The Determinants of Credit Access in the United States*

Trevor Bakker, U.S. Census Bureau
Eric English, U.S. Census Bureau
Jamie Fogel, Harvard University and Opportunity Insights
Nathaniel Hendren, MIT and Opportunity Insights
Daniel Herbst, University of Arizona

March 19, 2025

Preliminary and Incomplete: Please do not cite or quote

Abstract

We construct new population-level linked administrative data to study households' access to credit in the United States. We document large differences in credit access by race and class. On average, Black individuals and those who grew up in low-income families have persistently smaller balances, more credit inquiries, higher utilization rates, and lower credit scores than other groups throughout the life cycle. We then study the determinants of these disparities. Biases in credit scoring algorithms do not appear to contribute to these gaps, as the scores actually under-predict differences in realized loan delinquencies by race and class. For example, among those with a 650 credit score, Black borrowers are 30% more likely to fall delinquent on a loan than White borrowers, and those who grew up in the bottom quintile of parental income are 27% more likely to fall delinquent than those who grew up in the top quintile. Measures of individuals' income profiles and wealth do little to explain these repayment gaps. As a result, narrowing credit access disparities requires addressing the factors that lead to differences in loan repayment outcomes. Differences in repayment patterns emerge shortly after borrowers obtain credit in their early twenties, generally on credit cards and student loans, suggesting that childhood and early-life experiences play a role. Using a movers design, we show that childhood exposure to places with high repayment rates causes an increase in one's own likelihood of repayment. Places that promote high-repayment are similar to those that tend to promote upward income mobility but differ in some respects (e.g., they have lower divorce rates). Our results suggest that credit constraints in adulthood have roots in childhood.

*Any opinions and conclusions expressed herein are those of the authors and do not represent the views of the U.S. Census Bureau. The Census Bureau has ensured appropriate access and use of confidential data and has reviewed the results for disclosure avoidance protection (Project 7519874: CBDRB-FY24-0345, CBDRB-FY24-CES026-013, CBDRB-FY25-0024, CDRB-FY25-0155, and CDRB-FY25-0167). We thank Raj Chetty, Winnie van Dijk, John Friedman, Peter Hull, Jesse Shapiro, and seminar participants at the Federal Reserve Bank of New York for helpful comments and discussions. We are very grateful to Asher Baraban for outstanding research assistance.

I Introduction

Access to affordable credit is widely recognized as a catalyst for economic advancement, enabling households to smooth consumption, invest in education or housing, and accumulate wealth. Since Becker and Tomes (1986), credit access has long been hypothesized as a key pathway to intergenerational mobility. Therefore, understanding and mitigating existing barriers to credit could open doors to economic opportunity, especially for those from disadvantaged racial[1] and class[2] backgrounds. However, efforts to study these barriers have often been hindered by a lack of data sources linking credit records to demographic and income information. Such data limitations have prevented researchers from precisely quantifying disparities in credit access by race and class, much less investigating their determinants.

In this paper, we construct an anonymized longitudinal dataset that combines credit bureau records with information from Decennial Censuses, Census surveys, and federal income tax returns for more than 20 million individuals. Using this dataset, we provide new evidence on disparities in credit access in the United States, with particular attention to race and class background. We then investigate potential drivers of these disparities, with a particular focus on (i) algorithmic bias in credit scores, (ii) racial and class differences in income and wealth, and (iii) credit market behaviors in adulthood that are shaped by childhood environments.

We begin by providing new population-level measures of disparities in access to credit.[3] Over 90% of American adults have a credit file—a share that varies little across demographic groups. Among those who have credit files, however, credit scores differ dramatically by race and class background. In 2020, the average Vantage 4.0 credit score is 719 for White, 621 for Black, 641 for American Indian and Alaskan Native (AIAN), 670 for Hispanic, and 745 for Asian individuals. These gaps emerge before age 30 and remain largely stable throughout the life cycle. For example, the White-Black gap in mean credit scores is 97 points at age 30 and 92 points at age 60. The gaps are also larger in medians than in means. The median credit score is 743 for White, 604 for Black, 671 for Hispanic, and 768 for Asian individuals. In other words, the gap in median credit scores between Black and White individuals is around 140 points and exceeds 160 points when comparing Black and Asian individuals.

These disparities in credit scores by race are larger than those found in previous studies using zip code–based proxies for race (Garon, 2022) and are substantially larger than those found using data from mortgage applications or originations (Fuster et al., 2022; Bhutta, Hizmo and Ringo, 2024; Consumer Financial Protection Bureau Office of Research, 2021; Gerardi, Lambie-Hanson and Willen, 2021; Bhutta and Hizmo, 2021). We replicate both attenuation patterns in our data by aggregating geographies and restricting the sample to those who have a mortgage.

While previous research has examined racial gaps in credit, less attention has been paid to the role of class background. Following Chetty et al. (2020), we link individuals to their parents using dependent claiming on tax returns. We find modest disparities in having a credit score by class: 94.3% of individuals

[1]We measure race/ethnicity but refer to it as “race” throughout for brevity.

[2]For most analyses we define class based on parental income, though we also consider parental education levels.

[3]Our linkage approach lets us identify "credit-invisible" individuals who have no credit record with Experian, allowing us to examine the extensive margin of credit access.

with parents in the top quintile of the parental income distribution have a credit record at age 25 in 2020, compared to 90.1% in the bottom quintile. Among those with a credit record and a parent linkage, we find large disparities in credit scores: children from the bottom quintile have average credit scores of 615, compared to 725 for children from the top quintile. Similarly, children with a parent who completed a Bachelor's degree have credit scores that are, on average, 87.1 points higher than those whose parents did not complete high school.[4] These differences are roughly similar to the magnitudes of racial differences in credit scores but they do not fully account for these disparities. For example, Black individuals from the 90th percentile of parental income (around $150,000) have similar credit scores (around 660 points) as White individuals from the 25th percentile.

These disparities in credit scores align with large differences in balance sheet outcomes. On average, Black individuals in the 1978–1985 birth cohorts hold $32,720 in mortgage debt, $6,657 in auto debt, and $2,787 in credit card balances in 2020 (ages 35–42), compared to roughly double the average balances ($100,900, $8,804, and $5,078) for White borrowers of the same age. Likewise, 35-42 year-olds in 2020 born to parents in the bottom quintile of the parental income distribution hold average balances of $45,020 in mortgages, $7,078 in auto loans, and $3,118 in credit cards, compared to $154,200, $8,697, and $12,480 for those with parents in the top quintile of the income distribution.

These lower balances for disadvantaged groups are consistent with tighter constraints in the supply of credit, as opposed to lower demand for credit among these groups. We find higher rates of credit inquiries and higher credit card utilization rates for Black individuals and those from low-income families. We also find a starkly different pattern of borrowing when examining student loans, which are uniformly supplied to college enrollees regardless of credit score or other measures of creditworthiness. On average, Black borrowers held nearly twice as much student debt in 2020 as White borrowers ($18,880 versus $10,020), despite lower rates of college attendance. We also see evidence that these credit constraints arise when individuals first enter the credit market. Most individuals obtain their first line of credit or "tradeline" as either a student loan or credit card when they are between 18 and 20 years old. Comparing by race and class, we find that the total numbers of tradelines follow similar patterns in the early 20s as people initially obtain credit, but large disparities in the numbers of tradelines by race and class emerge by the mid-20s, aligning with the same time period when the gaps in credit scores widen. In sum, the large disparities in balances we observe appear to reflect supply constraints that are mediated by the credit score.

One possible driver of disparities in credit access by race and class background is that the credit scoring algorithm is biased against certain groups. Credit scores aim to predict the incidence of a future 90+ day delinquency. In forming these predictions, credit scores cannot be directly discriminatory because they do not rely on information about race or class.[5] Nonetheless, a recent literature has raised concerns that seemingly neutral prediction algorithms could generate unwarranted disparities (Elzayn et al., 2025; Black et al., 2022; Passi and Barocas, 2019; Barocas and Selbst, 2016; Kleinberg et al., 2018; Obermeyer et al., 2019).

[4]Children with parents who have higher levels of education also have higher credit scores, even after controlling for parental income. Children who have a parent who completed their Bachelor's degree have a credit score that is 53 points higher on average than for children whose parents did not complete high school for those with parental family income around the 25th percentile.

[5]Even if there is no direct discrimination in the credit score, it is important to note that there could be discrimination in the lending markets that combine the scores with human decisions about whether or not to lend (Butler, Mayer and Weston, 2023; Lanning, 2021).

The existing literature has focused on two distinct definitions of algorithmic bias (Kleinberg, Mullainathan and Raghavan, 2016). The first is *calibration*: conditional on a given credit score, do borrowers of different groups have the same rate of delinquency? The second is *balance*[6]: among those with the same realized delinquencies, did members of different groups receive similar credit scores? We measure each type of bias of the credit scoring system in the US.

From a *calibration* perspective, we find evidence that credit scores under-predict differences in delinquency by race and class. Among borrowers with a credit score of 650,[7] approximately 61% of Black borrowers fall delinquent in the subsequent four years, compared to just 47% for White borrowers and 40% for Asian borrowers.[8] In other words, Black borrowers are 30% more likely to fall delinquent relative to White borrowers and 54% relative to Asian borrowers. Thus, the already large racial gaps in credit scores would be even larger if the credit score accurately predicted delinquency within race. We find similar patterns by class: among individuals with a 650 credit score, 55% of borrowers from the bottom quintile of the parental income distribution fall delinquent in the following four years, compared to 42% for those from the top quintile.[9] In other words, from a calibration perspective, credit scoring appears to be biased against those from more advantaged backgrounds.

From a *balance* perspective, however, we find that credit scores are biased against those from disadvantaged backgrounds. Among borrowers with no future delinquencies, Black individuals are on average assigned credit scores that are roughly 40 points lower than White individuals. Similarly, those whose parents were in the bottom quintile of the income distribution are assigned credit scores that are roughly 60 points lower than those from the top quintile of the parental income distribution. These disparities persist even if we restrict to borrowers who have "spotless" credit records before and after the score is measured. Among those who never make a late payment between 2000 and 2020, Black individuals have credit scores that are on average 15 points lower in 2012 than White individuals, and those with parents in the bottom quintile have credit scores that are roughly 10 points lower than those from the top quintile. These patterns suggest that those from disadvantaged groups face a more difficult lived experience with the credit market (e.g. higher interest rates) even when restricting to those who perfectly repay their loans.

It is well known that no algorithm can be unbiased under both calibration and balance definitions (Kleinberg, Mullainathan and Raghavan, 2016). Our results highlight the importance of precisely defining the bias concept applied to a given setting, as the conflicting implications we find in our data are likely to manifest in other contexts: In the absence of direct information on race or parent background, prediction algorithms are likely to capture some, but not all, of the group-level differences in underlying risk. The

[6]The balance condition is sometimes referred to as "equal-odds" in the computer science literature (Berk et al., 2017; Zafar et al., 2017). This condition is satisfied when false negative and false positive rates are equal across groups.

[7]We condition on a 650 credit score as this falls roughly in between the average credit score for Black and White borrowers. The calibration biases we identify are smaller at lower credit score values and larger for higher credit score values.

[8]These findings echo those of Bhutta and Canner (2013) and Bayer, Ferreira and Ross (2016), who use linked Home Mortgage Disclosure Act (HMDA) data to show that Black borrowers have higher mortgage delinquency rates than White borrowers with similar credit scores, as well as Dobbie et al. (2021) who find evidence of lending bias against immigrant and older populations in the United Kingdom. By contrast, Blattner and Nelson (2021) argue that credit scores do not exhibit calibration bias for marginal home loans—a finding we do not reproduce in our population-level data.

[9]These results are robust to alternative scoring methods derived from predicted delinquencies, suggesting they are driven by the information in the credit file rather than the Vantage 4.0 scoring algorithm.

portion captured by the algorithm's predictions generates an imbalance in favor of the lower-risk group, whereas the residual difference generates miscalibration in favor of the higher-risk group. In other words, credit scores are destined to be "biased" according to one of these definitions if there are systematic differences in delinquency risk across groups. Therefore, the key to understanding sources of disparities in credit scores and their bias requires understanding the drivers of underlying disparities in loan repayment.

To assess this, we start by leveraging the linkage of our data to the panel of each individual's income taxes to explore the role of differences in annual incomes in driving repayment. As one might expect, those with higher income are more likely to repay, as are those who experience positive income shocks or greater future income growth. However, differences in incomes explain only a small share of the differences in repayment by race and class. Indeed, we find large race and class gaps in repayment conditional on income even within the set of people who are continuously employed at a single employer with stable income growth. We also show that differences in wealth, proxied using information from tax returns (e.g. 401k contributions), data on housing wealth, or measured wealth in the Survey of Income and Program Participation, explain only a small share of the racial and class differences in repayment. The fact that income and wealth cannot explain the gaps we observe is perhaps not too surprising considering that income and wealth disparities by race and class emerge later in the life cycle (e.g. 30s and 40s), while differences in repayment emerge early in the life cycle (e.g. 20s). As a result, the key remaining question is: what drives persistent differences in repayment conditional on income that emerges early in the life cycle?

Childhood environments provide a potential explanation. We find a strong correlation between a parent's credit score and both a child's future credit score and early-life delinquencies. Even conditional on a child's credit score, child's income, and parent's income, a 1 point higher parental credit score on average corresponds to a roughly 0.1 percentage point lower likelihood of having a 90+ day delinquency for the child in their early 20s. In other words, 90+ day delinquency rates are 50% higher for children who grew up in families with a 600 credit score than 800 credit score, holding their own income, parents' income, and even their own credit score fixed.

Building on the growing body of work examining exposure effects to places in childhood, we document large variation in loan repayment depending on where one grew up. The ultimate paper release will report statistics on this loan repayment for each Commuting Zone (CZ) in the US by race and parental income background. Importantly, we provide estimates of the unconditional rates of loan repayment alongside estimates of the likelihood of loan repayment likelihoods conditional on one's income.

We then show that much of this variation in repayment outcomes reflects the causal effect of exposure to these areas during childhood. Using a childhood movers design as in Chetty and Hendren (2018), we find that over half of the geographic variation in repayment outcomes reflects the causal effect of childhood exposure to those areas. This is true both for unconditional repayment and repayment conditional on income. Spending one more year growing up in a place where people are 1 percentage point more likely to repay conditional on income causes one to have roughly 0.02 percentage point higher repayment conditional on income.

In general, places that promote loan repayment conditional on income tend to also promote upward income mobility, suggesting that some of the same conditions create both reliable workers in the labor market and dependable borrowers in the credit market. However, we also find notable differences between places that promote upward income mobility and places that promote credit market repayment. Growing up in counties where people are more likely to become divorced is the strongest predictor of neighborhoods causing non-repayment, but is a smaller correlate of upward income mobility. In total, our results suggest that credit market disparities are driven by differences in repayment patterns that have their roots in childhood.

Our work relates to five related strands of existing literature. First, our findings complement several studies that document racial disparities in credit scores. These analyses tend to rely on location-based proxy measures for race (Garon, 2022; Martinchek, 2024; Brevoort, Grimm and Kambara, 2015) or restrict attention to self-selected subsamples, like mortgage applicants (Bhutta, Hizmo and Ringo, 2024; Consumer Financial Protection Bureau Office of Research, 2021; Gerardi, Lambie-Hanson and Willen, 2021; Bhutta and Hizmo, 2021). Notable exceptions include Warren et al. (2024) and Stavins (2020), who link credit-bureau records to survey data on income and demographics, as well as Federal Reserve Board of Governors (2007), which merges TransUnion data to records from the Social Security Administration for a representative sample of Americans with credit files between 2003 and 2004.[10] All three studies report descriptive findings that align with our results. Relative to this literature, our contribution is to study these disparities in the full population, their life-cycle trajectories, the relative role of race and class, and the childhood determinants of these disparities.[11]

A second body of work studies the fairness and optimal design of algorithmic scores (Elzayn et al., 2025; Black et al., 2022; Passi and Barocas, 2019; Barocas and Selbst, 2016; Kleinberg et al., 2018; Obermeyer et al., 2019). Our results show how the different definitions of algorithmic bias illuminate different patterns and experiences in the credit market. Black and Hispanic borrowers are more likely to fall delinquent than White borrowers with the same credit score; but among the set of people who do repay their loans ex-post, Black and Hispanic borrowers have received ex-ante lower credit scores than White borrowers.

Relative to previous literature on the determinants of credit scores by race, our results relate to Blattner and Nelson (2021) who show that because they have thinner files, many minority borrowers with high repayment rates tend to be scored similarly with those who have lower repayment rates. Our results align with this finding and show how these thinner files perhaps endogenously arise from credit constraints induced in early adulthood from non-repayment patterns. More closely related to our analysis, (Fuster et al., 2022) use credit files linked to mortgage application data to show that Black borrowers would be more harmed by increased information used in credit scoring. This finding is consistent with our finding on the population data that credit scores understate racial gaps in non-repayment.[12]

[10]The Federal Reserve Board of Governors (2007) study was mandated by Section 215 of the Fair and Accurate Credit Transactions Act of 2003 (FACT Act). We replicate their findings in Appendix Table A.1.

[11]There is also a literature studying racial differences in credit offerings and take-up conditional on credit scores. This literature finds evidence of worse offers to Black and Hispanic borrowers in the mortgage market (Munnell et al., 1996; Bayer, Ferreira and Ross, 2016; Bhutta and Hizmo, 2021) and take-up of higher interest rate mortgages conditional on an offer set, with minorities being less likely to use points to reduce interest rates (Bhutta and Hizmo, 2021) or to exploit lower interest rates by refinancing or moving (Gerardi, Willen and Zhang, 2023). Moreover, mortgage lenders have a history of steering Black and Hispanic borrowers into subprime mortgages when non-Hispanic White borrowers with similar credit profiles received prime loans (U.S. Department of Justice, Office of Public Affairs, 2012). While these factors could affect some of the disparities we observe, we note that we observe gaps within those who solely take out student loans at age 20, which have terms fixed by the federal government.

[12]Additional work shows that changes in reporting rules or the removal of derogatory marks can have meaningful effects on credit access, although these interventions rarely eliminate disparities in risk and borrowing terms (Dobbie, Keys and Mahoney,

A third strand of literature measures the relationship between credit and intergenerational mobility. Using panel data on credit reports and a householding strategy to match parents and kids, Hartley, Mazumder and Rajan (2019) document strong intergenerational persistence in credit scores. Braxton et al. (2024) use credit data linked to the LEHD to study the causal effect of credit access of the parents on incomes of children in adulthood. Relative to this literature, we focus on the interplay between parental background and a child's credit access and repayment behaviors, and the childhood determinants of repayment behaviors.

A fourth strand of research studies the determinants of credit market behavior. This literature has documented racial and socioeconomic disparities in measures of financial literacy (Lusardi and Mitchell, 2014, 2023) and found evidence that past experiences can shape behavior in credit markets (Malmendier and Nagel, 2011, 2016; Malmendier and Wachter, 2021).[13] Most closely to our work, Keys, Mahoney and Yang (2023) study the role of person versus place-based factors in generating late payments, collections, and bankruptcy. They study adults who move across locations and find minimal effects of adult location on late payments, and they conclude that most of the geographic variation is driven by person-specific factors. Our results suggest that these person-specific factors observed in adulthood are not innate characteristics of people, but rather are shaped through their childhood environments.

Finally, while this paper provides new evidence on algorithmic bias in credit scores, there is a related literature that finds biased or discriminatory lending decisions conditional on the credit score. It is well-documented that Black and Hispanic individuals have historically faced discrimination in US lending markets, leading to lower access to credit and higher interest rates than White borrowers of comparable creditworthiness (Munnell et al., 1996; Bostic, 1997).[14] Recent work also finds that minority applicants for auto loans face lower approval rates and higher interest rates than White applicants, perhaps because car dealerships exercise significantly more discretion over loan approval and rate-setting decisions than credit card or student loan originators do (Lanning, 2021; Butler, Mayer and Weston, 2023). Notably, Butler, Mayer and Weston (2023) do not find racial disparities in algorithmic credit card lending, even among the same borrowers for whom they find racial disparities in auto loans.[15] This finding adds to a growing body of evidence that racial disparities are more prevalent when there is more scope for human decision-making. Raymond (2024) finds that the entry of algorithmic investors into housing markets reduces racial dispari-

2017; Liberman et al., 2018).

[13] Several other papers further underscore that credit disparities arise from a complex interplay of financial structures, household resources, and behavioral responses. Fulford and Low (2024) highlight the importance of expense shocks as a driver of consumer loan delinquency and financial distress. Scott-Clayton and Li (2016) show that Black college graduates have larger student debt burdens than White graduates immediately after graduation, and this gap triples over the next few years due to differential repayment. Ganong and Noel (2019) document present-biased behavior among households facing UI benefit exhaustion. Finally, Brown et al. (2016) find that some financial literacy interventions modestly improve credit market outcomes such as delinquency and collections.

[14] Edelberg (2007) suggests that racial differences in loan terms among equally creditworthy borrowers have narrowed over time, and much of the remaining disparity may reflect average differences in borrowing choices conditional on the same set of offers. For example, Black mortgage borrowers may, on average, prefer higher interest rates over higher down payments relative to White mortgage borrowers. Bhutta and Hizmo (2021) also find that racial gaps in loan terms are explained by borrowers choosing different loan terms. Other research suggests these racial differences in credit choices may be driven by unobserved constraints, such as differences in the real or perceived costs of searching for favorable terms (Bartlett et al., 2022; Willen and Zhang, 2023; Gerardi, Willen and Zhang, 2023; Ricks and Sandler, 2022).

[15] A GAO report finds higher APR in majority Black zip codes: 2.17 points higher in majority Black neighborhoods than majority White neighborhoods. The gap falls to 0.72 points when controlling for credit score, presence of revolving debt, income, and state and month fixed effects (United States Government Accountability Office, 2023). See Table 7 on pages 86-87.

ties in home sales prices. Similarly, Bartlett et al. (2022) identify 40% less discrimination among FinTech lenders compared to traditional lenders. Most recently, Argyle et al. (2025) find that White bankruptcy trustees discriminate against minority filers more for Chapter 13 bankruptcies, where they have more discretion, than for Chapter 7 bankruptcies. Despite the roles played by both human decision-makers and credit scores in shaping borrower outcomes, our results suggest that changes to either lending practices or scoring algorithms will struggle to remove disparities in credit access. Eliminating racial and class disparities ultimately requires addressing delinquencies that emerge early in life and have roots in childhood environments.

The remainder of the paper is organized as follows. Section II provides more details on the data sources, the linkage, and the construction of the sample. Section III documents credit score gaps by race and class and their implications for household balance sheets. Section IV investigates how credit scores relate to delinquency, showing that they do not fully capture differences in default risk by race and class. Section V studies disparities in credit terms and their role in explaining disparities in repayment behavior. Section VI evaluates whether income or wealth can explain these disparities, and Section VII investigates the role of childhood environment and parental credit behavior. Section VIII concludes.

II Data

II.A Sample

We use an anonymized hashing procedure to link Experian credit bureau data with administrative and survey data housed at the U.S. Census Bureau. We focus our analysis on two primary samples of the US population: a population sample and an intergenerational sample.

Population Sample Our population sample is a random 1% sample of individuals with Social Security numbers (SSN) who appear in the US Census in 2000 or 2010.[16] We link these individuals to three Census administrative data sets: (1) the 2000 and 2010 Census short forms, which aim to cover the entire population; (2) the 2000 Census long form and 2005-2019 American Community Survey (ACS), which cover approximately one-sixth of households and 2.5% of households annually, respectively; and (3) federal income tax returns spanning 1979-2021, with complete coverage for 1994-1995 and 1998-2021 (U.S. Department of Commerce, Bureau of the Census, 2000, 2003, 2014). Then we use the procedure described in Appendix A to merge this sample to an individual panel of Experian credit records in 4-year intervals from June 2004-2020 using hashed SSNs. Importantly, Experian and Census agreed on the subset of SSNs corresponding to the sampling universe, which means that those who are unmatched are people who necessarily do not have a credit file in Experian.

[16]This drops individuals both because some SSN holders do not reside in the US and because the name and address information on the Census short forms are not able to be linked to Census administrative records. Census publications note that over 90% of people in the Decennial Census receive a Protected Identification Key (PIK) (Mulrow et al., 2011). PIKs are assigned by the Census Bureau using information such as SSNs, date of birth, names, and addresses. See Wagner and Layne (2014) for a detailed description of the PIK assignment process.

Intergenerational Sample The intergenerational sample mirrors the data construction in Chetty et al. (2020) to study intergenerational income mobility using Census data linked to tax return information. We start with all individuals with Social Security numbers who were born between 1978 and 1985. We again link Experian credit records in 4-year intervals from June 2004-2020 to the Census administrative data sets (1)-(3) discussed above. We then link each individual to their parents using the first parent to claim them as a dependent on a tax return in our data, again following Chetty et al. (2020).[17] To measure credit files for their parents, we use a 10% random sample of credit files for individuals born between 1935 and 1970 – a cohort range that includes the vast majority of the parents in our sample.

Both our intergenerational and population samples restrict to the subset of people who are captured by the decennial censuses or ACS and are able to be matched to their PIK. While the census rates are close to 100% for 2000 and 2010, only 90% of those who have a census form are able to be assigned a PIK that facilitates linkage to other datasets like tax returns. Relative to existing literature that often focuses on linkages to other datasets such as HMDA, the universe of people who obtain a PIK is a much more expansive sample. However, we note that the full set of "invisibles" in the credit market could be slightly larger than what we estimate in our paper. Nonetheless, we estimate that 16% of the U.S. adult population (18-85) did not have a credit file in 2020. Our estimates are broadly similar to Brevoort, Grimm and Kambara (2015), who estimate that in 2010 11% of the U.S. adult population lacked a credit file by comparing counts of credit records to population counts from the 2010 Census.

II.B Variable Definitions

Experian Credit Variables The data from Experian include a full snapshot of summary attribute information in four-year intervals pulled in June 2004 through June 2020. Each snapshot contains comprehensive information about credit balances, broken out by type of credit (credit cards, student loans, mortgages, auto loans, etc.). It also contains repayment information, including incidences of late payments of different levels (30 days, 60 days, 90 days, etc.), broken out by type of credit. We also observe information on inquiries, bankruptcy filings, ages of accounts, and other credit events.

In addition to information from the credit file, the Experian data include a VantageScore 4.0 credit score, which uses credit file information to measure individuals' creditworthiness on a scale of 300 to 850. VantageScore is one of several scoring methods intended to predict an individual's likelihood of falling 90 or more days delinquent on any line of credit within the next two years (Gibbs et al., 2024).[18] While the details of how these scores are constructed are not publicly disclosed, credit bureaus are legally prohibited from using information outside of the credit report, such as race, address, or even age, to determine one's score. To assess the robustness of our results to different scoring methodologies, we replicate our credit score analysis using our own predictions of 90+ day delinquencies and find similar patterns.[19] This suggests

[17]We limit our analysis to children born during or after 1978 because many children begin to leave the household at age 17 (Chetty et al., 2014) and the first year in which we have dependent claiming information is 1994. We then include the universe of birth cohorts up through 1985 to ensure all individuals have reached adulthood by 2004, the first year for which we have Experian data.

[18]For example, VantageScore 4.0 is closely related to the older, more commonly used FICO score, developed by the Fair Isaac Corporation (Gibbs et al., 2024).

[19]Note that, as with industry scoring methods, we predict a borrower's likelihood of falling 90 or more days delinquent on any

that the properties of the credit score that we identify are a structural feature of the predictive content of information in credit reports, not the exact scoring methodology.

Income, Education, and Race We follow Chetty et al. (2020) by measuring race using the information that individual respondents or a household member report on the 2000 and 2010 Census short forms and the ACS. We prioritize the 2010 Census short form, then the 2000 Census short form, and finally the ACS. We use these data to construct five main race groups—non-Hispanic White, non-Hispanic Black, Hispanic, non-Hispanic Asian, and non-Hispanic American Indian and Alaskan Natives (AIAN). In a similar sample to ours using individuals born between 1978 and 1992, Chetty et al. (2024) find that these 5 race groups comprise 97.3% of the children with non-missing race information in their similar sample. As in Chetty et al. (2024), there is considerable heterogeneity in outcomes within these five groups, and our conclusions may not apply uniformly to all subgroups within each of these populations. For simplicity, we use "White" to refer to non-Hispanic White individuals, "Black" to refer to non-Hispanic Black individuals, "AIAN" to refer to non-Hispanic American Indian and Alaskan Native individuals, and "Asian" to refer to Non-Hispanic Asian individuals.

In our intergenerational sample, we again follow Chetty et al. (2020) and measure parental income each year using the total pre-tax income of the primary tax filer and their spouse (if applicable), which we label family or household income. In years where a parent files a tax return, we define household income as the sum of Adjusted Gross Income, social security payments, and tax-exempt interest payments, as reported on their 1040 tax return. In years where a parent does not file a tax return, we define household income as W-2 income when available. Otherwise, we consider household income for non-filers to be zero. For our primary analysis, we define parental income in the child's youth as the mean household income over the five years in which the child is ages 13-17 (or the subset of those years for which we have tax data).

In addition to parental income, we also proxy for 'class' using measures of parental education. We obtain the educational status for children whose parents were surveyed by the ACS at some point between 2005 and 2019. We define parental education using the maximum of educational status across parents.

We also consider the role of an individual's adult income in shaping their repayment behavior. For those who file taxes, we define *household* income as the sum of Adjusted Gross Income, social security payments, and tax-exempt interest payments, as reported on their 1040 tax return. We define *individual* income as wage income reported on their W-2, in addition to self-employment and other non-wage income reported on their 1040 tax returns. We assign individuals who are married and filing jointly half of the self-employment and other non-wage income. For non-filers, we define both individual and household income as total wage earnings from W-2s, or as 0 if no W-2 is filed.

