Upward Mobility and Discrimination: The Case of Asian-Americans

Nathaniel G. Hilger*
Brown University and NBER

June 2016

Asian-Americans are the only non-white US racial group to experience long-term, institutional discrimination and subsequently exhibit high income. I re-examine this puzzle. I focus on California, where most Asians settled historically. First, high upward mobility of Asians stemmed primarily from gains in earnings conditional on education, as opposed to high educational attainment conditional on SES. Second, test score data from 1943 suggest Asians only faced taste-based discrimination, whereas blacks faced statistical discrimination. Results support the argument that competitive labor markets eliminate purely prejudice-based earnings gaps. National, market-level correlations of black-white test score and earnings gaps in 1940 reinforce this interpretation.

JEL: J62, J15, J7, J31

*I thank Anna Aizer, Sam Asher, Nathaniel Baum-Snow, Patrick Bayer, Leah Boustan, Kenneth Chay, Pedro Dal Bo, Eleanore Dillon, Andrew Foster, Oded Galor, Claudia Goldin, Rachel Friedberg, Justine Hastings, Brian Knight, Bao Lam, John Logan, Glenn Loury, Bhashkar Mazumder, Joana Naritomi, Emily Oster, Jesse Shapiro, Matthew Turner, David Weil and Danny Yagan, as well as seminar participants at Brown, Duke, Connecticut, Wesleyan and the "Exposure, Heritability and Opportunity" workshop at Universidad Carlos III de Madrid. Jenna Anders, Kian Ivey, and Jorge Perez provided excellent research assistance. Weiwei Zhang provided helpful assistance with the 100% 1940 data. Financial support from the Russell Sage Foundation (Grant No. 83-16-07) is gratefully acknowledged. All errors are my own.

Few white Americans today realize just how pervasive legal anti-Asian discrimination was before 1945.... In light of this history, the current problems of the Asian-American community seem relatively minor, and its success appears even more remarkable. Social scientists wonder just how this success was possible, and how Asian-Americans have managed to avoid the 'second-class citizenship' that has trapped so many blacks and Hispanics.

David Bell, *The New Republic*, July 15, 1985

1 Introduction

Asian-American history represents a unique and puzzling case study because Asians are the only American racial group to experience long-term, institutional discrimination, yet still achieve group income levels similar to whites by the late 1960s.[1] In this paper I re-examine and provide a new explanation for this puzzle. I make use of new data and new methods to compare intergenerational progress of Asians, blacks and whites in historical data. The new data include 100% 1940 census data (Minnesota Population Center and Ancestry.com, 2013), and Army General Classification Test (AGCT) score data for over 500,000 WWII enlistees in 1943 (Ferrie et al., 2012; Aaronson and Mazumder, 2011; Carruthers and Wanamaker, 2016).[2] New methods allow estimation of intergenerational mobility statistics on small groups in census data (Hilger, 2016). I focus my analysis on California (CA), which contained over 80% of Asians in 1940, and also contained a small black minority that had voluntarily migrated from the South in pursuit of economic opportunity. I then address four main questions.

Question 1: Does high Asian income reflect high dynastic income growth, or compositional effects of new immigration? To my knowledge even this basic

[1] While many other "white" immigrant groups such as Irish-Americans, Italian-Americans and Jewish-Americans encountered some prejudice historically and exhibit high incomes in the modern period, they did not experience the qualitative degree of institutional discrimination reserved for "non-white" groups including blacks, Native Americans, Asians and Hispanics, and described below in Section 2.1 in more detail (Jensen, 2002; Kenny, 2006; Diner, 2006; Mangione, 1993; Chang, 2004; Page, 2004; Gonzalez, 2011). For example, Kenny (2006) states "The Irish experience of race in the United States does not belong in the same category as black slavery or Asian exclusion," while Diner (2006) states "As women and men considered among the privileged by virtue of their whiteness, [Jews] enjoyed relative tolerance," and that they experienced "relatively full political and civil rights" from the end of the 18th century. While Hispanic Americans have faced substantial institutional discrimination (e.g. Gonzalez, 2011), I do not focus on them in this paper for several reasons. First, I cannot observe Hispanic Americans in my test score data described below. Second, IPUMS identifies Hispanic Americans in relatively complex ways related to nationality, language, and names that may be endogenous to some of the outcomes I study.

[2] I thank Bhashkar Mazumder for generously sharing his cleaned version of the WWII enlistee test score data.

question about Asian-American history has not been addressed in prior literature. Using pseudo-panels by year, race, and birth in CA, I identify parental income when children are age 1-17 in a base year (when most children still live with parents) and track incomes of these children in later years once they enter the labor and marriage markets. Under assumptions that I partially verify in the data, these intergenerational pseudo-panels yield two-generation group dynastic growth rates on balanced panels of dynasties.[3] I find that Asian dynastic growth rates exhibit upward "divergence" from blacks and upward "reversals of fortune" with respect to whites, and are therefore qualitatively inconsistent with neoclassical absolute convergence of groups to identical steady state incomes from different initial conditions (Ramsey, 1928; Solow, 1956). In this sense, Asians do exhibit unusually rapid dynastic growth relative to blacks and whites in every cohort born in CA since 1920. These high growth rates would have delivered high Asian incomes even in the absence of new high-skilled Asian immigration.

Question 2: Why did Asian dynastic income grow more rapidly than other groups? To shed light on this question I estimate an intergenerational decomposition of group earnings in each year into three terms: (1) parental income distributions, (2) children's educational attainment conditional on parental income, and (3) children's earnings conditional on education. I exploit the method developed in Hilger (2016) to estimate educational attainment conditional on parental income in cross-sectional census data. I find that all three components favor Asians over blacks historically, although these differences shrink dramatically when restricting to the CA-born. I develop a method to quantify the relative importance of these three components by imputing counterfactual, steady-state black-white earnings gaps for all children born 1920-1980, assigning blacks each of the three components of Asian and white earnings separately. Contrary to popular perception, large gains in earnings *conditional* on education have played the primary role in Asian earnings growth, alongside a secondary role for greater educational attainment conditional on parental income, and virtually no role for higher parental income.[4] The only *white* advantage over blacks in CA has been earnings conditional on education; greater educational attainment and parental income have played virtually no role.

Question 3: Why were CA-born Asians but not blacks able to close their

[3] I place no restrictions on where children born in CA live later in life.

[4] The belief that minorities get ahead by investing in more education, rather than obtaining greater pay conditional on education, is widespread. Kristof (2015) focuses on high educational attainment of Asian children, which he partially attributes to "East Asia's long Confucian emphasis on education." President John F. Kennedy implicitly adopted this theory of group progress in 1963 when he told an assembly of black civil rights leaders, "it seems to me...that we could emphasize... which I think the Jewish community has done, on educating their children, on making them study, making them stay in school and all the rest" (Branch, 1989).

conditional earnings gap? To shed light on this question, I examine determinants of conditional earnings gaps in 1940 by exploiting AGCT test scores for whites, blacks and Asians in CA. I find that Asians in 1943 already exhibit near-parity with whites in mean test scores both overall and within all education groups, while analogous black mean test scores lag behind both Asians and whites by nearly a full standard deviation, as has been found in more recent decades (Neal and Johnson, 1996; Johnson and Neal, 1998; Dickens and Flynn, 2006; Neal, 2006; Fryer, 2010). I quantify the contribution of these test score gaps to conditional earnings gaps by matching test score records to the 100% 1940 census to obtain a matched, national sample of 211,000 records containing test scores, earnings, and education for white and black men ages 18-35. Replicating specifications in Johnson and Neal (1998) I find that black-white skill gaps account for 40% of black-white earnings gaps in 1940, which is only slightly less than the 50% share of black-white earnings gaps accounted for by AFQT scores in NLSY data for the 1990s. As of 1940, these findings suggest a relatively larger role for *taste-based* or some other non-statistical discrimination in the Asian conditional earnings gap (Becker, 1957), and for *statistical* discrimination stemming from skill deficits in the black earnings gap (Aigner and Cain, 1977). Like most modern researchers, I interpret historical black skill gaps as a legacy of slavery and severe educational discrimination, a legacy not shared by Asians or whites.

These findings bear on the question of whether taste-based and other forms of non-statistical discrimination, by themselves, can generate persistent group earnings gaps in competitive labor markets (Arrow, 1972; Goldberg, 1982). To my knowledge, Asian-American history offers the most direct empirical evidence on this question to date, because Asians are the only persecuted non-white American minority to display test score parity–and hence plausibly skill parity–with whites in historical data. Asian-American history therefore provides an important case study supporting the notion that earnings gaps driven *entirely* by prejudice are not sustainable in competitive labor markets, although only after 1940 once severe institutional discrimination subsided. These results raise a fourth question.

Question 4: If taste-based discrimination no longer reduced Asian earnings by the 1960s, why might it continue to reduce black earnings? Prior research has documented large black-white pay gaps that are only partly explained by test scores and educational attainment in recent decades (Neal and Johnson, 1996; Johnson and Neal, 1998; Fryer, 2010), and shown these residual gaps plausibly stem from employer prejudice (Charles and Guryan, 2008).[5] I discuss several theories—based on stereotypes

[5]Estimates in Charles and Guryan (2008) suggest that a large share of the negative black wage residual not explained by AFQT scores or education in Neal and Johnson (1996) may stem from employer prej-

(Bordalo et al., 2016), pay compression (e.g., Frank, 1984; Acemoglu and Pischke, 1999), labor market tightness (Biddle and Hamermesh, 2013; Baert et al., 2015), and prevalence of minority employers—in which group skill levels affect prejudice-driven components of earnings gaps as well as those components driven by unbiased statistical discrimination. These theories all share in common a prediction that *market-level* group skills should have larger effects on group earnings than individual skills due to social multiplier effects. I present evidence consistent with this prediction using non-experimental variation in black-white test score and earnings gaps across labor markets defined by state and broad education categories. These findings reinforce the key lesson of Asian-American history as interpreted here: large group earnings gaps appear hard to maintain in competitive labor markets without large group skill gaps.

Prior researchers have suggested many *qualitative* explanations for high Asian incomes in the modern period including selective migration, intra-group spillovers, positive stereotypes, demographic imbalance, portability of human capital, and culture, among others (e.g., see arguments and literature reviews in Hirschman and Wong, 1986; Sue and Okazaki, 2009; Lee and Zhou, 2015). Chiswick (1983) documents advantageous labor market outcomes among Asian-American men in the 1970 census and concludes that discrimination need not always result in worse labor market outcomes, but makes no distinction between statistical and non-statistical discrimination. Borjas (1992, 1993, 1994) explores determinants of multi-generational convergence in education and income among ethnic groups, but due to data constraints these studies place greater empirical emphasis on "white" European immigrants who were spared most forms of institutional discrimination; do not examine test score variation; and do not adequately impose geographic overlap across groups within the US. Darity Jr. et al. (1997) use Oaxaca-Blinder decompositions to study impacts of observable ethnic group characteristics on occupational indices in national samples and anticipate some of the findings here. I build on this earlier work by exploiting test score data, 100% census data, intergenerational pseudo-panels, and intergenerational decompositions of group earnings estimated using new estimation methods (Hilger, 2016).

udice. Many studies in recent decades document persistence of racial prejudice against blacks (e.g., Bertrand and Mullainathan, 2004; Pager et al., 2009) and surveys also suggest persistent prejudice against Asians (Committee of 100, 2001; The Gallup Organization, 2005).

2 Background on Asian-Americans

2.1 Historical Discrimination Against Asian-Americans

Unlike blacks who only arrived in the U.S. involuntarily after being kidnapped and subjected to many traumas including starvation, torture, and rape (e.g., Rediker, 2008), Asians typically migrated to the U.S. voluntarily in search of economic opportunity as prospectors, laborers, merchants, skilled professionals and students.[6] Therefore the experiences of the dynasties composing blacks and Asians in the U.S. for many generations preceding the era studied here were profoundly and incomparably different. In addition to escaping any burdens associated with slavery, the vast majority of Asian (and black) children in CA also escaped the worst aspects of educational discrimination imposed on black children in the South during the late 19th and early 20th centuries (Hendrick, 1975; Wollenberg, 1995; Graaf et al., 2001). However, in many other respects Asians faced severe institutional discrimination that shared much in common with the southern (and in many ways national) Jim Crow laws and practices encumbering blacks well into the 20th century. I here review some key historical elements of this discriminatory regime (for reviews, see Daniels, 1990; Chan, 1991; Sandmeyer, 1991; Takaki, 1998; Chang, 2004; Azuma, 2005).

Foreign-born Asians were barred from citizenship and hence voting by the Naturalization Act of 1790. Asians experienced mob violence including lynchings and purges throughout the late 19th century (Pfaelzer, 2008), and hostility from anti-Asian clubs much like the Ku Klux Klan (the Asiatic Exclusion League, Chinese Exclusion League, Workingmen's Party of CA). Asians could not testify against a white witness in court (*People v. Hall*, 1853), effectively barring Asians from legal protections against white aggression. The Chinese Exclusion Act of 1882 and the "Gentlemen's Agreement" in 1907 codified racial hostility against Asians by ending immigration of laborers from China and Japan. Subsequently, Chinese Americans suffered arbitrary deportation and effective nullification of citizenship due to restrictions on habeas corpus petitions (Scott Act 1888, *United States v. Ju Toy* 1905). The Geary Act of 1892 required all Chinese Americans to register with the federal government and display residency certificates upon penalty of deportation. The U.S. Supreme Court considered withdrawing eligbility of *native-born* Asian-Americans for citizenship, but declined to do so (*Wong Kim Ark v. U.S.* 1898).

Many cities and states levied discriminatory taxes and fees on Asians (1852 For-

[6] Native Americans, as well as many Hispanics living on land taken by the U.S. during the Mexican-American War, also "entered" the U.S. involuntarily under circumstances of violence and material expropriation.

eign Miner's Tax, 1852 Commutation Tax, 1860 Fishing License, 1862 Police Tax, 1870 "queue" ordinance, 1870 sidewalk ordinance, and many others). CA's new constitution in 1879 prohibited all corporations and governments (state, county and municipal) from hiring Asians. Most professional schools and associations in CA excluded Asians, as did most labor unions, and many companies declined to hire Asians well into the 20th century. Many companies that did hire Asians–such as the Central Pacific Railroad Company in the 1870s–paid them lower wages and barred them from managerial positions. From 1913-23, virtually all western states passed increasingly strict Alien Land Acts that prohibited Asians from owning land or leasing land for extended periods, stifling rapid growth of Asian agricultural entrepreneurship and likely pushing many Asians into small family businesses for lack of better options. Restrictive covenants against Asians pervaded housing markets and restricted neighborhood choice. Asians also faced laws against marriage to whites (1905 amendment to Section 60 of the CA Civil Code) and U.S. citizens (Expatriation Act 1907, Cable Act 1922). Over 100,000 Chinese American immigrants experienced extended, prison-like confinement to Angel Island during immigration proceedings from 1910-1940, and the US forcibly relocated over 100,000 mainland Japanese Americans (but not German or Italian Americans) to detention camps during WWII from 1942-1946, in practice liquidating or destroying much of their wealth in the process.

Researchers generally agree that institutional discrimination against minorities in CA weakened rapidly in the years following WWII, well before the Civil Rights Act of 1964. These changes came in the wake of President Roosevelt's famous Executive Order 8802 prohibiting discrimination by race among government agencies and their contractors in defense industries after 1941. For example, the U.S. Supreme Court struck down a law requiring U.S. citizenship to obtain a commercial fishing license in 1948 (*Takahashi v. Fish and Game Commission*), and CA repealed all remaining Alien Land Laws in 1956.

2.2 California: The One Shared Place

Figure I maps the number of Asian and black children (age 0-18) living in every county in the continental U.S. using the 1940 100% decennial census (Ruggles et al., 2015). Historically, Asian-Americans have lived primarily in two states: Hawaii and CA. I do not focus on Hawaii in in this paper for several reasons. Hawaii was not included in the census in 1940-50; almost no blacks have ever lived in Hawaii; and Asians in Hawaii probably did not experience the same kinds of institutional discrimination as mainland

Asians.[7] I therefore focus on group members born in CA ("CA-born") throughout much of the analysis. In 1940, CA had ten times more native-born Asians than any other state, and still had four times more native-born Asians than any other state as of 1980. CA is the only state with sufficient numbers of Asians to estimate intergenerational relationships precisely in historical census data.

