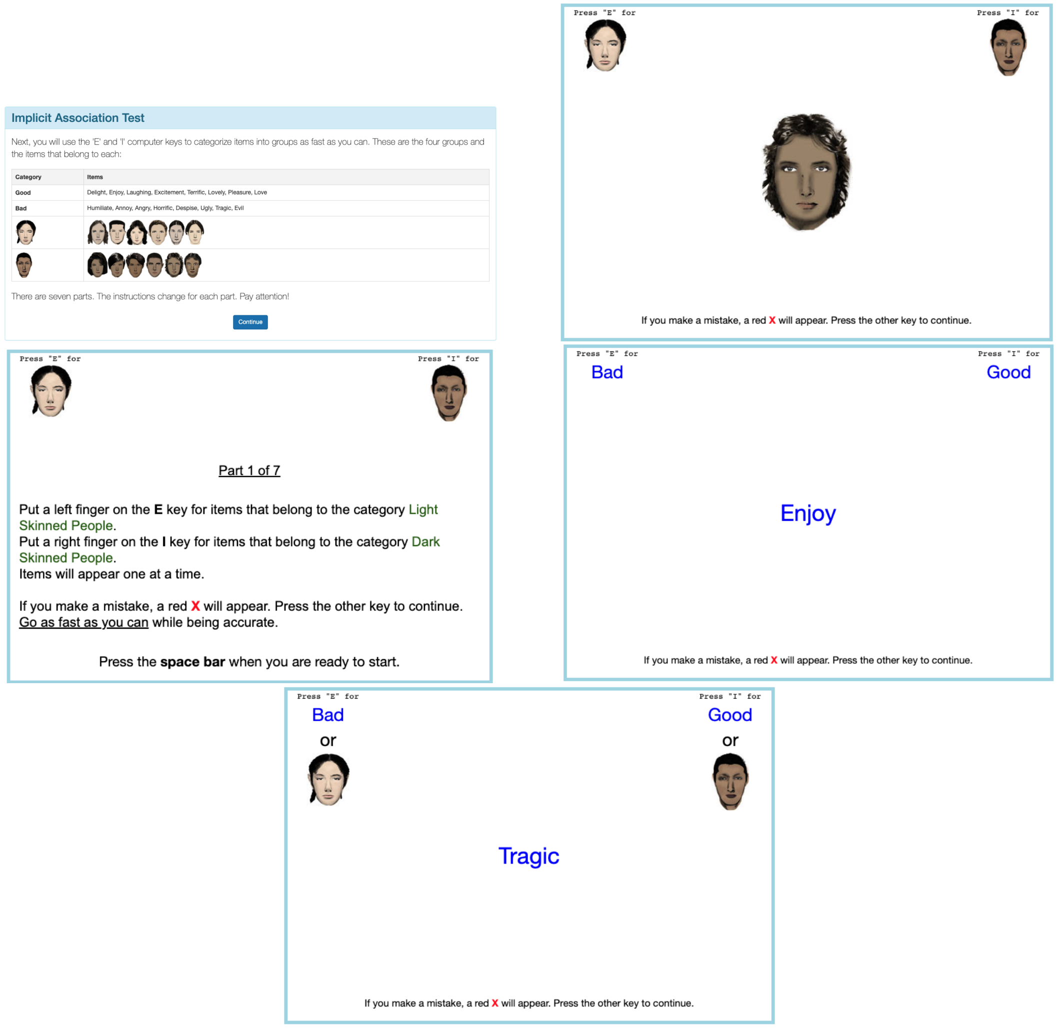
ONLINE APPENDIX

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

Hussain Hadah

A Data

Figure A.1: Examples of an Implicit Association Test



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

B Tables

Table A.1: Subjective Asian Identity and Asian Bias

	(1) A_i	(2) A_i	(3) A_i	(4) A_i	(5) A_i	(6) A_i	(7) A_i	(8) A_i
Bias	−0.04***	−0.14***	−0.02***	−0.02	−0.03***	−0.07**	−0.10***	−0.04
	(0.01)	(0.04)	(0.01)	(0.03)	(0.01)	(0.03)	(0.03)	(0.03)
Female	−0.01**	−0.01**	−0.01**	−0.01*	−0.01**	−0.01*	−0.01*	−0.01*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
College Graduate: Mother	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
College Graduate: Father	0.01**	0.01*	0.01**	0.01**	0.01**	0.01**	0.01**	0.01**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Both parents Asian	0.63***	0.63***	0.62***	0.62***	0.63***	0.63***	0.63***	0.62***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
First Gen	0.01	0.02	0.02*	0.02**	0.02*	0.02**	0.02*	0.02**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Second Gen	0.02	0.02	0.01	0.02	0.02	0.02	0.02	0.02
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
N	129078	129078	129078	129078	129078	129078	129078	129078
Region FE					X	X		
Year FE		X		X		X		
State FE			X	X				X
Year-Region FE							X	X

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

[1] I include controls for sex, quartic age, and parental education.