Wealth In addition to income data, we also use several measures of wealth. First, we use data from a data analytics provider which focuses on the real estate market. We use data from their tax assessment files, which are drawn from all public county tax assessments in the U.S. These data are then linked to valuation estimates, which are calculated based on recent transactions of similar properties. We use both

line of credit they currently hold. Presumably, prospective lenders find one's predicted repayment on existing loans informative of their likelihood of repayment on the marginal loan.

home value and home equity in our analysis. Second, we use wealth data from the Census Bureau's Survey of Income and Program Participation (SIPP), which is a nationally representative, longitudinal survey. The SIPP includes questions about all major asset categories, including home equity, retirement accounts, stocks, bonds, vehicles, etc., as well as corresponding questions on debt. We combine these measures to create net wealth values for all linked respondents in our sample. Following Choukhmane et al. (2024), we use deferred compensation from Box 12 of the IRS W-2 tax form, which includes savings for retirement like 401k and 403b contributions.

III Disparities in Credit Access

We begin by measuring disparities in who has credit in the US, their credit scores on their credit files, and explore how this translates into systematic differences in credit balances. We focus on differences by race and class background, but we also seek to quantify the wide unconditional variation in access to credit that we observe in the data.

III.A Credit "Invisibles"

We begin by documenting the share of individuals who have a credit file in the U.S. Figure I reports the fraction of the 2020 U.S. population that has a credit file at Experian by age, separately by racial group (Panel A) and parent income quintile (Panel B). Most individuals first obtain a credit file in their early 20s. Just 13.4% of individuals in the 2002 birth cohort have a credit file by 2020, but this rises to 90.7% for the 1990 birth cohort. These rates are persistent throughout the life cycle, rising to over 90% at age 65 in 2020.[20]

The rates of having a credit file varies by race and class. Figure I shows that Black and Hispanic individuals obtain credit files at similar rates in their 20s but plateau at a slightly lower fraction with a report in middle age: 95% of 35 year old Black and Hispanic individuals have credit files in contrast to 97% of 35 year old White and 96% of 35 year old Asian individuals in 2020. AIAN individuals have the lowest age 35 match rates, at 92%. Similarly, 93% of 35 year olds whose parents were in the bottom quintiles of the income distribution have a credit file at age 35 in 2020, in contrast to 98% for those whose parents were in the top quartile. 93.9% of 25 year olds in 2020 whose parents did not finish high school have a credit file, in contrast to 94.1% for 25 year olds in 2020 with a parent who completed a Bachelor's degree.

III.B Differences in Credit Scores

We next examine how credit scores vary by race and class. Not everyone with a credit file has a credit score, as some files are too "thin" for scoring. Appendix Figure A.2 reports the fraction of individuals who have a VantageScore 4.0 credit score. Among 35-year-olds in 2020, approximately 89% of AIAN individuals, 93% of Black individuals, 93% of Hispanics, 95% of Asians, and 96% of Whites had a credit score. For

[20] Appendix Figure A.1 reports these rates for other years (2004-2016). The fraction of people with credit files in the U.S. has increased slightly over time, but the broad patterns reinforce our interpretation of the patterns in Figure I as being a life cycle pattern.

individuals whose parents were in the top quintile, 98% have a credit score by age 35, in contrast to 91% for those with parents in the bottom quintile.[21] These patterns are broadly consistent with more limited credit use or access for those from disadvantaged backgrounds.

Turning to the set of people with credit scores, Figure II plots average 2020 credit scores by race and age in the 1% population sample. Panel A shows that the average credit score for White individuals is approximately 719, compared to about 621 for Black individuals —a gap of 98 points. Asians have the highest average credit scores with a mean of 745, while Hispanics and AIAN lie between Black and White individuals, with averages of 670 and 641, respectively.

Panel B of Figure II reports the average credit score by age and race. This shows that the gaps in credit scores emerge early in life and persist throughout the life cycle. For example, the mean difference in credit scores between White and Black individuals is 97 points at age 30 and 92 points at age 60.

Consistent with the strong persistence in scores we observe in the population across ages, Table I focuses on the credit score statistics for our intergenerational sample of individuals aged 35-42 in 2020. We find slightly lower scores for people aged 35-42 versus the overall population, consistent with Figure II. We also see that focusing on median scores, the race gaps are slightly larger. Among those with credit scores, we find medians of 719 for White borrowers, 581 for Black borrowers, 781 for Asian borrowers, and 655 for Hispanic borrowers. When focusing on medians, we can also extend our analysis to include those who do not have a credit score by assuming they have below-median access to credit. With this approach, we find median credit scores that are lower for all races, but similar gaps between races. White individuals have a median credit score of 712, in contrast to Black individuals with 571, Asian individuals with 776, and Hispanic individuals with 644. In other words, we find a 140 point gap in median credit scores between White and Black borrowers and a 205 point gap between Black and Asian borrowers at age 35-42.

The racial gaps we observe in Figure II and Table I are larger than those identified in previous literature. Several papers have found racial credit score gaps of roughly 10-60 points using HMDA data, which records both race and credit scores for mortgage applicants.[22] However, the set of borrowers who appear in HMDA are positively selected in terms of credit score because the population that applies for mortgages is a more creditworthy group than the general population. Indeed, Appendix Table A.1 reproduces the mean credit score for the subset of people with mortgages, generally finding smaller racial gaps (e.g. Black-White gap in medians closer to 79 points rather than the true gap of 139 points). More in line with our findings, Garon (2022) finds a median Vantage score gap of 105 points between 25 to 29 year-olds in majority Black and

[21]These findings echo those from the Consumer Financial Protection Bureau (CFPB), which uses credit bureau data and census tract demographic averages to study these patterns (Brevoort, Grimm and Kambara, 2015).

[22]Fuster et al. (2022) uses the subsample of individuals who have applied for mortgages and uses the HMDA filings to obtain race. They find median FICO scores of 774 for White non-Hispanic, 744 for Black and 775 for Asian borrowers. Although FICO and Vantage scores (which we use) are not directly comparable, these scores are naturally higher than our estimates — and the race gaps smaller — due to the sample selection into who applies for mortgages. Similarly, Bhutta, Hizmo and Ringo (2024) find a Black-White FICO score gap of 41 points using a HMDA sample of purchase and refinance applications in 2018 and 2019 for first-lien, 30-year fixed-rate mortgages, on owner-occupied single-unit homes for which an automated underwriting system recommendation was reported; Consumer Financial Protection Bureau Office of Research (2021) find Black-White credit score gaps of about 60 points on home mortgages in HMDA in 2018-2020[23]; Bhutta and Hizmo (2021) find a Black-White FICO score gap of 13 points in a data set of 30-year fixed rate FHA home purchase loans originated in 2014 and 2015 with loans amounts under 625,000 and LTVs over 90%, for single-family, site built, owner-occupied properties, Gerardi, Lambie-Hanson and Willen (2021) find a Black-White Vantage 3.0 score gap of 51 points using HMDA data.

majority White zip codes and Martinchek (2024) finds mean Vantage score gaps of 60-67 points among 20-29 year-olds in majority Black versus majority White zip codes. These raw gaps are qualitatively similar but slightly smaller than our results. Appendix Table A.1 replicates these patterns in previous literature and shows how these imputation methods generally attenuate measures of the race gaps in credit scores.

These disparities in credit scores that we observe across individuals at a given point in time persist within an individual throughout the life cycle. Appendix Figure A.3 splits individuals into credit score quintiles using their 2008 credit score and plots their average credit scores in the subsequent 16 years. While there is some mean reversion over time, roughly 73% of the gap between the first and fifth quintile persists after 12 years. The fact that credit score disparities emerge at early ages and persist throughout the life cycle suggests the gaps we observe in our intergenerational sample of individuals (who are 35-42 in 2020) linked to their parents are largely reflective of gaps over the lifetime. We therefore turn to this intergenerational sample to explore these gaps by not just race but also by class.

We find large gaps in credit scores by class that are of similar magnitude to those we find by race. Figure III presents credit score gaps by parental income for our intergenerational sample. The average credit scores of 615 for those from the bottom quintile of the parent income distribution and 725 from the top — and the gaps arise in one's early 20s and persist throughout the life cycle. Appendix Figure A.5 reports average credit scores by parental education level. We find credit scores of 638 for those with parents that have less than a high school education, in contrast to 725 and 742 for those with parents who have a bachelor's or graduate degree, respectively. In other words, the gap in credit scores by class are of the same order of magnitude as the gaps in credit scores by race.

In Figure IV, we explore the relative role of race and class. The horizontal axis presents the child's parental income rank, while the vertical axis measures the average credit score for each child with that parental income rank, with separate series for each race. Large racial gaps in credit scores persist conditional on parental income, albeit at an attenuated level. At age 25, White children from the twenty-fifth percentile of the parental income distribution have credit scores that are 69 points higher than their Black counterparts, compared to an unconditional Black-White gap of 95 points. White individuals from the bottom quintile have an average credit score of 655 in contrast to 762 for those from the top 5%. The gradients of credit scores with respect to parental income rank are similar for White, Black, Hispanic, and AIAN individuals, but diverge for Asian individuals: We find high credit scores for Asian individuals whose parents have relatively low incomes with an average around 730 even at the bottom of the income distribution, which remains higher than the 695 average credit score of children in Black families whose parents are in the top 5%. This phenomenon is largely attributable to the subset of first and second generation Asian immigrants in the US. Appendix Figure A.4 shows that when we restrict the sample to individuals who do not have immigrant mothers, the series for Asian individuals converges closely to the series for White individuals.

Several benchmarks can be useful to understand the magnitude and implications of these credit score disparities. A score of 600 corresponds to the cutoff between "near prime" and "subprime" and a score of 660 corresponds to the gap between "prime" and "near prime".[24] This means that the mean Black individual aged 35-42 is classified as near prime (and the median is subprime). In contrast, to the mean and median

[24]See https://www.vantagescore.com/the-complete-guide-to-your-vantagescore/

Asian and White borrowers aged 35-42 are classified as prime borrowers. The average interest rate for new car loans for borrowers with a near prime credit score is 9.83%, as compared to 6.87% for prime borrowers (Zabritski, 2024).[25] For a $20,000 auto loan with a term of 60 months, the difference between an interest rate of 9.83% versus 6.87% amounts to paying an extra $28 per month and a total difference in interest paid of $1,708. Similarly, for personal loans, borrowers in the 690–719 range pay average interest rates of 12.5% to 15.5% while those below 629 pay average rates of 28.5% to 32.0%.[26]

In sum, the results demonstrate large disparities in credit scores by race and class. We interpret these patterns as measuring heterogeneity in credit constraints, likely mediated by these credit scores. In the next subsection, we present evidence consistent with this interpretation.

III.C Differences in Debt Balances

Lower credit scores generally lead to higher interest rates or rejections of credit applications. To what extent do these differences in credit scores align with signs of higher credit constraints for borrowers? We start by studying differences in balances. Figure V Panel A explores differences in total debt in 2020 for our intergenerational sample. For each racial group, we present the average total debt held on their credit report in 2020, disaggregated by type of debt. Asian individuals have the highest balances at $190,500 followed by White individuals with $128,163 in loan balances. Mirroring the lower credit scores, Hispanic, Black, and AIAN borrowers have lower balances of $91,084, $62,770, and $52,064, respectively. Appendix Figure A.6 shows that these disparities persist throughout the life cycle: in 2004 we observe balances of $14,171 and $13,674 for White and Asian borrowers, in contrast to $10,920, $6,629 and $6,394 for Hispanic, Black, and AIAN borrowers.[27]

Credit market disparities emerge for each type of credit in a manner that is consistent with credit constraints. On average, Black and AIAN individuals in our intergenerational sample held $32,720 and $31,990 in mortgage debt in 2020, in contrast to $66,430 for Hispanic, $100,900 for White, and $161,000 for Asian borrowers. The lower auto and mortgage debts align with lower rates of homeownership and car ownership. For individuals in our intergenerational sample in the American Community Survey between 2015 and 2020, 89% of Black individuals with parents at the 25th percentile of household income owned a car and 41% owned a house, in contrast to 97% and 67% for White individuals.

We see similar levels of credit card debt across racial groups. However, Appendix Table A.2 shows that Black and Hispanic borrowers have much lower credit limits, so that their utilization rates are much higher than White and Asian borrowers (Appendix Figure A.7). On net, these patterns are consistent with credit constraints affecting the amount of auto, credit card, and mortgage balances taken out by Black and Hispanic relative to White and Asian borrowers.

The one exception to the pattern above is student loans. Black borrowers actually have higher student loan balances than any other racial group. Appendix Figure A.8 explores student debt in more detail and shows that this pattern is driven by an enormous rise in student debt loads for Black borrowers, espe-

[25] https://www.experian.com/blogs/ask-experian/average-car-loan-interest-rates-by-credit-score/

[26] https://www.bankrate.com/loans/personal-loans/interest-rate-statistics/

[27] Appendix Figure A.6 presents total balances by race and parental income, showing that we find lower balances for groups identified above to have lower credit scores.

cially Black women, that emerged from 2004-2020. The presence of higher student debt balances is again consistent with credit constraints in the formal credit market, as student loans are regulated by the federal government and are not underwritten using credit scores.

We also use data from the 1996 Beginning Postsecondary Survey to investigate the relative roles of initial borrowing amounts versus accumulated interest in explaining the different levels of student debt by race. Using a representative sample of student borrowers who entered college in 1996, Appendix Figure A.9 plots the average initial borrowing amounts along with the accumulated debt as of 2015 separately by race. While balances tend to decline for White, Asian, and Hispanic borrowers, we find that balances for Black individuals have increased, largely due to non-repayment. We return to this pattern when discussing non-repayment below.

Taken together, these higher debt loads held by Black borrowers are consistent with lower credit scores driving credit constraints: the one category of loan that is not underwritten by credit scores sees higher balances.

We see similar disparities in loan balances by class. Figure V Panel B shows individuals in the bottom quintile of the parental income distribution have $66,820 in balances, in contrast to $184,890 for those whose parents were in the top quintile of the income distribution — patterns that broadly mirror the differences in credit scores for these groups. Appendix Figure A.10 presents similar patterns by parental education. Finally, Appendix Figure A.11 shows that disparities in loan outcomes by race persist conditional on parental income — again consistent with the idea that individuals with lower credit scores face tighter credit constraints.

III.D Initial Credit

The life cycle path over which credit score balances disparities emerge also points towards those with lower scores facing larger credit constraints. Recall from Figures II and III that disparities in credit scores emerge early in the life cycle. This suggests that it is perhaps valuable to zoom into young adulthood to study what is happening at the entry points into the credit market and how these experiences vary by race and class. To begin this exploration, Figure VI plots the average number of tradelines by age, separately by race (Panel A) and parental income quintile (Panel B). People start entering the credit market in their late teens and early 20s, with large increases in borrowing at age 18 and 19. On average, individuals in 2020 have between 2-3 tradelines at age 21. However, the average number of tradelines diverge by race and class by age 30 in 2020. White individuals have 11 tradelines on average by age 30, in contrast to 10.3 for Black, 10.5 for Hispanic, and 11 for Asian borrowers.[28]

What are the types of credit that initially bring people into the credit market? Appendix Figure A.12 plots the fraction of the population that has a tradeline by a given age on the horizontal axis, separately by the type of credit they received as their first tradeline. Credit cards are effectively the only method of obtaining a tradeline before age 18, but by age 18 student loans become another common source of first tradelines (mortgages and auto loans are not common). Among the 38% of White individuals in 2020 who have received a tradeline by age 20, 55% of these individuals' first tradeline is a credit card and 33% is a

[28]The patterns by class are qualitatively similar, but the gaps are more pronounced.

student loan. In contrast, among the 42% of Black individuals in 2020 who have received a tradeline by age 20, 32% of these individuals' first tradeline is a credit card and 56% is a student loan. In summary, Black and low-income children are more likely to begin their credit histories with student loans, while White and high-income children are more likely to begin with credit cards.

The rise of student loans as a primary source of first tradeline is a relatively recent phenomenon. In 2004, only 15% of White individuals and 19% of Black individuals with a tradeline by age 20 had a student loan as their first tradeline. Alongside the rise of student loans, and perhaps as a result of them, we find that the Black-White disparities in credit score emerge slightly earlier in the life cycle in later cohorts where student loans are more likely to be present. Appendix Figure A.13, Panel A, shows this by repeating the graph of credit score by age and race, as in Figure II, but measures the credit score in 2004–2016 instead of 2020. Our race gaps in credit scores are similar at age 40 in both 2004 and 2020, but the gap among 25-year-olds is much wider in 2020. This is consistent with student loans leading to disparities in credit scores that emerge earlier in the life cycle. It also suggests that the rise of student loans led to credit score disparities arising earlier in the life cycle but does not appear to have large long-run impacts on credit scores.

Finally, we can explore how the demand or desire for credit — as proxied by credit inquiries — varies by race and class at early ages in the life cycle. Although the number of tradelines evolves similarly by race and class prior to age 30, Black and low-income individuals have higher rates of credit inquiries. Figure VII shows that at age 20 in 2020 the average Black individual has 1.2 inquiries in contrast to 0.8 for the average White individual. Children from the bottom quintile of the parent income distribution have 1.6 inquiries at age 21 in contrast to 0.7 for those from the top quintile of the parent income distribution. This further reinforces our interpretation above that the lower balances for Black individuals and those from low-income families are not driven by lower credit demand, but rather by credit constraints.

Summarizing our findings, we find large disparities in credit scores and credit balances that emerge early in the life cycle and are consistent with credit constraints mediated through credit scores. We focus the remainder of the paper on studying the determinants of these disparities in access to credit.

IV Algorithmic Bias: Credit Scores versus Loan Repayment

Credit scores appear to restrict access to credit, disproportionately for racial minorities and those from low-income and less educated backgrounds. It is therefore natural to ask whether these disparities in credit scores reflect bias on the part of credit scoring algorithms or true differences in the likelihood of loan repayment.

Credit bureaus are restricted in the type of information they can use as inputs into credit scores. They are not able to use race, income, address, or age and instead must rely on the information contained in people's credit files on things like types of credit, balances, inquiries, and repayment patterns. In this sense, credit scores are "race-blind" and cannot have "direct discrimination" in the language of Bohren, Hull and Imas (2022). However, there is growing recognition that seemingly neutral algorithms can generate unwarranted disparities (Elzayn et al., 2025; Black et al., 2022; Passi and Barocas, 2019; Barocas and Selbst, 2016; Kleinberg et al., 2018; Obermeyer et al., 2019; Bohren, Hull and Imas, 2022). For example, it is widely known that adding a child as an "authorized user" to a credit card on average increases the child's credit score (Bach et al., 2023). Appendix Figure A.14 replicates those findings in our setting and further shows that Black and Hispanic children are half as likely as White individuals to have these types of authorized user trades. On the other hand, structural racism faced by racial minorities could adversely affect future repayment outcomes conditional on their credit score. The direction of any potential bias in credit scores is therefore an empirical question.

In order to measure algorithmic bias towards a racial group or class background, it is important to precisely define what one means by algorithmic bias. As noted in Kleinberg, Mullainathan and Raghavan (2016), there are two definitions in the literature: calibration and balance. Each of these captures a distinct notion of bias, and as we show, these distinctions matter in practice.

Calibration We begin with *calibration*. A credit score is calibrated if rates of delinquency are the same between groups conditional on the credit score. To be precise, let $Y \in \{0,1\}$ denote a delinquency outcome, let R denote one's group membership, and let S denote the credit score. For expositional purposes, let's consider potential biases between White and Black borrowers: $R = Black$ versus $R = White$. The credit score is calibrated for Black and White groups if for all possible values of the credit score, s , we have that the rate of delinquency is the same for each group conditional on the score:

$$\Pr\{Y = 1|S = s, R = Black\} = \Pr\{Y = 1|S = s, R = White\}. \quad (1)$$

This definition conforms to notions of unbiasedness from Becker (1971) and subsequent empirical work using outcomes tests to detect bias in both algorithms and human decision-making (Obermeyer et al., 2019; Knowles, Persico and Todd, 2001; Simoiu, Corbett-Davies and Goel, 2017; Anwar and Fang, 2006; Benson, Li and Shue, 2023; Huang, Mayer and Miller, 2024; Arnold, Dobbie and Yang, 2018; Dobbie et al., 2021; Canay, Mogstad and Mountjoy, 2024). We reiterate that we focus here on the calibration bias of the algorithm (where Y is the target estimand for S – i.e. without loss of generality one could normalize the score to be $\Pr\{Y = 1|S = s\} = s$). Decisions made by decision-makers using the algorithm could introduce additional directions of bias due to (a) other information they might introduce into that decision and (b) they might be using the algorithm to judge applicants on an outcome that differs from Y (e.g. a new loan as opposed to the stock of current tradelines).

Figure VIII Panel A tests for the calibration of the credit score by race using our intergenerational sample in 2020. On the vertical axis, we plot the incidence of 90+ day delinquencies in the subsequent four years against the credit score on the horizontal axis, separately by race. Before focusing on any bias in the score, note that the broad patterns are as one might expect: higher credit scores predict lower future delinquency rates for all racial groups. Delinquency rates fall from nearly one with credit scores below 500 to close to zero with credit scores above 800. However, the non-overlapping lines for each group reveal that the credit score is not calibrated by race: Black borrowers have consistently higher delinquency rates than White and Asian borrowers with the same credit score. Hispanic and AIAN borrowers' delinquency rates fall between Black and White borrowers. Among individuals with a credit score of 650 (in between the Black and White median credit scores in 2016), 60% of Black borrowers have a 90+ day delinquency, in contrast to 45% for White, 40% for Asian, and 50% for Hispanic and AIAN borrowers. This means that among those with a 650 credit score, Black borrowers are ~30% more likely to fall delinquent on their loans than White borrowers and ~50% more likely to fall delinquent than Asian borrowers. Another way to interpret these patterns is that if credit scores were hypothetically modified to include the predictive information embodied in race but repayment patterns were held fixed, the Black-White gap in credit scores would rise from 100 points to ~130 points. This finding mirrors the conclusions of Fuster et al. (2022) who use HMDA data to argue more that predictive credit scores would disproportionately hurt Black borrowers.

We find a similar pattern when exploring patterns by class, as measured by either parental income or education level. Figure VIII Panel B plots the incidence of 90+ day delinquencies against the credit score, separately by parental income group. Roughly 42% of those with a 650 credit score from top quintile families have a subsequent 90+ day delinquency, in contrast to 55% of those whose parental income is in the bottom quintile. Appendix Figure A.15 plots the incidence of 90+ day delinquencies against the credit score, separately by parental education. We find that 53% of people with a 650 credit score with parents who lack high school degrees have a subsequent 90+ day delinquency, in contrast to 43% of those with at least one parent that has a Bachelor's degree. In short, credit scores systematically overestimate (underestimate) the repayment likelihood of those from disadvantaged (advantaged) groups.

Figure IX shows how the calibration bias varies by age. We regress 90+ day delinquency (2016-2020) on the credit score in 2016 along with indicators of racial group (using our population sample in Panel A) and parent income quintile (using our intergenerational sample in Panel B), separately by age. The figure then reports the coefficients at each age for Black, Hispanic, and Asian (White is the omitted category; we do not display the coefficient for AIAN due to sample sizes). Panel B reports the coefficients for quintiles 1, 2, 4, and 5, with the 3rd quintile as the omitted category.

In both panels, these results show that the credit score is less calibrated for younger ages when people are first obtaining credit. This is consistent with people having shorter credit files so that credit file information is relatively less predictive than one's socioeconomic background. Throughout the 20s as people build credit files, the predictive power of race conditional on the credit score declines while the observed gaps in credit scores expand. Intuitively, the repayment behaviors observed in the credit file start to capture some but not all information about one's likelihood of future delinquency that was contained in one's class and racial background.

The fact that the credit file becomes more informative throughout one's early 20s is consistent with further evidence in Figure VI. Recall that this figure showed that tradelines evolve similarly at young ages, but diverge later in one's 20s. Appendix Figure A.16 studies repayment patterns over those ages. We plot the incidence of 30-day late payments by race and age. At age 25 in 2020, 75% of Black borrowers have 30+ day late payments on their report, in contrast to 25% for Asian and 41% for White borrowers. 67% of individuals from the bottom quintile of the parental income distribution have a 30-day late payment, in contrast to 23% for those from the top quintile of the parental income distribution.

Broadly, our results suggest that credit scores are under-calibrated against the race and class heterogeneity in repayment, especially at young ages. These patterns are consistent with the credit score having greater difficulty in forecasting non-repayment for those in their 20s, and further aligns with the fact that we see smaller differences in initial credit availability by race and class. However, by one's late 20s the credit score captures some of the differences in repayment outcomes. The greater calibration of credit scores by race throughout the 20s coincides with the expanding racial disparities in credit scores. Consistent with credit scores then leading to credit constraints, we see disparities in balances and tradelines emerge at the same time as the calibration bias is reduced and credit score disparities increase in the late 20s.

All of our results pertain to the calibration of the Vantage 4.0 score vis a vis 90+ day delinquency in the subsequent 4 years. However, we find that our broad conclusions continue to hold for other race-blind scoring systems one could use to forecast future non-repayment,[29] and other measures of delinquency or non-repayment (e.g. 30-day, 60-day, collections indicators, different follow-up time frames, etc.). This suggests that the calibration properties of the credit score that we observe are not unique to the Vantage 4.0 score or any other race-blind scoring method, but rather reflect a fundamental property about the information contained in the credit file.

Balance The second definition of algorithmic bias is equal-odds or *balance* (Berk et al., 2017; Zafar et al., 2017; Bohren, Hull and Imas, 2022; Hardt, Price and Srebro, 2016; Chouldechova, 2017; Arnold, Dobbie and Hull, 2022). A credit score is balanced if it assigns similar scores on average to each group among those who ultimately have the same realized creditworthiness. To be precise, we condition on a particular repayment outcome, $Y = y$, and ask whether the algorithm assigned systematically different scores to different groups.[30] Mathematically, this corresponds to

$$E[S|Y = y, R = Black] = E[S|Y = y, R = White], \tag{2}$$

which can be evaluated within the set who repay $Y = 0$ and do not repay $Y = 1$. This definition has its roots in computer science and statistics and seeks to measure whether the error rates of prediction algorithms vary by group.[31]

To assess this, we consider the subset of borrowers who do not have a 90+ day delinquency between 2016 and 2020, and we study the heterogeneity in the credit scores they were assigned in 2016. Figure X plots the mean credit score by race, separately by age. Among those who do not experience a 90+ day delinquency at age 30 in 2016, Black borrowers are assigned credit scores that are 78 points lower than White borrowers, 91 points lower than Asian borrowers, and 37 points lower than Hispanic borrowers.

We find similar patterns by class (Figure XI): those from the bottom quintile at age 30 in 2016 are assigned credit scores that are 80 points lower than those from the top quintile of the parental income distribution. This means that within the set of people who do not make 90+ day late payments, those from more disadvantaged backgrounds have likely experienced worse terms in the credit market.

This measure of balance conditions on a particular outcome — in our case, the 2020 incidence of

[29]We ran a range of different prediction models to forecast these repayment outcomes using the credit file information and generated predictive scores for each. We find broadly similar patterns when controlling for these scores and studying their associated repayment outcomes. For example, Appendix Figure A.17 shows these patterns are similar when constructing our own prediction of 90+ day repayment in 2020 using the information in the 2016 credit file.

[30]For binary signals, this corresponds to testing whether rates of false negative and false positive classifications are the same across groups.

[31]If S were binary, this test would correspond to asking whether rates of false negative and false positive classifications are the same across groups.

any 90+ day late payment. The incidence of 90+ day delinquency is a relatively extreme adverse credit outcome. In practice, there might be differences by race and class in other outcomes such as 30 or 60 day delinquencies, which might suggest that conditioning on having no 90+ day delinquency does not equalize the ex-post creditworthiness of the borrower. Similarly, it could be the case that demonstrated repayment outcomes are required for the credit score to more accurately forecast future repayment and reduce potential biases. To explore this, in Figure XII we restrict to the subset of individuals in our data who have active credit through all years of our sample (2004-2020) but never have any recorded late payments or delinquencies. We then plot the average credit score in 2012, which means they have at least 8 years of perfect repayment before measuring the score and 8 years of perfect repayment after measuring the score. Within this set of borrowers, we find significant, but greatly attenuated, racial and class disparities in credit scores. Black borrowers have an average credit score of 773 at age 35, in contrast to 788 for White borrowers, 788 for Asian, and 779 for Hispanic borrowers. These gaps are similar across the age distribution. These patterns suggest that from the perspective of a borrower who perfectly repays, it is difficult to escape their group-level disparities in credit scores.

This finding that credit scores are calibrated against those from advantaged groups but biased from an equal-odds perspective against those from disadvantaged groups highlights the importance of being clear about one's definition of "bias". It is known that it is not feasible to have an algorithm that satisfies both calibration and balance (Kleinberg, Mullainathan and Raghavan, 2016). However, our results provide a case in which the ultimate conclusions about the presence of bias in the score depend on the framing of the question. Are White and Asian individuals less likely to fall delinquent than Black individuals with the same credit score? Yes. Among those who had perfect repayment of their loans, did Black and Hispanic people achieve this despite receiving lower credit scores along the way? Yes. Both are true. The key structural reason for these two facts is the underlying disparities in the extent to which people repay their loans on time. We devote the remainder of the paper to studying the determinants of loan repayment.

V Credit Terms

One factor that could lead to differences in repayment is that credit terms could systematically differ across groups. If people of different backgrounds face higher interest rates or other adverse lending terms, these could drive differences in repayment behavior.