Fortunately, Figure I indicates that CA also contained a significant minority of black children over the 1940-2000 period. The comparison between Asians and blacks living in CA is fortuitous for several reasons. Like Asian-American families, black families in CA represent a small group of voluntary migrants from distant and culturally foreign locations in pursuit of economic opportunities (Graaf et al., 2001).[8] This selection process potentially controls for some unobservable characteristics of Asian immigrants; below I document that Asians and blacks born in CA have much more similar observed characteristics than Asians and blacks nationally. Restricting to CA-born groups also partly controls for institutional factors that may account for dramatic variation in upward mobility across places in the the US (Chetty et al., 2014; Chetty and Hendren, 2015; Hilger, 2016). For example, CA ranked high among states in public school quality in the early 20th century (Ayres, 1920), and most Asians and blacks in CA attended racially integrated schools by the late 19th century, although after WWII restrictive covenants yielded *de facto* racially segregated schools in many urban districts (Hendrick, 1975; Wollenberg, 1995; Graaf et al., 2001).[9] A final advantage of restricting to CA-born groups is that both blacks and Asians historically represent small population shares in CA. Margo (1990), as well as Card and Krueger (1992b), document that larger black population shares across states and counties in the South were historically associated with greater segregation and school quality disparities for clear reasons related to school budgeting.[10]

As I document further below, mid-20th century CA can be viewed as a place where

[7] Hawaii did develop a two-tiered educational system that favored whites starting in 1924 with the advent of "English Standard Schools" (Hughes, 1993). However, the Hawaiian Department of Public Instruction claimed to adopt the policy that only English language examinations–not race–would determine entry into these schools. English Standard Schools indeed appear to have been heavily racially integrated at the time of a 1948 report (Meller, 1948), and the Department of Public Instruction claimed that per-student expenditures were strictly equal across types of schools.

[8] The parents in these black families had migrated to CA primarily from Texas and Louisiana, and to a lesser extent Arkansas, Mississippi and Alabama. Most blacks residing in CA had migrated to CA from other states until 2000, when a majority of black adults were CA-born for the first time.

[9] California was considered an extremely attractive destination for blacks throughout the early 20th century due to its relatively open and tolerant institutional environment (Graaf et al., 2001).

[10] For example, in 1920 class sizes at white and black schools were nearly identical in states with black population shares under 10%, while class sizes were twice as large in black schools in states with black population shares over 50% (Card and Krueger, 1992b, Figure II). For comparison, in 1920 CA was 3.4% Asian, and the two counties in CA with the largest population of Asians (Los Angeles and San Francisco) were only about 4% Asian, with similarly small shares of blacks.

small, advantageously selected subsets of blacks and Asians migrated in pursuit of similar economic opportunities and faced relatively similar institutional and social environments. The key difference is that blacks, unlike Asians, brought with them to CA long dynastic histories of slavery, educational and labor market discrimination, and persecuted minority status. Comparisons of CA-born Asians and blacks therefore offer a unique opportunity to shed light on the relative importance of these dynastic legacies, as opposed to contemporary institutional discrimination, in determining group income trajectories.

3 Data

The decennial census is the only data set large enough and extending back far enough in time to conduct detailed historical comparisons of Asians with other groups.[11] I rely on census data from 1940-2000, when income and education are both available (Ruggles et al., 2015). Critically, I rely on newly a digitized 100% sample 1940 census data, making it possible to examine minorities in CA in these early years, and to match census data with test score data. I rely on census data spanning the longer period of 1850-2000 in order to examine longer-term aggregate group trends. I define "Asians" broadly as Chinese, Japanese, and "Other Asian or Pacific Islander"; almost all Asians up through 1970 were Chinese or Japanese. Asians have been identified in the census race variable through "enumerator observation" (1850-1950) and self-reporting (1960-2000) in every year back to 1850.

I focus on household annual labor earnings (head + spouse) as my primary measure of income for several reasons: non-labor income is not available in the 1940 census, hourly wages not suffer from measurement error in reported hours[12] (Baum-Snow and Neal, 2009); both earnings of head and hourly wages do not capture total resources available for investments on children's education; and household wages allow pooling of male and female children on a comparable footing in order to maximize sample size. I exclude parents reporting zero income from my primary analyses. In Section 9, I show key results are robust to imputing incomes for these households, and also robust to use of male earnings rather than total household earnings. In Hilger (2016), I address several additional concerns related to the adjustment of statistics for independent children that cannot be linked directly to parents after ages of school completion.

[11] The Consumer Population Survey is another large, long-standing survey. For this paper, census data are preferable to CPS data because the March CPS begins in 1962, only introduces "Asian/Pacific Islander" to its racial classification in 1988, and excludes military and incarcerated individuals from its sample.

[12] Neal (2006) imputes hours from CPS data. The CPS does not separately identify Asians in its race variable before 1988 and is too small to provide useful imputations for Asians in later years.

I make extensive use of recently-discovered World War II enlistment data containing Army General Classification Test (AGCT) scores for a large sample of enlistees in 1943 (Ferrie et al., 2012). These data are discussed in Ferrie et al. (2012); Aaronson and Mazumder (2011); Carruthers and Wanamaker (2016). The AGCT was intended to measure "ability to learn" in the army environment (not innate intelligence) and contained 140-150 multiple-choice questions on vocabulary, arithmetic, and block counting. The test was shown to correlate strongly with IQ scores, to display high reliability and validity, and to strongly predict in-service and post-service occupations. Some authors claim that enlistment in the army was conditional on a minimum literacy standard, suggesting the tests would not have been racially biased for literacy reasons, but other sources paint a more complex picture of testing practices for illiterate and non-English-speaking enlistees (Bingham, 1946). The large size of these data allow me to conduct three novel empirical exercises. First, I separately examine test scores of Chinese-American, black and white enlistees born in CA in the 1920s, during a period when many Jim Crow laws were still in effect.[13] Second, I match these test score data to 100% individual census data in 1940 to assess cross-sectional effects of individual test scores on earnings. Third, I use test score and earnings data separately on full unmatched samples to compare black-white test score and earnings gaps across labor markets defined by state of residence and broad education category.

For simplicity, I match the AGCT and census data on exact state of birth, race, first name and last name, and year of birth plus or minus one year.[14] I obtain a match rate of nearly 40%, which is high by the standards of census matching, most likely due to the short time interval between the two datasets. Summary statistics for men in the 1940 census sample, the 1943 enlistment records sample, and the matched data, for both US and CA residents are presented in Table II. The table indicates that the test score data are reasonably representative of the US and CA population as contained in census data. Chinese American men ages 18-38 represent 0.8% of all CA residents in both the 1940 census and the test score data, consistent with overall composition of WWII servicemen (Smith, 1947). Blacks are over-represented in the test score data among CA residents, but under-represented at the national level as among servicemen generally (Smith, 1947). Some of this discrepancy between US and CA samples may reflect rapid migration of blacks to CA between 1940 and 1943 (Graaf et al., 2001). Overall, the test

[13] Japanese Americans are almost entirely unrepresented among WWII enlistees due to the Japanese Internment policy.

[14] I drop individuals who would be under age 23 in 1940 to assure that most individuals are no longer in school, and individuals with reported education under 5 years, which is the 2nd percentile of education in this year.

score data contain about 3% of men ages 18-38 in the 1940 census, both for the US and CA. Table II also indicates that the matched sample is a fairly representative sample of the AGCT data. Test scores are slightly higher in the matched sample, although standard deviations are similar. Age, education, and race are also similar in the matched sample, with the exception that I match almost no Chinese-Americans, most likely due to difficulties matching Chinese names. I therefore only examine Asians in unmatched test score data.

4 Basic Historical Trends

In this section I plot national aggregate outcomes by race over all available years of data, reweighted to match the white age and gender distribution in each year and restricting to ages 25-65. I focus on national trends with no further restrictions, rather than trends restricting to those born in the US or in CA, because national trends may have informed broader perceptions of Asian-American history. I also incorporate Native Americans in this section as a second "involuntary immigrant" group that has also been subjected to multiple centuries of institutional discrimination in human capital and labor markets (Page, 2004).

These comparisons furnish broad historical context, but they confound intergenerational group mobility with compositional changes from migration (Borjas, 1987; Chen, 2011; Suzuki, 1995, 2002). Figure II illustrates the magnitude of this problem by plotting gross immigration flows into the U.S. from various Asian countries since 1820. The Chinese Exclusion Act, the Gentlemen's Agreement, and the 1965 Immigration Act are all discernable. On the right-hand axis, the dashed line labeled "Share" plots total Asian immigrant flows as a share of the total Asian population stock in the US in the previous decade, and indicates that migration flows were large relative to stocks even before the 1965 Immigration Act, and enormous thereafter. I am not aware of comparable data on return-migration flows which also affect the composition of Asians between censuses (Suzuki, 1995). Below I develop comparisons that address these problems.

Figure IIIa plots literacy rates by race and year.[15] Asians had much higher literacy rates than blacks and Native Americans in 1870, but this gap had closed by 1900. Figure IIIb plots average educational attainment by race and year. In every year 1940-2000, Asians exhibit significantly higher education than all other groups, followed by whites, followed by blacks and Native Americans.

[15] Only free blacks were asked about literacy by census enumerators through 1860, and only "taxed" Native Americans were asked about literacy up through 1870. I therefore drop these observations from the figure due to concerns about selection. A fire destroyed the 1890 census.

Figure IV plots log earnings of men 1940-2000. All three minorities reduce their earnings gaps with whites after 1940. Earnings of blacks and Native Americans relative to Asians have fluctuated around 60-80% for 60 years. I also impute log earnings in earlier periods based on occupation.[16] Figure IVb displays imputed log earnings of men from 1880-2000. All minorities exhibit convergence toward whites over most of the last 150 years. However, blacks and Native Americans do not display any clear convergence toward Asians.

Table I presents the most common occupation for each race by year, and sketches out the different occupational trajectories underlying Figure IVb. In 1860-80, many Asians were "mine operatives and laborers." Later Asians tended to work on farms as wage laborers, rather than tenants, again reflecting differences in regional economies. Asians then worked in restaurants, laundries and other service industries before shifting into more white-collar positions in late 20th century. In comparison, blacks and Native Americans tended to work on farms as tenants after the Civil War, and have remained in lower-skilled agriculture and manufacturing throughout the 20th century.

These aggregate national trends characterize Asians as higher-skilled than native U.S. minorities at every point over the past 150 years. The perception that Asians began their history in the U.S. disadvantaged by lower skills and earnings could potentially be true relative to whites, but is less plausible when comparing Asians to blacks and Native Americans at the national level. In contrast, trends in human capital and earnings among the CA-born are more consistent with this perception, as shown in Appendix Figures A.1-A.2. Among the CA-born, Asians appear highly disadvantaged in 1940, but rapidly overtake both blacks and whites in education and earnings over subsequent decades. Unfortunately, these figures also confound effects of intergenerational group mobility with time-varying selective migration, now to an even greater extent due to inter-state migration of blacks and whites in addition to international migration of Asians.

5 Intergenerational Group Mobility: Pseudo-Panels

I now present historical outcome trends that isolate variation in group intergenerational mobility, excluding changes in group composition due to migration. To do so I construct pseudo-panels that link adult outcomes to parental characteristics during childhood,

[16] I follow Smith (1984) and Margo (1990) and impute earnings back to 1860 based on earnings in occupations in 1940, allowing earnings to differ by native-born status and restricting to men ages 25-65. I do not allow earnings to differ by race within occupations in this imputation. By fixing earnings within occupation the imputation provides a simple index of occupational quality. To harmonize occupations across years I rely on the IPUMS variable OCC1950.

exploiting the fact that most children live with their parents until age 17. Consider children age 1-17 in 1940 with known state of birth in the US. For this 17-year cohort block of children we can observe parental characteristics such as income and education. We can then observe outcomes of these children at ten-year intervals in later censuses using self-reported state of birth. This strategy delivers balanced pseudo-panels if individuals report race, age, and place of birth consistently across decades, and if families with U.S.-born children rarely emigrate. Pseudo-panels permit calculation of "group mobility" as $\frac{\mathbb{E}[y_{i,r,t}]}{\mathbb{E}[y_{i,r,t-1}]}$, where $y_{i,r,t}$ denotes household earnings of person i in group r in generation t. I construct these pseudo-panels for 17-year cohort blocks still living with parents in each decade 1940-2000.[17]

I can partially test the assumptions required to obtain valid pseudo-panels by testing for anomalous changes in the size and gender composition of cohorts defined by place of birth.[18] Figures V-VI plot log frequencies for these cohort-blocks born in the US and CA, respectively. In a truly balanced pseudo-panel, cohort size weakly declines over time due to death and out-migration. While this restriction is approximately satisfied in most cases, cohort size does increase between some censuses for some races. These violations may reflect inconsistencies in census sampling techniques or individuals' self-reported age, place of birth or race. However, the violations are typically small in comparison to the massive changes in Asian population size and composition displayed in Figure II.[19]

I also assess the validity of these pseudo-panels by tracking their gender composition over time. Earlier cohorts of Asians reporting U.S. birth exhibit "excess" males due to mass falsification of U.S. nativity records by largely male Asian migrants after the destruction of immigration records in the 1906 San Francisco earthquake, and possibly due to widespread incentives to avoid restrictions on foreign-born Asians through false nativity papers (Bureau of the Census, 1914; Chang, 2004). If this "paper sons" phenomenon somehow continued into later cohorts we would expect to see excess men or excess volatility of gender ratios. Appendix Figures A.3 and A.4 plot the share of men in each of these cohorts and races for native-born and CA-born cohorts, respectively. The figures document a male share very close to 50%, falling slightly as cohorts age, which is exactly the pattern that would arise from valid pseudo-cohorts due to the greater longevity of women. While there are some anomalies for certain cohorts in certain years, the selected

[17] I pool all children ages 1-17 to maximize statistical power, weighting families by number of children in household. I omit cohorts age 1-17 in 1950 because earnings and education in 1950 are only observed for one member of each household.

[18] These are imperfect tests because consistent cohort size may conceal changes in the composition of the cohort.

[19] I have experimented with reweighting the pseudo-cohorts to maintain a fixed age distribution over time, with no significant change in the results.

pseudo-cohorts appear reasonable based on both stable frequencies and gender balance over time.

These pseudo-panels yield one parental income observation for each cohort, and one child earnings observation for each cohort-year after children have reached ages of labor market entry. This combination of multiple observations on every cohort serves as a further test of internal consistency. Figure VII illustrates how I present these data to compare dynastic growth rates across groups parsimoniously. The figure plots parental household earnings ratios with respect to Asians on the X-axis, and children's household earnings ratios with respect to Asians on the Y-axis. The 45-degree line represents the benchmark of identical earnings ratios across generations, and divides the pseudo-panel estimates into evidence distinguishing two broad families of models. In the "neo-classical absolute convergence" region of this figure, Asian IM can be rationalized by inter-group mean-reversion of groups with identical preferences and technologies from different initial conditions (Ramsey, 1928; Solow, 1956). In the "Divergence" region, relatively rich Asian parents have children who are relatively even richer, or poor Asian parents have children who surpass previously richer groups. Points in this region suggest that Asians are converging to higher steady-state income levels than comparison groups.[20] Differential steady states across groups are consistent with (1) "conditional convergence" models with group variation in preferences and technologies (Barro and Sala-i Martin, 1992), and (2) "new growth" models with identical groups under departures from neo-classical assumptions. Leading examples of non-neoclassical growth models include human capital externalities (Azariadis and Drazen, 1990; Borjas, 1992), knowledge spillovers (Romer, 1986), and credit constraints (Galor and Zeira, 1993).

Figure VIII presents this figure with data for black, white and Asian cohorts age 1-17 in 1940, 1960, 1970 and 1980. Panel (a) restricts to children born in the U.S., and panel (b) restricts to children born in CA. Nationally, Asian cohorts overtake whites, and do not exhibit any significant convergence toward poorer blacks. Results for children born in CA strongly reject neo-classical absolute convergence with respect to both blacks and whites in every cohort born in CA since 1920. These results suggest that Asian dynasties raising children in CA either benefit from more advantageous preferences or technologies than other groups, or benefit from some growth externality or non-convexity that violates the assumptions of the neo-classical growth model. I now turn to understanding what factors might account for this unusually rapid dynastic earnings growth among Asians. Given the extraordinary pace of this growth, explanatory factors should involve

[20] The upper unlabeled region of the graph would suggest that Asians are diverging to *lower* steady-state income, and is never empirically relevant.

parametric differences affecting group steady state incomes, or group-level externalities or non-convexities that depart from the neoclassical model.

6 Intergenerational Earnings Decomposition

Why have Asian dynasties in CA been converging toward higher group earnings than blacks and whites? To shed light on this question I estimate a simple, intergenerational decomposition of group earnings in the spirit of Conlisk (1974). Let $y_{r,t}$ indicate average adult log earnings in group r in generation t, and $h_{r,t}$ indicate average adult education. Let $f_r(y_{t-1})$ indicate the probability density function of parental income in group r. Mean earnings of group r in generation t can then be written non-parametrically as:

$$\mathbb{E}[y_{r,t}] = \int_{y_{t-1}} y_{r,t}(h_{r,t}(y_{t-1}))\, f_{r,t}(y_{t-1})\, dy_{t-1}. \tag{1}$$

This decomposition breaks mean group earnings into three terms. The term $f_{r,t}(y_{t-1})$ captures a group's parental income distribution and can be thought of as resource "endowments." The term $h_{r,t}(y_{t-1})$ captures educational attainment conditional on parental income. This "investment" relation can vary across races due to many factors including, for example, discrimination in human capital markets, anticipated discrimination in labor markets, information and beliefs about the value of education, and parental preferences. This relation could also be highly nonlinear if families with low incomes face sharp liquidity constraints on educational investments. The term $y_{r,t}(h_{r,t})$ captures children's earnings conditional on education. These "earnings functions" can differ across races due to factors such as school quality, labor market discrimination, or family skills not captured by educational attainment.

