

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

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.

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 ancestry to construct an ostensibly objective measure of identity. I find self-reported identity to be negatively correlated with individual and parental

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 (Frieze, Bluemke, and Wänke 2007), and health (Leitner et al. 2016).

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 An 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

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)

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 associated 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).

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

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 :

$$U_i = U_i(\mathbf{a}_i, \mathbf{a}_{-i}, I_i) \quad (1)$$

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:

$$I_i = I_i(\mathbf{a}_i, \mathbf{a}_{-i}; \mathbf{B}_{e_i}) \quad (2)$$

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

A person i chooses action \mathbf{a}_i that maximizes their utility function given racial 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.):

5. Akerlof and Kranton (2000) refer to \mathbf{B}_{e_i} as proscription.

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

Whose solution a_i^* yields utility U_i^* . Now, suppose a person can choose their racial 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 race.

That is i will change identities when the benefits of doing so $\tilde{U}_i^* - U_i^*$ 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 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) from the 2004-2021, 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.

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

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

7. I restrict first-generation immigrants whose parents were born in an Asian 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.

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 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 from the 2004-2021. 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.

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

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.

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

10. Research showed that the IAT tests are correlated with economic outcomes (Chetty et al. 2020; Glover, Pallais, and Pariente 2017), voting behavior (Frieze, 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).

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.

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 hate 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).

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."

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

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 :

$$A_{ist}^g = \beta_1^g \text{Bias}_{st} + \beta_2^g \text{DadCollegeGrad}_{ist} + \beta_3^g \text{MomCollegeGrad}_{ist} + \beta_4^g \text{Women}_{ist} + X_{ist}^g \pi + \gamma_{rt} + \varepsilon_{ist}; \text{ where } g \in \{1, 2, 3\} \quad (4)$$

Where A_{ist}^g 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 , $\text{DadCollegeGrad}_{ist}$, and $\text{MomCollegeGrad}_{ist}$ 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_1^g coefficient summarizes individual's i responsiveness to bias changes in the state which they live.

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.

In other words, β_1^g 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_1^g 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_1^g will be equal to zero.

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.

I report the main results of estimating equation (4) in Figure (5). I present the results of estimating the main specification for all generations in panel (A) and for sub-samples of first-, second-, and third-generations 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 iden-

tity. 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.

6 Robustness Checks and Discussions

In this section, I explore the empirical relationship between bias and interracial marriages, as well as the migration patterns of second-generation Asian immigrants as robustness checks to my main analysis and the effect of proxy response on my results. 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:

$$\text{interracial}_{ist}^2 = \beta_1^2 \text{Bias}_{st} + X_{ist}^2 \pi + \gamma_{rt} + \varepsilon_{ist} \quad (5)$$

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

I present the results of estimating equation (5) in Table (4). I find that a one standard deviation increase in bias increases the probability of having 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 An Asian husband marrying a White

wife. A one standard deviation increase in bias is associated a 3 percentage points increase in the chances of An 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.

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:

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

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

Where $\text{BirthPlaceMigration}_{ist}^2$ is an indicator variable equal to one if person i in state s at the interview t lives in a state that is different from his or her birth state and zero otherwise. $\text{BirthPlaceMigration}_{ilb}^2$ is an indicator variable that is equal to one if person i in birthplace l does not currently live in the same state he

or she lived in at the year of birth b and zero otherwise. The analysis, restricted to second-generation Asian immigrants with both parents born in An 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 $\text{BirthPlaceMigration}_{ist}^2$. 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 $\text{BirthPlaceMigration}_{ilb}^2$.

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:

$$Y_{ist} = \beta_0 + \beta_1^2 \text{Asian}_{ist} + X_{ist}^2 \pi + \varepsilon_{ist} \quad (8)$$

Where $Y_{ist} \equiv \text{Bias}_{ist} - \text{Bias}_{ilb}$, Bias_{ist} is i 's bias in state s at the time of interview t , and 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 An 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 Hispanic 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.¹⁵

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 (Frieze, 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

15. 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.

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.

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.