There is a large literature discussing evidence of discrimination in lending markets, with evidence of declining but still significant racial discrimination in mortgage and, especially, auto lending markets. For example Lanning (2021) finds that Black borrowers pay significantly higher interest rates on auto loans and Butler, Mayer and Weston (2023) find that Black and Hispanic auto loan applicants' approval rates are 1.5 percentage points lower, even after controlling for creditworthiness. Similarly, Bartlett et al. (2022) find that lenders charge Latinx/African-American borrowers 7.9 and 3.6 basis points more for purchase and refinance mortgages respectively, costing them $765M in aggregate per year in extra interest. In Appendix Table A.3, we replicate these findings of disparities for auto and mortgage markets using data from the Survey of Consumer Finances (SCF). Black borrowers tend to face higher interest rates of 1.24 percentage points on auto loans and 0.63 percentage points on mortgage loans, even after conditioning on income (Appendix Table A.3, columns 5 and 6). Note that these gaps do not condition on credit scores (which are not available in the SCF), and are thus likely an overestimate of the gaps conditional on creditworthiness.

While there are differences in terms in the auto and mortgage lending markets, the non-repayment disparities we observe in our data emerge early in the life cycle before most borrowers have auto or mortgage loans. Non-repayment is concentrated among those with student loans and credit cards. We do not observe loan terms in our data, but terms on most student loans are set by the government. Similarly, credit card offerings are generally automated and not subject to loan officer discretion. Indeed, Butler, Mayer and Weston (2023) do not find racial disparities in credit card lending, even among the same borrowers for whom they find racial disparities in auto loans. Consistent with this, we find essentially no differences in credit card interest rates by race in the SCF, as shown in Figure XIII. Indeed, we find that Black borrowers pay slightly lower interest rates conditional on income and wealth. As a result, differences in credit terms do not explain the differences in non-repayment that we observe in the data.

VI Income and Wealth

A natural reason two people with the same loan might differ in their repayment is that one individual may not have the money to repay. There are well-known differences in incomes by race and parental income (Davis and Mazumder, 2018; Chetty et al., 2020). Previous literature has found settings in which differences in income or wealth can mediate credit market and consumption smoothing behavior (Ganong et al., 2020). In this section, we use the richness of our administrative tax data to explore the role of different incomes, income shocks, and income growth in explaining repayment.

We begin with our intergenerational sample and consider the repayment patterns that emerge early in the life cycle. We begin by asking whether these patterns are explained by differences in early-life income. Figure XIV Panel A focuses on Black and White individuals, plotting delinquency rates in the 2008 credit file (occurring between 2004-2008) as a function of the individual's 2004 income on the x-axis, conditional on their credit score. The pattern is largely flat, with a slight negative slope above the median income. Delinquency rates hover around 60% for Black individuals below the 80th household income percentile, in contrast to around 45% delinquency rates for White borrowers below the 80th percentile. While, we do observe delinquencies declining with income, especially above the median, large racial gaps in delinquency rates persist at every income level. Consistent with this, a simple re-weighting exercise shows that differences in income do not explain a significant portion of the Black-White gap in delinquency. We take the income distribution for White individuals and assign the Black repayment rates at each income level. This counterfactual average is reported in the dotted line on Figure XIV. Re-weighting the set of Black borrowers to have the same incomes as White borrowers (but maintaining their same repayment rate conditional on income), yields roughly the same delinquency rate. In this sense, differences in early-life incomes do not explain the differences in delinquency rates.

Panel B asks whether differences in income *shocks* can explain these differences. We know Black individuals and those from low-income backgrounds are more likely to experience negative income shocks over the life cycle (Morduch et al., 2019; Hoynes, Miller and Schaller, 2012; Kaila et al., 2021). Figure XIVB compares Black and White delinquency rates as a function of the percent change in incomes from 2004 to 2008 (excluding those with negative income in 2004 or 2008), conditional on their credit score. Those experiencing larger negative income shocks have higher delinquency rates than those who experience income increases. Moreover, Black individuals are more likely to experience negative income shocks, as shown by the histograms in Figure XIV. We repeat the same reweighting exercise we performed for income levels in Panel A. Re-weighting the set of Black borrowers to experience the same income shocks as White borrowers (but maintaining their same repayment rate conditional on income changes), yields roughly the same delinquency rate. Again, this is consistent with the fact that Black borrowers have higher rates of delinquency for any level of income shock.

Appendix Figure A.18 further extends this analysis to explore the role of longer-run future income growth from 2004-2020 on the horizontal axis, plotted against repayment in 2008 on the vertical axis. Here, we see that those who experience future income growth are more likely to repay their loans and White individuals are more likely to see greater future income growth. However, re-weighting Black borrowers' repayment rates by the income growth profile of White individuals barely changes the Black delinquency rate. In this sense, our measures of yearly incomes, 4-year income shocks, and 16-year differences in income growth do not appear to be driving the differences in loan repayment.

Table II extends this analysis to include the full vector of incomes experienced by individuals between 2003 and 2021. We also present results for all race and parental income groups. Broadly, we find similar patterns: the race and class gaps in repayment are not explained by different income trajectories.

Our tax data focus on annual incomes. This means that we omit potential non-repayment that is driven by variation in incomes within a year. Differences in within-year variability in incomes across race and class could potentially lead to differences in non-repayment. While we cannot observe within-year income variation directly in our data, Appendix Table A.4 presents results analogous to Table II restricting to the subsample of individuals who are consistently employed with the same employer from 2015 to 2021 with the income rank in each year staying within 10% of their income rank in 2015. On this sample, we continue to see large differences in repayment by race and class conditional on the credit score.

While differences in income cannot explain the race and class gaps in repayment, it is well known that there are class and race differences in wealth conditional on income (Darity Jr et al., 2018; Aliprantis, Carroll and Young, 2022; Thompson and Suarez, 2015). Thus, another hypothesis is that wealth, not income, is driving differences in repayment. We therefore consider two approaches to controlling for wealth in our analysis in Table II. First, we use administrative data to construct measures of wealth. This includes data on housing equity and cumulated 401k contributions from W-2s. Column (6) shows the effects of including these controls. Second, we construct a direct measure of wealth and assets using data from the SIPP. The SIPP contains information on all financial and non-financial assets. Column (8) shows that controlling for these measures of wealth does not meaningfully affect the gaps in repayment relative to controlling for credit score alone.

In summary, we find that differences in repayment by race and class cannot be explained by differences in measured income using tax data or administrative or survey-based wealth data. Beyond the race and class gaps, income in general does not explain much of the variation in repayment outcomes after controlling for credit score ranks.

VII Childhood

Differences in the budget constraint in adulthood are unable to explain a significant portion of the differences in repayment behavior that we observe. If it is not the budget constraint, what explains these large differences in repayment behavior we observe? One clue towards answering these questions is that these differences emerge early in life. As we showed in Figures II, III, and IX above, the credit score and repayment gaps emerge early in the life cycle. This suggests perhaps factors in childhood might drive some of these patterns.

As a first test of this, we study whether the parent's credit score predicts repayment conditional on the child's credit score and measured incomes. Figure XV presents a binned scatterplot of the relationship between having a 90+ day delinquency in young adulthood between 2004 and 2008 and parental credit score (measured in our sample in 2004) for the 1981 birth cohort, conditional on the child's income and credit score and parent's income, all measured in 2004.[32] This shows that children of parents with higher credit scores are significantly less likely to fall delinquent on their loans, even after conditioning on their own income and that of their parents. For children with average income and credit scores in 2004, and average parental incomes in childhood, 60% of those whose parents have credit scores around 600 experience a 90+ day delinquency, in contrast to 40% for children whose parents had credit scores near 800.

Table III shows how the slope in Figure XV varies with additional controls. We find similar slopes even if one adds further controls for parental education and measures of 401k and housing wealth. Broadly, these patterns suggest that the factors generating repayment behavior may have their roots in childhood environments.

Our next step towards exploring the role of childhood moves to the level of place, with the aim of identifying differences in childhood experiences. It turns out the likelihood of repaying a loan depends significantly on where one grew up. In measuring this variation, we build on earlier work that documents wide variation in intergenerational income mobility based on where children grew up (Chetty et al., 2014, forthcoming). Like that work, the final version of this paper will contain maps that illustrate the likelihood of repayment based on where children grow up. We plan to present results for a range of credit market outcomes, including credit scores, repayment likelihood, and repayment conditional on income. We will release each of these measures as a function of race and as a continuous function of parental income using the nonlinear spline methods in Chetty et al. (forthcoming). This will enable us to discuss in more detail both (a) how rates of repayment vary by geography in general and (b) how race and class gaps vary across place based on where people grew up.

We are not the first to document geographic variation in delinquency rates. Keys, Mahoney and Yang (2023) shows that loan delinquency rates vary significantly across areas of the US. They then study people who move across areas of the US when they are adults, and they show that people do not change their repay-

[32]See Hartley, Mazumder and Rajan (2019) for earlier work documenting the positive relationship between parent and child credit scores.

ment behavior very much. Moving to a place with a 1 percentage point higher delinquency rate increases delinquency by 0.05-0.10pp. This suggests that 90-95% of the geographic variation reflects a selection effect that does not change when adults move across areas.

We return to this idea but focus on exposure to place during childhood. We employ a movers design following the methodology of Chetty and Hendren (2018) and Chetty et al. (2024). This approach leverages variation in the age at which children move across counties to identify the causal effects of childhood exposure to a place on later-life outcomes. Relative to Keys, Mahoney and Yang (2023), our results ask whether the selection effect identified in their work is perhaps a causal effect of childhood exposure to where they grew up.

To conduct this childhood movers analysis, we split our sample into those that move across counties exactly once during our sample window and those who do not. For the non-one-time movers, we construct the average outcome, $\bar{y}_{cps}$, for people who grew up in county c with parental income p_i in birth cohort s_i. We construct these separately for a range of credit outcomes, including 90+ day delinquency, credit score, and the residual after regressing 90+ day delinquency on income.

Following Chetty et al. (2024), we take the set of people who move one time in childhood from origin o to destination d, and we consider a regression of the following form:

$$y_i = \sum_{m=-9}^{35} I(m_i = m)[\beta_m \Delta_{odps} + \alpha_m + \phi_m \bar{y}_{ops} + \zeta_m p_i] + \varepsilon_i \tag{3}$$

where y_i is the credit outcome of interest, m_i is the child's age at the time of moving, Δ_{odps} is the difference in mean outcomes, $\Delta_{odps} = y_{dps} - y_{ops}$ between the destination and origin counties for non-one-time movers in cohort s with parental income rank p.

The key idea of this regression is that where people move may not be orthogonal to other determinants of y_i. But under the assumption that these selection effects do not vary with the age of the child at the time of the move, the difference $\beta_{m-1} - \beta_m$ captures the causal effect of spending age m in a location with a one unit higher predicted outcome, as measured by Δ_{odps}. This assumption could of course be violated in practice. For example, perhaps good parents go to good places disproportionately when their children are young. Additional robustness analyses, such as including family fixed effects so that one compares exposure between siblings will help support our conclusions below.

Figure XVI Panel A presents the coefficients β_m for the outcome of repayment around age 30. We observe parental moves in our sample ranging from 9 years prior to birth through age 35. The coefficients largely hover around 0.9 to 1 prior to birth, consistent with movers looking similar to non-movers and no effects on children prior to their birth. However, we observe exposure effects that emerge after birth and persist through the early 20s. The results suggest that 1 year of exposure to a place with a 1 percentage point higher repayment rate causes the child to have a 0.015 higher repayment rate at 30 years of age.

One channel through which places might be affecting repayment is by increasing their incomes, as opposed to changing repayment conditional on income. To understand this, we let y_i denote the residual from regressing non-repayment measured in a four year window around age 30 on the individual's income rank in the year immediately prior to the four year window, and we analogously construct $\bar{y}_{ops}$ and Δ_{odps} using this outcome. Figure XVI Panel B then presents the resulting coefficients for β_m. We find a very similar pattern relative to looking at delinquency alone. Overall, this suggests that the presence of a causal effect of childhood exposure to places where other children grow up to be more likely to repay loans, with a causal effect not just on repayment in general but also repayment conditional on one's income in adulthood.

Characteristics of Places Promoting Repayment The variation in repayment based on where people grow up largely reflects the causal effect of childhood exposure to these places. In this final section, we study the characteristics of places that tend to promote loan repayment with a lens towards understanding the mechanisms driving these effects.

We do find some interesting differences in the types of places that promote upward income mobility versus loan repayment. The strongest correlate of loan repayment we have found is the divorce rate of people growing up in the area. This remains true even if we look at the repayment rates of people who are not married (and thus not facing direct effects of divorce). Figure XVII Panel A presents a binned scatter plot of the relationship between the average 90+ day delinquency rate and the divorce rate both estimated at the 25th percentile of the parent household income distribution in the county.[33] We measure delinquency rates among individuals who never marry to rule out a mechanical relationship between divorce and delinquency. A 1 percentage point higher divorce rate leads to a 0.72 percentage point increase in delinquency. While other place characteristics correlated with divorce rates could be driving the causal effect of these places, it is suggestive that a potential mechanism is household stability or commitment to dynamic contracts more broadly.

In future work, we will release all data by the county level and provide correlations with a wide range of other measures including social capital measures and financial literacy measures by place.

VIII Conclusion

We provide new measures of population disparities in access to credit. We document wide disparities in access to credit by race and class that emerge at young ages and persist throughout the life cycle. Access appears to be mediated through credit scores, yet our results suggest credit scores actually understate true differences in repayment rates by race and class. Differences in repayment are correlated with income, but income only explains a relatively small fraction of repayment. Instead, we show that these differences in repayment are causally affected by differences in childhood environment.

From a policy perspective, our results suggest that there is no easy change to the credit scoring system that would reduce barriers in access to credit. Indeed, more "accurate" credit scores would likely exacerbate racial gaps in credit access. Instead, our results point to a greater focus on childhood and early-life periods that seem to be shaping future repayment outcomes. More broadly, our results align with conclusions from broader work on intergenerational mobility that suggest a key role of social capital and trust in shaping economic outcomes. Places that promote greater repayment in credit conditional on income are also places

[33]Panel B restricts to White individuals, yielding similar patterns.

that promote greater upward mobility in general. In light of these shared attributes, efforts to improve upward economic mobility may therefore reduce disparities in access to credit, and vice versa.

In conclusion, we hope our findings help inform policy aimed at broadening access to financial markets and underscore the role that childhood factors play in shaping lifelong disparities in credit access.

References

Aliprantis, Dionissi, Daniel Carroll, and Eric R Young. 2022. "The dynamics of the racial wealth gap."

Anwar, Shamena, and Hanming Fang. 2006. "An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence." *American Economic Review*, 96(1): 127–151.

Argyle, Bronson, Sasha Indarte, Benjamin Iverson, and Christopher Palmer. 2025. "Racial Disparities and Bias in Consumer Bankruptcy." National Bureau of Economic Research Working Paper 33575.

Arnold, David, Will Dobbie, and Crystal S Yang. 2018. "Racial bias in bail decisions." *Quarterly Journal of Economics*, 133(4): 1885–1932.

Arnold, David, Will Dobbie, and Peter Hull. 2022. "Measuring Racial Discrimination in Bail Decisions." *American Economic Review*, 112(9): 2992–3038.

Bach, Helena, Pietro Campa, Giacomo De Giorgi, Jaromir Nosal, and Davide Pietrobon. 2023. "Born to be (sub) prime: An exploratory analysis." Vol. 113, 166–171, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Barocas, Solon, and Andrew D Selbst. 2016. "Big data's disparate impact." *Calif. L. Rev.*, 104: 671.

Bartlett, Robert, Alan Morse, Richard Stanton, and Nigel Wallace. 2022. "Consumer-Lending Discrimination in the FinTech Era." *Journal of Financial Economics*, 143(1): 188–223. Also available as NBER Working Paper No. 25943.

Bayer, Patrick, Fernando Ferreira, and Stephen L. Ross. 2016. "The Vulnerability of Minority Homeowners in the Housing Boom and Bust." *American Economic Journal: Economic Policy*, 8(1): 1–27.

Becker, Gary S. 1971. "The Economics of Discrimination." *University of Chicago Press Economics Books*.

Becker, Gary S., and Nigel Tomes. 1986. "Human Capital and the Rise and Fall of Families." *Journal of Labor Economics*, 4(3): S1–S39.

Benson, Alan, Danielle Li, and Kelly Shue. 2023. "Potential and the Gender Promotion Gap."

Berk, Richard, Hoda Heidari, Shahin Jabbari, Matthew Joseph, Michael Kearns, Jamie Morgenstern, Seth Neel, and Aaron Roth. 2017. "A convex framework for fair regression." *arXiv preprint arXiv:1706.02409*.

Bhutta, Neil, and Aurel Hizmo. 2021. "Do minorities pay more for mortgages?" *The Review of Financial Studies*, 34(2): 763–789.

Bhutta, Neil, and Glenn B Canner. 2013. "Mortgage market conditions and borrower outcomes: Evidence from the 2012 HMDA data and matched HMDA-credit record data." *Federal Reserve Bulletin*, 99(4).

Bhutta, Neil, Aurel Hizmo, and Daniel Ringo. 2024. "How much does racial bias affect mortgage lending? Evidence from human and algorithmic credit decisions." *FRB Philadelphia Consumer Finance Institute Working Paper WP*.

Black, Emily, Hadi Elzayn, Alexandra Chouldechova, Jacob Goldin, and Daniel Ho. 2022. "Algorithmic fairness and vertical equity: Income fairness with irs tax audit models." *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, 1479–1503.

Blattner, Laura, and Scott Nelson. 2021. "How costly is noise? Data and disparities in consumer credit." *arXiv preprint arXiv:2105.07554*.

Bohren, J Aislinn, Peter Hull, and Alex Imas. 2022. "Systemic Discrimination: Theory and Measurement." *NBER Working Paper w29820*.

Bostic, Raphael W. 1997. "The role of race in mortgage lending: Revisiting the Boston Fed study." *Board of Governors of the Federal Reserve System Finance and Econ. Disc. Series*, 97–2.

Braxton, J Carter, Nisha Chikhale, Kyle F Herkenhoff, and Gordon M Phillips. 2024. "Intergenerational mobility and credit." *NBER Working Paper w32031*.

Brevoort, Kenneth P., Philipp Grimm, and Michelle Kambara. 2015. "CFPB Data Point: Credit Invisibles." Consumer Financial Protection Bureau. This publication is part of the Consumer Financial Protection Bureau's Office of Research series.

Brown, Meta, John Grigsby, Wilbert Van Der Klaauw, Jaya Wen, and Basit Zafar. 2016. "Financial education and the debt behavior of the young." *The Review of Financial Studies*, 29(9): 2490–2522.

Butler, Alexander W, Erik J Mayer, and James P Weston. 2023. "Racial disparities in the auto loan market." *The Review of Financial Studies*, 36(1): 1–41.

Canay, Ivan A, Magne Mogstad, and Jack Mountjoy. 2024. "On the use of outcome tests for detecting bias in decision making." *Review of Economic Studies*, 91(4): 2135–2167.

Chetty, Raj, and Nathaniel Hendren. 2018. "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects." *Quarterly Journal of Economics*, 133(3): 1107–1162.

Chetty, Raj, John N. Friedman, Nathaniel Hendren, Maggie R. Jones, and Sonya R. Porter. forthcoming. "The Opportunity Atlas: Mapping the Childhood Roots of Social Mobility." *American Economic Review*.

Chetty, Raj, Nathaniel Hendren, Maggie R Jones, and Sonya R Porter. 2020. "Race and economic opportunity in the United States: An intergenerational perspective." *Quarterly Journal of Economics*, 135(2): 711–783.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States." *Quarterly Journal of Economics*, 129(4): 1553–1623.

Chetty, Raj, Will S Dobbie, Benjamin Goldman, Sonya Porter, and Crystal Yang. 2024. "Changing opportunity: Sociological mechanisms underlying growing class gaps and shrinking race gaps in economic mobility." *NBER Working Paper w32697*.

Choukhmane, Taha, Jorge Colmenares, Cormac O'Dea, Jonathan L Rothbaum, and Lawrence DW Schmidt. 2024. "Who benefits from retirement saving incentives in the US? Evidence on gaps in retirement wealth accumulation by race and parental income." NBER Working Paper w32843.

Chouldechova, Alexandra. 2017. "Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction." *Big Data*, 5(2): 153–163. Available at https://arxiv.org/abs/1703.00056.

Consumer Financial Protection Bureau Office of Research. 2021. "Data Point: 2020 Mortgage Market Activity and Trends." Consumer Financial Protection Bureau Research Report. Accessed on January 09, 2025.

Darity Jr, William, Darrick Hamilton, Mark Paul, Alan Aja, Anne Price, Antonio Moore, and Caterina Chiopris. 2018. "What we get wrong about closing the racial wealth gap." *Samuel DuBois Cook Center on Social Equity and Insight Center for Community Economic Development*, 1(1): 1–67.

Davis, Jonathan, and Bhashkar Mazumder. 2018. "Racial and ethnic differences in the geography of intergenerational mobility." *Available at SSRN 3138979*.

Dobbie, Will, Andres Liberman, Daniel Paravisini, and Vikram Pathania. 2021. "Measuring bias in consumer lending." *The Review of Economic Studies*, 88(6): 2799–2832.

Dobbie, Will, B Keys, and Neale Mahoney. 2017. *Credit market consequences of credit flag removals.* SSRN.

Edelberg, Wendy. 2007. "Racial dispersion in consumer credit interest rates." *Finance and Economics Discussion Series (FEDS) Working Paper*.

Elzayn, Hadi, Evelyn Smith, Thomas Hertz, Cameron Guage, Arun Ramesh, Robin Fisher, Daniel E Ho, and Jacob Goldin. 2025. "Measuring and mitigating racial disparities in tax audits." *Quarterly Journal of Economics*, 140(1): 113–163.

Federal Reserve Board of Governors. 2007. "Report to the congress on credit scoring and its effects on the availability and affordability of credit." Federal Reserve Board of Governors Research Report. Accessed on January 19, 2025.

Fulford, Scott, and David Low. 2024. "Expense shocks matter." *Consumer Financial Protection Bureau Office of Research Working Paper*, , (2024-08).

Fuster, Andreas, Paul Goldsmith-Pinkham, Tarun Ramadorai, and Ansgar Walther. 2022. "Predictably unequal? The effects of machine learning on credit markets." *The Journal of Finance*, 77(1): 5–47.

Ganong, Peter, and Pascal Noel. 2019. "Consumer spending during unemployment: Positive and normative implications." *American Economic Review*, 109(7): 2383–2424.

Ganong, Peter, Damon Jones, Pascal J Noel, Fiona E Greig, Diana Farrell, and Chris Wheat. 2020. *Wealth, race, and consumption smoothing of typical income shocks.* National Bureau of Economic Research Cambridge, MA.

Garon, Thea. 2022. "Young Adults' Credit Trajectories Vary Widely by Race and Ethnicity." *Urban Wire*. Accessed on January 09, 2025.

Gerardi, Kristopher, Lauren Lambie-Hanson, and Paul Willen. 2021. "Racial differences in mortgage refinancing, distress, and housing wealth accumulation during COVID-19." *Federal Reserve Bank of Boston Current Policy Perspectives*.

Gerardi, Kristopher, Paul S. Willen, and David Hao Zhang. 2023. "Mortgage prepayment, race, and monetary policy." *Journal of Financial Economics*, 147(3): 498–524.

Gibbs, Christa N, Benedict Guttman-Kenney, Donghoon Lee, Scott T Nelson, Wilbert H van der Klaauw, and Jialan Wang. 2024. "Consumer credit reporting data." NBER Working Paper w32791.

Hardt, Moritz, Eric Price, and Nathan Srebro. 2016. "Equality of Opportunity in Supervised Learning." *Advances in Neural Information Processing Systems (NeurIPS)*. Available at https://arxiv.org/abs/1610.02413.

Hartley, Daniel, Bhash Mazumder, and Aastha Rajan. 2019. “How similar are credit scores across generations.” *Chicago Fed Letter*.

Hoynes, Hilary, Douglas L Miller, and Jessamyn Schaller. 2012. “Who suffers during recessions?” *Journal of Economic perspectives*, 26(3): 27–48.

Huang, Ruidi, Erik J. Mayer, and Darius P. Miller. 2024. “Gender Bias in Promotions: Evidence from Financial Institutions.” *Review of Financial Studies*, 37(5).

Kaila, Martti, Emily Nix, Krista Riukula, et al. 2021. “Disparate impacts of job loss by parental income and implications for intergenerational mobility.” *Federal Reserve Bank of Minneapolis*.

Keys, Benjamin J, Neale Mahoney, and Hanbin Yang. 2023. “What determines consumer financial distress? Place-and person-based factors.” *The Review of Financial Studies*, 36(1): 42–69.

Kleinberg, Jon, Jens Ludwig, Sendhil Mullainathan, and Cass R Sunstein. 2018. “Discrimination in the Age of Algorithms.” *Journal of Legal Analysis*, 10: 113–174.

Kleinberg, Jon, Sendhil Mullainathan, and Manish Raghavan. 2016. “Inherent Trade-Offs in the Fair Determination of Risk Scores.” *arXiv preprint arXiv:1609.05807*.

Knowles, John, Nicola Persico, and Petra Todd. 2001. “Racial bias in motor vehicle searches: Theory and evidence.” *Journal of Political Economy*, 109(1): 203–229.

Lanning, Jonathan A. 2021. “Testing models of economic discrimination using the discretionary markup of indirect auto loans.” Working paper.

Liberman, Andres, Christopher Neilson, Luis Opazo, and Seth Zimmerman. 2018. “The equilibrium effects of information deletion: Evidence from consumer credit markets.” NBER Working Paper w25097.

Lusardi, Annamaria, and Olivia S Mitchell. 2014. “The economic importance of financial literacy: Theory and evidence.” *American Economic Journal: Journal of Economic Literature*, 52(1): 5–44.

Lusardi, Annamaria, and Olivia S Mitchell. 2023. “The importance of financial literacy: Opening a new field.” *Journal of Economic Perspectives*, 37(4): 137–154.

Malmendier, Ulrike, and Jessica A Wachter. 2021. “Memory of past experiences and economic decisions.” *Available at SSRN 4013583*.

Malmendier, Ulrike, and Stefan Nagel. 2011. “Depression babies: do macroeconomic experiences affect risk taking?” *Quarterly Journal of Economics*, 126(1): 373–416.

Malmendier, Ulrike, and Stefan Nagel. 2016. “Learning from inflation experiences.” *Quarterly Journal of Economics*, 131(1): 53–87.

Martinchek, Kassandra. 2024. “Young Adults’ Use of Debt and Credit During the Pandemic: Racial Disparities and the Effectiveness of State Initiatives to Protect Young Adults Living in Communities of Color.” *Available at SSRN 4768409*.

Morduch, Jonathan, Bradley Hardy, William Darity, and Darrick Hamilton. 2019. “Wealth inequality, income volatility, and race.” SASE 31st Annual Meeting.

Mulrow, Edward, et al. 2011. “Assessment of the U.S. Census Bureau’s Person Identification Validation System.” NORC Technical Report. Accessed: 2025-03-05.

Munnell, Alicia H, Geoffrey MB Tootell, Lynn E Browne, and James McEneaney. 1996. "Mortgage lending in Boston: Interpreting HMDA data." *American Economic Review*, 25–53.

Obermeyer, Ziad, Brian Powers, Clara Vogeli, and Sendhil Mullainathan. 2019. "Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations." *Science*, 366(6464): 447–453. Available at https://doi.org/10.1126/science.aax2342.

Passi, Samir, and Solon Barocas. 2019. "Problem formulation and fairness." *Proceedings of the conference on fairness, accountability, and transparency*, 39–48.

Raymond, Lindsey. 2024. "The market effects of algorithms." Working Paper.

Ricks, Judith, and Ryan Sandler. 2022. "Effects of Entering the Credit Market in a Recession." In . [Sl]: SSRN.

Scott-Clayton, Judith, and Jing Li. 2016. "Black-white disparity in student loan debt more than triples after graduation." *Evidence Speaks Reports*, 2(3).

Simoiu, Camelia, Sam Corbett-Davies, and Sharad Goel. 2017. "The problem of infra-marginality in outcome tests for discrimination." *The Annals of Applied Statistics*, 11(3): 1193–1216.

Stavins, Joanna. 2020. "Credit Card Debt and Consumer Payment Choice: What Can We Learn from Credit Bureau Data?" *Journal of Financial Services Research*, 58: 59–90.

Thompson, Jeffrey P., and Gustavo A. Suarez. 2015. "Exploring the Racial Wealth Gap Using the Survey of Consumer Finances." Board of Governors of the Federal Reserve System Finance and Economics Discussion Series 2015-076, Washington, DC.

United States Government Accountability Office. 2023. "Credit Cards: Pandemic Assistance Likely Helped Reduce Balances, and Credit Terms Varied among Demographic Groups." United States Government Accountability Office GAO-23-105269.

U.S. Department of Commerce, Bureau of the Census. 2000. "United States Census 2000: Informational Copy." *Form D-61B*.

U.S. Department of Commerce, Bureau of the Census. 2003. "Chapter 5: Sample Design and Estimation; 2000 Census of Population and Housing: Public-Use Microdata Samples Technical Documentation." *Technical Report*.

U.S. Department of Commerce, Bureau of the Census. 2014. "Chapter 4: Sample Design and Selection; American Community Survey Design and Methodology (January 2014)." *Technical Report*.

U.S. Department of Justice, Office of Public Affairs. 2012. "Justice Department Reaches Settlement with Wells Fargo Resulting in More Than $175 Million in Relief for Homeowners to Resolve Fair Lending Claims." *https://www.justice.gov/opa/pr/justice-department-reaches-settlement-wells-fargo-resulting-more-175-million-relief*, Press release number: 12-869.

Wagner, Deborah, and Mary Layne. 2014. "The Person Identification Validation System (PVS): Applying the Center for Administrative Records Research and Applications' (CARRA) Record Linkage Software." *Center for Administrative Records Research and Applications Working Paper*.