After examining these terms non-parametrically, I also make use of a linearized version of this decomposition. Write educational investments as $h_{f,t}(y_{r,t-1}) = \theta_{r,t} + \gamma_{r,t} y_{r,t-1}$ and adult earnings functions as $y_{r,t}(h_{r,t}) = \alpha_{r,t} + \beta_{r,t} h_{r,t}$, implying

$$\mathbb{E}[y_{r,t}] = \alpha_{r,t} + \beta_{r,t}\theta_{r,t} + \beta_{r,t}\gamma_{r,t} E[y_{r,t-1}] \tag{2}$$

as well as the steady state relation

$$y_{r,SS} = \frac{\alpha_r + \beta_r \theta_r}{1 - \beta_r \gamma_r}. \tag{3}$$

These decompositions allow me to state how group outcome gaps would mechanically be affected by replacing each of these three components for one group with the corre-

sponding component of another group in a given generation. For example, I can estimate the share of the black-white earnings gap that would be closed over one generation, or in steady-state, if blacks adopted Asian investment behavior $h_{asian,t-1}(y_{t-1})$ or white parental income $f_{white,t-1}(y_{t-1})$. These counterfactuals provide a simple way to quantify the "importance" of group differences in three broad components of group mean earnings.[21]

I estimate these three components in each year for whites, blacks and Asians born in CA.[22] Parental income distributions and children's earnings conditional on schooling can be estimated directly in census data. I rely on the method developed in **?** to estimate children's final education conditional on parental income, which addresses the longstanding problem that many children can no longer be linked to their parents at ages of school completion (e.g., Cameron and Heckman, 1993). The key assumption required to make this adjustment, verified in detail in **?**, is that dependent and independent children in their mid-to-late 20s exhibit similar relationships between final schooling and parental income. While this assumption cannot be verified directly for Asians due to small samples in panel datasets, it appears to be a reasonable approximation for whites, blacks, men, women, and all time periods spanning 1940-2000.

Even as a purely descriptive, reduced-form exercise, this decomposition has many limitations. Some of the more important examples are that two-generation mobility statistics likely overstate multi-generational mobility (e.g., Clark, 2014; Olivetti et al., 2014; Stuhler, 2014; Braun and Stuhler, 2015; Solon, 2015); final educational attainment is a highly imperfect measure of human capital as I discuss in more detail below; and annual group earnings variation likely understates *lifetime* group earnings variation due to reversion toward different group means (Rothstein and Wozny, 2014). Nonetheless, the exercise provides a useful diagnostic exercise for assessing the most likely potential causes of variation in group dynastic income growth rates.

[21] It is also straightforward to solve for the transition path of the linear decoposition for any generation T as

$$\mathbb{E}[y_T] = (\alpha + \beta\theta)\sum_{j=1}^{T}(\beta\gamma)^{j-1} + (\beta\gamma)^T \mathbb{E}[y_0]. \tag{4}$$

[22] In practice it is important to estimate these relationships over bounded regions of income and education variables, because the linearity assumptions break down outside the main support. I therefore drop children with education below the bottom 2% of the population education distribution in each year as in Card and Krueger (1992a), and I use mean log of parental income within population parental income deciles in each year, bounding income at the mean of the top and bottom deciles. Therefore in practice I estimate $\mathbb{E}[y_{r,t}] = \alpha_{r,t} + \beta_{r,t}\theta_{r,t} + \beta_{r,t}\gamma_{r,t}E[y_{r,t-1}] - \beta_r\gamma_r y_{\min} - \beta_r h_{\min}$ and $y_{r,SS} = \frac{\alpha_r+\beta_r\theta_r-\beta_r\gamma_r y_{\min}-\beta_r h_{\min}}{1-\beta_r\gamma_r}$, where $y_{\min}$ and $h_{\min}$ are as described.

6.1 Estimates

Figure IX plots parental income distributions by race in 1940 for children born in the US and CA. Parental incomes are grouped into population deciles. Panel (a) shows that Asian and black children in 1940 grew up with extremely different parental income distributions at the national level, with black children concentrated in the bottom deciles. Panel (b) shows this contrast is much milder when restricting to black and Asian children born in CA, likely due to positive selection of black CA-born children's parents into voluntary long-distance migration.

Figure X plots educational investments conditional on parental income, and again illustrates the key role of geography in Asian-American history. Panel (a) shows that, among all native-born in 1940, Asian dynasties invest in higher levels of children's schooling than whites, whites invest in higher levels of schooling than blacks, and these patterns are especially pronounced among lower-income families. However, panel (b) shows that as of 1940 these differences completely vanish when restricting to CA-born. Among the CA-born, all races display high and income-insensitive educational investment relative to to national trends. Panel (c) shows that CA-born Asians do exhibit a higher investment schedule than other CA-born groups in later years, while white and black educational investment schedules remain virtually identical. Given that these relations are approximately linear, Table III presents estimated intercepts and slopes of linear investment schedules for the CA-born in order to summarize these time trends parsimoniously.[23] As the figure suggests, Asians do not display any advantage in 1940, but display significantly higher investment schedules than both whites and blacks in later years.

Finally, I estimate group earnings conditional on educational attainment. Figure XI displays log earnings for men by educational attainment in 1940 and 1980, restricting to the CA-born. In 1940, Asians and blacks both received about $0.4 - 0.6$ log points lower pay than whites at every level of education. By 1980, Asians had closed this gap entirely while blacks had only made significant progress at higher education levels, and even there continued to lag behind. Table III presents intercepts and slopes of linearized conditional earnings functions by race and year in Columns (7)-(12), and show that Asians caught up to white earnings levels by 1970. The slopes of these lines, i.e. the Mincerian return to schooling, rise rapidly after 1980 for all groups, especially for blacks.[24]

[23] There is insufficient data on CA-born blacks and Asians to fit lines to these curves with any precision in 1960, barely enough in 1970, and the curves cannot be estimated at all in 1950 due to collection of census data from only one member of each household.

[24] Non-parametric earnings function reveal that all three races exhibit sharp "convexification" of the returns to schooling after 1970 (Lemieux, 2006; Heckman et al., 2006).

6.2 Counterfactual Black-White Earnings Gaps

All three elements of the decomposition—parental income, educational investments, and earnings conditional on education—favor Asians relative to blacks historically in CA. In order to assess the relative contribution of these three components I construct counterfactual estimates of the black-white earnings gap over time. I consider counterfactuals in which I assign to black dynasties each of these three components from whites and Asians separately, still restricting to the CA-born.

I permute Asian components to whites and blacks using linear estimates of intercepts and slopes reported in Table III in two ways. Table IV reports actual and counterfactual log earnings by race based on Equation (2), which takes only one generation of transmission into account. Table V reports actual and counterfactual log earnings by race based on steady-state income in Equation (3), which takes all future generations of transmission into account. The results are easiest to understand in terms of impacts on counterfactual black-white earnings gaps. Figure XIII displays one-generation counterfactual black-white earnings gaps constructed from estimates in Table IV as well as similar estimates permuting white components to blacks rather than Asian components. Results for steady-state estimates are nearly identical. Panel (a) shows that the overwhelmingly most important black disadvantage relative to *whites* is lower earnings conditional on education. Lower parental income and differentla propensities to invest in education out of parental resources play almost no role. Panel (b) repeats this exercise but imputes Asian components to black dynasties. Once again, conditional earnings gaps are the most quantitatively important factor, though high educational attainment of Asian children conditional on parental income play a large secondary role.

7 Why Did Asians but not Blacks Close Conditional Earnings Gaps?

Estimates above indicate that gains in earnings conditional on education—rather than high educational attainment of Asian children conditional on parental income—play the largest role in accounting for rapid Asian earnings growth since 1940. Why did Asians, but not blacks, close their conditional earnings gaps with whites?

The discrimination literature distinguishes two broad explanations for group earnings gaps: "statistical" discrimination based on skills, and "taste-based" discrimination based on prejudice.[25] Prior work on more recent data has documented that skill gaps as re-

[25]Prejudice of employers, customers, or workers can all generate racial pay gaps under different sets of

flected in test scores emerge early in childhood (Fryer and Levitt, 2013) and account for a large share of black-white pay gaps (Neal and Johnson, 1996; Johnson and Neal, 1998; Fryer, 2010), but also suggests a significant role for taste-based labor market discrimination (Charles and Guryan, 2008). This distinction is important because only statistical discrimination maximizes profits. For this reason theory tends to predict that competitive labor markets should eliminate earnings gaps driven by taste-based or mistaken statistical discrimination, but not statistical discrimination based on genuine skill gaps (Arrow, 1972, though see Goldberg, 1982).

One reasonable conjecture, therefore, is that Asians closed their earnings gaps more quickly than blacks after 1940 because they faced primarily taste-based, rather than statistical, discrimination. To explore this possibility I make use of AGCT scores in WWII enlistment records from 1943. These data provide a measure of skills conditional on education and, remarkably, are large enough to compare Asians, blacks and whites living in CA.[26]

Figure XII plots the distribution of normalized test score residuals by race from an OLS regression of test z-scores on dummies for education and age. Asians and whites have similar conditional skill distributions, while the black skill distribution lags behind by nearly a full standard deviation. Table VI shows that this pattern holds separately within broad educational categories.[27]

Can black-white test score gaps in 1943 account for large share of conditional earnings gaps in 1940, as in the more recent period? To address this question I turn to the matched national sample of AGCT scores linked to earnings and educational attainment in census data, restricting to individuals over age 23 in 1940 to allow school completion and labor market entry. Table VII reports estimates from regressions of log earnings on a dummy for black and AGCT scores, replicating specifications in Johnson and Neal (1998) on data fifty years earlier in time. Column (1) documents a large 0.57 log point raw black-white earnings gap controlling for age. Column (2) indicates that AGCT scores reduce this gap by 40% to 0.37 log points, which is only slightly less than the 50% share of black-white earnings gaps accounted for by AFQT scores in the early 1990s in Johnson and Neal (1998). Column (2) also documents that returns to AGCT scores are

plausible assumptions.

[26] I restrict to residence rather than birth in CA due to the greater quality of the state of residence variable in the test score data. Results are similar using state of birth.

[27] Anecdotally, historical animus toward Asian-Americans often acknowledges their high skill levels. V.S. McClatchy, Secretary of the CA Joint Immigration Committee, expressed this common sentiment in a 1920 hearing: "It will be agreed...that the facts now before us conclusively establish that the Japanese are undesirable as immigrants and as citizens, not because they are of inferior race, but because they are superior in certain characteristics...." (Bell, 1935).

approximately linear in the full sample. Column (3) shows that educational attainment is likely an important channel by which AGCT scores increase earnings, but does not further reduce the unexplained share of the black-white earnings gap. Columns (4)-(7) show that, relative to whites, blacks exhibit lower returns to age and education and similar returns to AGCT scores.[28]

These results suggest that Asian-white conditional earnings gaps in the 1940s reflected taste-based or mistaken statistical discrimination, while earnings gaps of blacks to some extent reflected profit-maximizing statistical discrimination based on skill gaps. Many theories of discrimination therefore predict less persistence for Asian conditional earnings gaps than black conditional earnings gaps (e.g., Arrow, 1972). Indeed, I find that after 1940, as institutional labor market discrimination against Asians subsided, Asians rapidly closed their conditional earnings gaps. To my knowledge, Asian-American history as interpreted here represents the first direct empirical test of this prediction. Asians in 1940 provide a novel test of this prediction because their earnings disadvantage cannot be explained by education or test scores, unlike other groups previously studied in the discrimination literature, including blacks (Neal and Johnson, 1996; Johnson and Neal, 1998; Fryer, 2010; **?**), Mexican Americans (Trejo, 1997; Johnson and Neal, 1998), and in some respects women (e.g., Goldin, 2014; Kleven et al., 2015).

8 Can Other Groups Close Earnings Gaps by Closing Skill Gaps?

If Asians apparently eliminated earnings gaps based on non-statistical discrimination, why does some research suggest that taste-based discrimination still contributes to black-white earnings gaps (see reviews in Charles and Guryan, 2011; Lang and Lehmann, 2012)? Why would earnings gaps driven by taste-based discrimination persist for some groups but not others? This question is important to understand whether elimination of group skill gaps will eliminate group earnings gaps.

One explanation is that labor market prejudice against Asians in CA has always been weaker than labor market prejudice against blacks nationally.[29] While plausible, the

[28] Additional controls for hours and weeks worked in these regressions are highly significant but do not reduce the black-white earnings gap further in my sample. The exact mechanisms underlying black-white test score gaps remain poorly understood by social scientists, and do not seem easily explained by differences in parental socioeconomic status, neighborhoods, schools, or teachers (Neal, 2006; Fryer and Levitt, 2006; Fryer, 2010).

[29] Note I focus on *labor market* prejudice here; as discussed above, blacks nationally have definitely faced much greater *human capital market* prejudice than Asians in CA, both due to slavery and due to severe Jim Crow schooling policies that were never implemented as forcefully against blacks or Asians

severe labor market discrimination against Asians in CA discussed above, and survey evidence of continuing anti-Asian prejudice in the modern period (Committee of 100, 2001; The Gallup Organization, 2005), makes this explanation both hard to evaluate and not obviously compelling. I therefore explore an alternative explanation. As documented above, Asians display dramatically higher skill distributions than blacks as early as 1943. For many reasons, it is possible that high group skill levels would tend to alleviate effects of both statistical *and* taste-based or mistaken statistical discrimination on group earnings. For example:

- Racially prejudiced beliefs may represent exaggerated stereotypes causally dependent on a "kernal of truth" about lower-skilled groups (Bordalo et al., 2016).
- Employers may find it more costly to *categorically* exclude higher-skilled groups if firms earn greater profits on higher-skilled workers due to pay compression (e.g., Frank, 1984; Acemoglu and Pischke, 1999), or if markets for higher-skilled workers are tighter (Biddle and Hamermesh, 2013; Baert et al., 2015).[30]
- Higher-skilled groups may contain more managers and business owners who act as unprejudiced employers of their own group members.

Interestingly, all of these theories entail intra-group spillovers in which a group member's earnings depend on her group's *market-wide* or *aggregate* skill levels rather than just her own. A key implication of such theories is that variation in group skills at the market level should account for a greater share of group earnings than variation in skills at the individual level due to "social multiplier" effects. These theories therefore represent a clear departure from the neoclassical growth model.

This point can be formalized heuristically using a linear-in-means peer effects framework (e.g., Glaeser et al., 2003; Sacerdote, 2011). For log earnings y_{ijr} and test score x_{ijr} of individual i in market j and race $r \in \{w, b\}$, let log earnings be determined by the true relation

$$y_{ijr} = \alpha + \gamma_b 1\{r = b\} + \beta x_{ijr} + \delta \bar{x}_{jr,-i} + e_{ijr}, \tag{5}$$

where γ_b denotes the residual black-white log earnings gap, and $\bar{x}_{jr,-i} \equiv E[x_{ijr}|j, r, -i]$ or mean test scores of an individual's own-race peers within a market. I assume $e_{ijr} \perp$

in CA.

[30] For example, Bain (2000) argues that railroad employers were highly prejudiced against Asians in the late 19th century, but nonetheless rapidly hired large numbers of Asian workers after observing their high productivity relative to competiing workers in other racial and ethnic groups.

j, r to abstract from endogenous sorting. Consider the short linear predictor $y_{ijr} = \tilde{\alpha} + \tilde{\gamma}_b + \tilde{\beta} x_{ijr} + \tilde{e}_{ijr}$, as estimated above and in Neal and Johnson (1996); Johnson and Neal (1998) on national samples. If blacks tend to have lower-skilled peers such that $Cov\left(1\left\{r = b\right\}, \bar{x}_{jr,-i}\right) < 0$, then two results obtain. First, $\tilde{\gamma}_b < \gamma_b (\leq 0)$ implying the short regression overstates the share of black-white earnings gaps not accounted for by measured skills. Second, $\tilde{\beta} < \beta + \delta$, implying that the coefficient on skills in the short regression understates the impact of a mean-shift in group skills on group earnings due to social multipliers. Two approaches can recover γ_b and $\beta + \delta$. First, I can run a version of the short regression on data aggregated to the level of market and race: $\bar{y}_{jr} = \check{\alpha} + \check{\gamma}_b + \check{\beta} \bar{x}_{jr} + \check{e}_{jr}$, which yields $\check{\gamma} = \gamma$ and $\check{\beta} = \beta + \gamma$. This approach does not require microdata, and therefore allows me to use the full, unmatched versions of the test score and earnings data. Second, I can include $\bar{x}_{jr,-i}$ in the regression on matched microdata to approximate the true relation.[31]

I am unable to rigorously test for social multiplier effects due to a lack of exogenous variation in group skills, implying potential violations of the maintained assumption $e_{ijr} \perp j, r$. However, I can assess whether non-experimental correlations are consistent with an important role for social multipliers. I first divide the U.S. into "labor markets" defined by state and four broad education categories: no high school, some high school, high school degree, and any college. I then aggregate earnings by markets and race for men age 23-38 in the 1940 100% census, and WWII enlistment test z-scores based on the national score distribution for men ages 23-38. Figure XIV plots earnings gaps against test score gaps across all labor markets along with the estimated regression line. The implied relationship predicts very small black-white earnings gaps in hypothetical markets without black-white skill gaps, although this prediction is far out of sample. To my knowledge, even this cross-sectional correlation has not been documented previously due to data limitations.[32]

Table VIII presents fixed effects regressions of log earnings on a black dummy variable and AGCT scores in a variety of specifications. In Columns (1)-(3), I estimate a simplified version of the regression in Table VII in the matched microdata. Column (1) documents a slightly larger black-white earnings gap without age controls. Column (2) controls for market fixed effects (state × broad education group), and shows that black-white earnings gaps fall slightly. Column (3) once again shows that in this simplified specification, AGCT

[31] Borjas (1992) takes the first approach to test for evidence of social multipliers across many US ethnic groups, and interprets these multipliers as "ethnic capital."