[2] Standard errors are clustered on the state level.

Table A.2: Relationship Between Bias and Self-Reported Asian Identity: By Generation

	(1) A_i	(2) A_i^1	(3) A_i^2	(4) A_i^3
Bias	-0.09***	-0.05	-0.08**	-0.08**
	(0.03)	(0.04)	(0.03)	(0.03)
Female	-0.01*	-0.01	0.00	-0.01
	(0.00)	(0.01)	(0.00)	(0.01)
College Graduate: Mother	0.01	-0.01	0.01	0.03
	(0.01)	(0.01)	(0.01)	(0.02)
College Graduate: Father	0.01**	0.02**	0.00	-0.01
	(0.01)	(0.01)	(0.01)	(0.02)
Observations	129,078	15,499	80,137	33,442
Mean	0.65	0.44	0.22	0.66
Year × Region FE	X	X	X	X

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

[1] Each column is an estimation of a heterogeneous effect of regression (4) by generation with region × year fixed effects. I include controls for sex, quartic age, fraction of Asians in a state, and parental education. I also added parents' (AA, AW, and WA) and grandparents' (AAAA, AAAW, AAWA, etc.) type dummy variables to the regression on second and third generation immigrants, where A is objectively Asian (born in a Asian country) and W is objectively White (native-born). Standard errors are clustered on the state level.

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

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

Table A.3: Relationship Between Bias and Self-Reported Asian identity Among Second-Generation Asian Immigrants: By Parental Type

Parents Type	All	Both Parents from Asian Country (AA)	Father from Asian Country (AW)	Mother from Asian Country (WA)
	(1) A^2	(2) A^2	(3) A^2	(4) A^2
Bias	-0.08**	-0.05	-0.15	-0.10
	(0.03)	(0.03)	(0.10)	(0.06)
Female	0.00	-0.01	-0.03*	0.01
	(0.00)	(0.00)	(0.02)	(0.02)
College Graduate: Mother	0.01	-0.01*	0.10***	0.03
	(0.01)	(0.01)	(0.03)	(0.02)
College Graduate: Father	0.00	0.01	-0.05**	0.01
	(0.01)	(0.01)	(0.02)	(0.03)
Observations	80,137	50,835	9,055	20,247
Year × Region FE	X	X	X	X
Mean	0.73	0.97	0.39	0.3

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

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

[2] The samples include second-generation Asian children ages 17 and below who live in intact families. Native-born second-generation Asian immigrant children with at least one parent born in a Asian country.

[3] Column (1) includes the results to regression (4) on all second-generation immigrants, column (2) includes the results to regression (4) on second-generation immigrants that who has a father and mother that were born in a Asian country (AA), column (3) includes the results to regression (4) on second-generation immigrants that who has a father that was born in a Asian country and a native-born mother (AW), and column (4) includes the results to regression (4) on second-generation immigrants that who has a native-born father and a mother that was born in a Asian country (WA).

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

Table A.4: Logistic Regression Analysis of Bias and Interracial Marriages

		Asian Men	Asian Women
	(1) Interracial	(2) Interracial	(3) Interracial
Bias	1.47***	0.83	1.39**
	(0.16)	(0.14)	(0.20)
College Graduate: Wife	1.42***	1.55***	1.75***
	(0.06)	(0.09)	(0.09)
College Graduate: Husband	0.94	0.97	0.86***
	(0.04)	(0.06)	(0.04)
Observations	69,800	52,032	60,171
Year × Region FE	X	X	X

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

[1] This is the result to estimating (5) as a logistic regression. The coefficients are exponentiated, thus should be interpreted as odds ratios.

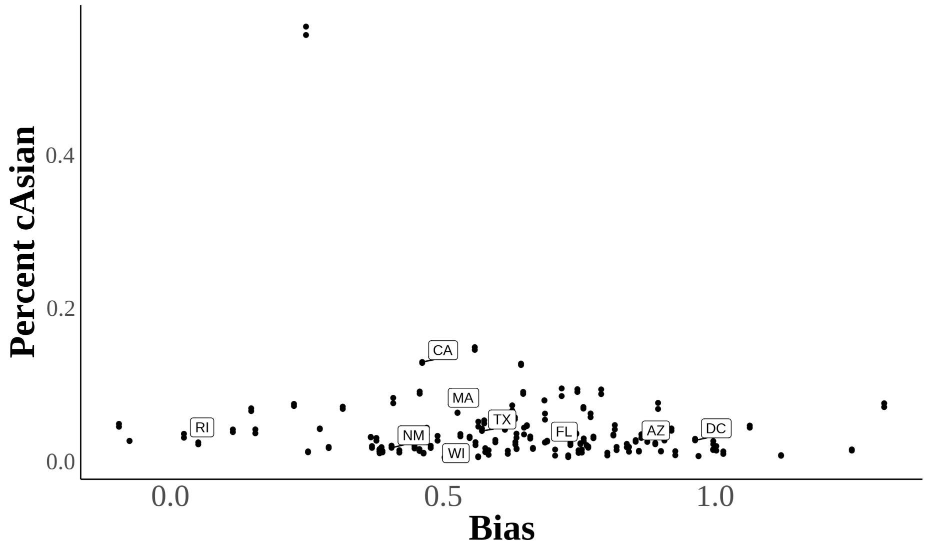
[2] I include controls for partners' sex, age, education, and years since immigrating to the United States. Standard errors are clustered on the household level.