Warren, Andrew, Wanjira Chege, Kennan Cepa, and Necati Celik. 2024. "Pulse Points: Disparities in Credit Scores and Length of Credit History." Financial Health Network. Accessed: 2025-02-24.

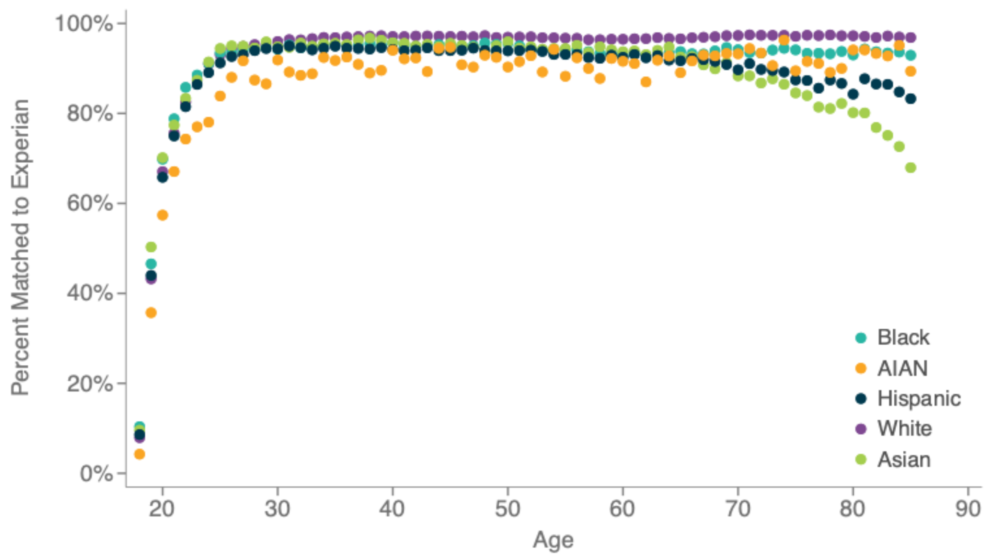
Willen, Paul, and David Zhang. 2023. “Testing for discrimination in menus.” Working Paper.

Zabritski, Melinda. 2024. “State of the Automotive Finance Market: Q3 2024.” Experian Automotive. Public report.

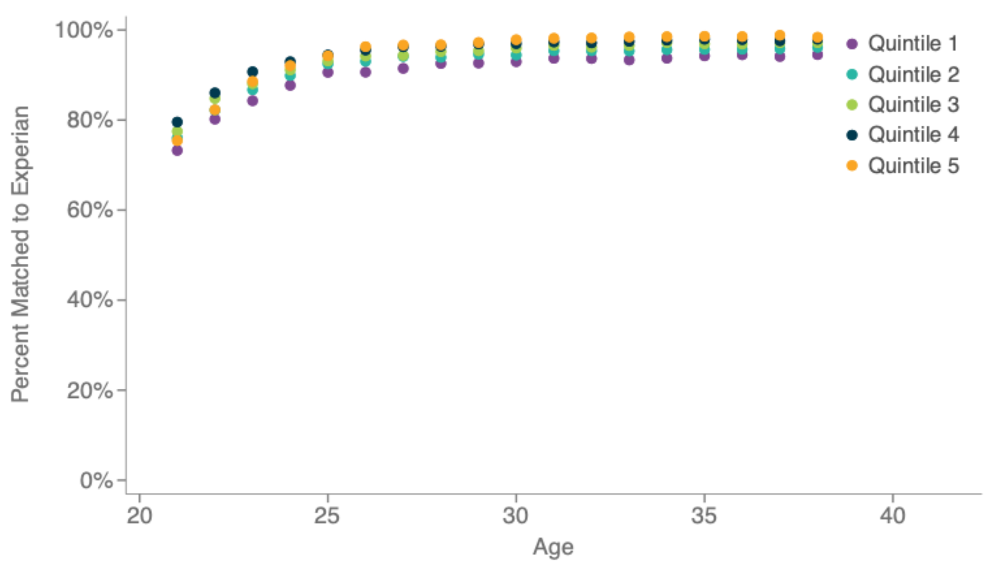
Zafar, Muhammad Bilal, Isabel Valera, Manuel Gomez Rogriguez, and Krishna P Gummadi. 2017. “Fairness constraints: Mechanisms for fair classification.” 962–970, PMLR.

FIGURE I
Fraction of Population with Credit Record

A. Fraction with Credit File by Age, Separately by race



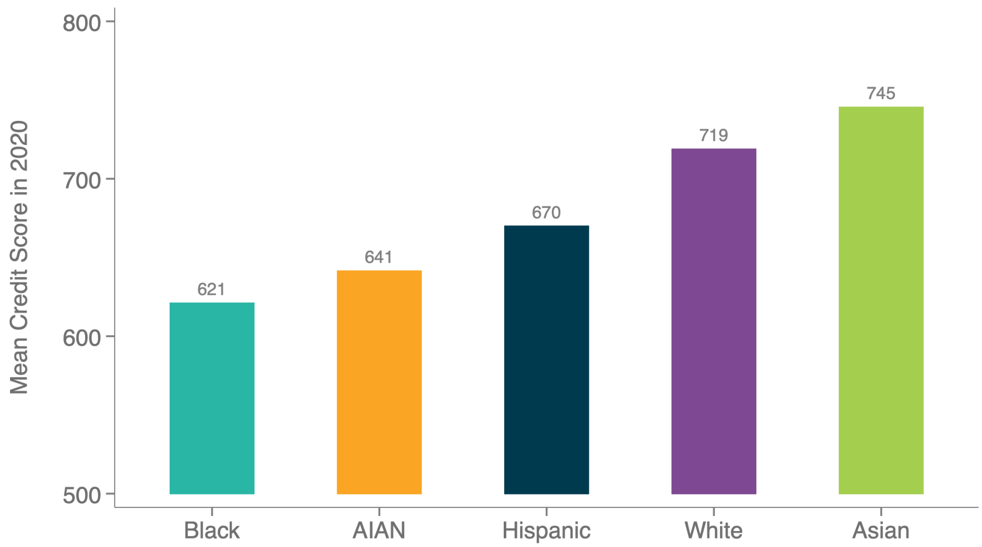
B. Fraction with Credit File by Age, Separately by Parental Income



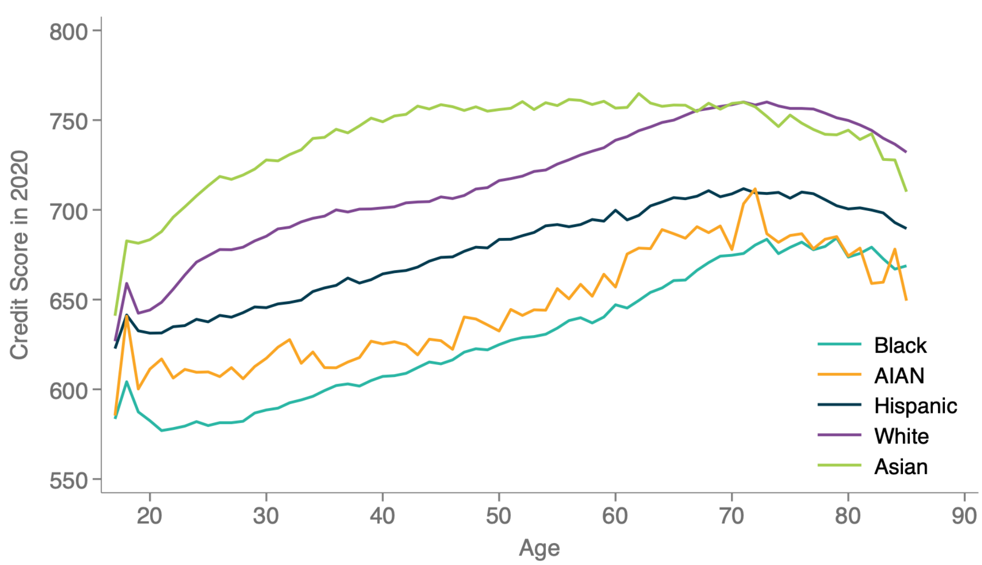
Notes: This Figure presents the fraction of people who have a credit file by age, race, and parental income. Panel A reports the fraction of people in our population sample who have a credit file in any year in our Experian sample (2004, 2008, 2012, 2016, 2020), separately by age and race. Panel B reports the fraction of people in our intergenerational sample (1978-1985 birth cohorts linked to their parents) who have a credit file, separately by age and parental income quintile.

FIGURE II
Credit Scores by Race

A. Average Credit Scores by Race



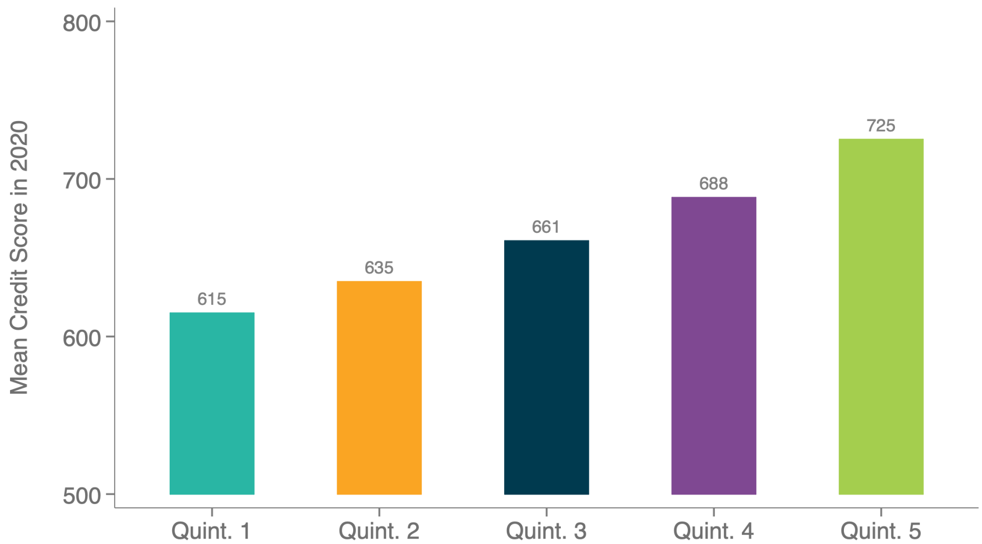
B. Average Credit Scores by Age and Race



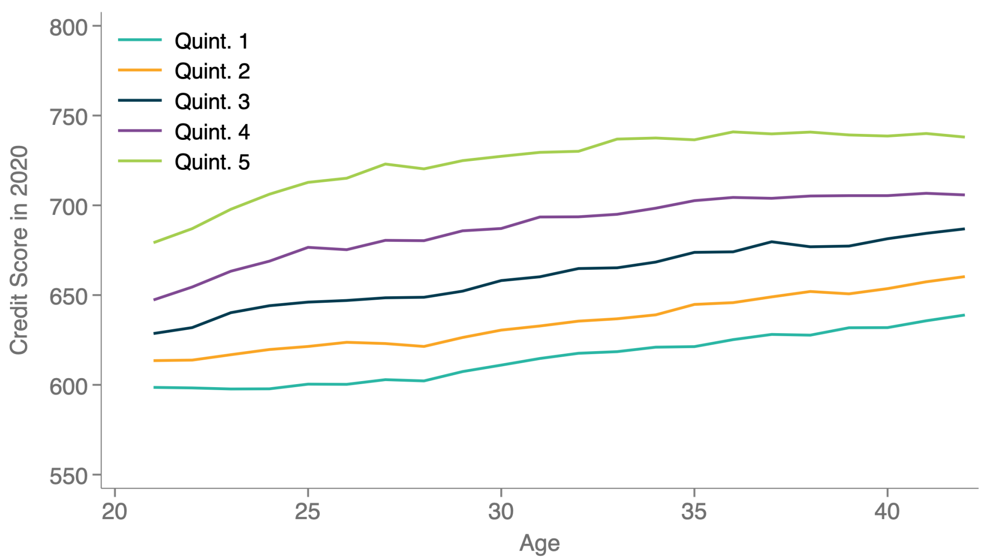
Notes: This figure presents the average credit scores in 2020 for our population sample by race. Panel A reports the average credit scores by race. Panel B reports the average credit scores by age, separately by race. All averages are taken over the subset of individuals with credit scores. The credit score is binned into categories by lenders. An "Excellent" score is between 781 and 850. A "Good" score is between 661 and 780. A "Fair" score is between 601 and 660. A "Poor" score is between 500 and 600. A "Very Poor" score is from 300 to 499. These ranges are from https://www.chase.com/personal/credit-cards/education/credit-score/credit-score-ranges-and-what-they-mean.

FIGURE III
Credit Scores by Class

A. Average Credit Scores by Parent Income Quintile

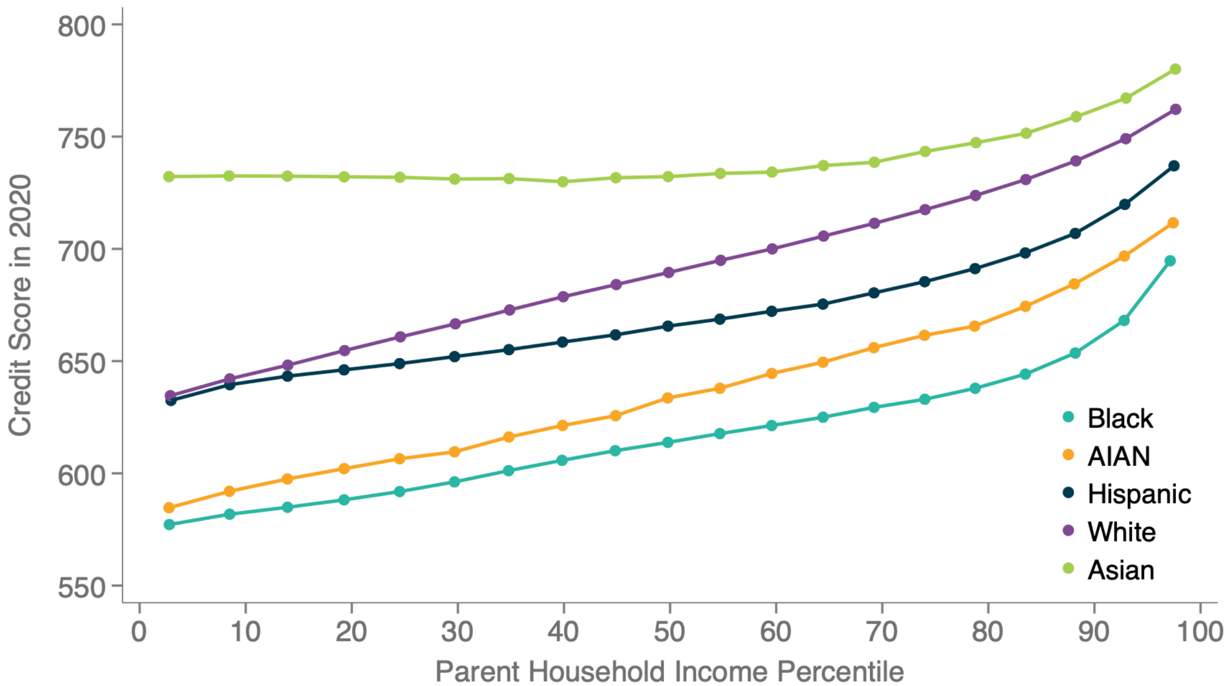


B. Credit Scores by Age and Parent Income Quintile



Notes: This figure presents the average credit scores in 2020 for our population sample by parental income quintile, restricting to ages over which we are able to link to parents. Panel A reports the average credit scores by parental income quintile. Panel B reports the average credit scores by age, separately by parental income quintile. All averages are taken over the subset of individuals with credit scores. The credit score is binned into categories by lenders. An "Excellent" score is between 781 and 850. A "Good" score is between 661 and 780. A "Fair" score is between 601 and 660. A "Poor" score is between 500 and 600. A "Very Poor" score is from 300 to 499. These ranges are from https://www.chase.com/personal/credit-cards/education/credit-score/credit-score-ranges-and-what-they-mean.

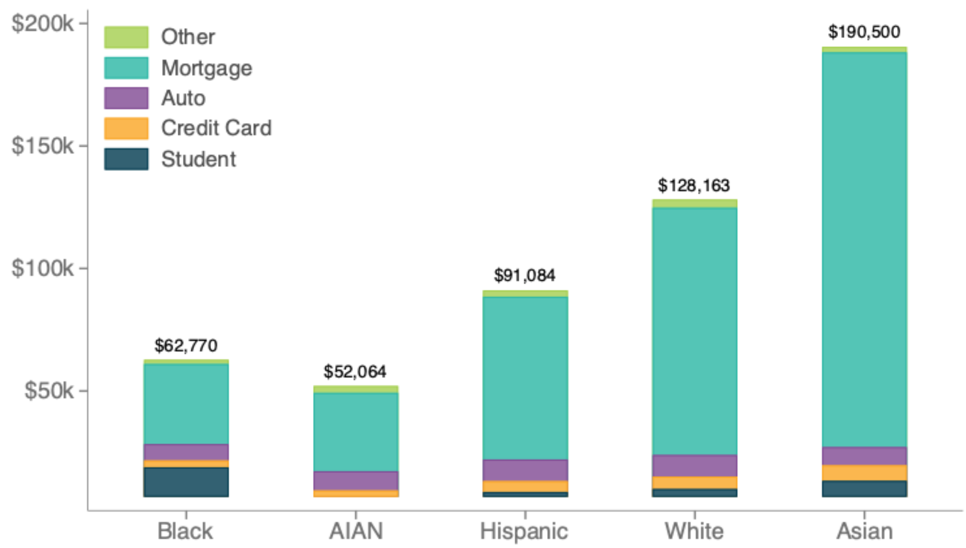
FIGURE IV
Credit Scores by Race and Parent Household Income



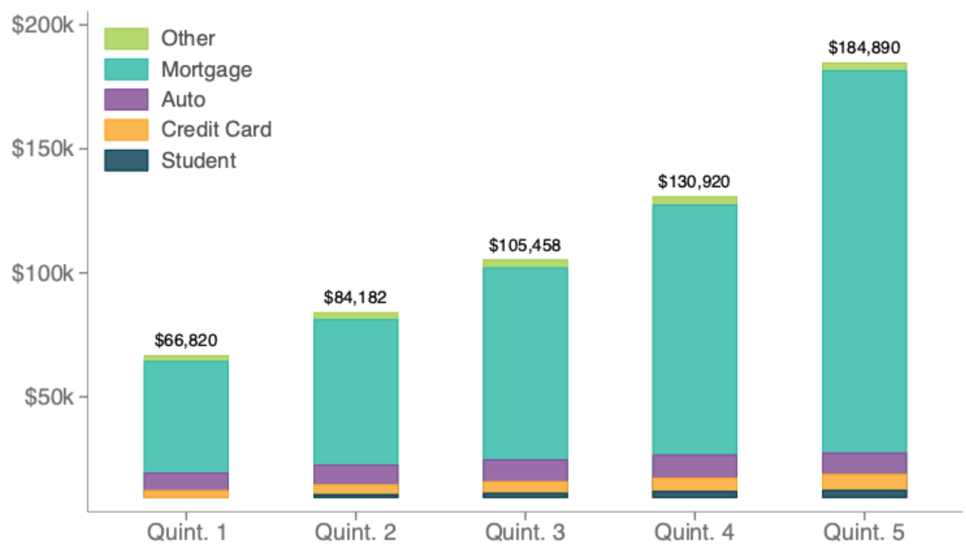
Notes: This figure presents the average credit scores in 2020 for our intergenerational sample by parental income percentile, separately by race. See Appendix Figure A.4 for a version of this figure that restricts the sample to those with US-born mothers.

FIGURE V
Composition of Total Debt

A. Race



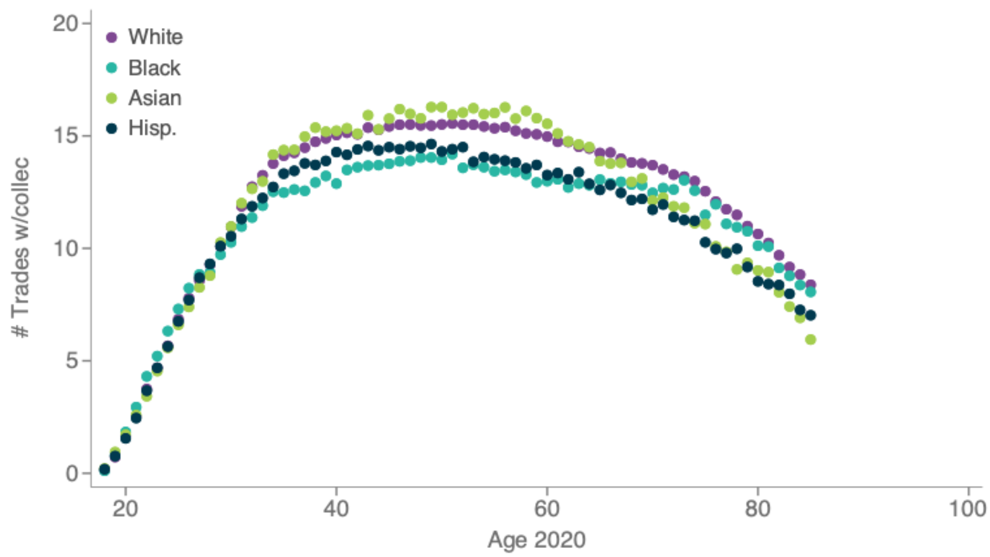
B. Parent Income Quintile



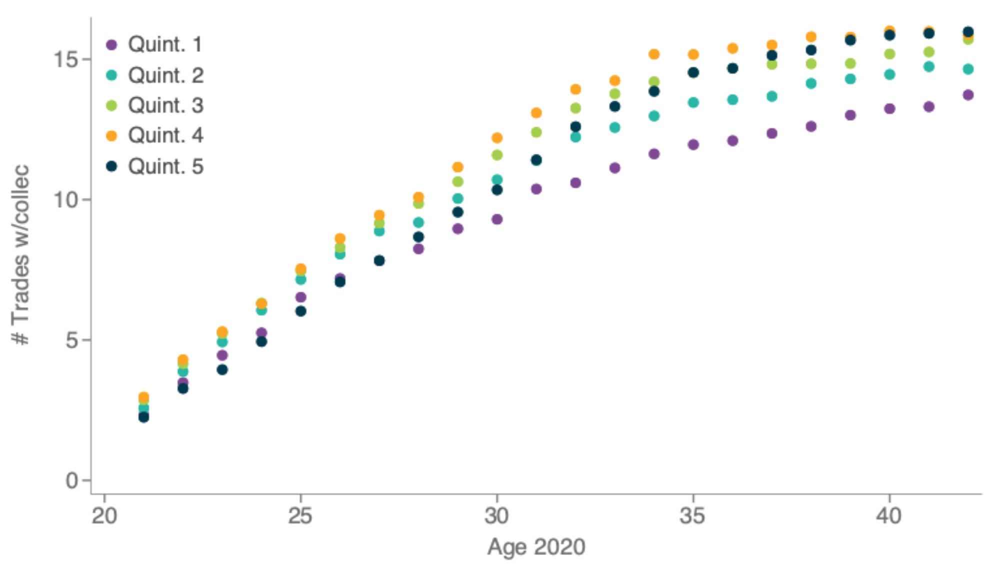
Notes: This figure presents a stacked bar chart of average debt holdings by type of credit. We present 4 credit types: mortgages, auto loans, credit card balances, and student loans, which comprise nearly all debt on credit reports. Panel A reports average debt by race. Panel B reports average debt by parental income quintile.

FIGURE VI
Number of Tradelines by Age

A. Race



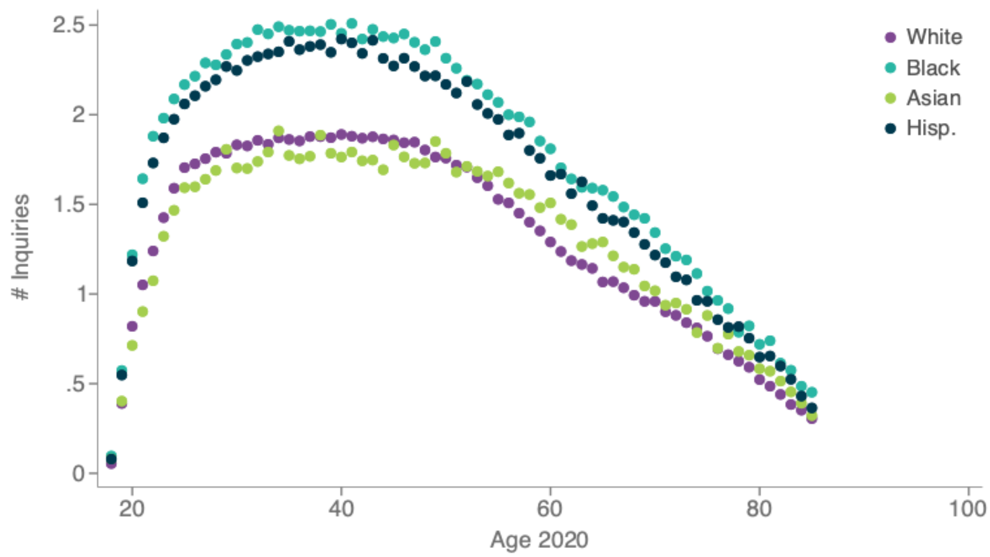
B. Parent Income Quintile



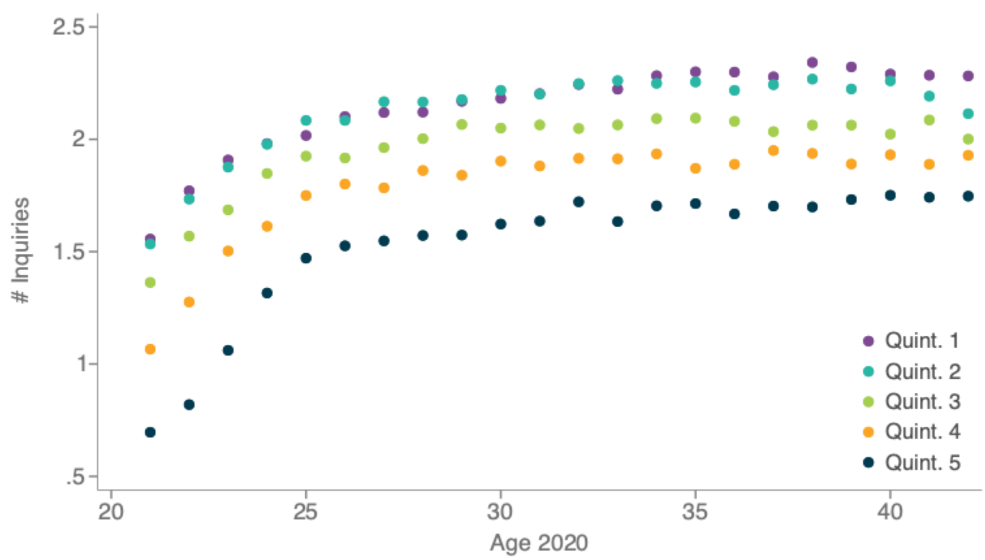
Notes: This figure presents the number of tradelines by age. Panel A presents these results by race. Panel B presents these patterns by parental income quintile, restricting to the ages we are able to match to parents.

FIGURE VII
Number of Inquiries by Age

A. Race



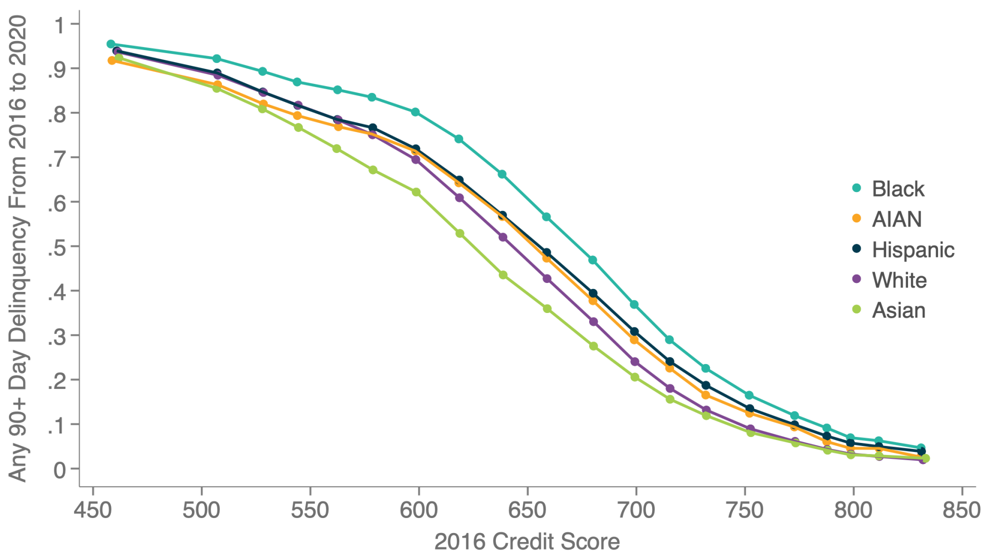
B. Parent Income Quintile



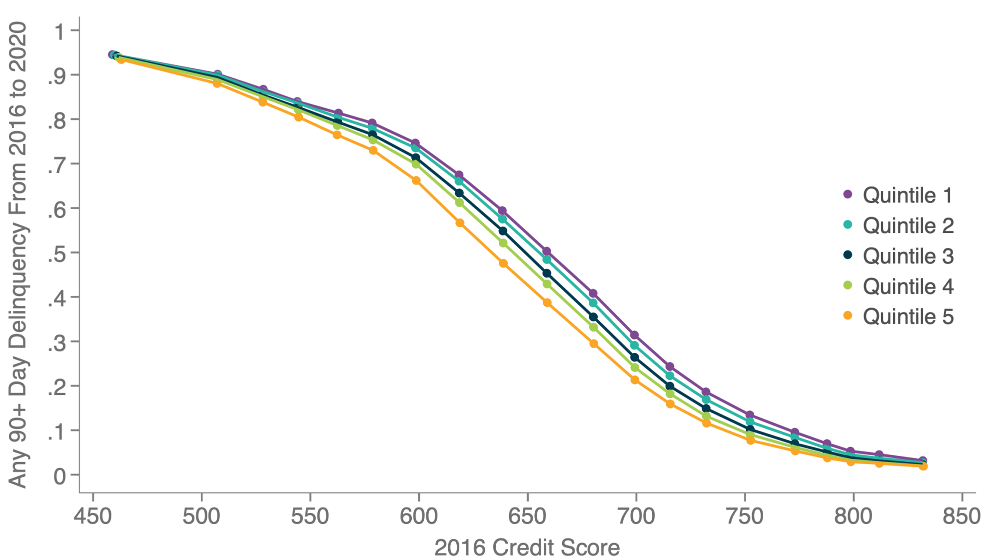
Notes: This figure presents the number of credit inquiries across different loan categories. Panel A splits out the series by race and Panel B splits out the series by parent income quintile.