[32] The NLSY is too small to estimate black-white test score gaps by state. NAEP data is large enough to estimate black-white test scores by state, but is only recorded in grades 4, 8 and 11 and therefore does not contain final educational attainment.

scores account for a slightly smaller share of black-white earnings gaps: about 25% rather than the 40% reported above. Column (4) adds mean peer test scores to the specification. Consistent with intra-group spilloverse and social multipliers, the black dummy declines and becomes insignificant, and the coefficient on peer scores is very large even conditional on own scores. Columns (5)-(7) replicate columns (1)-(3) on data aggregated to the level of race and market. Again consistent with an important role for intra-group spillovers, the coefficient on AGCT score nearly triples in the aggregate specification, and now accounts for over 60% of black-white earnings gaps. However, these specifications are limited by the smaller size and potentially less representative nature of the matched subsample, which only contains a subset of all potential markets due to missing data. Therefore, in columns (8)-(10) I once again estimate specifications from columns (1)-(3), but now on the full samples of test score and earnings data without restricting to the matched sample. These columns strengthen the findings from the matched sample: the coefficient on group-level AGCT in column (10) is now well over three times its microdata analogue in column (3) and accounts for 78% of black-white earnings gaps.[33]

Of course, the patterns documented in this section are descriptive and could be driven by unobserved variables rather than intra-group human capital spillovers. The findings could also be driven by measurement error in individual human capital, with market-level test scores simply proxying for individual skills (Borjas, 1992); this would be consistent with some recent work finding individual skills can explain a much larger share of black-white conditional wage gaps in the South in 1940 than previously thought (Carruthers and Wanamaker, 2016). Unfortunately, my data do not allow me to distinguish these theories from human capital spillovers.

9 Robustness

Throughout the analysis I have excluded households with zero and missing total earnings (head + spouse) from the analysis. Many of these zeros represent self-employed families with positive labor supply and business income, and many others likely represent measurement error or transitory earnings. An alternative approach is therefore to im-

[33]It is puzzling that group test score gaps can almost fully explain group earnings gaps for blacks in 1940, given that Asians exhibit similarly large earnings gaps but no skill gaps. This result seems to suggest that Asians faced *more* taste-based discrimination than blacks in 1940, which seems unlikely. However, recall that all main results in earlier sections restrict to Asians and blacks born in CA, while the results in this section examine variation across all states. In CA there is very little variation in black-white skill gaps across education groups, so it is not possible to identify the relationship between market-level group skills and market-level group earnings in CA by itself, and therefore not possible to assess whether it follows the national pattern.

pute positive household earnings for these households and include them in the analysis. I implement this by calculating average household earnings among all individuals with non-zero earnings in cells defined by year, race, age, sex, education, marital status, and state or country of birth. I assign cell means to individuals with zero and missing earnings based on this set of characteristics, roughly following Autor et al. (1998). In households with two earners I take the maximum of these two predictions. For households with zero earnings I follow Neal (2006) and adjust for selection by multiplying imputed household earnings by 0.6. While this method is somewhat *ad hoc*, it provides a rough check on whether households with zero and missing earnings are likely to be driving the main results. Using this new income variable I re-estimate the main results from pseudo-panels in Section 5 and counterfactual black-white earnings gaps in Section 6. Appendix Figures (A.5)-(A.6) document that the main results are virtually unchanged.

Focusing on household earnings may confound group variation in earnings with group variation in rates of single parenthood. To examine this I re-estimate group dynastic growth rates as in Section 5 using fathers' and sons' individual earnings rather than total household earnings. Appendix Figure A.7 documents that the results are virtually unchanged.

Duncan and Trejo (2016) document that 20% of second-generation Asian Americans in the 2000s report their race to census enumerators as "white," and that these Asians tend to have lower schooling than other Asians. However, I find that Asians exhibit unusually high dynastic growth rates in pseudo-panels for every cohort born in the CA after 1920, and that Asians experienced rapid declines in conditional earnings gaps by the late 1960s. While I cannot rule out a role for endogenous ethnic identification in these results, it seems likely that a much lower share of Asians would have identified themselves as white to census enumerators in these earlier decades. Note that intermarriage of Asians with whites increased dramatically in decades after 1970, suggesting widespread cultural assimilation occurred over the decades after the main results documented in this paper.

10 Conclusion

In this paper I address four main questions. By focusing on the CA-born, I obtain more meaningful comparisons between Asians, blacks and whites that control for broad institutional environment and selection of parents or grandparents into voluntary, long-distance migration in pursuit of economic gain.

Question 1: Does high Asian income reflect high dynastic income growth, or compositional effects of new immigration? I exploit pseudo-panels linking

parental income to future income of their CA-born children to distinguish dynastic group income growth from compositional effects of new high-skilled migration. I find that Asian dynasties exhibit rapid intergenerational growth consistent with higher steady-state group income than both whites and blacks for every cohort born in CA since 1920. Asian dynasties since 1940 do possess some growth advantage over both white and black dynasties born in CA.

Question 2: Why did Asian dynastic income grow more rapidly than other groups? In order to shed some light on what this growth advantage might be, I estimate a simple intergenerational decomposition of group earnings into parental income distributions, final educational attainment conditional on parental income, and children's earnings conditional on education. I quantify the importance of these three components by permuting them across groups to construct counterfactual black-white earnings gaps in the next generation and in dynastic steady state. Contrary to public perception, Asian dynastic income has grown faster primarily due to large increases in earnings *conditional* on education, with a large but secondary role for higher educational attainment conditional on parental income, and no role for higher parental income. The key feature of post-1940 Asian success is that Asians, unlike blacks, fully and rapidly eliminated their large conditional earnings gaps.

Question 3: Why were CA-born Asians but not blacks able to close their conditional earnings gap? Using a large and fairly representative sample of WWII enlistee test scores from 1943 both on their own and matched to the 1940 census, I document the striking fact that these test scores can account for a large share of the black, but not Asian, conditional earnings gap in 1940. This result suggests that Asians earnings gaps in 1940 stemmed primarily from *taste-based* or some other non-statistical discrimination, in sharp contrast with the black earnings gap which largely reflected statistical discrimination based on skill gaps inherited from centuries of slavery and educational exclusion. The rapid divergence of conditional earnings between CA-born Asians and blacks after 1940–when CA rapidly abandoned its most severe discriminatory laws–provides the first direct empirical evidence in support of the hypothesis of Arrow (1972) and others that competitive labor markets tend to eliminate earnings gaps based purely on taste-based but not statistical discrimination.

Question 4: If taste-based discrimination no longer reduced Asian earnings by the 1960s, why might it continue to reduce black earnings? Finally, I discuss several reasons why taste-based discrimination may persistently amplify group earnings gaps based on skill gaps, and point out that these theories tend to involve intra-group spillovers and hence predict larger impacts of market-wide skills than individual skills

on earnings. I test and strongly confirm this prediction on test score and earnings data in microdata in the matched sample, and on aggregate market-level data. Controlling for market-level black skills, rather than individual skills as in the standard empirical models of discrimination (Neal and Johnson, 1996; Johnson and Neal, 1998), results in much larger predicted impacts of skills and much smaller residual black-white earnings gaps.

Taken together, the results in this paper suggest that blacks and Native Americans face a much more difficult problem than anything ever faced by Asians. However, the results are optimistic in the sense that they suggest equality of group skills do historically tend to suffice for equality of group earnings, at least in the absence of severe institutional discrimination.

References

Aaronson, D. and B. Mazumder (2011): "The Impact of Rosenwald Schools on Black Achievement," *Journal of Political Economy*, 119, 821–888.

Acemoglu, D. and J.-S. Pischke (1999): "The Structure of Wages and Investment in General Training," *Journal of Political Economy*, 107, 539–572.

Aigner, D. J. and G. G. Cain (1977): "Statistical Theories of Discrimination in Labor Markets," *Industrial and Labor Relations Review*, 30, 175–187.

Arrow, K. (1972): "Some Mathematical Models of Race in the Labor Market," in *Racial Discrimination in Economic Life*, ed. by A. H. Pascal, Lexington, MA: Lexington Books.

Autor, D. H., L. F. Katz, and A. B. Krueger (1998): "Computing Inequality: Have Computers Changed the Labor Market?" *Quarterly Journal of Economics*, 113, 1169–1214.

Ayres, L. P. (1920): *An Index Number for State School Systems*, Department of education, Russell Sage foundation.

Azariadis, C. and A. Drazen (1990): "Threshold Externalities in Economic Development," *The Quarterly Journal of Economics*, 105, 501–526.

Azuma, E. (2005): *Between Two Empires: Race, History, and Transnationalism in Japanese America*, New York: Oxford University Press, 1 edition ed.

Baert, S., B. Cockx, N. Gheyle, and C. Vandamme (2015): "Is There Less Discrimination in Occupations Where Recruitment Is Difficult?" *ILR Review*, 0019793915570873.

Bain, D. H. (2000): *Empire Express: Building the First Transcontinental Railroad*, New York: Penguin Books.

Barro, R. J. and X. Sala-i Martin (1992): "Convergence," *Journal of Political Economy*, 100, 223–251.

Baum-Snow, N. and D. Neal (2009): "Mismeasurement of Usual Hours Worked In the Census and ACS," *Economics letters*, 102.

Becker, G. (1957): *[1971] The Economics of Discrimination*, Chicago: University Of Chicago Press, 2nd edition ed.

Bell, R. (1935): *Public school education of second-generation Japanese in California*, California: Stanford University press.

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

Biddle, J. E. and D. S. Hamermesh (2013): "Wage discrimination over the business cycle," *IZA Journal of Labor Policy*, 2, 7.

Bingham, W. V. (1946): "Inequalities in Adult Capacity From Military Data," *Science*, 104, 147–152.

Bordalo, P., K. Coffman, N. Gennaioli, and A. Shleifer (2016): "Stereotypes," *Quarterly Journal of Economics*, Forthcoming.

Borjas, G. J. (1987): "Self-Selection and the Earnings of Immigrants," *The American Economic Review*, 77, 531–553.

——— (1992): "Ethnic Capital and Intergenerational Mobility," *The Quarterly Journal of Economics*, 107, 123–150.

——— (1993): "The Intergenerational Mobility of Immigrants," *Journal of Labor Economics*, 11.

——— (1994): "Long-Run Convergence of Ethnic Skill DIfferentials: The Children and Grandchildren of the Great Migration," *Industrial and Labor Relations Review*, 47, 553–573.

Branch, T. (1989): *Parting the Waters : America in the King Years 1954-63*, New York: Simon & Schuster, reprint edition ed.

Braun, S. and J. Stuhler (2015): "The Transmission of Inequality Across Multiple Generations: Testing Recent Theories with Evidence from Germany," Tech. rep., Mimeo, November 2014, Universidad Carlos III de Madrid.

Bureau of the Census (1914): "Bulletin 127: Chinese and Japanese in the United States, 1910," Tech. rep.

Cameron, S. V. and J. J. Heckman (1993): "Comment on Trends in College Entry among Whites, Blacks, and Hispanics by Robert M. Hauser," in *Studies of Supply and Demand in Higher Education*, ed. by C. T. Clotfelter and M. Rothschild, University of Chicago Press, January, 61–120.

Card, D. and A. B. Krueger (1992a): "Does School Quality Matter ? Returns to Education and the Characteristics of Public Schools in the United States," *The Journal of Political Economy*, 100, 1–40.

——— (1992b): "School Quality and Black-White Relative Earnings: A Direct Assessment," *The Quarterly Journal of Economics*, 107, 151–200.

Carruthers, C. K. and M. H. Wanamaker (2016): "Separate and Unequal in the Labor Market: Human Capital and the Jim Crow Wage Gap," *Journal of Labor Economics*.

Chan, S. (1991): *Asian Americans: An Interpretive History*, Boston: Twayne Publishers, 1st edition ed.

Chang, I. (2004): *The Chinese in America: A Narrative History*, London: Penguin Books, reprint edition ed.

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

——— (2011): "Studying Discrimination : Fundamental Challenges and Recent Progress," .

Chen, J. (2011): "Chinese Immigration to the United States: History, Selectivity and Human Capital," *Working Paper*.

Chetty, R. and N. Hendren (2015): "The Impacts of Neighborhoods on Intergenerational Mobility: Childhood Exposure Effects and County-Level Estimates," *Unpublished Manuscript*.

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

Chiswick, B. R. (1983): "An Analysis of the Earnings and Employment of Asian-American Men," *Journal of Labor Economics*, 1, 197–214.

Clark, G. (2014): *The Son Also Rises: Surnames and the History of Social Mobility*, Princeton University Press.

Committee of 100 (2001): "American Attitudes Toward Chinese Americans and Asian Americans," Tech. rep.

Conlisk, J. (1974): "Can Equalization of Opportunity Reduce Social Mobility?" *The American Economic Review*, 64, 80–90.

Daniels, R. (1990): *Asian America: Chinese and Japanese in the United States since 1850*, Seattle: University of Washington Press, reprint edition ed.

Darity Jr., W., J. Dietrich, and D. K. Guilkey (1997): "Racial and Ethnic Inequality in the United States: A Secular Perspective," *The American Economic Review*, 87, 301–305.

Dickens, W. T. and J. R. Flynn (2006): "Black Americans Reduce the Racial IQ Gap Evidence From Standardization Samples," *Psychological Science*, 17, 913–920.

Diner, H. R. (2006): *The Jews of the United States, 1654 to 2000*, Berkeley: University of California Press.

Duncan, B. and S. J. Trejo (2016): "The Complexity of Immigrant Generations: Implications for Assessing the Socioeconomic Integration of Hispanics and Asians," Working Paper 21982, National Bureau of Economic Research.

Ferrie, J. P., K. Rolf, and W. Troesken (2012): "Cognitive disparities, lead plumbing, and water chemistry: Prior exposure to water-borne lead and intelligence

test scores among World War Two U.S. Army enlistees," *Economics & Human Biology*, 10, 98–111.

Frank, B. R. H. (1984): "Are Workers Paid their Marginal Products ?" *The American Economic Review*, 74, 549–571.

Fryer, R. G. (2010): "Racial Inequality in the 21st Century : The Declining Significance of Discrimination," *Handbook of Labor Economics*, 4.

Fryer, R. G. and S. D. Levitt (2006): "The Black-White Test Score Gap Through Third Grade," *American Law and Economics Review*, 8, 249–281.

——— (2013): "Testing for Racial Differences in the Mental Ability of Young Children," *American Economic Review*, 103, 981–1005.

Galor, O. and J. Zeira (1993): "Income Distribution and Macroeconomics," *The Review of Economic Studies*, 60, 35–52.

Glaeser, E. L., J. A. Scheinkman, and B. I. Sacerdote (2003): "The Social Multiplier," *Journal of the European Economic Association*, 1, 345–353.

Goldberg, M. S. (1982): "Discrimination, Nepotism, and Long-Run Wage Differentials," *The Quarterly Journal of Economics*, 97, 307–319.

Goldin, C. (2014): "A Grand Gender Convergence: Its Last Chapter," *American Economic Review*, 104, 1091–1119.

Gonzalez, J. (2011): *Harvest of Empire: A History of Latinos in America*, New York: Penguin Books, revised edition ed.

Graaf, L. B. D., K. Mulroy, and Q. Taylor, eds. (2001): *Seeking El Dorado: African Americans in California*, Los Angeles : Seattle: University of Washington Press, first edition ed.

Heckman, J. J., L. J. Lochner, and P. E. Todd (2006): "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond," *Handbook of the Economics of Education*, 1.

Hendrick, I. G. (1975): "Public Policy Toward the Education of Non-White Minority Group Children in California, 1849-1970. Final Report." Tech. rep., California University Riverside School of Education and National Institute of Education (DHEW), Washington D.C.

Hilger, N. G. (2016): "The Great Escape: Intergenerational Mobility in the United States Since 1940," Working Paper 21217, National Bureau of Economic Research.

Hirschman, C. and M. G. Wong (1986): "The Extraordinary Educational Attainment of Asian-Americans: A Search for Historical Evidence and Explanations," *Social Forces*, 65, 1.

Hughes, J. R. (1993): "The Demise of the English Standard School System in Hawai'i," *The Hawaiian Journal of History*, 27, 65–89.

Jensen, R. (2002): "'No Irish Need Apply': A Myth of Victimization," *Journal of Social History*, 36, 405–429.