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

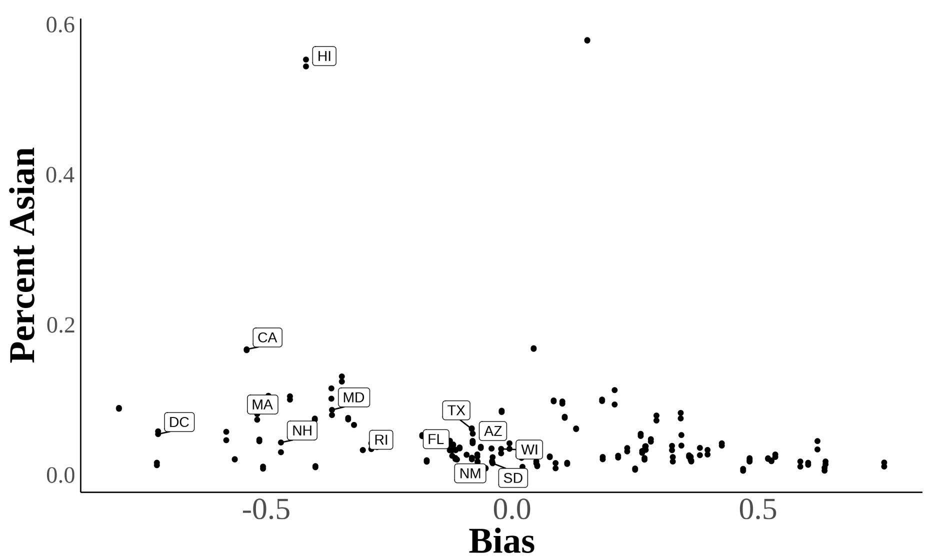
C Figures

Figure A.2: Scatter Plot of Proportion Subjectively Asian on Bias

(a) Year < 2015



(b) Year $\geqslant$ 2015



Here are two scatter plots showing the relationship between bias and subjective Asian population in a state. Each dot represents a state in a certain year. Percent subjectively Asian = $\frac{\#\text{Asian}}{\text{Population}}$
Source. 2004-2021 Current Population Survey.

C.1 Using Lubotsky and Wittenberg (2006) to Construct Bias Index

In Lubotsky and Wittenberg (2006), the authors propose a method to reduce measurement error in proxies by constructing a composite index. The Lubotsky-Wittenberg (henceforth LW) consider a model where a covariate is unobserved. Therefore, they use two proxies in its place, which will have measurement error. Thus, the LW method allows researchers to use two proxies that are error-ridden.

LW consider a setup with the following model:

$$\begin{aligned} y &= \alpha + \beta x^* + \epsilon \\ x_1 &= x^* + \mu_1 \\ x_2 &= x^* + \mu_2 \end{aligned}$$

Where x_i^* is the unobserved covariate, x_{1i} and x_{2i} are the proxies, and the measurement errors μ_1 and μ_2 are assumed to be classical and allowed to covary. The covariance matrix of the errors is given by:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$$

Replacing the unobserved x^* with x_1 or x_2 yields the following expectations of the OLS estimates:

$$\mathbb{E}\left[\hat{\beta}_1\right] = \beta \frac{\sigma_{x^*}^2}{\sigma_{x^*}^2 + \sigma_1^2} \quad ; \quad \mathbb{E}\left[\hat{\beta}_2\right] = \beta \frac{\sigma_{x^*}^2}{\sigma_{x^*}^2 + \sigma_2^2}$$

Both estimates are biased; the one with the smaller variance of the measurement error being less biased.

LW then propose defining a new proxy x_3 as a weighted average of x_1 and x_2:

$$x_3 = \lambda x_1 + (1-\lambda) x_2$$

To minimize the attenuation bias in the OLS estimate of β, they solve for the optimal value of λ:

$$\lambda^* = \frac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}}$$

This optimal value of λ is not directly useful because the variances of the measurement errors and their covariance are unobserved. However, if you estimate a bivariate regression using OLS (i.e., regress y on x_1 and x_2), then the expectation of the sum of the two coefficient estimates is identical to the expectation of the OLS coefficient estimate on x_3 in a univariate regression using the optimal choice of λ:

$$\mathbb{E}\left[\hat{\beta}_1 + \hat{\beta}_2\right] = \mathbb{E}\left[\hat{\beta}_{x_3}\right]$$

Thus, OLS produces an estimate of β with the least bias by optimally combining the information in x_1 and x_2.