FIGURE VIII
Subsequent Non-Repayment versus Credit Score

A. Race



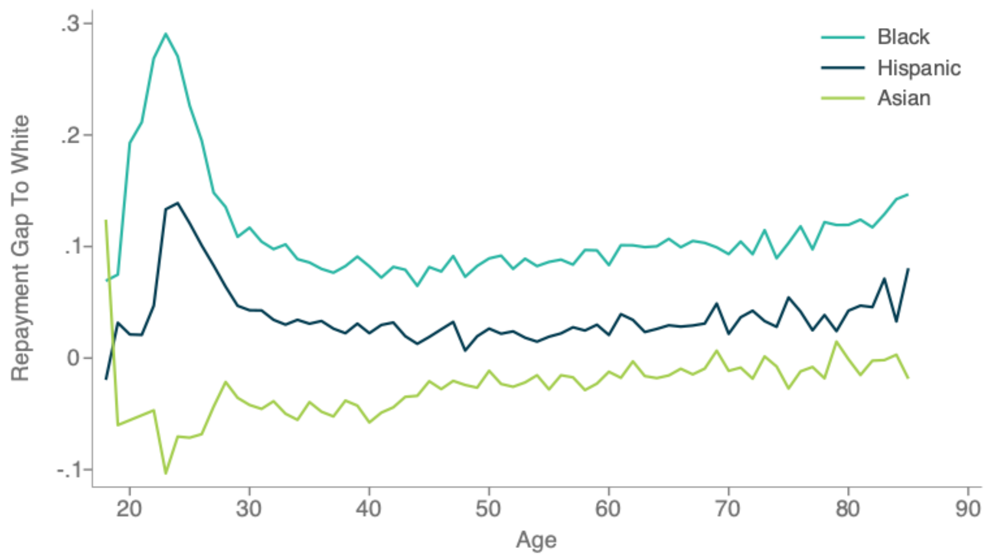
B. Parent Income Quintile



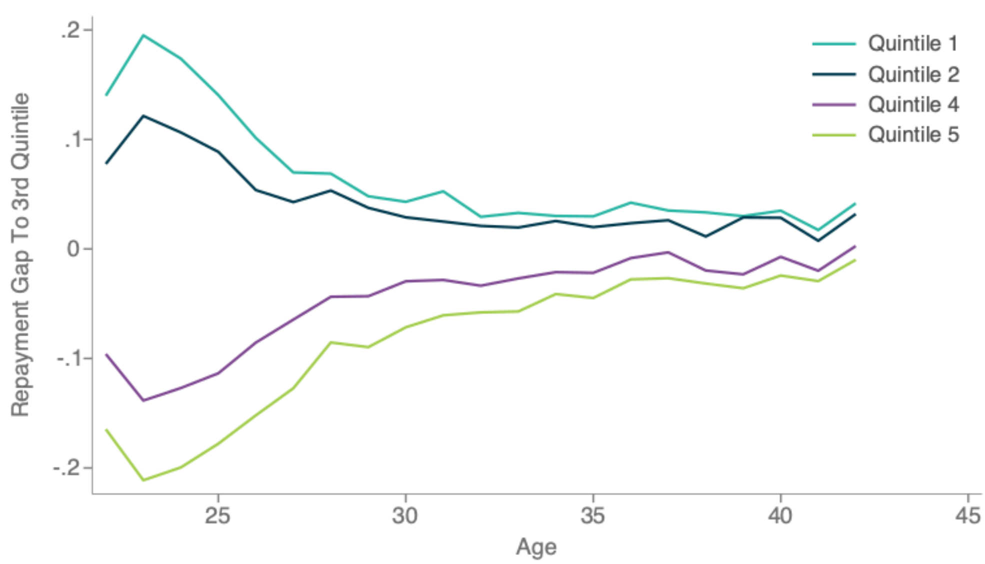
Notes: This figure presents the average 90+ day delinquency rate in the 2020 credit file as a function of the 2016 credit score on the horizontal axis, separately by group. Panel A reports separate series by race. Panel B reports separate series by parental income quintile. Both series use the intergenerational sample of 1978-85 birth cohorts.

FIGURE IX
Predictive Content of Race and Parental Income for 90+ Day Delinquency, by Age

A. Race (White Omitted Group)



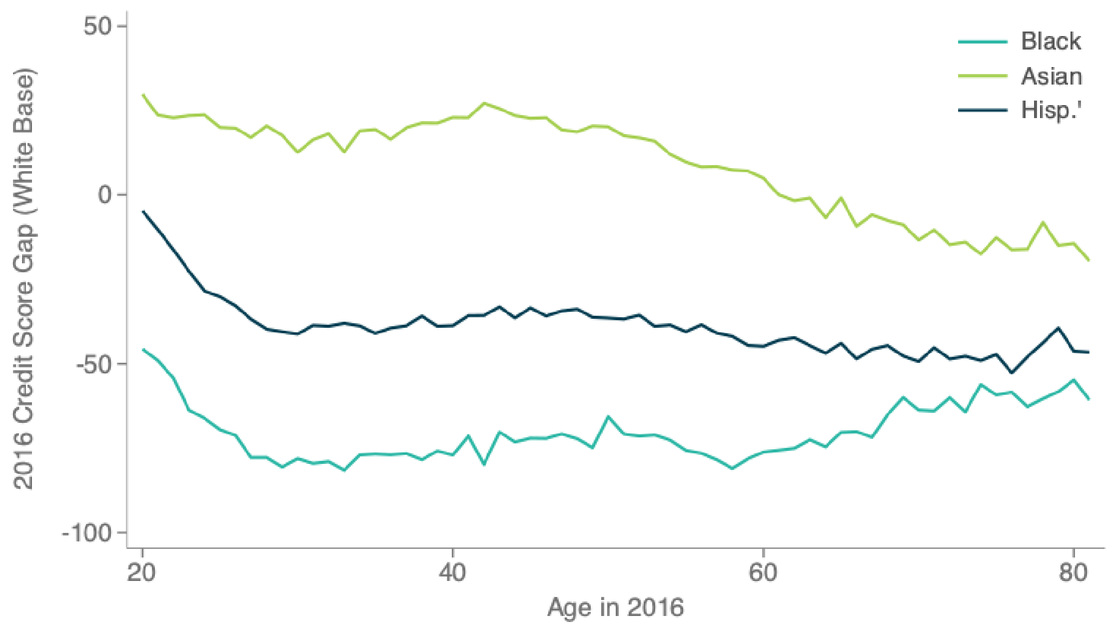
B. Parent Income Quintile (3rd Quintile Omitted)



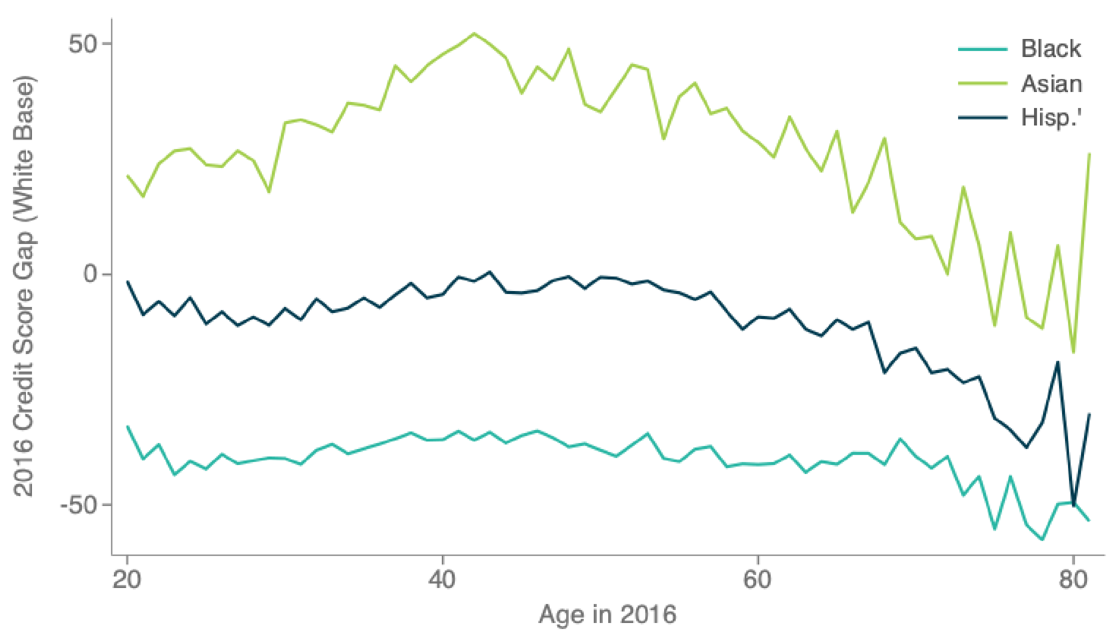
Notes: This figure presents coefficients from a regression of 90+ Day Delinquency on the credit score along with (Panel A) race indicators or (Panel B) parental income quintiles. Panel A plots race fixed effects, omitting the White category. Panel B plots the parent income quintile fixed effects, omitting the middle (third) quintile.

FIGURE X
2016 Credit Score Race Gaps by 90+ Day Delinquency in 2020 and Age

A. Individuals With A 90+ Day Delinquency in 2020



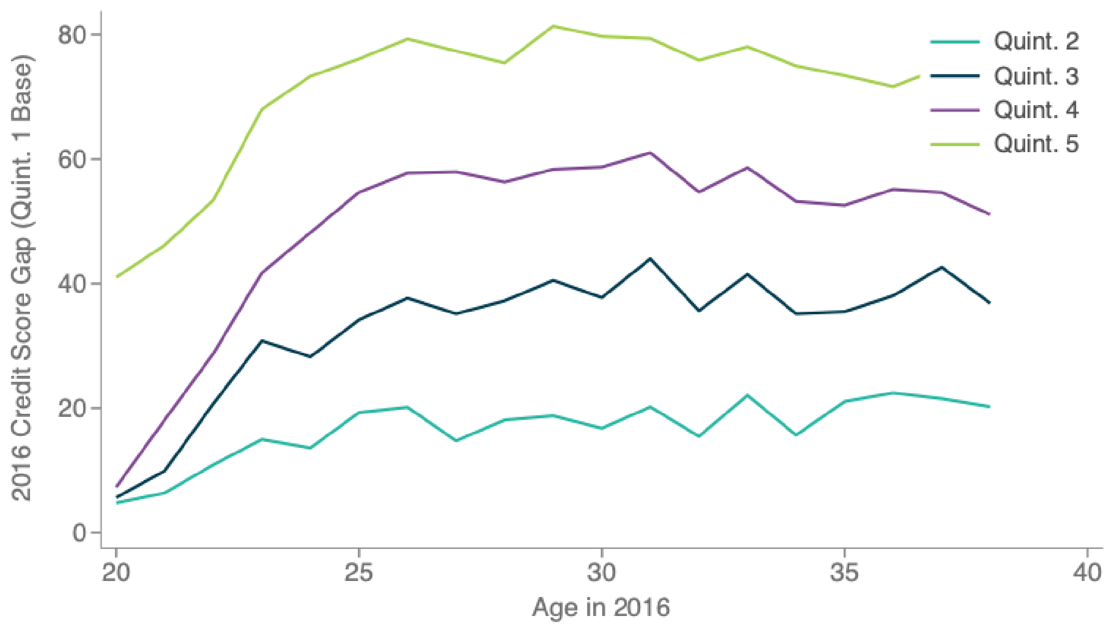
B. Individuals Without a 90+ Day Delinquency in 2020



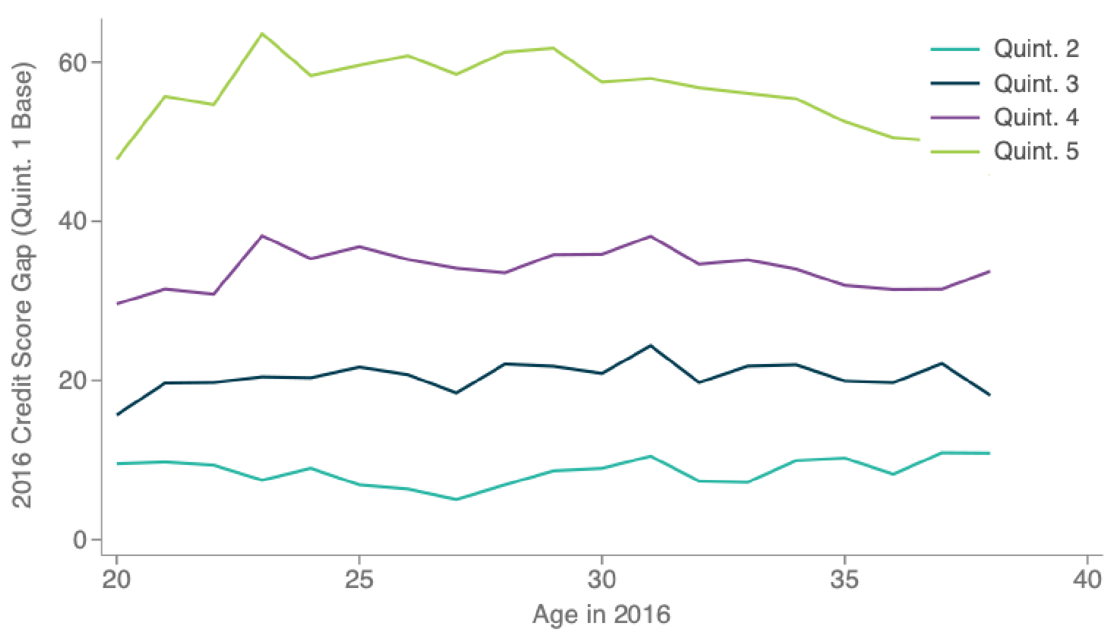
Notes: This figure presents a balance test for bias in credit scores by race. For each age, we present the coefficients on race indicators from a regression of the 2016 credit score on race indicators, omitting White borrowers. Panel A restricts to the set of individuals who do not have a 90+ day delinquency. Panel B restricts to the set of individuals who do have a 90+ day delinquency.

FIGURE XI
2016 Credit Score Parent Income Gaps by 90+ Day Delinquency in 2020 and Age

A. Individuals With A 90+ Day Delinquency in 2020



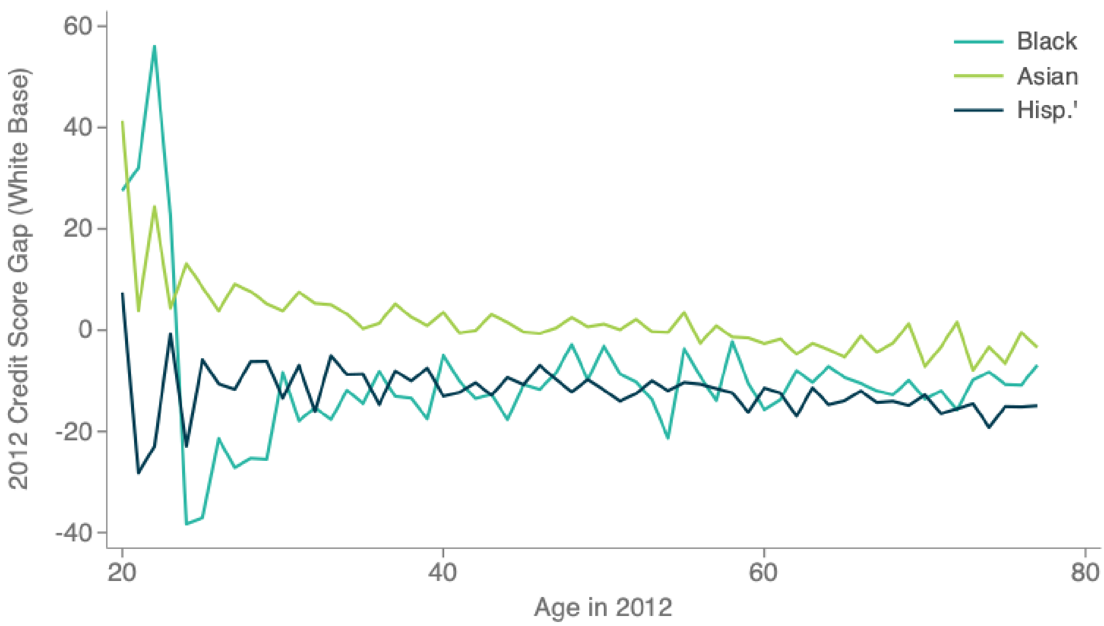
B. Individuals Without a 90+ Day Delinquency in 2020



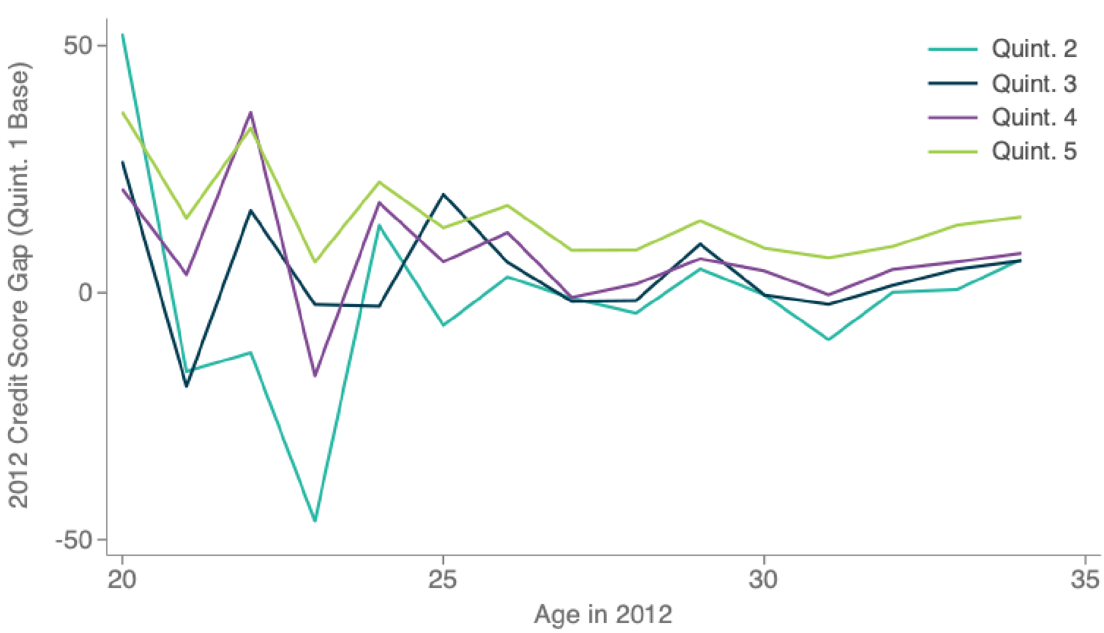
Notes: This figure presents a balance test for bias in credit scores by parental income quintile. For each age, we present the coefficients on parental quintile indicators from a regression of the 2016 credit score on parental quintile indicators, omitting the lowest quintile. Panel A restricts to the set of individuals who do not have a 90+ day delinquency. Panel B restricts to the set of individuals who do have a 90+ day delinquency.

FIGURE XII
Credit Score Race and Class Gaps for Those Without A Late Payment, by Age

A. Race

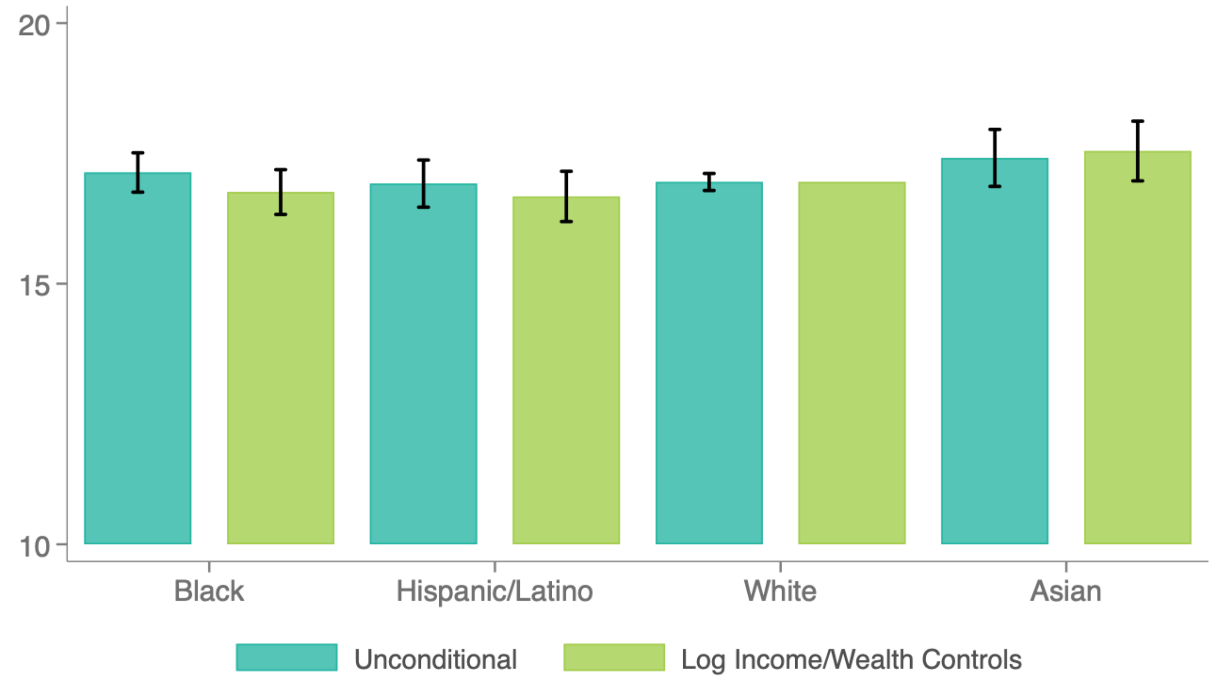


B. Parental Income Quintile



Notes: This figure presents differences in credit scores at each age by race and parental income among the subset of borrowers in our population sample who never made a payment over 30 days late at any point between 1994 and 2020. Panel A present the coefficients on race indicators from a regression of the 2016 credit score on race indicators, omitting White borrowers. Panel B presents the coefficients on parental income quintile, omitting the lowest quintile.

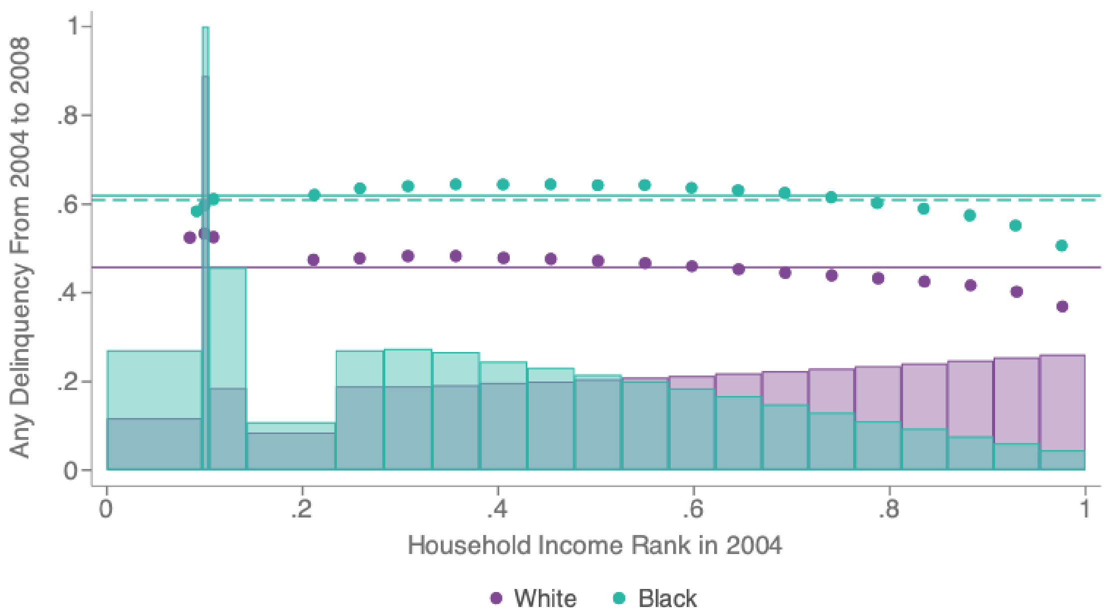
FIGURE XIII
Credit Card Interest Rates by Race



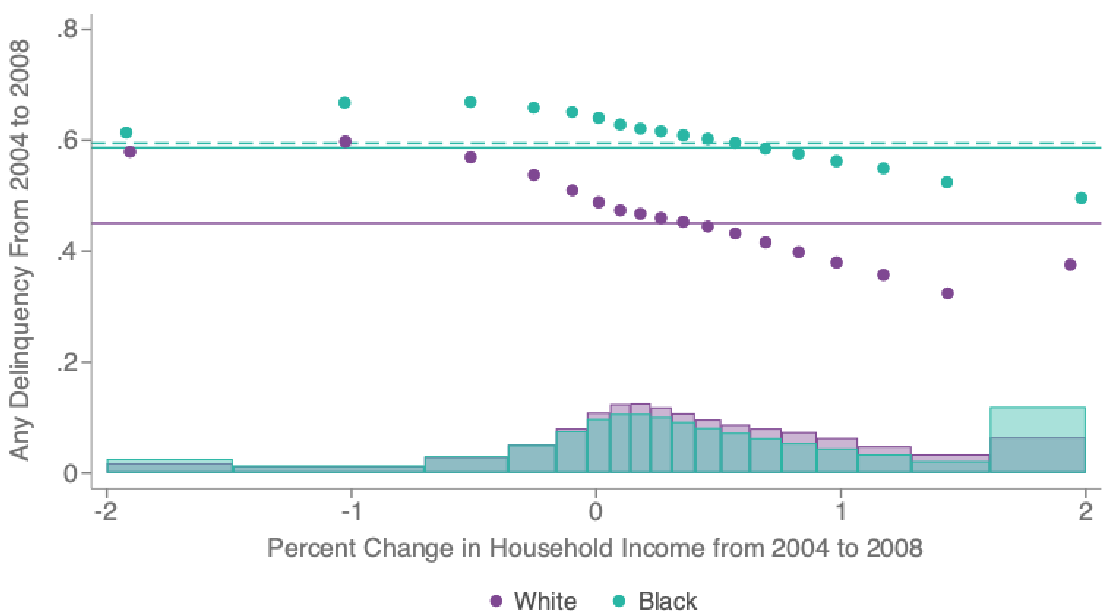
Notes: This figure presents average credit card interest rates by race as reported in the 2022 Survey of Consumer Finances. All regressions use SCF sample weights. "Unconditional" bars present raw means by race. "Log Income/Wealth Controls" bars are computed by taking the unconditional mean for White borrowers and then adding the coefficients on race dummy variables from a regression of credit card interest rates on race dummy variables (with White as the omitted category), log income, and log net wealth. Net wealth is computed following the method used for Federal Reserve Bulletin articles: https://www.federalreserve.gov/econres/files/bulletin.macro.txt. Excludes those who do not have a credit card or have a credit card with no interest rate.