Johnson, W. R. and D. Neal (1998): "Basic skills and the black-white earnings gap," *The Black-White test score gap*, 480–497.

Kenny, K. (2006): "Violence, Race, and anti-Irish Sentiment in the Nineteenth Century," in *Making the Irish American: History and Heritage of the Irish in the United States*, ed. by J. J. Lee and M. Casey, New York: NYU Press, 289–301.

Kleven, H. J., C. Landais, and J. E. Sogaard (2015): *Children and gender inequality: Evidence from Denmark*, Unpublished.

Kristof, N. (2015): "The Asian Advantage," *The New York Times*.

Lang, K. and J.-Y. K. Lehmann (2012): "Racial Discrimination in the Labor Market: Theory and Empirics," *Journal of Economic Literature*, 50, 959–1006.

Lee, J. and M. Zhou (2015): *The Asian American Achievement Paradox*, New York: Russell Sage Foundation.

Lemieux, T. (2006): "The "Mincer Equation" Thirty Years after Schooling, Experience and Earnings"," in *Jacob Mincer: A Pioneer of Modern Labor Economics*.

Mangione, J. (1993): *La Storia: Five Centuries of the Italian American Experience*, New York: Harper Perennial.

Margo, R. A. (1990): *Race and Schooling in the South, 1880-1950: An Economic History*, Chicago: University Of Chicago Press, 1 edition ed.

Meller, N. (1948): "Hawaii's English Standard Schools," Tech. Rep. 3, Legislative Reference Bureau, University of Hawaii.

Minnesota Population Center and Ancestry.com (2013): "IPUMS Restricted Complete Count Data: Version 1.0 [Machine-readable database]," Tech. rep., University of Minnesota, Minneapolis.

Neal, D. (2006): "Chapter 9: Why Has Black-White Skill Convergence Stopped?" in *Handbook of the Economics of Education*, Elsevier, vol. 1, 511–576.

Neal, D. A. and W. R. Johnson (1996): "The Role of Premarket Factors in Black-White Wage Differences," *The Quarterly Journal of Economics*, 104(5): 869–895.

Nordyke, E. C. and Y. S. Matsumoto (1977): "Japanese in Hawaii: a Historical and Demographic Perspective," .

Olivetti, C., M. D. Paserman, and L. Salisbury (2014): "Intergenerational Mobility Across Three Generations in the 19th Century : Evidence from the US Census," *Unpublished Manuscript*.

Page, J. (2004): *In the Hands of the Great Spirit: The 20,000-Year History of American Indians*, New York: Free Press, new edition edition ed.

Pager, D., B. Western, and B. Bonikowski (2009): "Discrimination in a Low-Wage Labor Market: A Field Experiment." *American Sociological Review*, 74, 777–799.

Pfaelzer, J. (2008): *Driven Out: The Forgotten War against Chinese Americans*, Berkeley: University of California Press.

Ramsey, F. P. (1928): "A Mathematical Theory of Saving," *The Economic Journal*, 38, 543–559.

Rediker, M. (2008): *The Slave Ship: A Human History*, New York: Penguin Books, reprint edition ed.

Romer, P. M. (1986): "Increasing Returns and Long-Run Growth," *Journal of Political Economy*, 94, 1002–1037.

Rothstein, J. and N. Wozny (2014): "Permanent Income and the Black- White Test Score Gap," *Journal of Human Resources*, 48, 509–544.

Ruggles, S. J., K. Genadek, R. Goeken, J. Grover, and M. Sobek (2015): *Integrated Public Use Microdata Series: Version 6.0 [Machine-readable database].*, Minneapolis: University of Minnesota.

Sacerdote, B. (2011): "Chapter 4 - Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?" in *Handbook of the Economics of Education*, ed. by S. M. a. L. W. Eric A. Hanushek, Elsevier, vol. 3, 249–277.

Sandmeyer, E. (1991): *The Anti-Chinese Movement in California*, Urbana: University of Illinois Press, reprint edition ed.

Smith, M. (1947): "Populational Characteristics of American Servicemen in World War II," *The Scientific Monthly*, 65, 246–252.

Smith, P. J. (1984): "Race and Human Capital," *The American Economic Review*, 74, 685–698.

Solon, G. (2015): "What Do We Know So Far about Multigenerational Mobility?" Working Paper 21053, National Bureau of Economic Research.

Solow, R. M. (1956): "A Contribution to the Theory of Economic Growth," *The Quarterly Journal of Economics*, 70, 65–94.

Stuhler, J. (2014): "Mobility Across Multiple Generations," *Unpublished Manuscript*.

Sue, S. and S. Okazaki (2009): "Asian-American educational achievements: A phenomenon in search of an explanation." *Asian American Journal of Psychology*, S, 45–55.

Suzuki, M. (1995): "Japanese Immigrant Economic Achievement and Return Migration, 1920-1930," *The Journal of Economic History*, 55, 889–901.

——— (2002): "Selective Immigration and Ethnic Economic Achievement: Japanese Americans before World War II," *Explorations in Economic History*, 39, 254–281.

Takaki, R. (1998): *Strangers from a Different Shore: A History of Asian Americans, Updated and Revised Edition*, Boston: Little, Brown and Company, revised and updated edition ed.

The Gallup Organization (2005): "Employee Discrimination in the Workplace," Tech. rep.

Trejo, S. J. (1997): "Why Do Mexican Americans Earn Low Wages?" *Journal of Political Economy*, 105, 1235–1268.

Wollenberg, C. (1995): ""Yellow Peril" in the Schools (I and II)," in *The Asian American Educational Experience: A Sourcebook for Teachers and Students*, Routledge.

	White		Black	
Year	Top Occupation	Share	Top Occupation	Share
1860	Farmers (owners and tenants)	0.395	Laborers (nec)	0.273
1870	Farmers (owners and tenants)	0.368	Farm laborers, wage workers	0.455
1880	Farmers (owners and tenants)	0.362	Farmers (owners and tenants)	0.335
1900	Farmers (owners and tenants)	0.270	Farmers (owners and tenants)	0.375
1910	Farmers (owners and tenants)	0.230	Farmers (owners and tenants)	0.357
1920	Farmers (owners and tenants)	0.205	Farmers (owners and tenants)	0.326
1930	Farmers (owners and tenants)	0.158	Laborers (nec)	0.257
1940	Laborers (nec)	0.118	Laborers (nec)	0.333
1950	Operative and kindred workers (nec)	0.100	Laborers (nec)	0.223
1960	Operative and kindred workers (nec)	0.108	Laborers (nec)	0.204
1970	Operative and kindred workers (nec)	0.095	Operative and kindred workers (nec)	0.155
1980	Managers, officials, and proprietors (nec)	0.126	Operative and kindred workers (nec)	0.139
1990	Managers, officials, and proprietors (nec)	0.149	Operative and kindred workers (nec)	0.108
2000	Managers, officials, and proprietors (nec)	0.172	Operative and kindred workers (nec)	0.096
	Asian		Native American	
Year	Top Occupation	Share	Top Occupation	Share
1860	Mine operatives and laborers	0.753	Other non-occupation	0.578
1870	Mine operatives and laborers	0.457	Farmers (owners and tenants)	0.322
1880	Mine operatives and laborers	0.294	Laborers (nec)	0.289
1900	Laborers (nec)	0.221	Farmers (owners and tenants)	0.263
1910	Laborers (nec)	0.212	Farmers (owners and tenants)	0.373
1920	Farm laborers, wage workers	0.154	Farmers (owners and tenants)	0.388
1930	Farm laborers, wage workers	0.197	Farmers (owners and tenants)	0.358
1940	Farm laborers, wage workers	0.249	Laborers (nec)	0.323
1950	Farm laborers, wage workers	0.175	Laborers (nec)	0.179
1960	Cooks, except private household	0.085	Laborers (nec)	0.195
1970	Cooks, except private household	0.075	Laborers (nec)	0.115
1980	Managers, officials, and proprietors (nec)	0.105	Operative and kindred workers (nec)	0.085
1990	Managers, officials, and proprietors (nec)	0.127	Laborers (nec)	0.081
2000	Managers, officials, and proprietors (nec)	0.147	Managers, officials, and proprietors (nec)	0.107

Table I: Top Occupation by Race, 1860-2000

Notes: Occupation shares calculated for men age 25-65, excluding residents of Alaska and Hawaii. Each race reweighted to match age distribution of whites in each year.

	U.S. Residents					
	AGCT Sample		Census Sample		AGCT-Census Match	
Variable	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
AGCT	97.681	22.707			99.058	22.228
Age	23.144	5.915	27.498	6.031	23.064	5.355
White	0.907	0.291	0.898	0.302	0.937	0.243
Black	0.068	0.251	0.096	0.294	0.062	0.241
Chinese	0.002	0.042	0.001	0.028	0.000	0.011
Japanese	0.000		0.001	0.033	0.000	
Less Than HS	0.254	0.435	0.436	0.496	0.243	0.429
Some HS	0.333	0.471	0.226	0.418	0.328	0.470
HS Graduate	0.308	0.462	0.217	0.412	0.318	0.466
Any College	0.106	0.308	0.121	0.326	0.111	0.314
N	525,792		19,975,888		211,103	

	California Residents					
	AGCT Sample		Census Sample		AGCT-Census Match	
Variable	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
AGCT	100.561	21.503			103.293	20.678
Age	24.142	6.988	28.011	5.966	24.547	5.960
White	0.896	0.305	0.943	0.232	0.957	0.204
Black	0.040	0.195	0.018	0.131	0.041	0.199
Chinese	0.009	0.095	0.007	0.086	0.001	0.031
Japanese	0.000		0.014	0.119	0.000	
Less Than HS	0.176	0.381	0.257	0.437	0.168	0.374
Some HS	0.324	0.468	0.245	0.430	0.318	0.466
HS Graduate	0.328	0.470	0.304	0.460	0.329	0.470
Any College	0.172	0.377	0.193	0.395	0.185	0.389
N	34,604		1,186,305		8,509	

Table II: Summary Statistics

Notes: Table presents summary statistics for three samples for all U.S. residents and CA residents. All samples restrict to men. "AGCT Sample" is the sample of WWII enlistment records from 1943. "Census Sample" is 100% IPUMS census microdata for 1940. "AGCT-Census Match" is a match of these two prior datasets on first name, last name, state of birth, race and year of birth plus or minus one year.

	θ = Investment Intercept			γ = Investment Slope			α = Earnings Intercept			β = Earnings Slope		
	White	Black	Asian	White	Black	Asian	White	Black	Asian	White	Black	Asian
Year	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
1940	9.566	9.525	9.529	0.752	0.757	0.604	9.222	8.786	8.942	0.078	0.065	0.050
	(0.307)	(0.476)	(0.427)	(0.13)	(0.214)	(0.192)	(0.037)	(0.03)	(0.032)	(0.005)	(0.004)	(0.004)
1960	12.224	17.178	13.429	0.243	-0.262	1.155	10.113	9.779	10.065	0.054	0.059	0.019
	(0.557)	(1.589)	(1.944)	(0.225)	(1.666)	(1.118)	(0.035)	(0.228)	(0.13)	(0.006)	(0.039)	(0.022)
1970	11.184	15.620	17.038	0.860	-0.569	-0.006	10.307	9.832	10.223	0.061	0.073	0.069
	(0.251)	(1.137)	(2.473)	(0.127)	(0.556)	(0.775)	(0.024)	(0.124)	(0.081)	(0.005)	(0.023)	(0.015)
1980	12.096	12.489	14.592	0.442	0.273	0.335	10.204	9.596	10.245	0.059	0.130	0.062
	(0.364)	(0.317)	(0.775)	(0.159)	(0.154)	(0.418)	(0.031)	(0.047)	(0.042)	(0.007)	(0.01)	(0.009)
1990	12.239	12.208	14.295	0.313	0.284	-0.172	10.115	9.611	10.074	0.133	0.198	0.158
	(0.212)	(0.346)	(0.493)	(0.092)	(0.154)	(0.206)	(0.048)	(0.061)	(0.077)	(0.012)	(0.015)	(0.019)
2000	12.492	12.063	14.009	0.378	0.380	0.129	10.090	9.483	10.147	0.150	0.233	0.157
	(0.219)	(0.239)	(0.309)	(0.106)	(0.099)	(0.139)	(0.078)	(0.12)	(0.124)	(0.019)	(0.03)	(0.031)

Table III: Linear Decomposition Estimates, Born in CA

Notes: "Investment" intercepts and slopes estimated from linear regressions of children's highest grade attained by ages 22-29 on parental log income, using data grouped at the year by race by parental log income decile level. "Earnings" intercepts and slopes estimated from linear regressions of log household earnings for heads age 25-65 on highest grade of schooling attained in cells defined by year, race, and highest grade attained. Restricts to children born in CA. Probability weights are used to construct cell means, but no weights are used in the regressions. All races reweighted to age and sex distribution of whites in each year. Earnings deflated to 2011 dollars using CPI-Urban before taking logs. "Earnings" regressions drop bottom 2% of relevant education distribution. Intercepts reflect estimated values at minimum observed value of independent variable in collapsed, restricted sample, as opposed to estimated values at zero.

	Next Gen - Actual			Next Gen - Asian Investments		Steady State - Asian Returns		Next Gen - Asian Endowments	
	White	Black	Asian	White	Black	White	Black	White	Black
Year	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1940	9.723	9.175	9.230	9.692	9.157	9.265	9.241	9.695	9.181
1960	10.430	10.370	10.222	10.626	10.217	10.175	10.251	10.417	10.356
1970	10.610	10.322	10.846	10.856	10.494	10.568	10.685	10.674	10.253
1980	10.451	10.113	10.635	10.583	10.400	10.505	10.493	10.440	10.121
1990	10.515	10.163	10.680	10.629	10.393	10.551	10.516	10.526	10.202
2000	10.587	10.147	10.825	10.733	10.478	10.669	10.596	10.598	10.172

Table IV: Implications of Intergenerational Decomposition for Next Generation, Born in CA

Notes: Presents estimated log household earnings in steady state using parameter estimates in Table (III) and equation (2). "Asian Investments" assigns blacks the estimated intercept and slope of Asians' conditional expectation of children's education with respect to parental income. "Asian Returns" assigns blacks the estimated intercept and slope of Asians' conditional expectation of children's earnings with respect to their education. "Asian Endowments" refers to Asian mean parental income. Estimates based on sample restricted to children born in CA.

	Steady State - Actual			Steady State - Asian Investments		Steady State - Asian Returns	
	White	Black	Asian	White	Black	White	Black
Year	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1940	9.732	9.177	9.230	9.697	9.157	9.252	9.245
1960	10.430	10.364	10.232	10.638	10.233	10.174	10.250
1970	10.605	10.315	10.846	10.856	10.494	10.559	10.665
1980	10.446	10.108	10.634	10.582	10.407	10.501	10.497
1990	10.505	10.150	10.687	10.632	10.392	10.541	10.522
2000	10.574	10.129	10.826	10.732	10.482	10.660	10.613

Table V: Implications of Intergenerational Decomposition in Steady State, Born in CA

Notes: Presents estimated log household earnings in steady state using parameter estimates in Table (III) and equation (3). "Asian Investments" assigns blacks the estimated intercept and slope of Asians' conditional expectation of children's education with respect to parental income. "Asian Returns" assigns blacks the estimated intercept and slope of Asians' conditional expectation of children's earnings with respect to their education. Estimates based on sample restricted to children born in CA.

Education Level	White	Black	Asian
No High School	81.6	67.7	79.5
	(0.256)	(0.798)	(1.61)
	[4,942]	[353]	[88]
Some High School	96.4	80.3	91.8
	(0.191)	(0.713)	(1.73)
	[9,956]	[570]	[92]
High School Graduate	109	90.8	104
	(0.158)	(0.941)	(2.55)
	[10,576]	[322]	[86]
Any College	117	97.2	116
	(0.216)	(1.8)	(2.5)
	[5,540]	[126]	[48]

Table VI: Mean WWII Enlistee Test Scores by Race and Schooling in 1943, CA

Notes: Table presents raw means of WWII enlistment test scores by race and broad education category, restricting to enlistees reporting CA residence. Standard errors of means in parantheses, sample sizes in brackets.