FIGURE XIV
Relationship between Delinquency and Income

A. Income

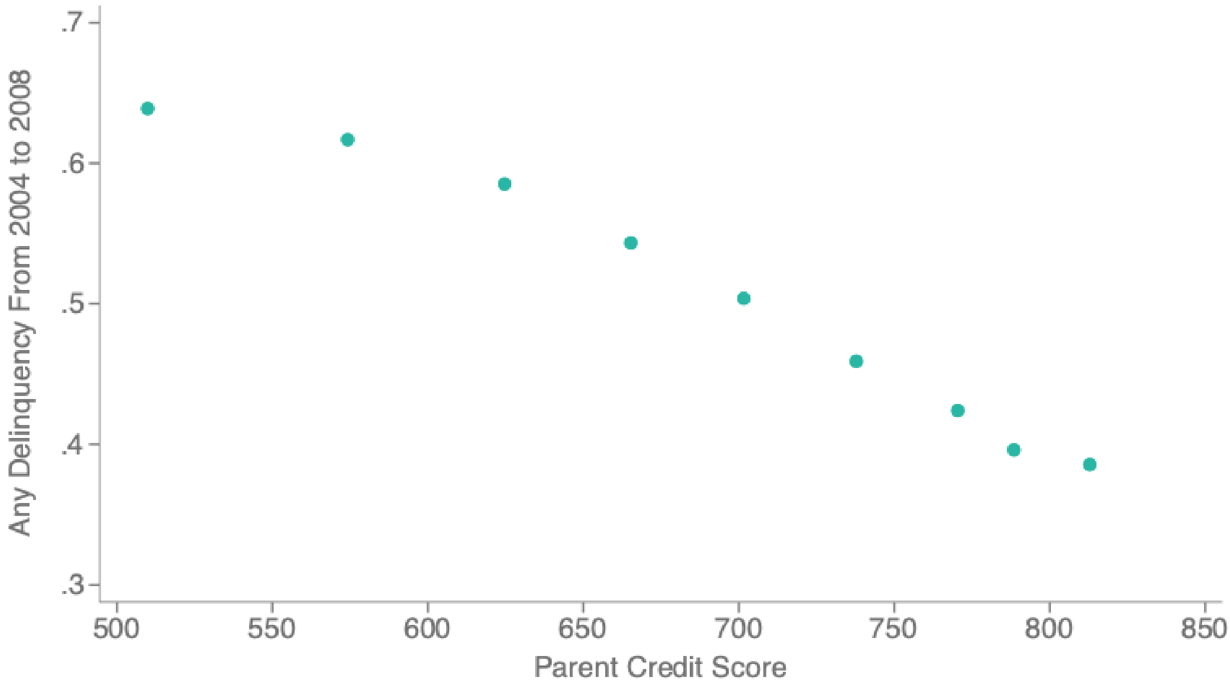


B. Income Shocks



Notes: This figure presents the relationship between income and delinquency using our intergenerational sample. Panel A presents the average 90+ day delinquency rate in 2004 by household income percentile in 2004 conditional on credit score rank, separately for Black and White individuals. The histograms present the distribution of household income percentiles by race. We cap the height of the bin, which corresponds to the mass point at 0 income. 3.1% of White individuals are in that bin and 7.6% of Black individuals. The dashed horizontal line presents the mean delinquency rates for Black individuals if they had the same delinquency conditional on income but faced the White income distribution. Panel B repeats this exercise but using income change between 2004 and 2008 on the horizontal axis, defined as $\frac{Income_{2008}-Income_{2004}}{0.5(Income_{2008}+Income_{2004})}$.

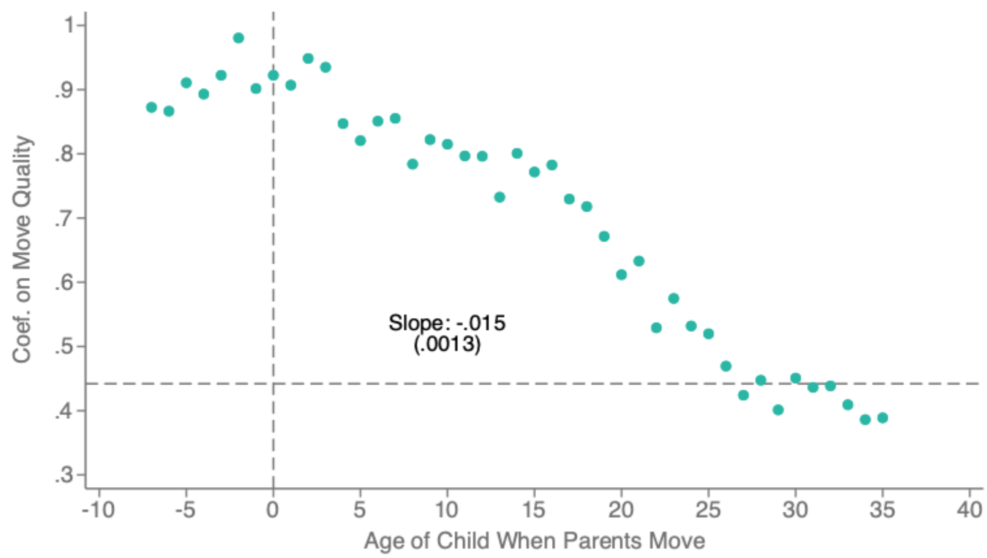
FIGURE XV
Relationship between 90+ Day Delinquency and Parent Credit Score, with Controls



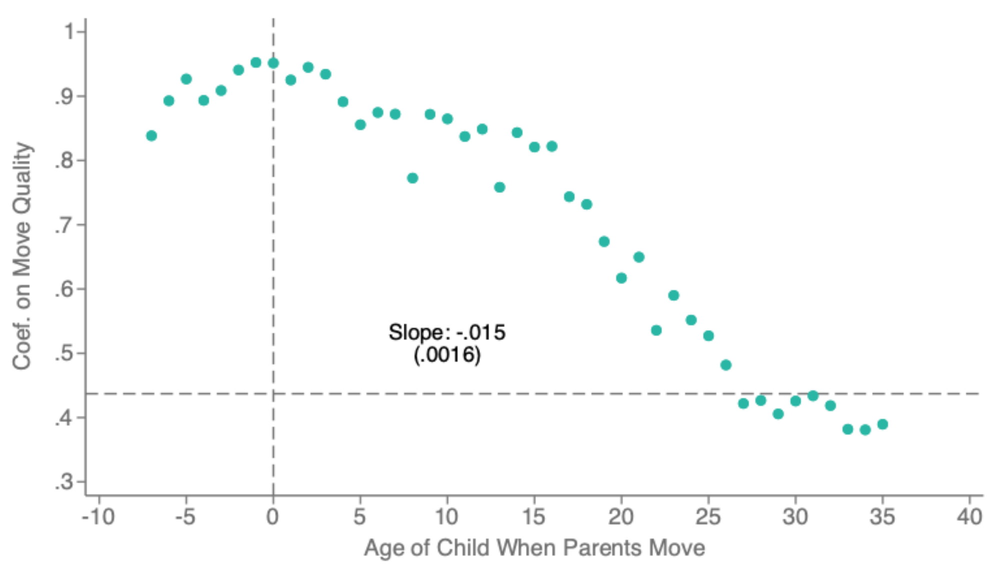
Notes: This figures presents a binned scatter plot of 90+ day delinquency from 2004 to 2008 on the vertical axis plotted against the parent's credit score in 2004 on the horizontal axis, controlling for the child's 2004 credit score, 2004 household income rank, and their parent's household income rank. The sample is individuals born in 1981.

FIGURE XVI
Childhood Exposure Effects for 90+ Day Delinquency

A. 90+ Delinquency



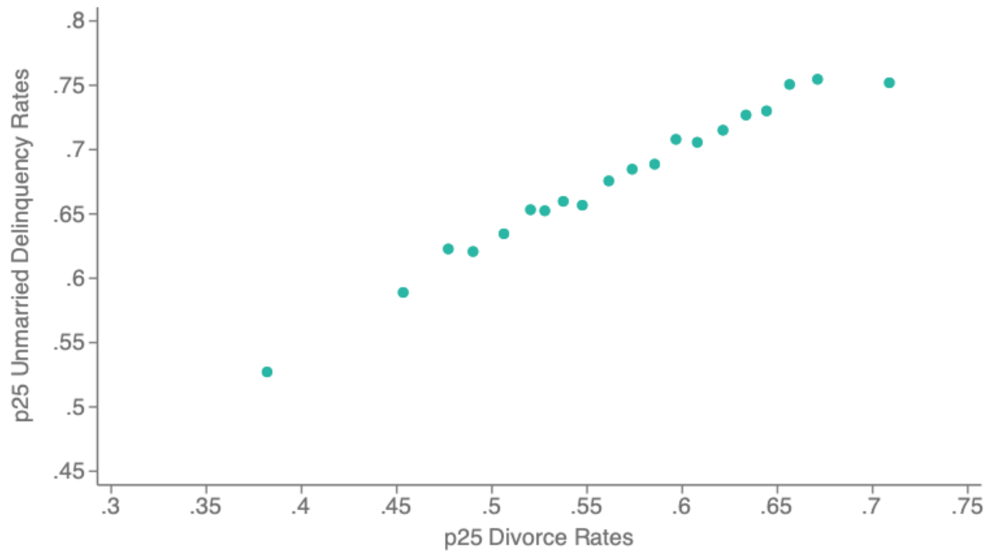
B. 90+ Day Delinquency Controlling for Income



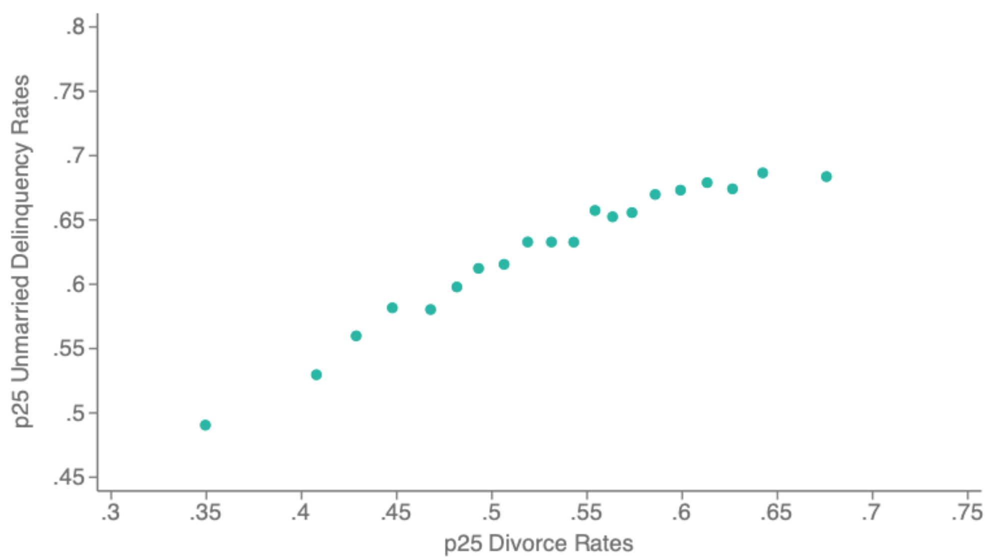
Notes: Panel A plots estimates of the coefficients β_m vs. the child's age when the parents move (m) using specification in equation 3, for the case where the outcome is whether a child had a 90+ day delinquency in a four year window around age 30. We interpret the difference in coefficients, $\beta_{m-1} - \beta_m$, as the causal effect of spending age m of childhood in a place where outcomes of non-1-time-movers are 1 percentile higher. The figure reports the slope of a regression of β_m on age, m, between ages 0 and 23, along with its standard error. Panel B does the same but with delinquency residualized on the child's income in the year immediately before the four year repayment window, so that it captures the effect of moving to a place where individuals are 1 pecentile more likely to repay conditional on their income.

FIGURE XVII
Correlation Between Divorce Rates and Unmarried Delinquency

A. Pooled Sample



B. White Individuals



Notes: This figure presents a binned scatter plot of the relationship between the average 90+ day delinquency rate and the divorce rate both estimated at the 25th percentile of the parent household income distribution in the county. We measure delinquency rates among individuals who never marry and divorce rates among the people who have been married. Panel A presents results for the full sample, while Panel B restricts to White individuals.

TABLE I
Median Credit Scores by Race (Including 0s)

	Avg. Credit Score (Population)	Avg. Credit Score	Median Credit Score	Median Credit Score w/Zeros
White	719	701	719	712
Black	621	601	581	571
Asian	745	741	781	776
Hispanic	670	659	655	644
AIAN	641	622	602	583

Notes: This table presents the mean and median credit scores in 2020 separately by race. Column (1) reports the mean credit score on the subset of the population sample that have a credit score. Column (2) reports the mean credit score on the subset of the intergenerational sample that have a credit score. Column (3) reports the median score among those who have a credit score in the intergenerational sample. Column (4) reports the median score after imputing low scores for those with missing files in the intergenerational sample.

TABLE II
Relationship between 90+ Day Delinquency and Income and Wealth

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				90+ Day Delinquency				
Black	0.084	0.083	0.083	0.085	0.079	0.079	0.078	0.118
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.019)
Asian	-0.022	-0.022	-0.022	-0.009	-0.014	-0.014	-0.015	0.030
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.037)
Hispanic	0.034	0.033	0.033	0.033	0.026	0.026	0.025	0.052
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.019)
Credit Score Rank	-1.136	-1.121	-1.121	-1.104	-1.096	-1.094	-1.095	-1.041
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.031)
Income		-0.027	-0.028	0.023	0.024	0.025	0.008	0.058
		(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.069)
Parent Income					-0.040	-0.039	-0.039	-0.016
					(0.000)	(0.000)	(0.000)	(0.024)
Housing Wealth						0.001		0.001
						(0.000)		(0.000)
Log Bank Assets								-0.017
								(0.004)
N (100s)	249,700	249,700	249,700	249,700	249,700	249,700	249,700	41
R^2	0.425	0.425	0.425	0.434	0.434	0.435	0.434	0.438

Notes: This table presents a regression of any delinquency from 2016 to 2020 on race dummies. The columns add controls. Column (2) adds 2016 household income rank. Column (3) adds savings summed from 2004 to 2016. Column (4) adds a vector of household income rank from 2003 to 2021. Column (5) adds parent income rank. Column (6) adds housing wealth which is difference between the value of the house and their mortgage in 2016. Column (7) controls for an individual income vector from 2003 to 2021. Column (8) adds in log bank assets from the SIPP and an indicator for if the variable is zero or negative.

TABLE III
Parent Credit Score Predicts Child Delinquency

	(1)	(2)	(3)	(4)	(5)
			90+ Day Delinquency		
Parent Credit Score Rank 2004	-0.641	-0.437	-0.367	-0.356	-0.333
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Income Rank in 2016		-0.608	-0.583	-0.534	-0.528
		(0.001)	(0.001)	(0.001)	(0.002)
Parent Income Rank			-0.136	-0.128	-0.076
			(0.001)	(0.001)	(0.002)
N	4,147,000	4,147,000	4,147,000	4,147,000	1,050,000
R^2	0.102	0.209	0.214	0.218	0.214

Notes: This table presents a regression of any delinquency from 2016 to 2020 on parent credit score rank in 2004. Column (1) reports a univariate regression of delinquency on parent credit score rank. The next columns add controls. Column (2) adds child income in 2016. Column (3) adds parent income rank. Column (4) adds savings, which is the sum of deferred compensation from 2004 to 2016 and housing wealth. Column (5) adds parent education fixed effects.

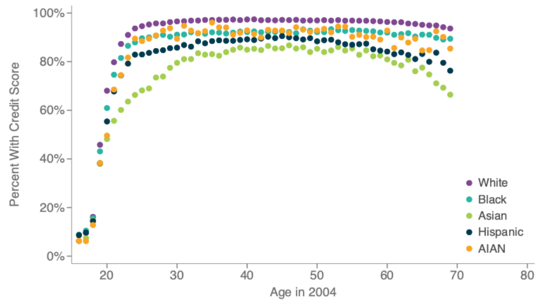
A Data Appendix

A.A Linking Experian Data to Census Data

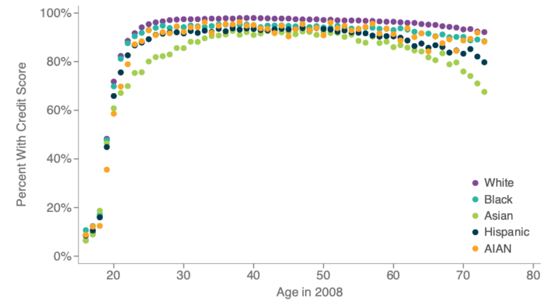
We link the Experian data to the administrative data housed at the Census Bureau as follows. For the intergenerational sample, the 1978-1985 birth cohorts, Experian identifies all individuals in their data who were born during those years and securely transfers the data, including a hashed version of the SSN, to the Census Bureau. Census and Experian agreed upon a hashing algorithm that is not known to researchers or anyone interacting with the data at Census, allowing us to perform this match while ensuring confidentiality. The Census Bureau replaces the hashed SSNs from the Experian data with anonymized Protected Identification Keys (PIK) that facilitate linkages with other data housed at the Census Bureau, including the decennial censuses, ACS, and federal income tax returns. We then merge the Experian data to the Census and tax records for all individuals in the 1978-1985 birth cohorts using the PIK. Because we observe our full target population in the Census and tax data, we are therefore able to identify individuals who do not appear in the Experian data due to missing credit histories, thereby including so-called "credit invisibles" in our study.

FIGURE A.1
Fraction with a Credit Score by Race/Class and Age and Year

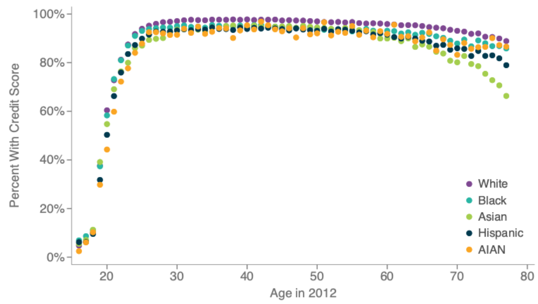
A. Race, 2004



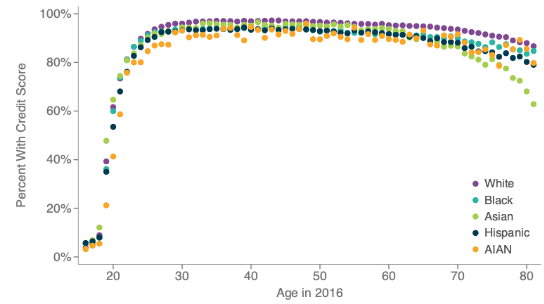
B. Race, 2008



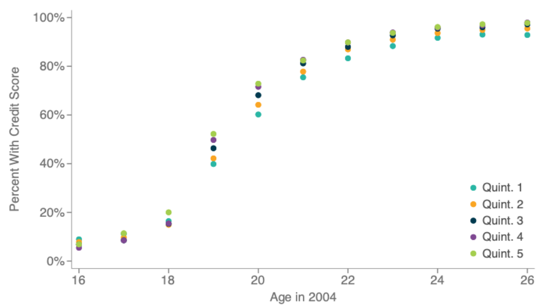
C. Race, 2012



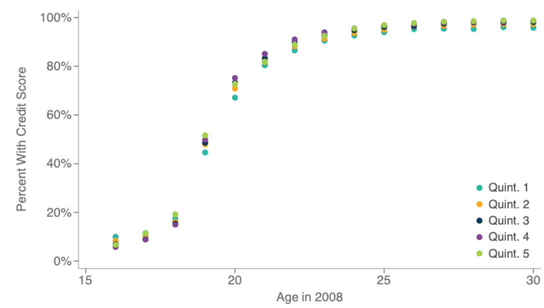
D. Race, 2016



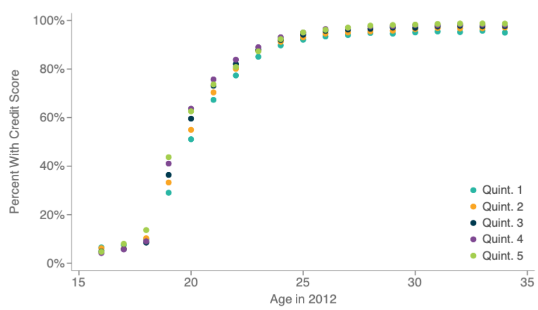
E. Parent Income, 2004



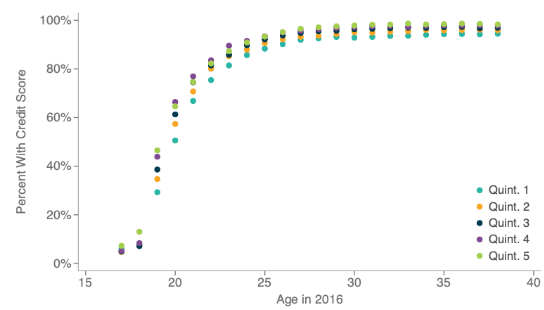
F. Parent Income, 2008



G. Parent Income, 2012



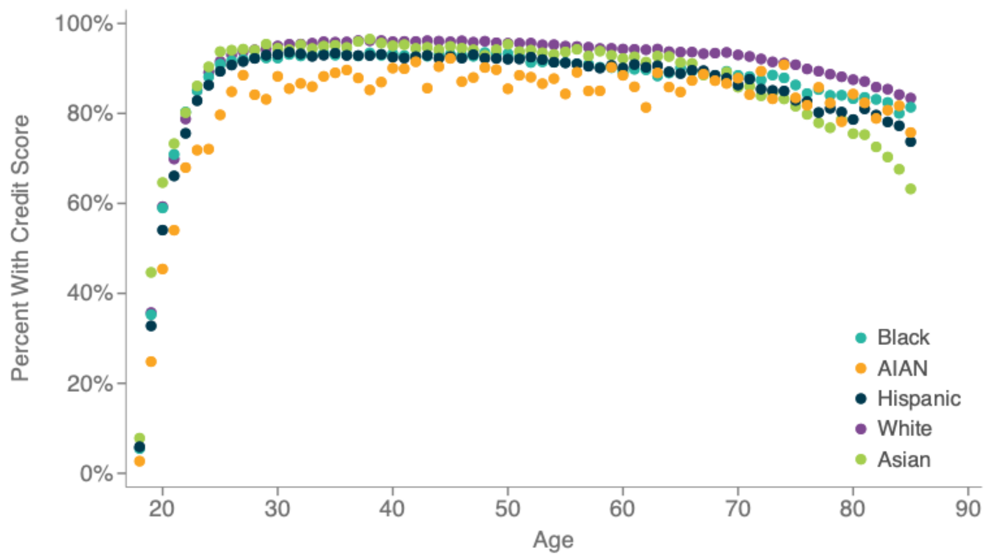
H. Parent Income, 2016



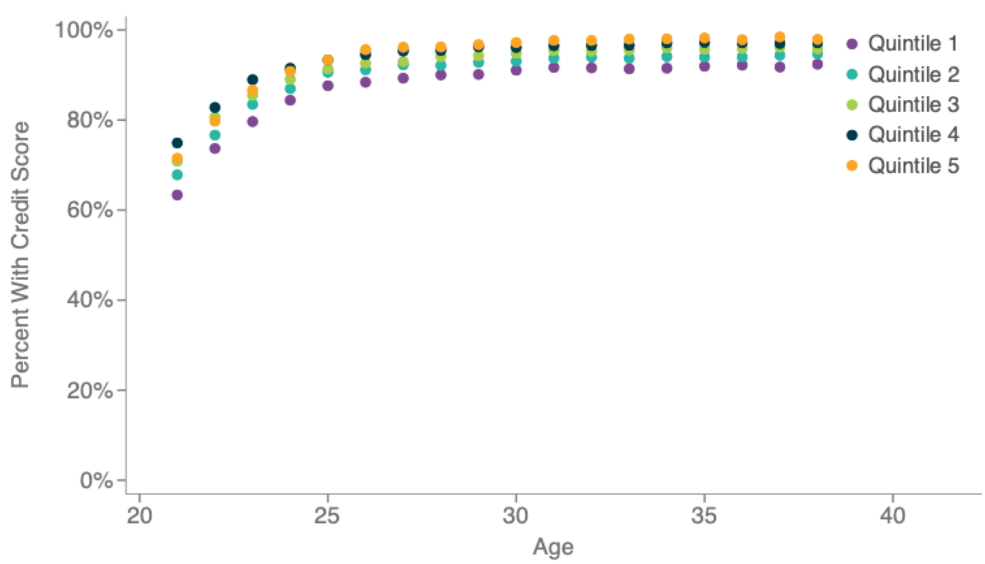
Notes: This figure presents the fraction of individuals in our population sample in different year who have a Vantage 4.0 credit score. Panels A-D break this out by race and Panels E-H split by parent income.

FIGURE A.2
Rates of Having a Vantage 4.0 Score by Age

A. Race

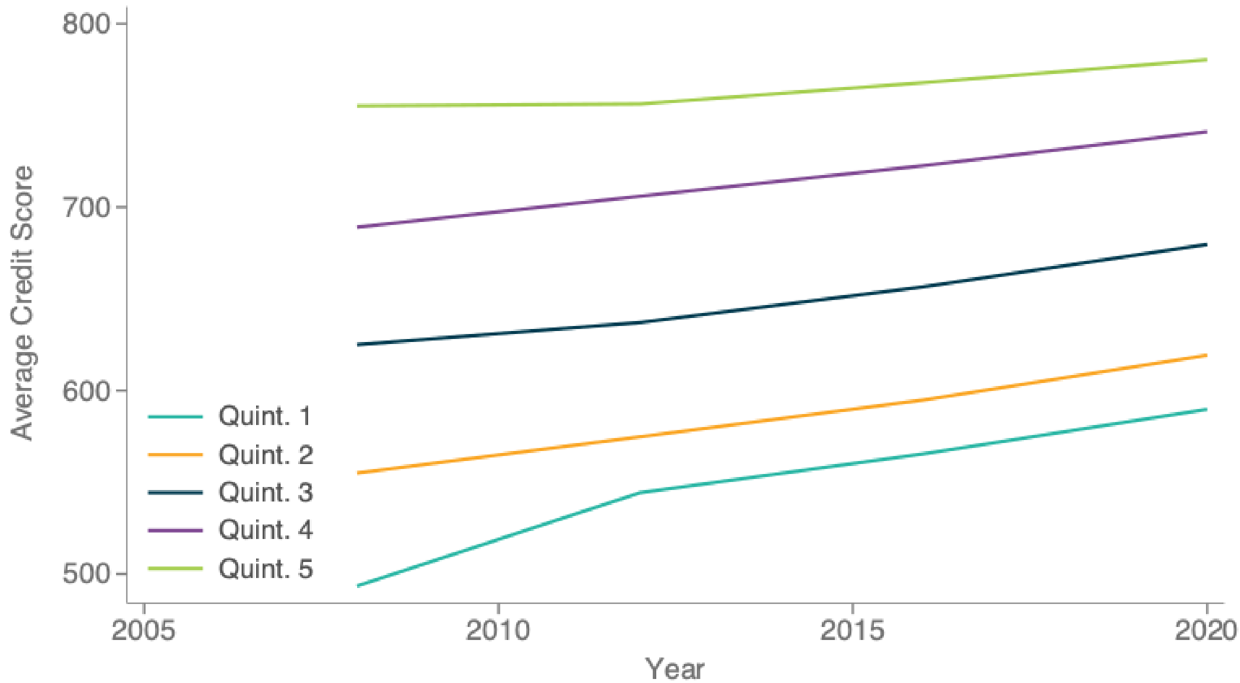


B. Parent Income Quintile



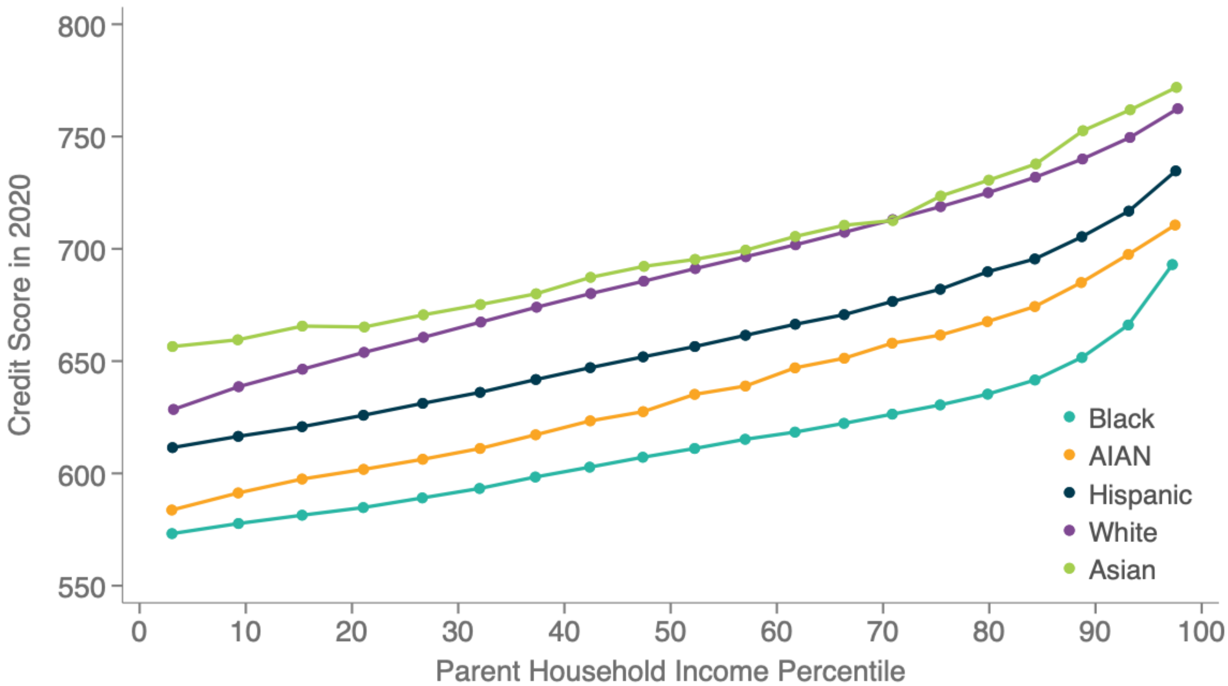
Notes: This figure presents the fraction of individuals in our population sample in 2020 who have a Vantage 4.0 credit score. Panel A plots this rate by race. Panel B plots this rate by parent household income percentile.

FIGURE A.3
Credit Score Within-Person Persistence



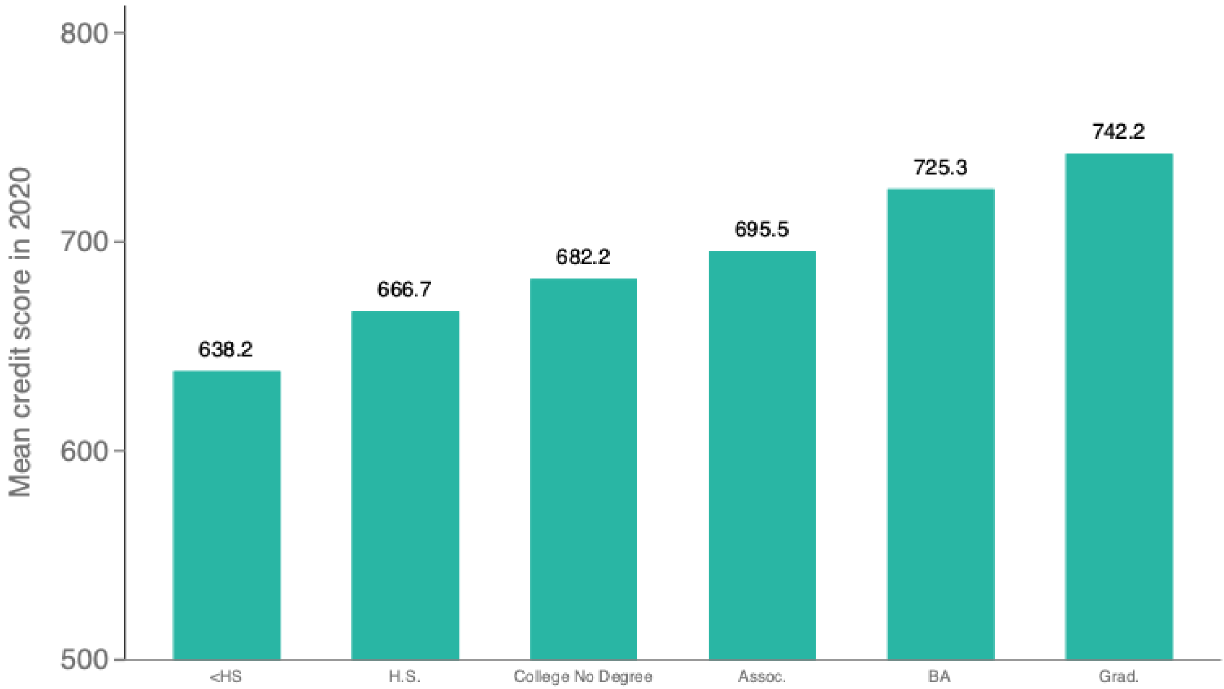
Notes: This figure presents the average credit scores by year in our intergenerational sample by the quintile of credit score in 2008.

FIGURE A.4
Credit Scores by Race and Parent Income 2020, Children of Native-born Mothers



Notes: This figure presents the average credit scores in 2020 by race and parental income percentile as in Figure IV, but restricts to the subset of children whose mothers were born in the US.

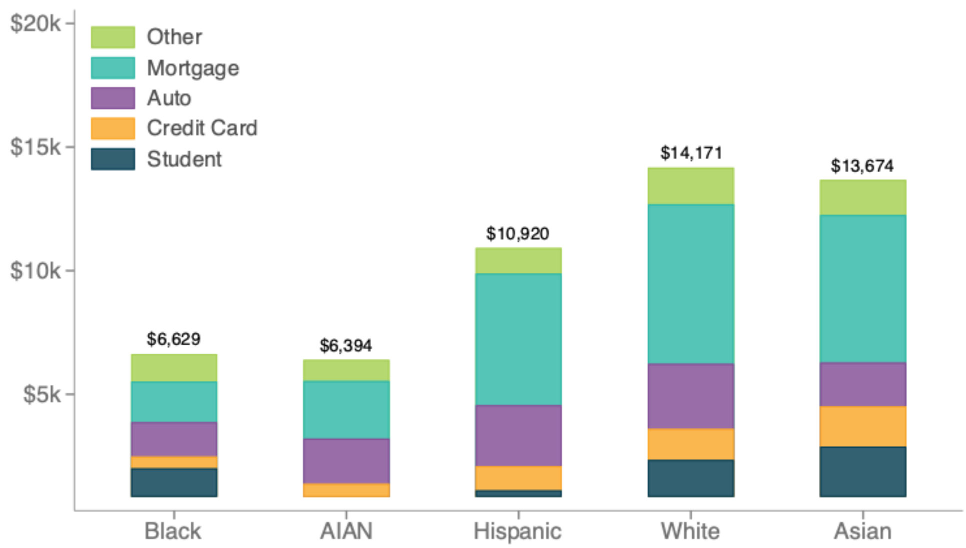
FIGURE A.5
Credit Scores by Parent Education



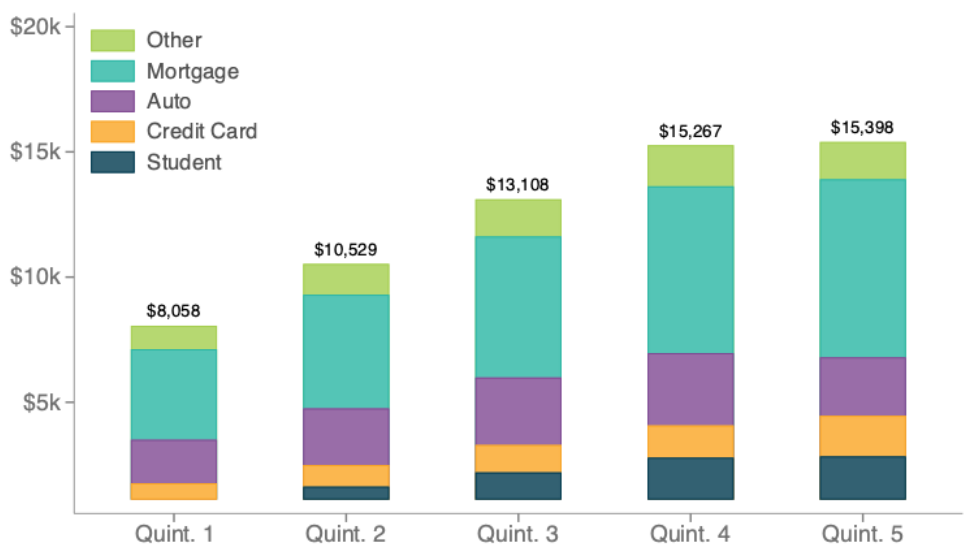
Notes: This figure presents the average credit scores in 2020 by parent education for our intergenerational sample. We define parent education as the most education received by either parent from the ACS.

FIGURE A.6
Debt Composition By Race and Class in 2004

A. Debt Composition by Race

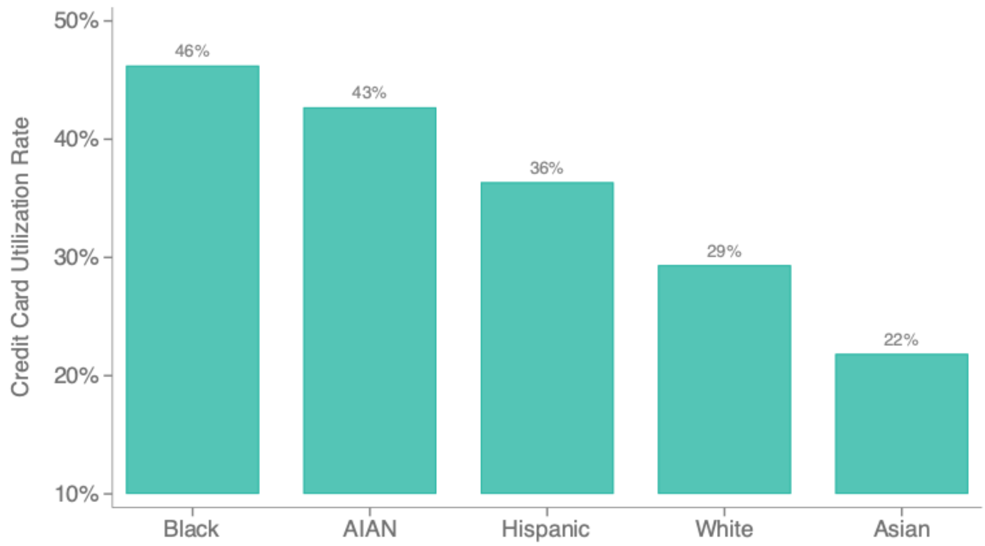


B. Debt Composition by Parent Income



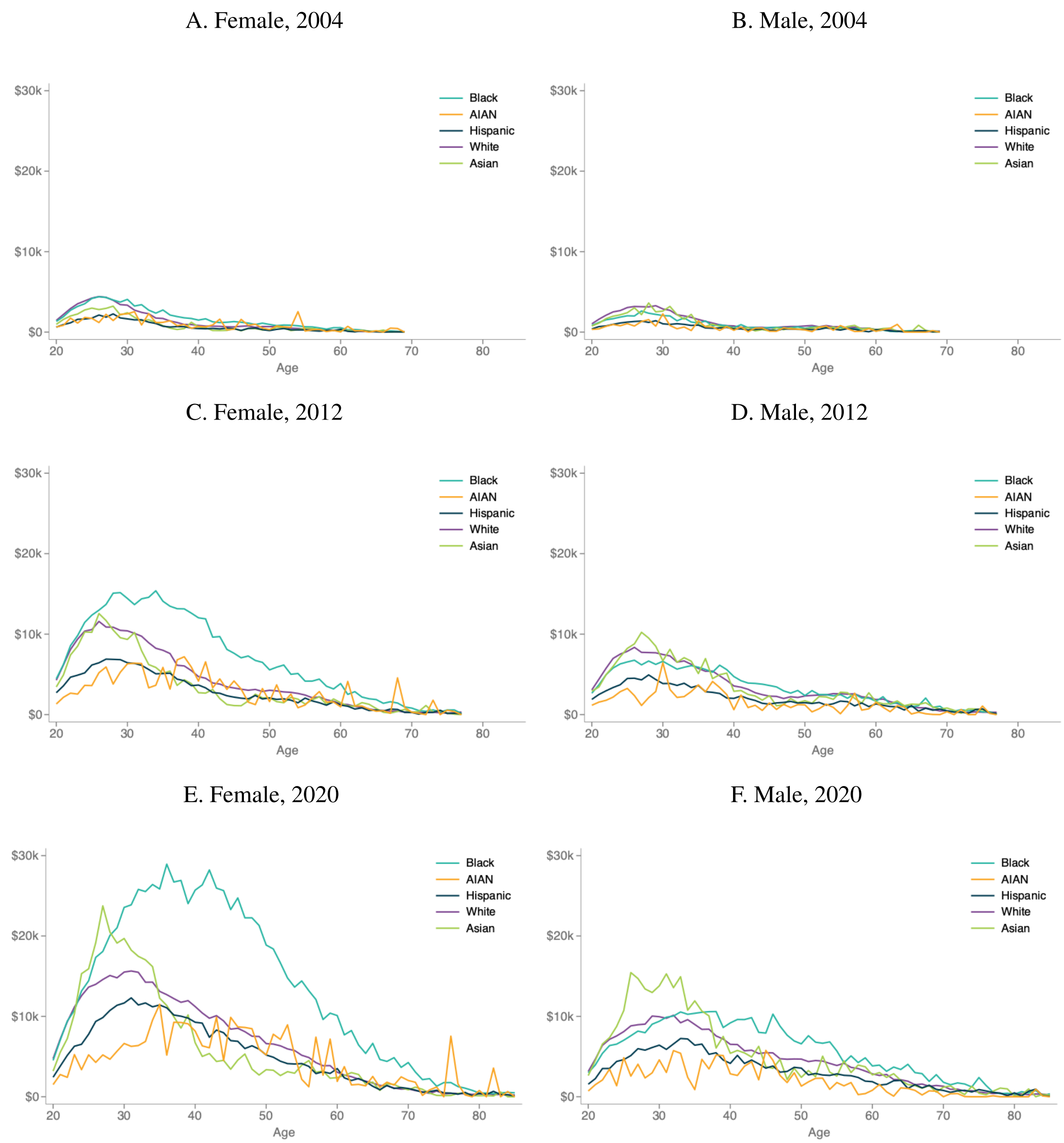
Notes: This figure shows the composition of debt balances by race and class in 2004. Panel A depicts the data breakdown by race, and Panel B shows the breakdown by class quintiles.

FIGURE A.7
Credit Card Utilization by Race



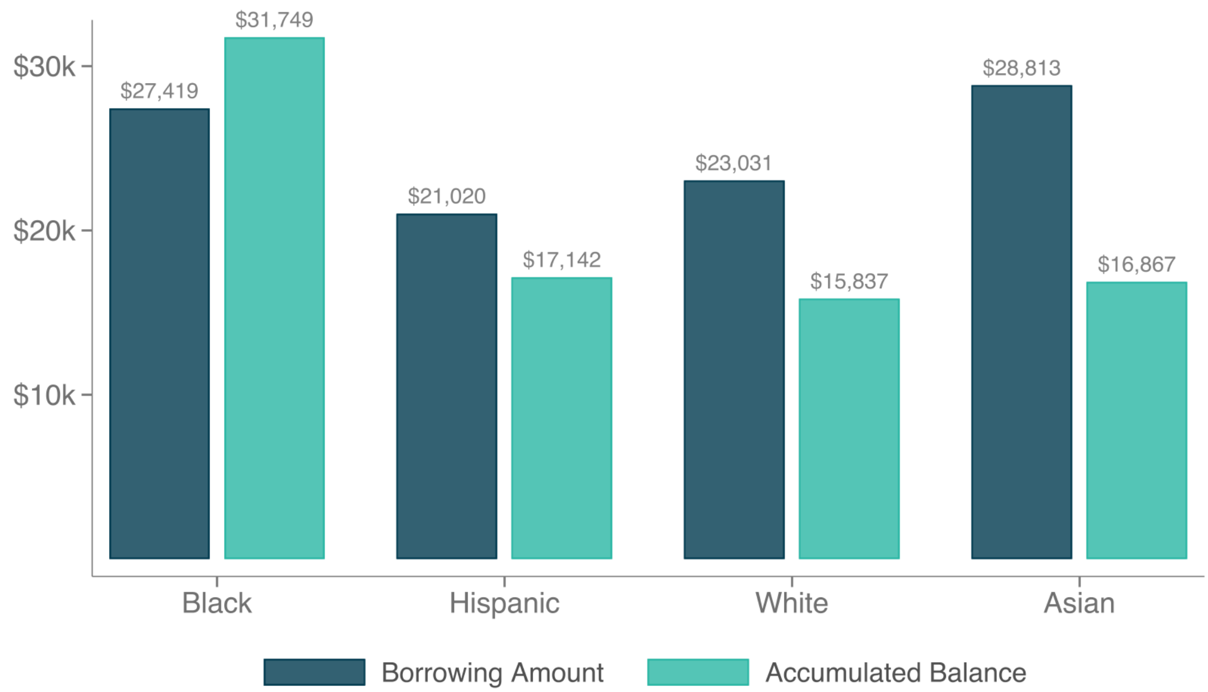
Notes: This figure presents the average individual level credit card utilization in our intergenerational sample in 2020 by race.

FIGURE A.8
Student Debt by Year and Race



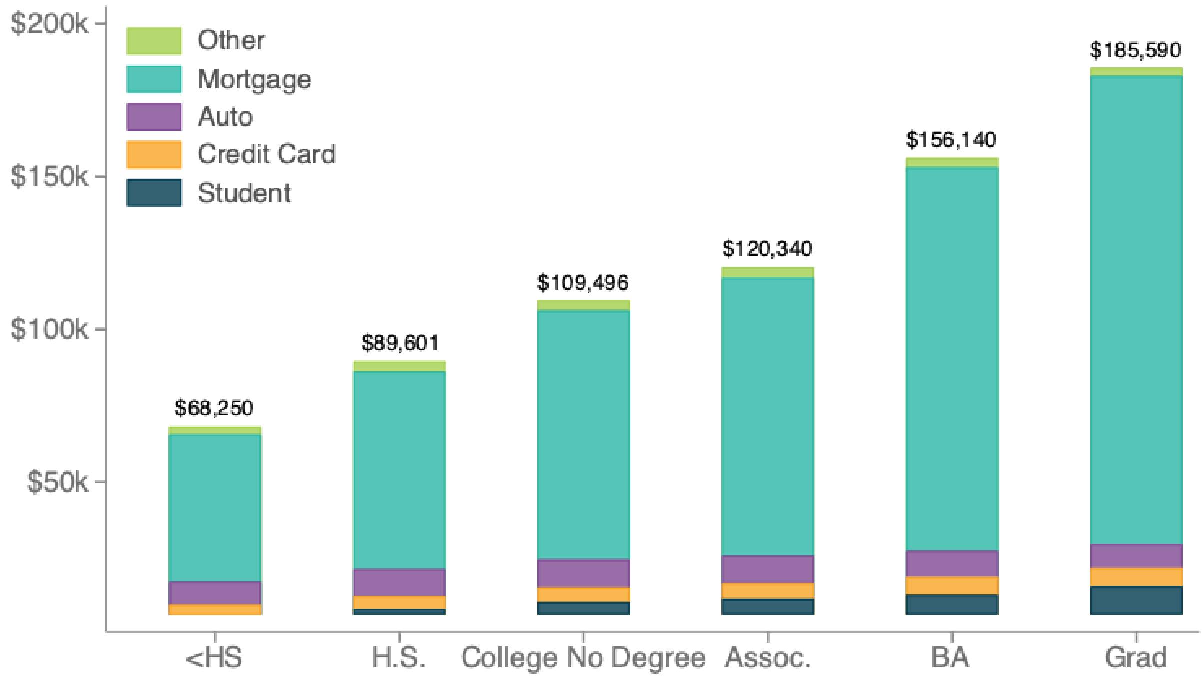
Notes: This figure presents the average student loan balance in our population sample by age broken out by year and gender.

FIGURE A.9
Student Debt Borrowing Amounts and Balances



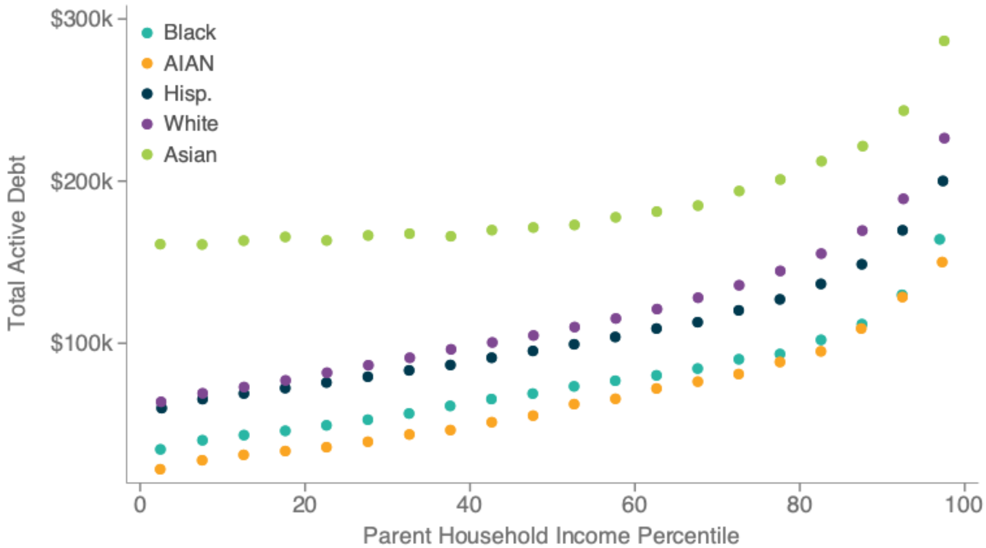
Notes: This figure plots average borrowing amounts and accumulated balances by race as of 2015 among student borrowers who entered college in 1996. Averages are calculated using data from the 1996 Beginning Postsecondary Students (BPS) study, which merges survey responses with administrative records for a representative sample of 12,400 first-time college enrollees. Source: U.S. Department of Education, National Center for Education Statistics, 1996 Beginning Postsecondary Students (BPS) study, authors' calculations.

FIGURE A.10
Composition of Total Debt by Parent Education



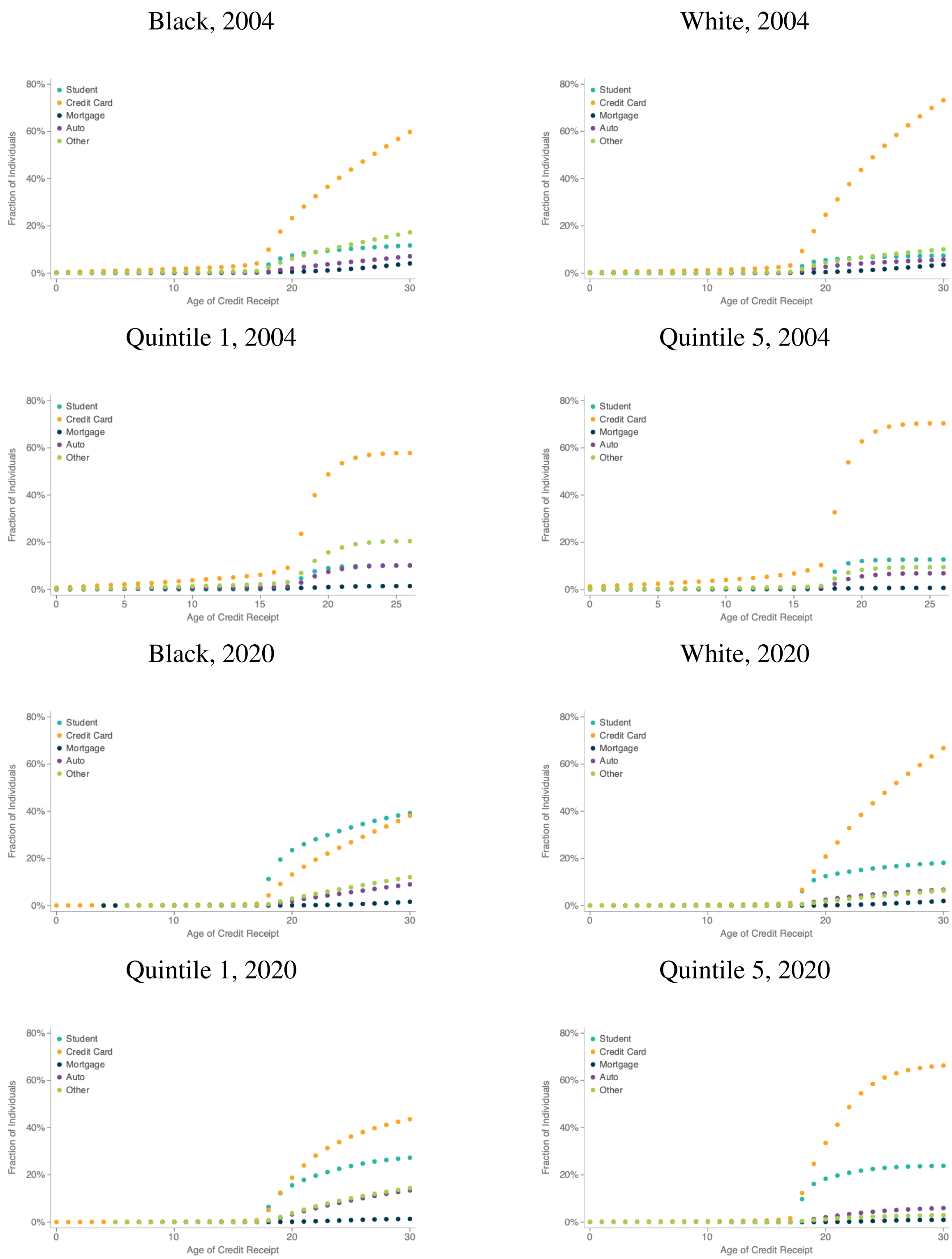
Notes: This figure presents a stacked bar chart of average debt holdings by type of credit split out by parental education. We present 4 credit types: mortgages, auto loans, credit card balances, and student loans, which comprise nearly all debt on credit reports.

FIGURE A.11
Total Debt by Race and Parental Income Percentile



Notes: This figure presents a binned scatter plot of total active debt in 2020 on parent household income percentile by race in our intergenerational sample.

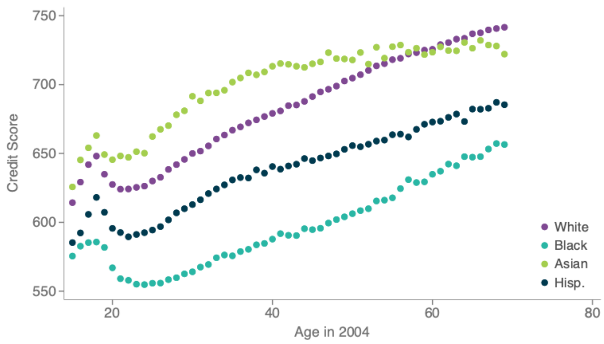
FIGURE A.12
Type of First Tradeline by Age



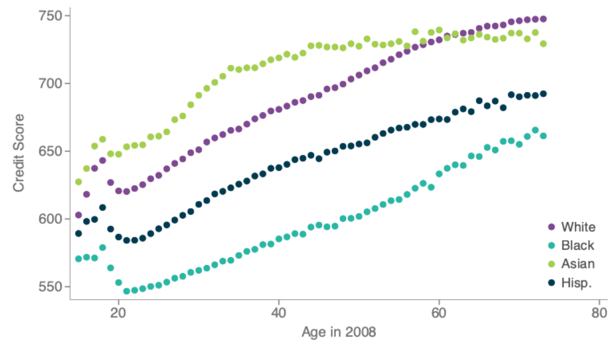
Notes: This figure presents the type of first tradeline by age and category of debt. For each individual, we measure the age at which they first obtained a tradeline in the credit file and the type of tradeline they received. Then, for each age, we plot the fraction of people whose first tradeline was a given type amongst those who have received a tradeline by that age. At older ages, the graph therefore presents the fraction of borrowers in the population who received different types of credit as their first tradeline.

FIGURE A.13
Credit Scores by Age and Race, by Year

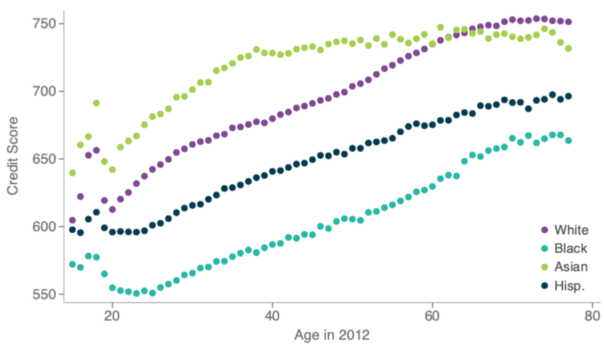
A. 2004



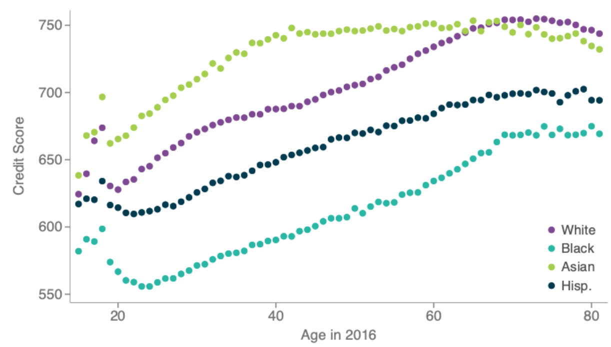
B. 2008



C. 2012

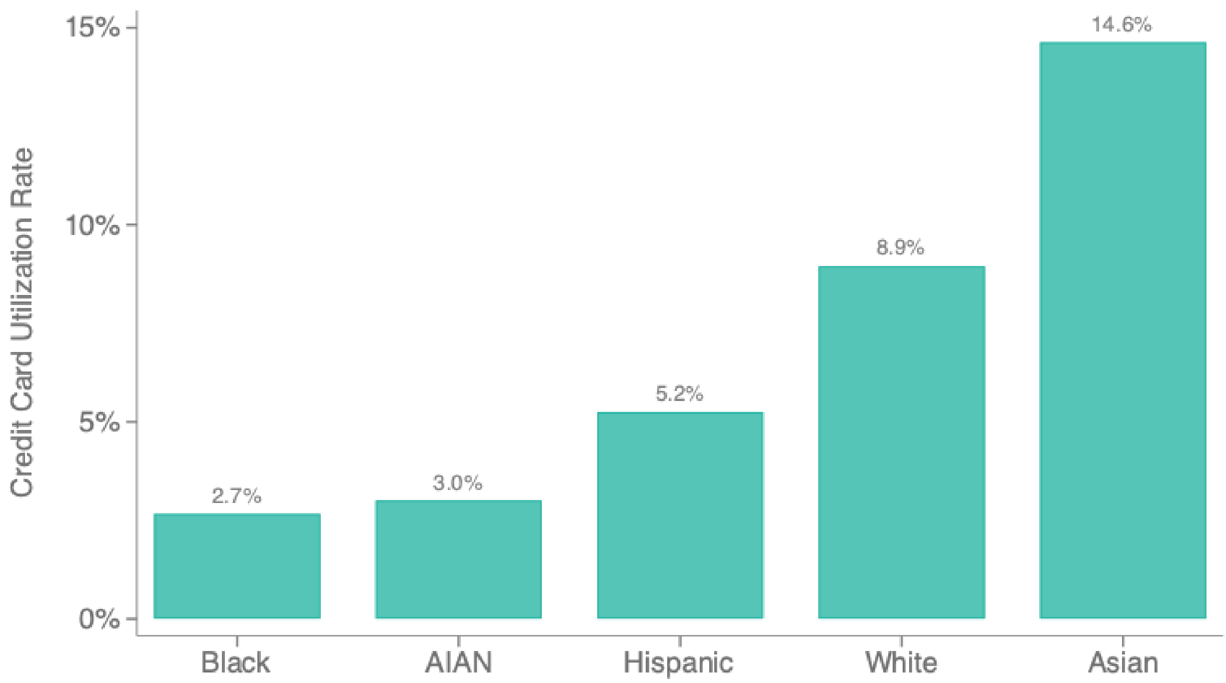


D. 2016



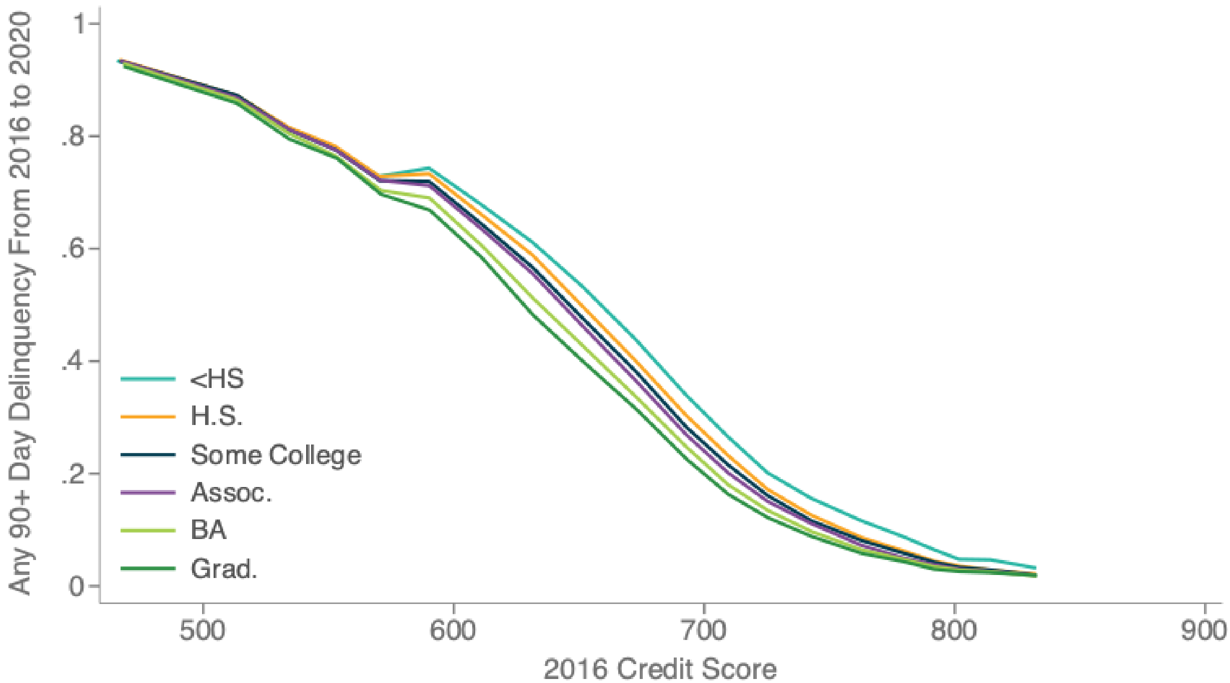
Notes: This figure presents the average credit score by age in years 2004, 2008, 2012, and 2016 by race.