Variables	(1) All Races	(2) All Races	(3) All Races	(4) White	(5) White	(6) Black	(7) Black
Black	-0.585**	-0.365**	-0.368**				
	(0.0362)	(0.0299)	(0.0291)				
Age	0.0741**	0.0746**	0.0759**	0.0752**	0.0765**	0.0618**	0.0628**
	(0.00177)	(0.00161)	(0.00187)	(0.00150)	(0.00177)	(0.00617)	(0.00608)
AGCT		0.164**	0.0909**	0.167**	0.0930**	0.173**	0.111**
		(0.00550)	(0.00709)	(0.00544)	(0.00772)	(0.0335)	(0.0348)
$AGCT^2$		-0.00376	-0.00792*	-0.00812*	-0.0119**	0.0297	0.0211
		(0.00356)	(0.00326)	(0.00356)	(0.00319)	(0.0160)	(0.0142)
Education			0.0506**		0.0512**		0.0367**
			(0.00454)		(0.00499)		(0.00832)
Constant	6.956**	6.904**	6.367**	6.890**	6.346**	6.856**	6.466**
	(0.0335)	(0.0340)	(0.0464)	(0.0316)	(0.0494)	(0.191)	(0.194)
Observations	42,139	42,139	41,609	40,180	39,669	1,959	1,940
R-squared	0.162	0.195	0.217	0.179	0.203	0.113	0.126

Standard errors in parentheses
** p<0.01, * p<0.05

Table VII: Earnings Regressions on AGCT Scores in 1940

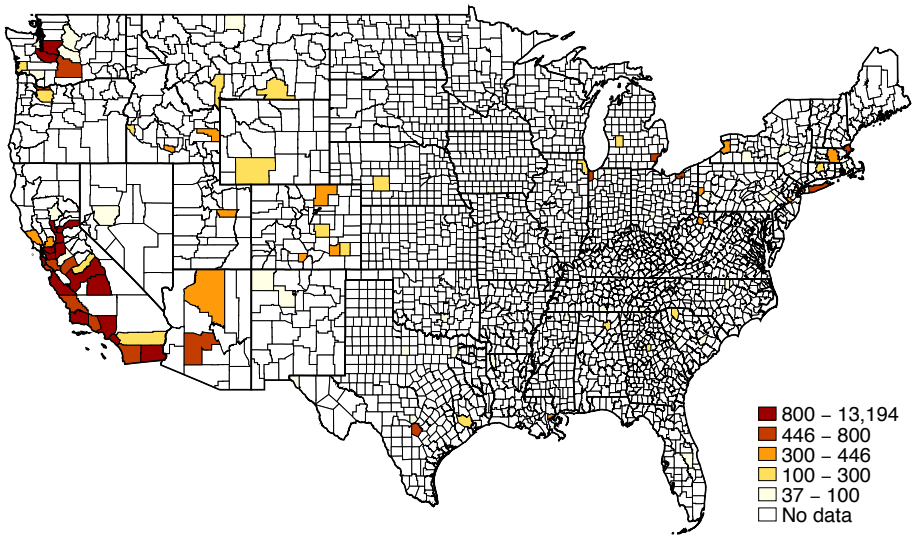
Notes: Dependent variable is log of annual earnings in all regressions. Sample matches AGCT data from 1943 WWII enlistment records to earnings data from 1940 census 100% sample. Restricts to men over age 22 in 1940; almost all men in sample between ages 22 and 35. Standard errors clustered at the state level.

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					Ln(Earnings)					
Black	-0.590**	-0.524**	-0.399**	-0.204*	-0.590**	-0.524**	-0.201	-0.686**	-0.487**	-0.105
	(0.0443)	(0.0300)	(0.0249)	(0.0915)	(0.0444)	(0.0436)	(0.133)	(0.0581)	(0.0281)	(0.109)
AGCT			0.137**	0.136**			0.359*			0.444**
			(0.0123)	(0.0125)			(0.133)			(0.115)
AGCT_Market				0.218*						
				(0.0904)						
Constant	9.153**	9.327**	9.427**	9.572**	9.153**	9.327**	9.566**	9.461**	9.380**	9.602**
	(0.0462)	(0.00931)	(0.0146)	(0.0598)	(0.0463)	(0.0135)	(0.0871)	(0.0475)	(0.000929)	(0.0572)
Observations	35,311	35,311	35,311	35,311	209	209	209	354	354	354
R-squared	0.022	0.064	0.078	0.078	0.329	0.951	0.956	0.284	0.986	0.990
Market FE	No	Yes	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Data Level	Micro	Micro	Micro	Micro	Market	Market	Market	Market	Market	Market
Dataset	Matched	Matched	Matched	Matched	Matched	Matched	Matched	Separate	Separate	Separate

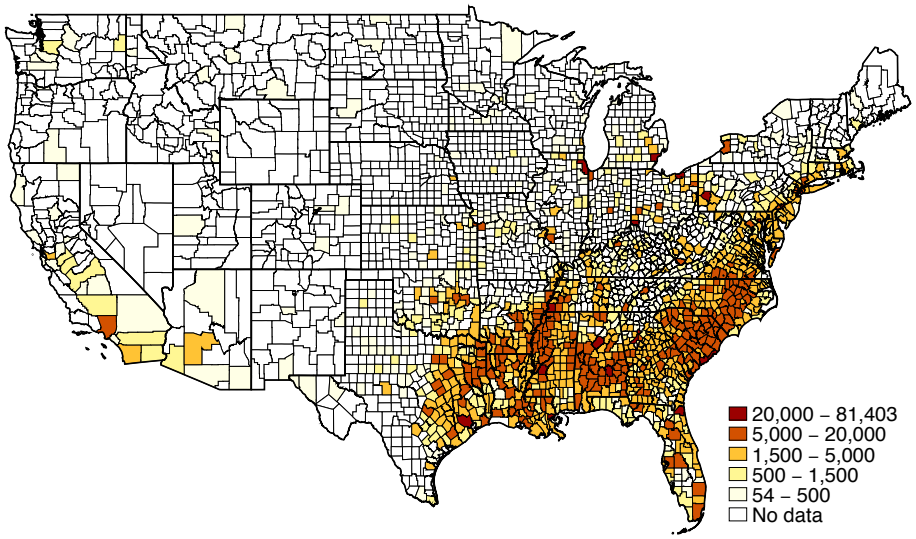
Robust standard errors in parentheses
** p<0.01, * p<0.05

Table VIII: Earnings Regressions on AGCT Scores for State-Education Aggregates in 1940

Notes: Dependent variable is log of annual earnings in all regressions. "Market FE" indicates whether market fixed effects are included in the regression, where "market" is defined by state of residence and broad educational category (no high school, some high school, high school graduate, any college). "Data Level" indicates whether regression is run on individual microdata or data that has been collapsed to the level of market × race. "Dataset" indicates whether earnings and AGCT scores taken from subsample of WWII enlistment records matched to 1940 100% census microdata, or whether earnings and AGCT scores taken from complete unmatched datasets separately. All regressions restricts to men over age 22 in 1940; almost all men in sample between ages 22 and 35. Standard errors clustered at the state level.



(a) Asian



(b) Black

Figure I: Population of Native-Born Children Across U.S. Counties, 1940

Notes: Children age 0-18, excluding Alaska and Hawaii.

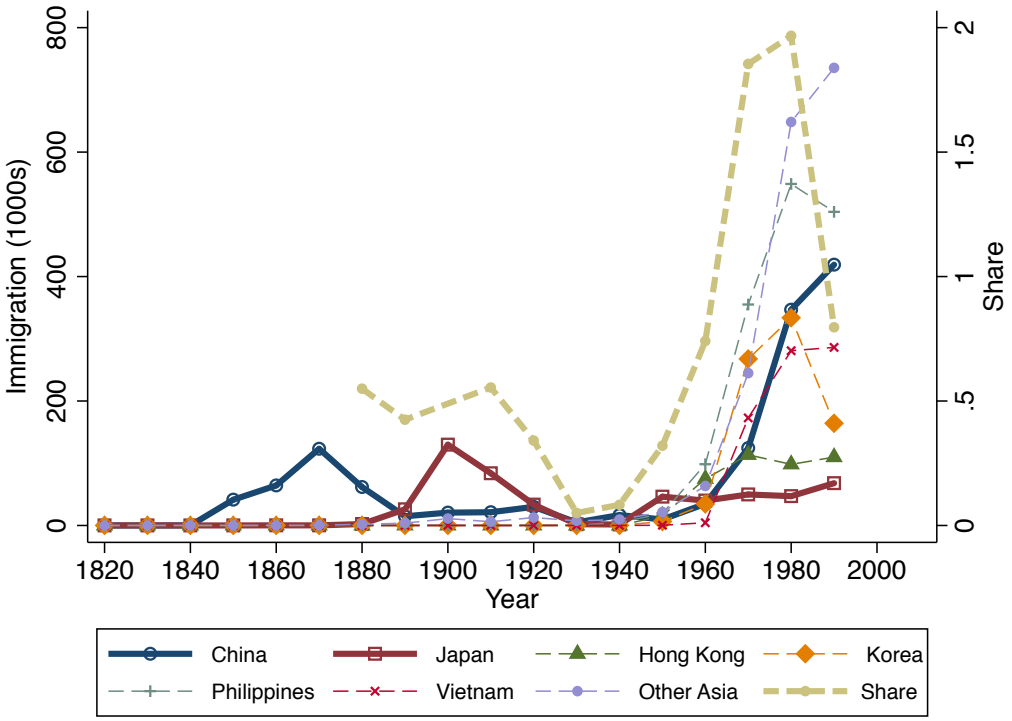
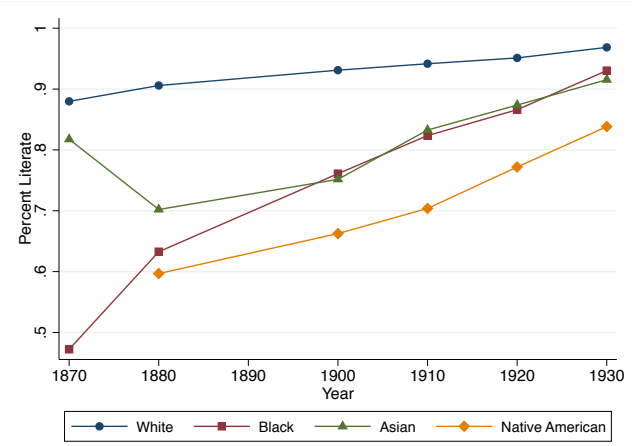
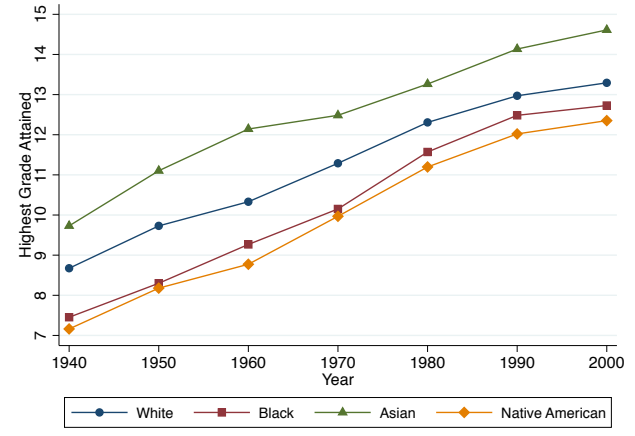


Figure II: Gross Immigration into U.S. from Various Asian Countries, 1821-1991

Notes: Data on immigration flows taken from Department of Homeland Security, Yearbook of Immigration Statistics 2003. Data on stock of Asians from census data, adjusted to include Japanese in Hawaii as reported in Table I in Nordyke and Matsumoto (1977). Immigration totals include "foreign nationals who, during a fiscal year, were granted lawful permanent residence (i.e., admitted as immigrants or became legal permanent residents), were admitted into the United States on a temporary basis (e.g., tourists, students, or workers), applied for asylum or refugee status, or were naturalized." No adjustment made for undocumented immigration. Stock of Asians cut off after 1970 for readability of figure.



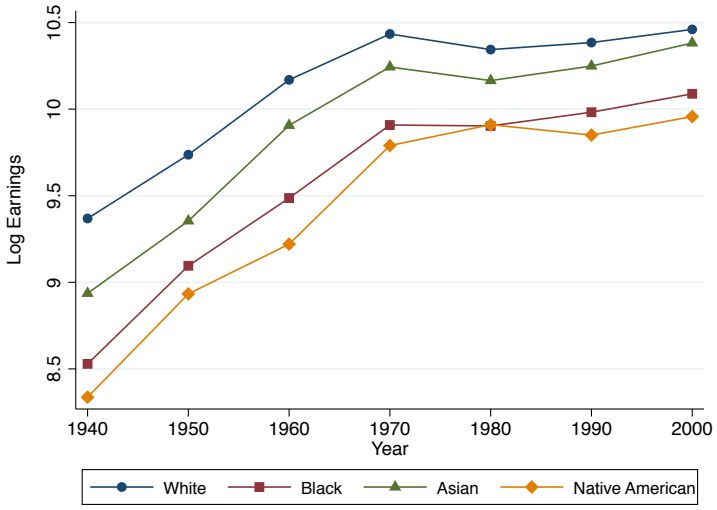
(a) Percent Literate, 1880-1930



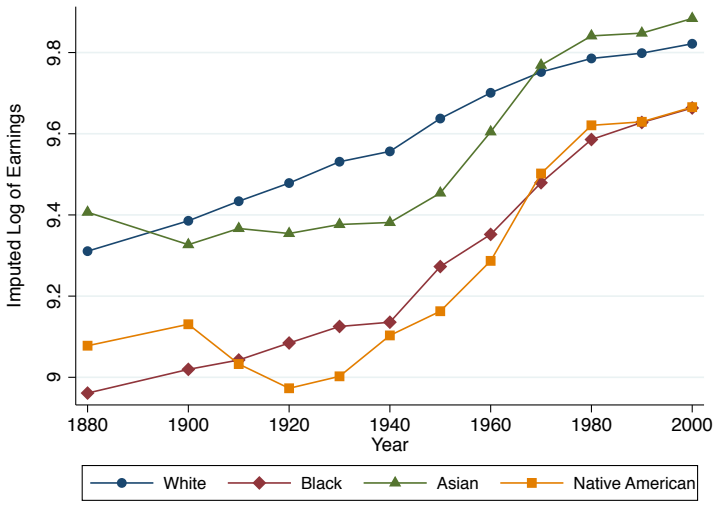
(b) Highest Grade Attained, 1940-2000

Figure III: Human Capital by Race, 1880-2000

Notes: Literacy defined as ability to read and write in any language. Figure restricts to ages 25-65 and excludes residents of Alaska and Hawaii. All races reweighted to match age and sex distribution of whites in every year.



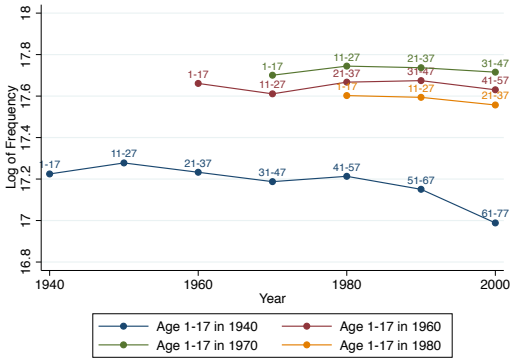
(a) Log Earnings, 1940-2000



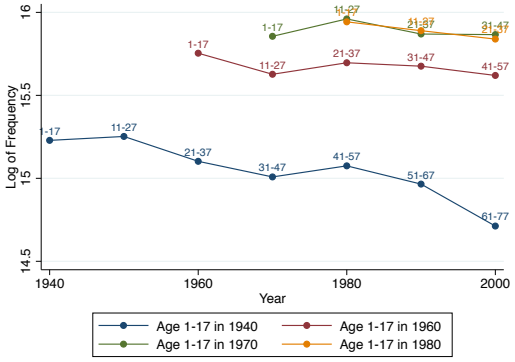
(b) Imputed Log Earnings, 1880-2000

Figure IV: Log Earnings of Men, 1880-2000

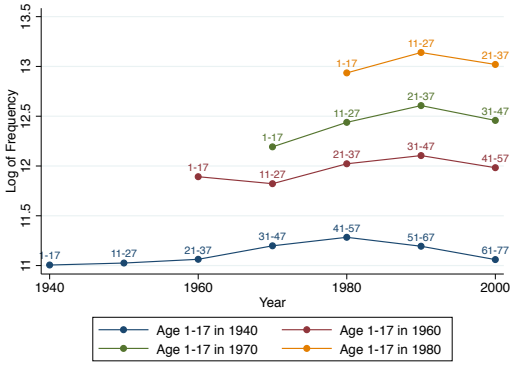
Notes: Panel (a) plots average log male earnings age 25-65 by race and year. Panel (b) plots average imputed log male earnings age 25-65 by race and year, with imputation based on earnings in 1940 averaged by cells defined by OCC1950 and native-born status, excluding observations with zero earnings or missing occupation, and restricting to household heads. Residents of Hawaii and Alaska excluded and races reweighted to match age distribution of whites in every year.



(a) White



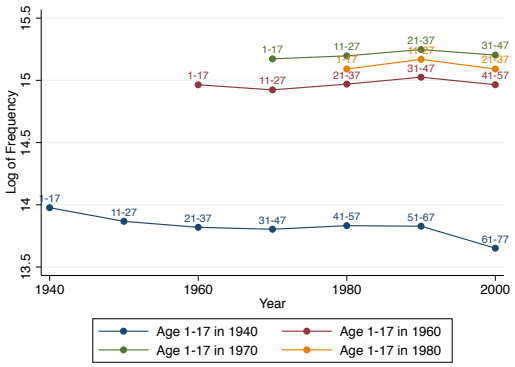
(b) Black



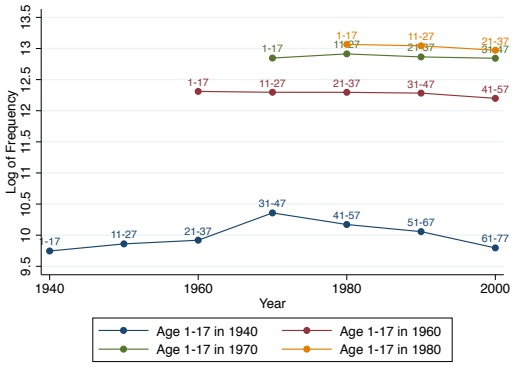
(c) Asian

Figure V: Log Frequencies of Pseudo-Cohorts: Born in U.S.

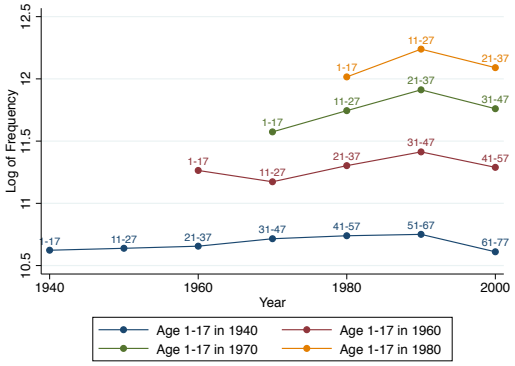
Notes: Figure plots log of frequencies by race for cohorts of native-born individuals age 1-17 in in 1940, 1960, 1970, 1980.



(a) White



(b) Black



(c) Asian

Figure VI: Log Frequencies of Pseudo-Cohorts: Born in CA

Notes: Figure plots log of frequencies by race for cohorts of individuals born in CA ages 1-17 in 1940, 1960, 1970, 1980.

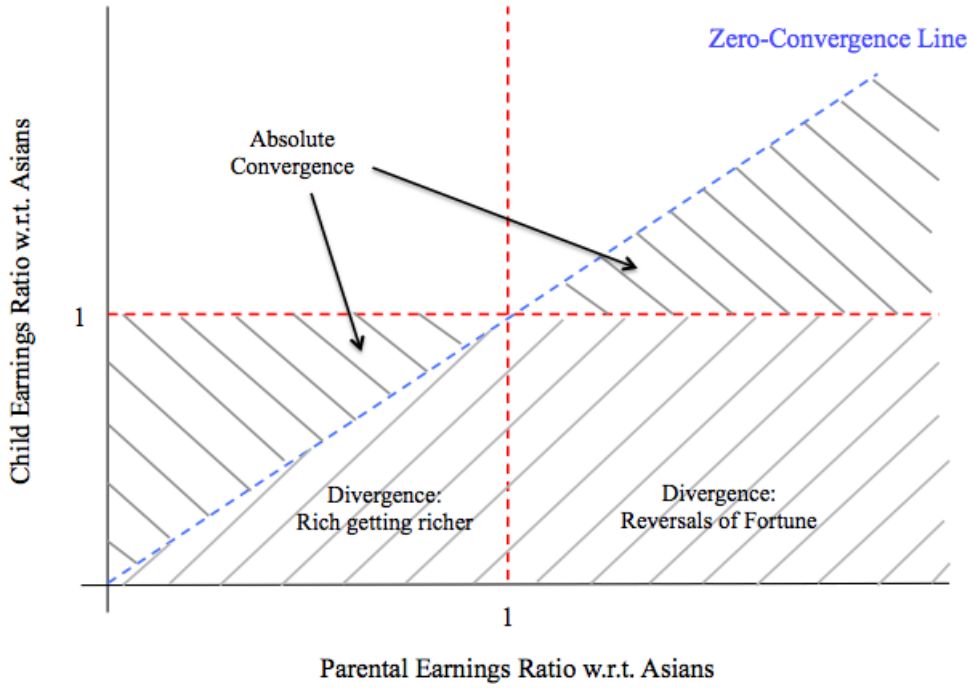
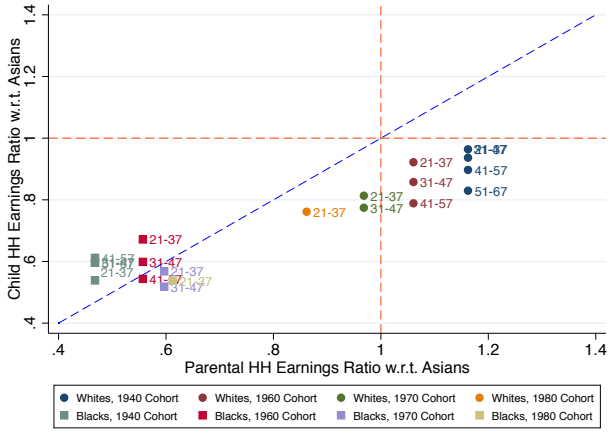
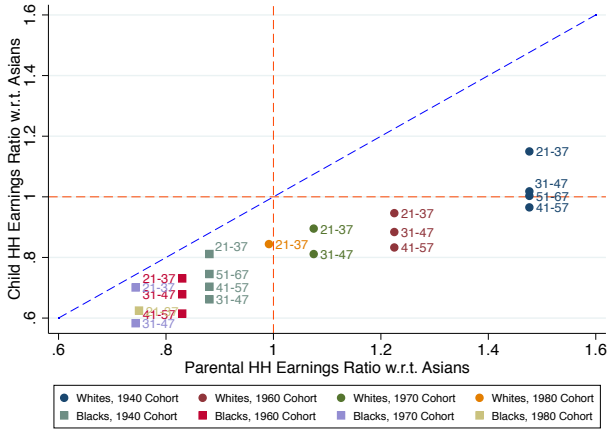


Figure VII: Earnings Ratio Convergence Figures: Illustration

Notes: Diagram illustrates interpretation of earnings ratios of parents and children for various groups with respect to Asians. Points above 1 on the x-axis imply parents in a group are richer than Asian parents, and points above 1 on the y-axis imply children in a group are richer than Asian children.



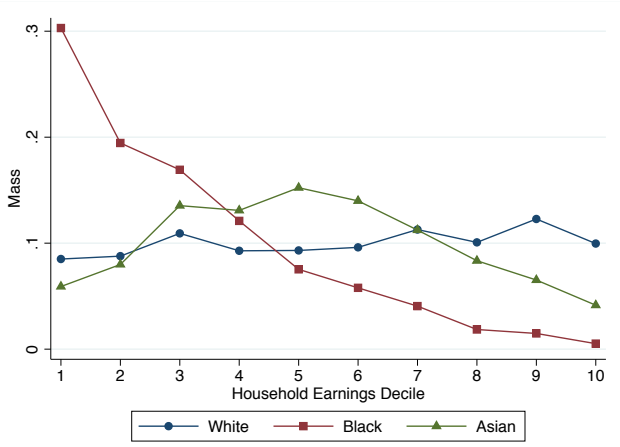
(a) Born in US



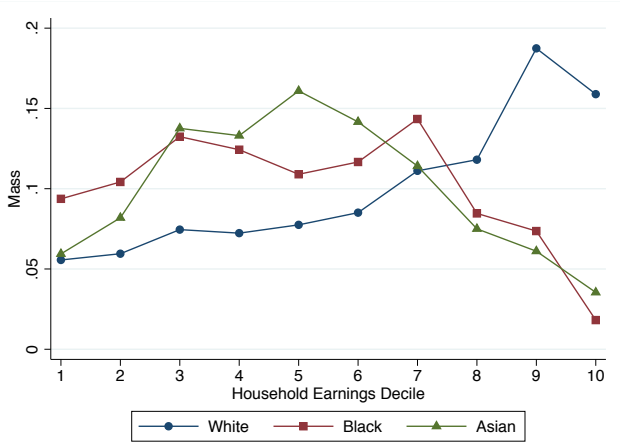
(b) Born in CA

Figure VIII: Intergenerational Change in Earnings Ratios

Notes: Figure plots earnings ratios for parents of children age 1-17 in 1940 and 1960 on the x-axes, and earnings ratios for children in these cohorts at later ages in subsequent censuses. All earnings ratios are plotted with respect to Asians of the same ages in the same parent and child groups. Household earnings plots earnings of head and spouse, counting absence of spouse as zero spousal earnings. Households with zero total earnings excluded from all calculations.



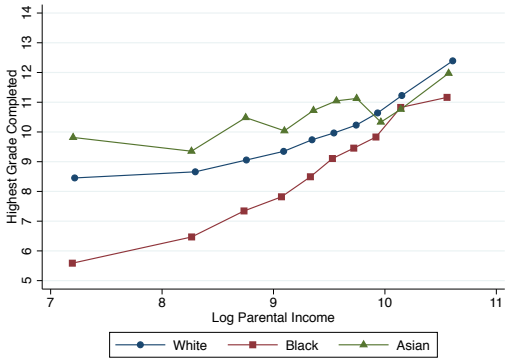
(a) Born in U.S.



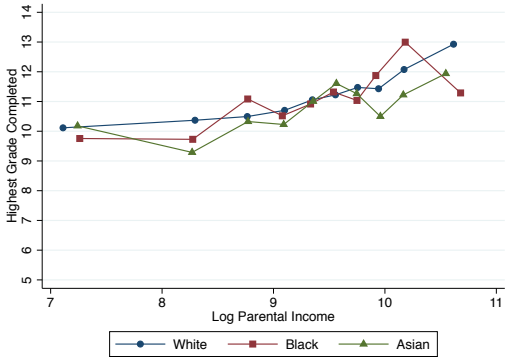
(b) Born in CA

Figure IX: Parental Income Distribution for Native-Born by Race, 1940

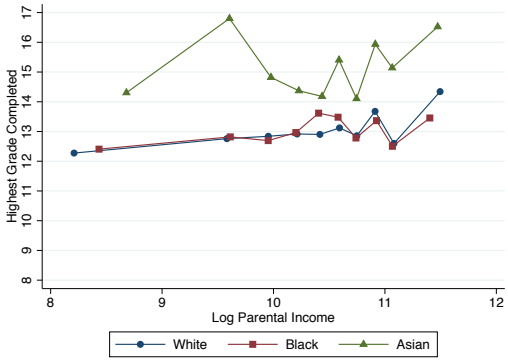
Notes: Figures plot the probability mass functions for total parental income deciles of native-born children under age 18 in 1940. Deciles calculated over full US population, and therefore held fixed across races and locations.



(a) Born in U.S., 1940



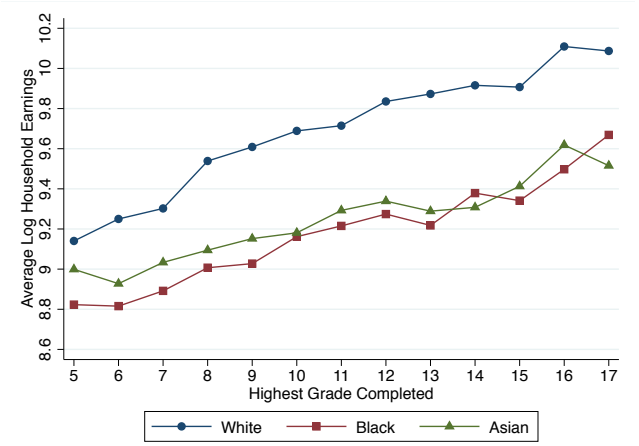
(b) Born in CA, 1940



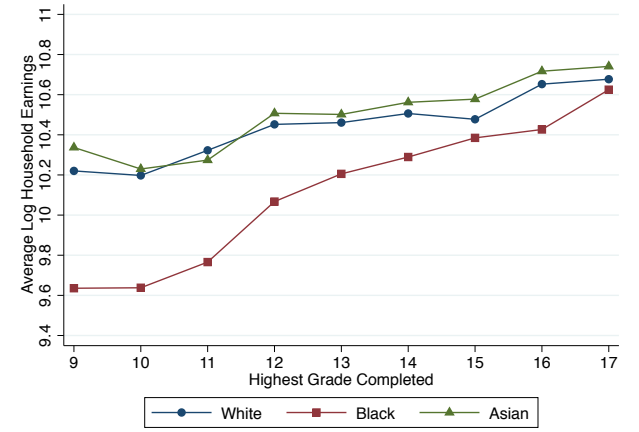
(c) Born in CA, 1980

Figure X: Educational Attainment by Log Parental Income

Notes: Hawaii and Alaska excluded. Figure adjusts for independent children and pools ages 22-29. Log parental income calculated as sum of head and spouse earnings over full population age 25-65.



(a) 1940



(b) 1980

Figure XI: Log Household Earnings by Highest Grade Attained: Born in CA

Notes: Restricts to CA-born ages 25-65. Reweights all groups to age and sex distribution of whites. Earnings deflated using CPI-Urban to 2011 dollars.

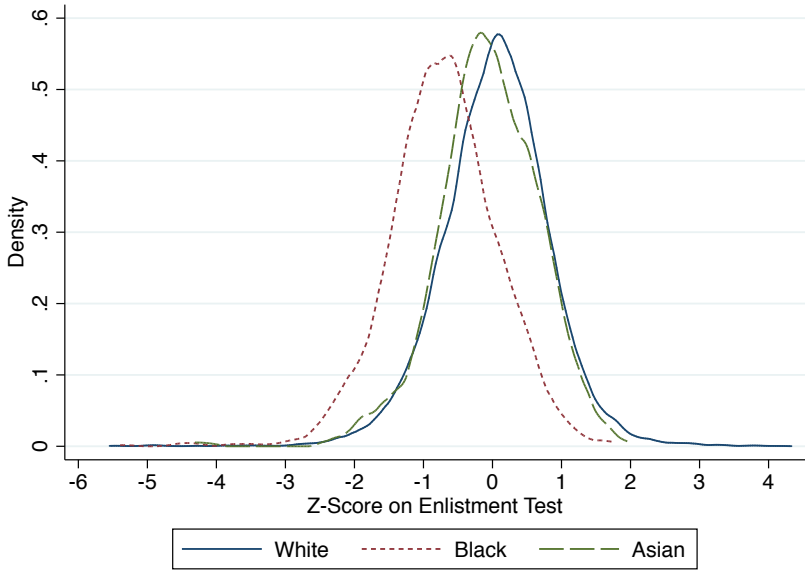
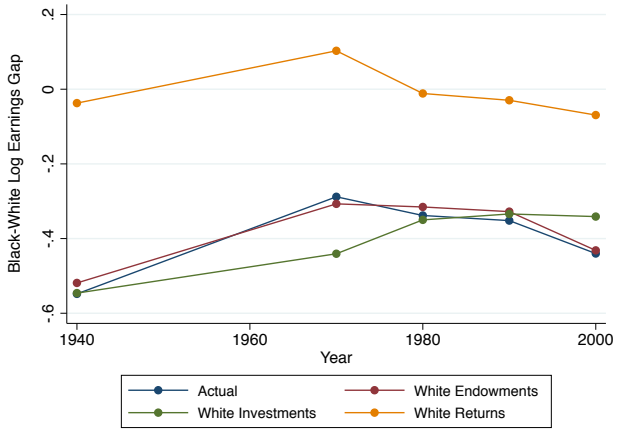
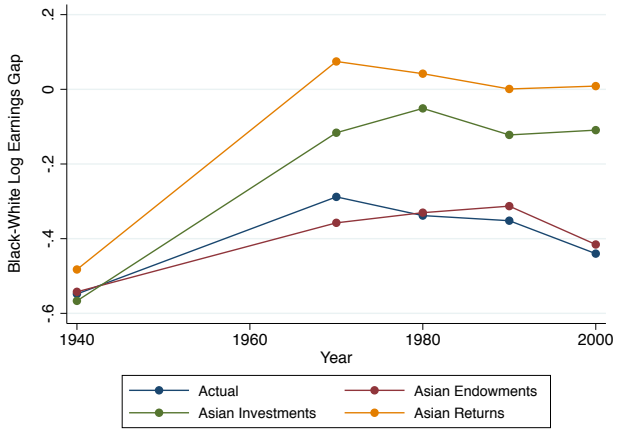


Figure XII: WWII Enlistment Test Score Distributions by Race in 1943, CA

Notes: Figure plots distributions of residuals from regression of normalized test scores on complete sets of education and age dummies. Restricts to native-born men ages 25-65 living in CA.



(a) Imputing White Earnings Components to Blacks



(b) Imputing Asian Earnings Components to Blacks

Figure XIII: Counterfactual Black-White Log Earnings Gaps in CA, 1940-2000

Notes: Figure presents simulated black-white log earnings gaps using estimates of the four parameters in Equation (2) for each race r and each generation t, $\alpha_{r,t}, \beta_{r,t}, \gamma_{r,t}, \theta_{r,t}$ as shown in Table IV, as well as mean parental income $E\left[y_{r,t-1}\right]$. "Earnings" refers to log of household earnings (head + spouse). "Actual" predicts black-white earnings gaps using estimated parameters for each racial group. "White Endowments" assigns the white parental income distribution to blacks. "White Investments" assigns the white conditional expectation of children's education with respect to parental income to blacks. "White Returns" assigns the white conditional expectation of household earnings to blacks. Panel (b) repeats this but assigns these respective components from Asians to blacks. All estimates restrict to "children" born in CA.