FIGURE A.14
Incidence of Authorized User Trades



Notes: This figure presents percentage of individuals in our intergenerational sample in 2004 who have have an authorized user trade by race.

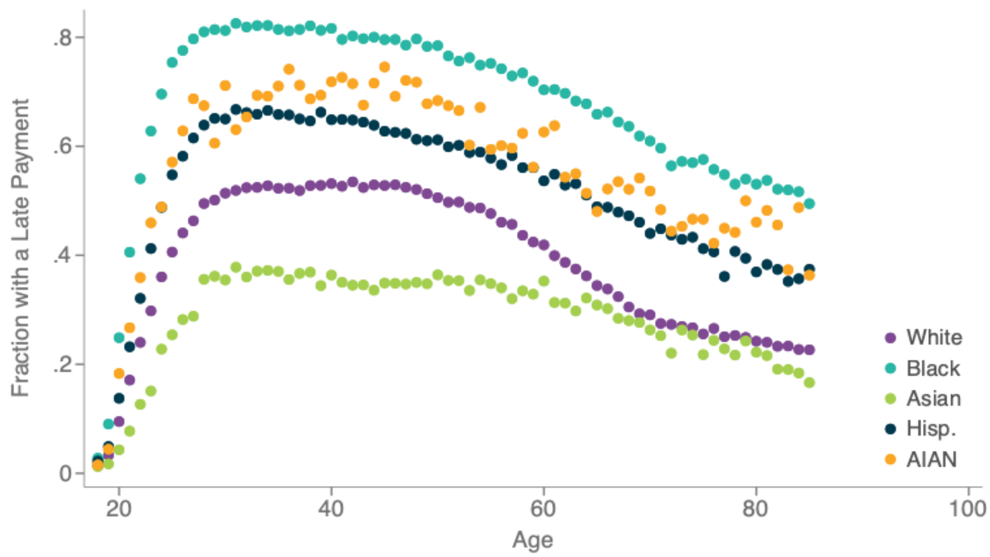
FIGURE A.15
90+ Day Delinquency versus Credit Score by Parental Education



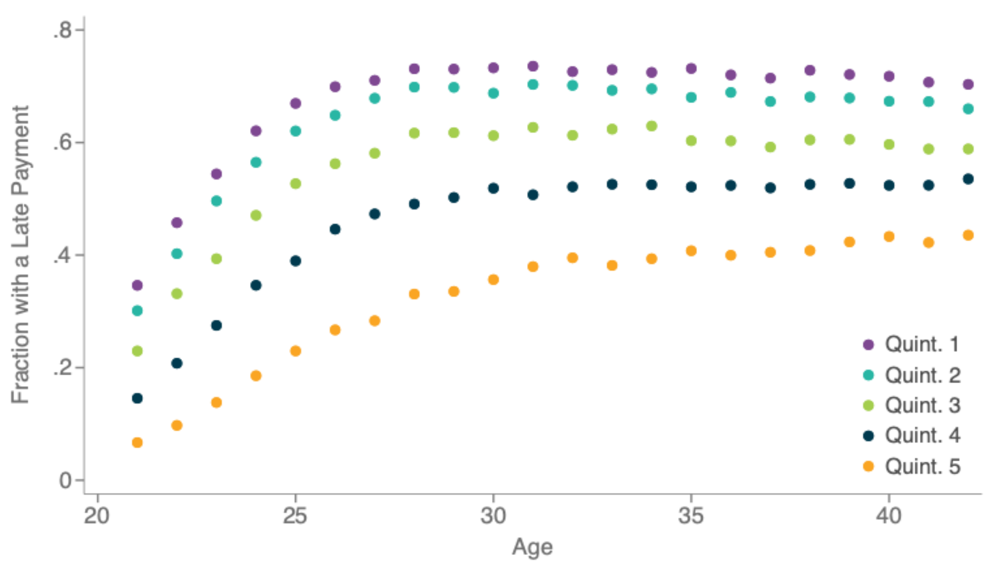
Notes: This figure presents the average 90+ day delinquency rate in the 2020 credit file as a function of the 2016 credit score on the horizontal axis, separately by parental education.

FIGURE A.16
Any Late Payments Reported to the Credit Bureau by Age and Race (30+ Days)

A. Race



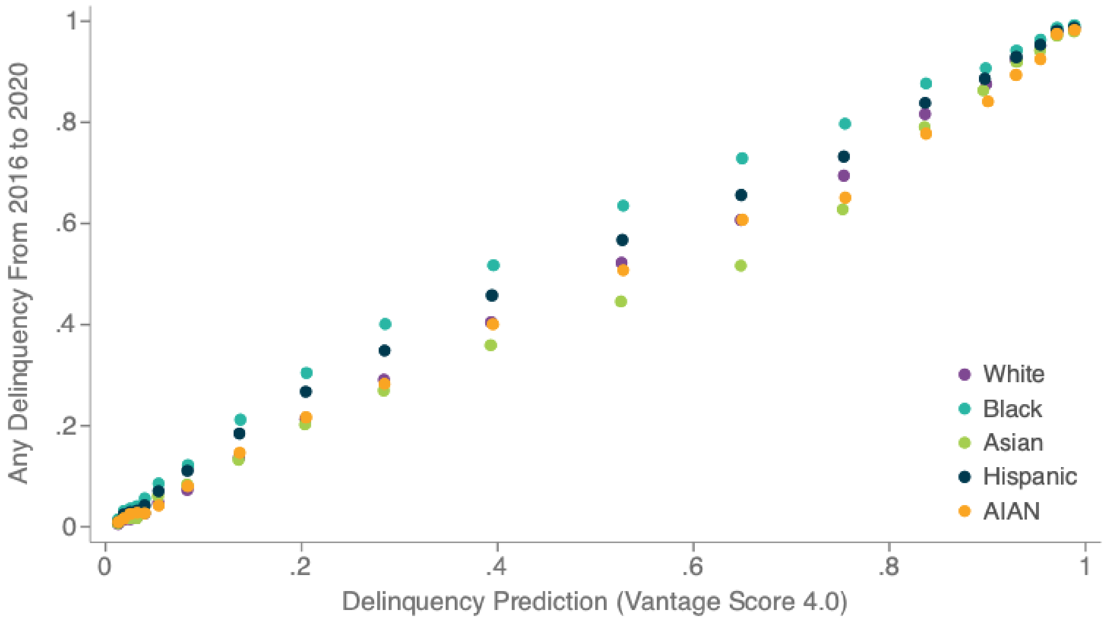
B. Parent Income Quintile



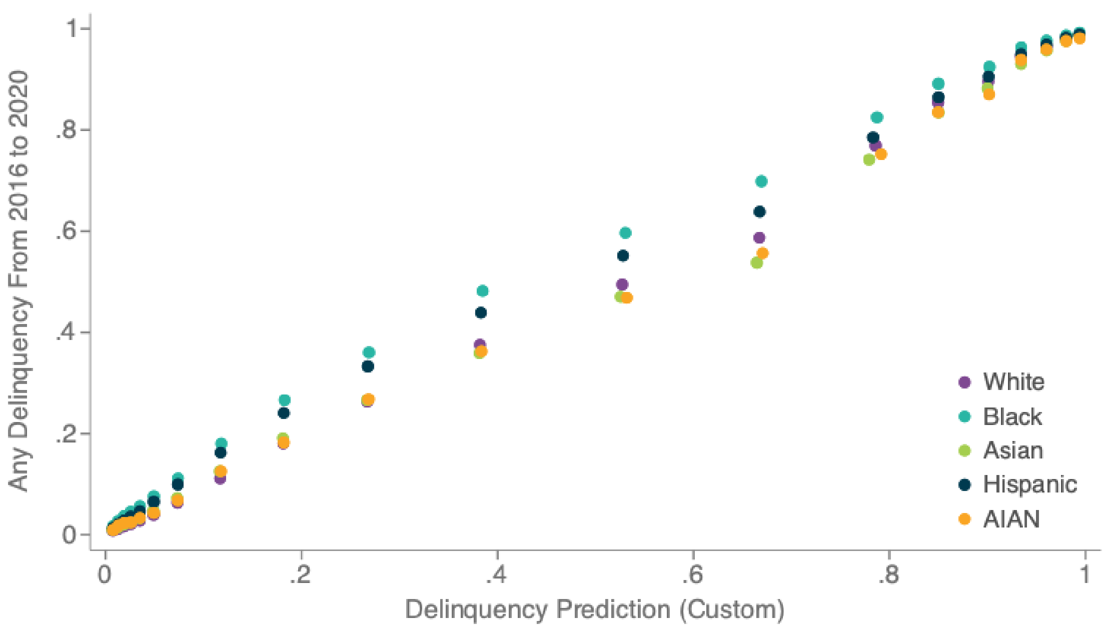
Notes: This figures shows the fraction of individuals with a late payment (30+ day late) by age. Panel A splits out by race and Panel B breaks it out by class.

FIGURE A.17
Alternative Delinquency Predictors

A. Experian's Prediction

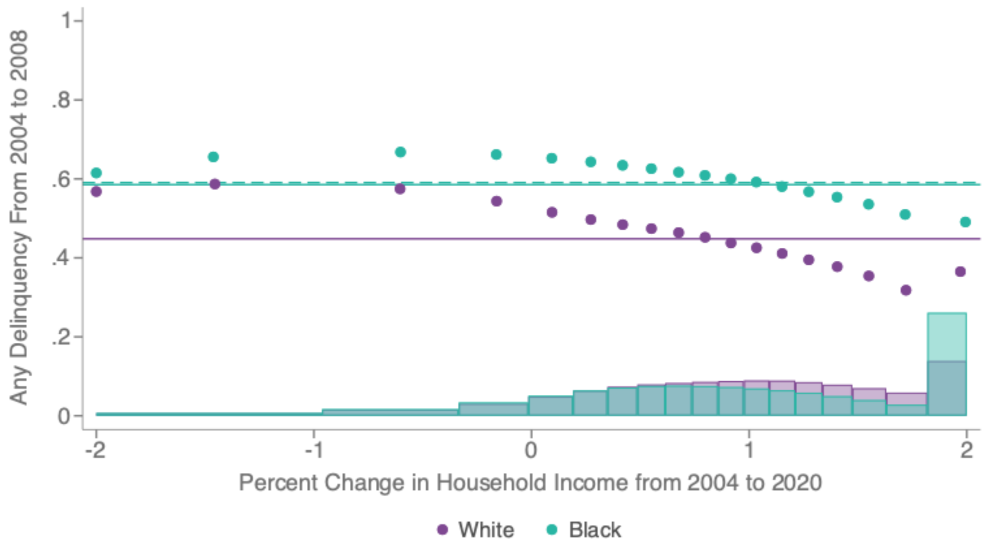


B. Custom Prediction



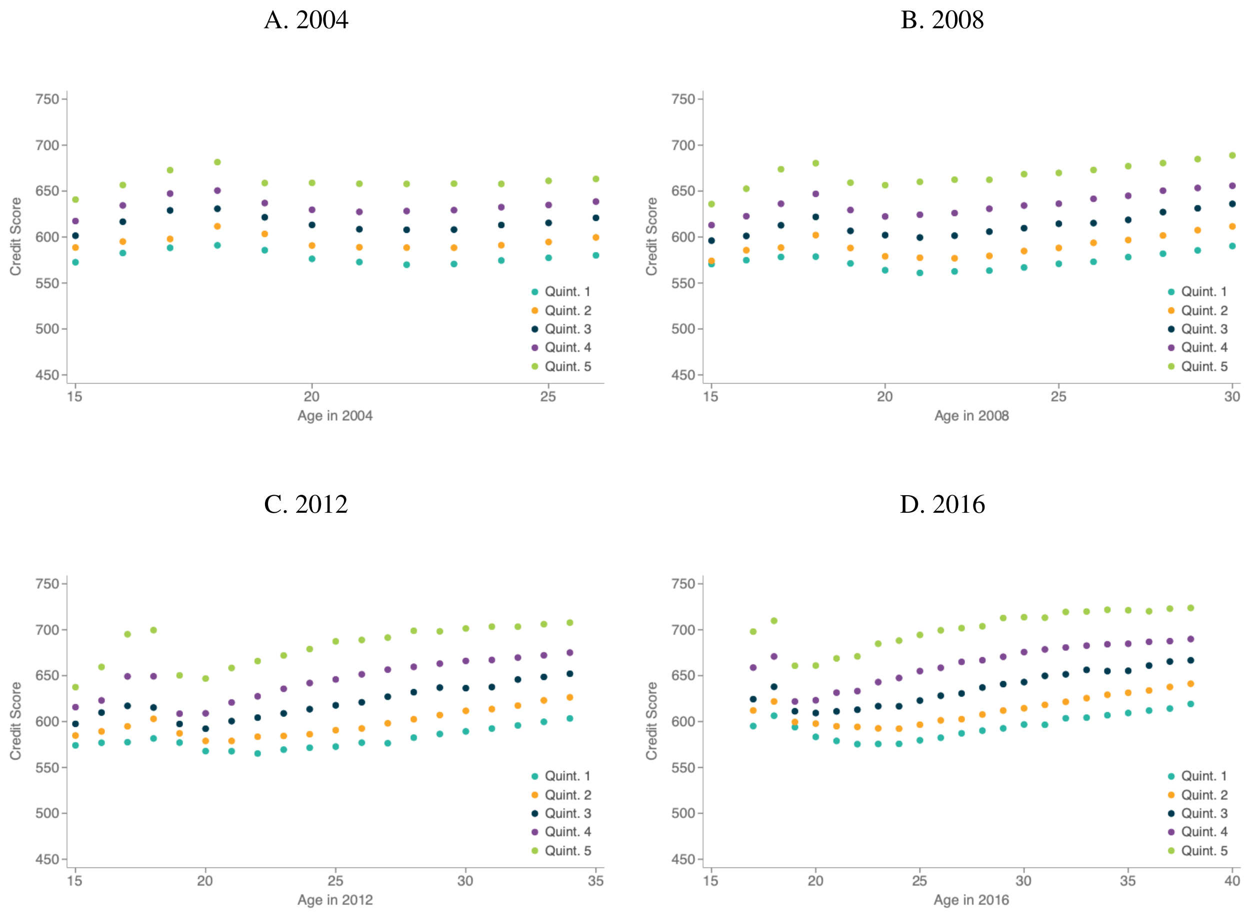
Notes: This figure compares our conclusions for calibration bias by race when using the Vantage 4.0 score versus a simple logit regression using a vector of credit score attributes. The horizontal axis presents the predicted probability of observing a 90+ day late payment on the 2020 credit score as a function of the 2016 credit file information. Panel A uses the 2016 Vantage 4.0 score to predict delinquency while Panel B uses our logit regression prediction using 2016 credit file variables, showing that the calibration bias patterns are similar when using other prediction algorithms from the information in the credit file.

FIGURE A.18
90+ Day Delinquency versus Income Growth (2004-2020)



Notes: This figure presents This figure reproduces XIV Panel B, except with longer term income growth. The figure presents the average 90+ day delinquency rate in 2004 by income change bin from 2004-2020 conditional on credit score rank, separately for Black and White individuals. The histograms present the distribution of household income percent changes by race. The horizontal lines present the mean delinquency rates for Black individuals if they had the same delinquency conditional on income changes but faced the White distribution of income changes. Income change is defined as $\frac{Income_{2020} - Income_{2004}}{0.5(Income_{2020} + Income_{2004})}$.

FIGURE A.19
Credit Scores by Age and Parent Income Quintile

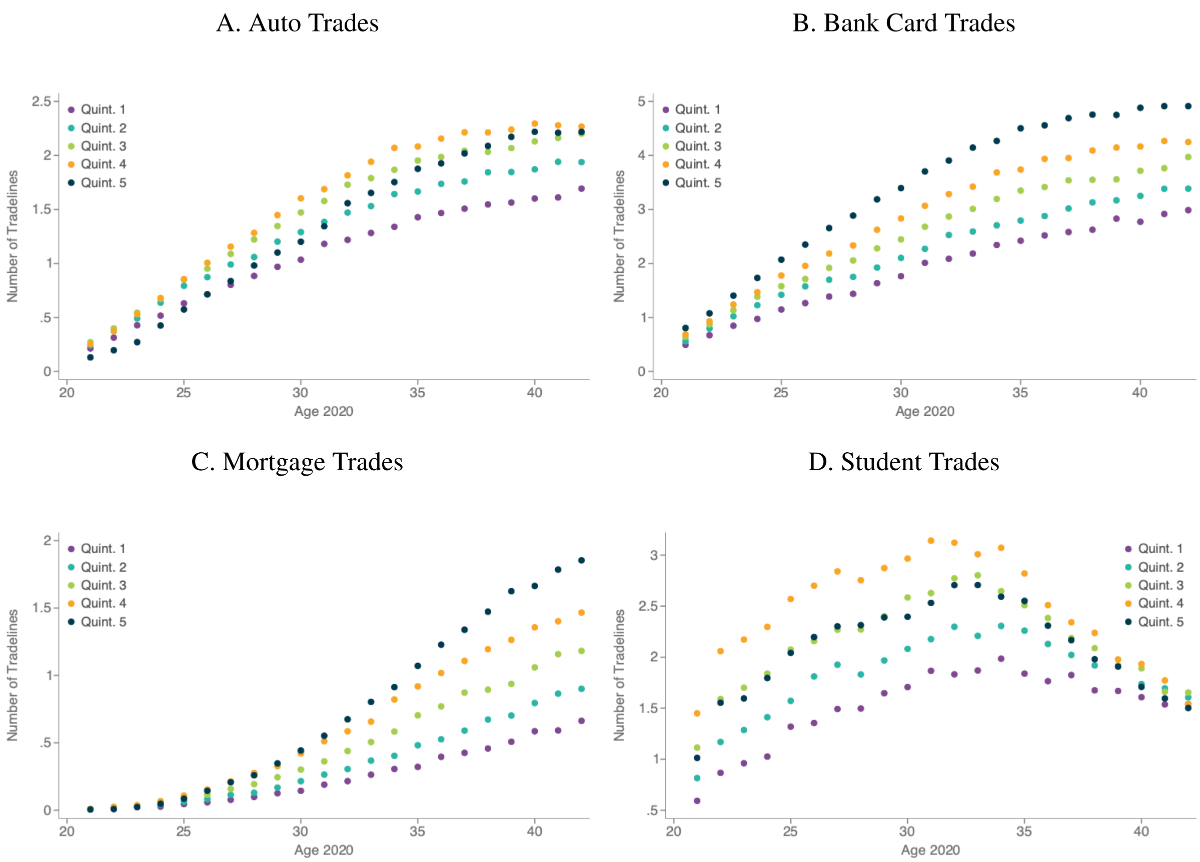


Notes: This figure presents the average credit score by age in years 2004, 2008, 2012, and 2016 by parent household income quintile.

FIGURE A.20
Number of Tradelines by Age and Race

Notes: This figure presents the number of tradelines across different categories of debt in our population sample by race/ethicity.

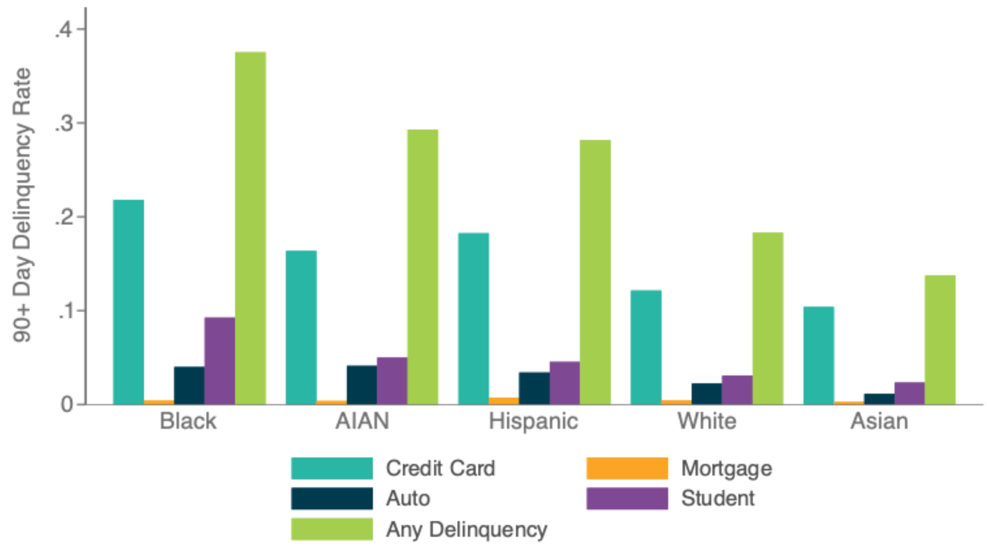
FIGURE A.21
Number of Tradelines by Age and Parental Income Quintile



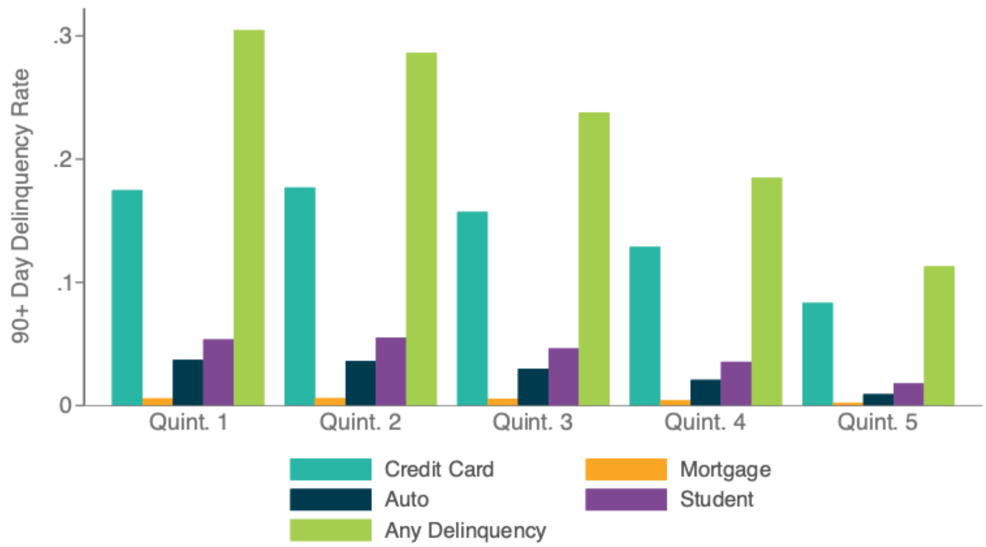
Notes: This figure presents the number of tradelines across different categories in our population sample by class.

FIGURE A.22
90+ Day Late Payment Breakdown of 18-25 Yr Olds by Debt Type

A. Race



B. Parent Income Quintile



Notes: This figure shows the fraction of individuals in our intergenerational sample in 2004 who are 90+ day delinquent by the type of tradeline. Panel A breaks it out by race and Panel B breaks it out by parent household income quintile.

TABLE A.1
Estimation of Median Credit Scores by Race/Ethnicity

Race	True Median	Zip Code Weighted By Population	Median Credit Score Mortgage Holders
Asian	768		781
Black	604	602.4	669.5
Hispanic	671	676	710.6
White	743	737.1	749

Notes: This figure compares the estimates of the median credit score by race using zip codes where more than 60% of the population identifies as a given racial group.

TABLE A.2
Average Credit Card Credit Limits by Race

Black-White Gap	-16538.6***	-7026.4***	-3935.8***
	(562.7)	(541.5)	(508.2)
Hispanic-White Gap	-14332.1***	-7069.9***	-3535.9***
	(662.3)	(596.2)	(576.2)
Asian-White Gap	12803.4***	8042.4***	5650.9**
	(2234.0)	(1982.1)	(1914.1)
Log Family Income		16525.0***	11285.1***
		(294.4)	(695.3)
All loan payments made as scheduled in last year			6822.3***
			(507.0)
Observations	22728	22728	22728
Controls	None	Log Income	Full

Notes: This table presents average credit limits on credit cards, separately by race, as reported in the 2022 Survey of Consumer Finances. All regressions use SCF sample weights. "Full" controls consist of log income, log net wealth, income quintile indicators, net wealth quintile indicators, an age quadratic, and an indicator for experiencing any delinquency in the past year.

TABLE A.3
Differences in Average Interest Rate by Race, by Tradeline

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Credit Card	Auto Loan	Mortgage	Credit Card	Auto Loan	Mortgage	Credit Card	Auto Loan	Mortgage
Black-White	0.225	1.904***	0.792*	0.0310	1.237***	0.625	-0.363	0.696**	1.032**
	(0.210)	(0.277)	(0.368)	(0.212)	(0.272)	(0.359)	(0.223)	(0.260)	(0.317)
Hispanic-White	-0.0425	1.352***	0.287	-0.192	0.889***	0.132	-0.486*	0.537*	-0.150
	(0.246)	(0.245)	(0.314)	(0.246)	(0.240)	(0.325)	(0.244)	(0.243)	(0.723)
Asian-White	0.465	-1.488***	0.634	0.587*	-0.699**	0.864	0.525	-0.569*	0.153
	(0.292)	(0.255)	(0.834)	(0.293)	(0.256)	(0.710)	(0.286)	(0.236)	(0.487)
Log Family Inc.				-0.430***	-1.947***	-0.329*	0.303*	-0.242	0.937*
				(0.0701)	(0.115)	(0.132)	(0.134)	(0.218)	(0.376)
Observations	17573	5556	291	17573	5556	291	17573	5556	291
Controls	None	None	None	Log Income	Log Income	Log Income	Full	Full	Full

Notes: This table presents gaps in interest rates for credit cards, auto loans, and mortgages by race as reported in the 2022 Survey of Consumer Finances. All differences are relative to White borrowers. All regressions use SCF sample weights. "Unconditional" bars present raw differences. 'Full" controls consist of log income, log net wealth, income quintile indicators, net wealth quintile indicators, an age quadratic, and an indicator for experiencing any delinquency in the past year. Net wealth is computed following the method used for Federal Reserve Bulletin articles: https://www.federalreserve.gov/econres/files/bulletin.macro.txt. Each column includes only borrowers who have the relevant tradeline.

TABLE A.4
90+ Day Delinquency versus Income and Wealth for Stably Employed Sub-Sample

	(1)	(2)	(3)	(4)	(5)
			90+ Day Delinquency		
Black	0.084	0.105	0.096	0.098	0.089
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Asian	-0.022	0.002	0.009	0.010	0.014
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Hispanic	0.034	0.032	0.038	0.038	0.031
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Credit Score Rank	-1.136	-0.765	-0.877	-0.869	-0.860
	(0.000)	(0.001)	(0.001)	(0.001)	(0.002)
Income			-0.064	-0.060	-0.064
			(0.002)	(0.002)	(0.002)
N	24,970,000	1,158,000	1,526,000	1,428,000	1,242,000
R^2	0.425	0.243	0.318	0.315	0.383

Notes: This table presents a regression of any delinquency from 2016 to 2020 on race dummies. Column (1) controls for credit score rank in 2016. Column (2) presents the same regression on the subset of individuals who have a positive mortgage balance in 2016 and a new mortgage trade in the last 6 months. Column (3) restricts to individuals who are stably employed from 2015 to 2021 and controls for household income rank in 2016. We define stably employed as having one w2 in each year from 2015 to 2021 from the same employer. Also, the individual's income rank in every year from 2016 to 2021 must stay within 10% of their income rank in 2015. Column (4) adds the restriction that individuals maintain the same marital status from 2015 to 2021. Column (5) adds employer fixed effects to the spec in Column (4).