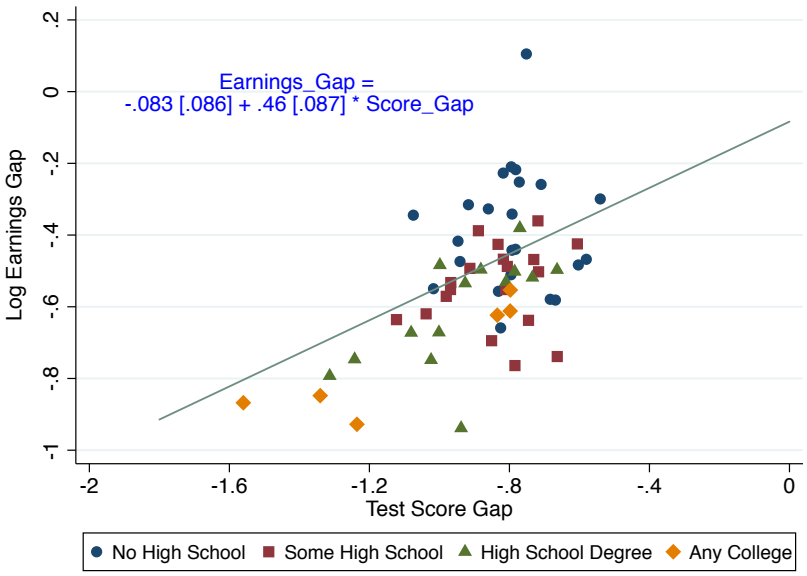
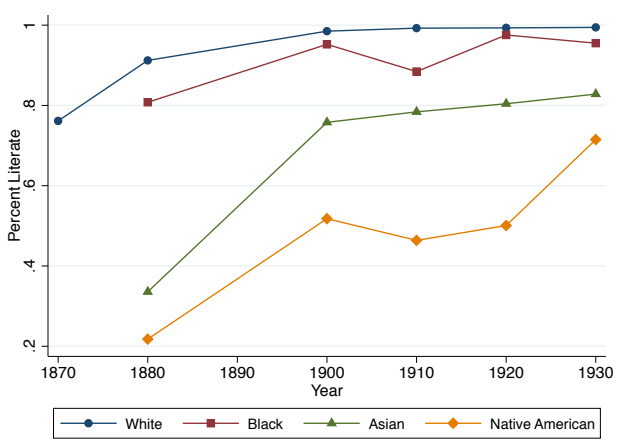
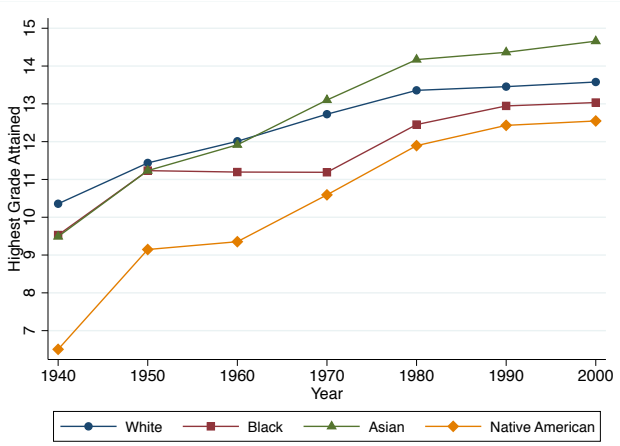


Figure XIV: Black-White Earnings and Skill Gaps by State and Education Level in 1940

Notes: Figure plots log earnings gaps by skill gaps at the level of broad educational group and state of residence for men ages 18-38. Earnings gaps defined as log earnings of whites minus log earnings of blacks. Cells with fewer than 30 individual blacks omitted from figure. Education groups are no high school, some high school, high school degree, and any college. Test scores normalized into z-score in microdata before construction of score gaps at the state-education level.



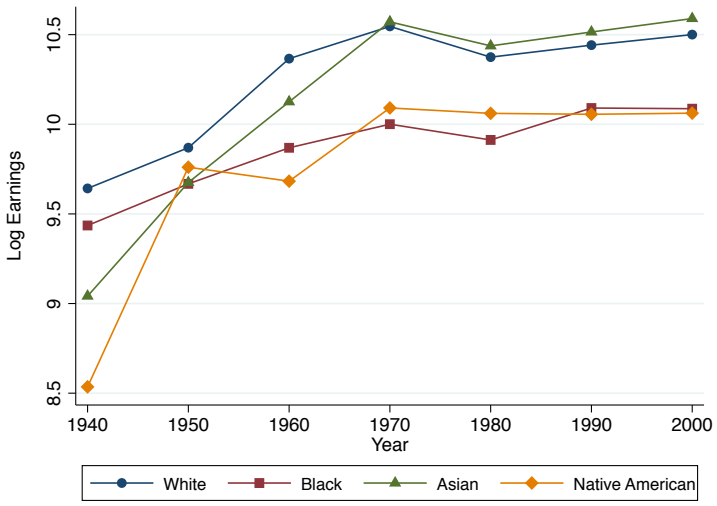
(a) Percent Literate, 1880-1930



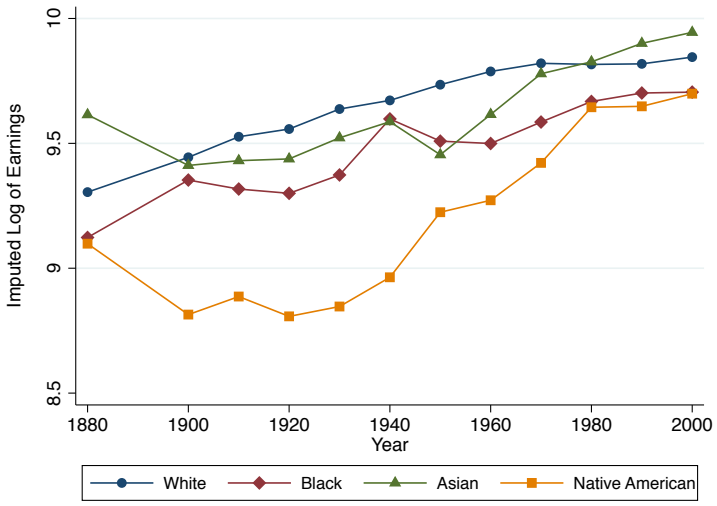
(b) Highest Grade Attained, 1940-2000

Figure A.1: Human Capital by Race, CA-Born 1880-2000

Notes: Restricting to individuals born in California. Literacy defined as ability to read and write in any language. Figure restricts to ages 25-65 and excludes residents of Alaska and Hawaii. All races reweighted to match age and sex distribution of whites in every year.



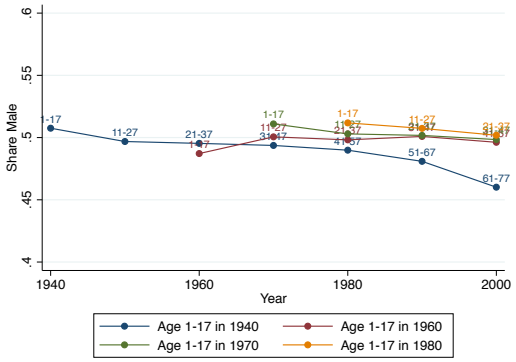
(a) Log Earnings, 1940-2000



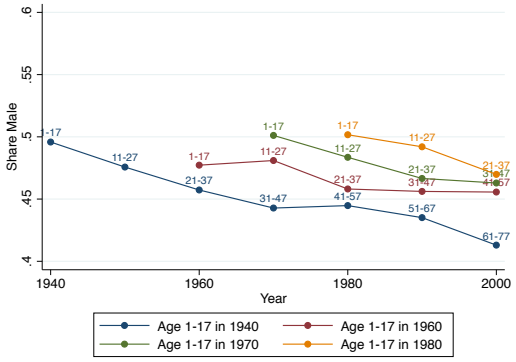
(b) Imputed Log Earnings, 1880-2000

Figure A.2: Log Earnings of Men, CA-Born 1880-2000

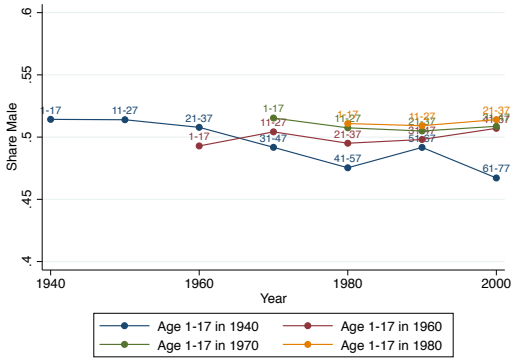
Notes: Restricting to individuals born in California. Panel (a) plots average log male earnings age 25-65 by race and year. Panel (b) plots average imputed log male earnings age 25-65 by race and year, with imputation based on earnings in 1940 averaged by cells defined by OCC1950 and native-born status, excluding observations with zero earnings or missing occupation, and restricting to household heads. Residents of Hawaii and Alaska excluded and races reweighted to match age distribution of whites in every year.



(a) White



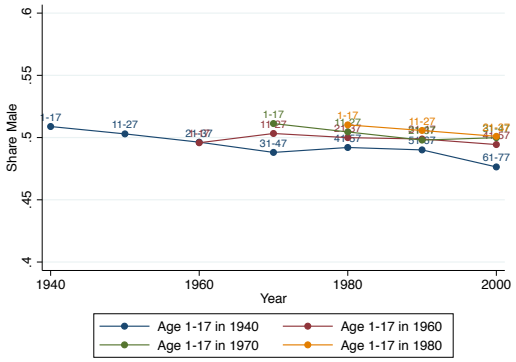
(b) Black



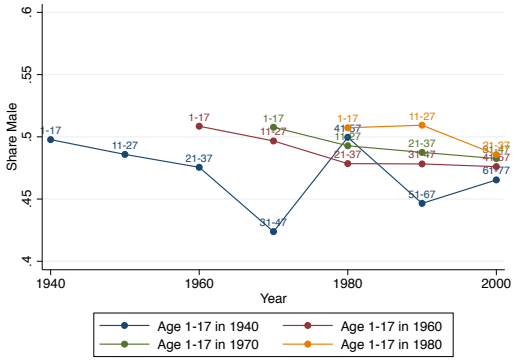
(c) Asian

Figure A.3: Male Share in Pseudo-Cohorts: Born in U.S.

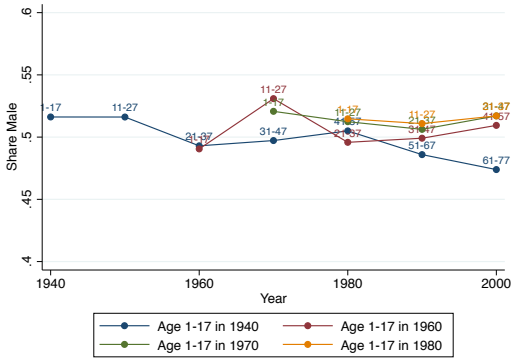
Notes: Figure plots male share by race for cohorts of native-born individuals age 1-17 in 1940, 1960, 1970, 1980.



(a) White



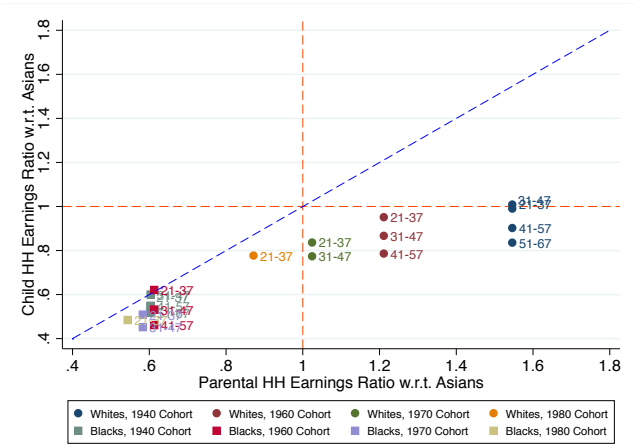
(b) Black



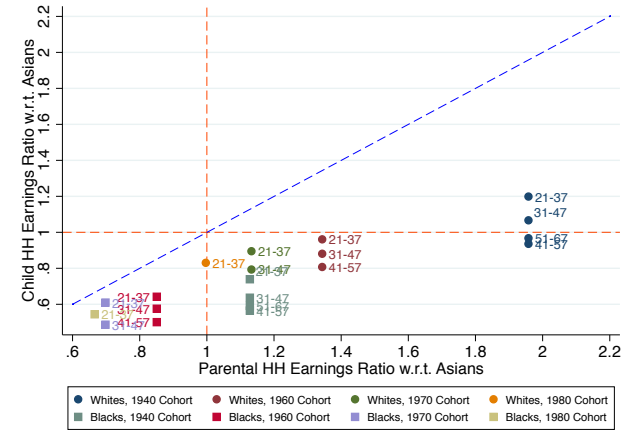
(c) Asian

Figure A.4: Male Share in Pseudo-Cohorts: Born in CA

Notes: Figure plots male share by race for cohorts of individuals born in CA ages 1-17 in 1940, 1960, 1970, 1980.



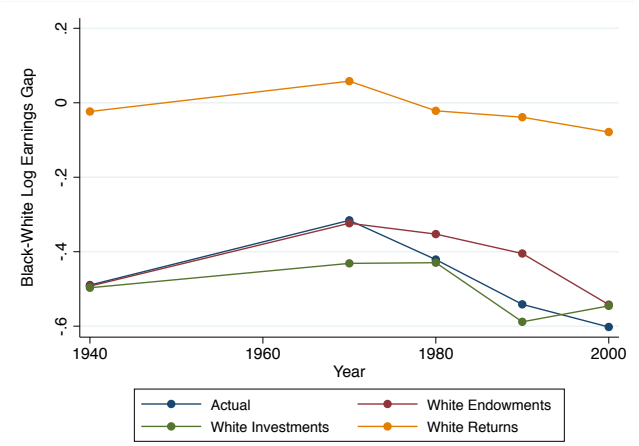
(a) Born in US



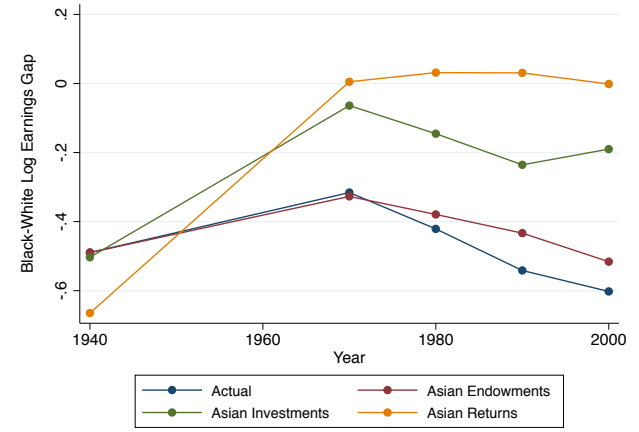
(b) Born in CA

Figure A.5: Intergenerational Change in Earnings Ratios, with Imputations

Notes: Replicates Figure VIII using household earnings with imputations for zero and missing values as described in text.



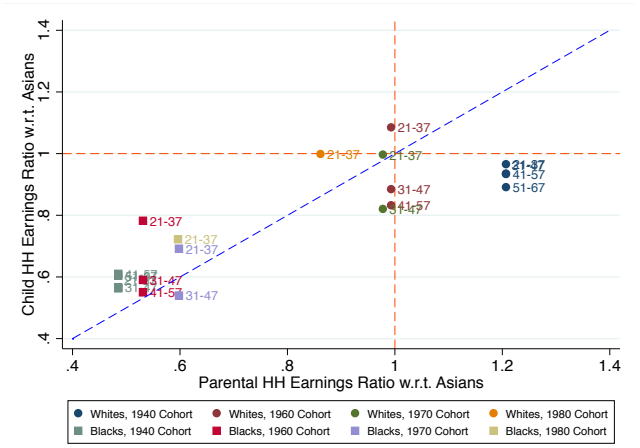
(a) Assign White Components to Blacks



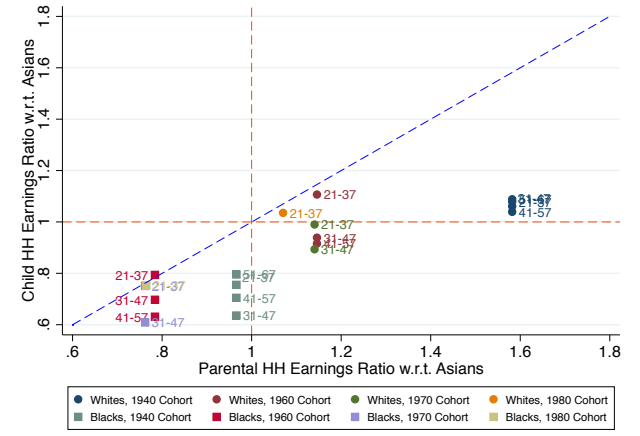
(b) Assign Asian Components to Blacks

Figure A.6: Counterfactual Black-White Log Earnings Gaps in CA, 1940-2000

Notes: Replicates Figure XIII using household earnings with imputations for zero and missing values as described in text.



(a) Born in US



(b) Born in CA

Figure A.7: Intergenerational Change in Individual Earnings Ratios

Notes: Replicates Figure VIII using fathers' and sons' earnings rather than total household earnings.