CAGE

World War II and Black economic progress

CAGE working paper no. 387

September 2018 (Revised May 2021)

Andreas Ferrara

Formerly titled: World War II and African American socioeconomic progress

WARWICK
THE UNIVERSITY OF WARWICK

UKRI
Economic and Social Research Council

World War II and Black Economic Progress

Andreas Ferrara*

University of Pittsburgh

12th March, 2021

Abstract

During the 1940s, a substantial share of Southern Black men moved from low-skilled to much better paying semi-skilled jobs. Using newly digitized military data, I show that counties with higher World War II casualty rates among semi-skilled White soldiers saw an increase in the share of semi-skilled Black workers. These deaths opened new employment opportunities for Black Southerners and, together with learning effects by employers, can explain up to 22.6% of the occupational upgrading at mid-century. I provide evidence that the casualty-induced labor shortages reduced racial barriers to entry, leading to a positive selection of Black workers into semi-skilled employment.

Keywords: AFRICAN-AMERICANS; OCCUPATIONAL CHOICE; WORLD WAR II.

JEL Classification: J15, J24, N42

*University of Pittsburgh, Department of Economics, and CAGE. Email: a.ferrara@pitt.edu

I thank Cihan Artunç, Martha Bailey, Sascha O. Becker, Leah Boustan, Clément de Chaisemartin, James Fenske, Price Fishback, Carola Frydman, Stephan Heblich, Taylor Jaworski, Ian Keay, Christoph König, Felix König, Allison Shertzer, Luigi Pascali, Steve Pischke, Patrick Testa, Nico Voigtländer, Fabian Waldinger, and seminar participants at the University of Arizona, Australian National University, Carlos III, Collegio Carlo Alberto, London School of Economics, Monash University, University of New South Wales, University of Pittsburgh, Universitat Pompeu Fabra, Queen's University, University of Warwick, and at the 3rd ASREC Europe conference, 56th Cliometric Society Conference, 29th EALE conference, EEA-ESEM Congress 2018, 78th EHA annual meeting, EHS conference 2018, IZA World Labor Conference 2018, NBER DAE Summer Institute 2020, RES conference 2018, 18th World Economic History Congress, 23rd Spring Meeting of Young Economists, 7th IRES Graduate Student Workshop, 3rd RES Symposium of Junior Researchers, and the 20th IZA Summer School for valuable comments and discussions.

1 Introduction

The 1940s marked a major turning point in Black Southerners' economic progress. A key driver of this development was the shift of Black workers from low- into semi-skilled occupations (Wolfbein, 1947). In 1940, only 15% of Southern Black men worked in semi-skilled jobs, which paid significantly better than low-skilled jobs that the majority of Black men held at the time.[1] However, Southern Black workers faced significant racial barriers to entry into these types of jobs prior to the 1940s (Myrdal, 1944; Collins, 2001).[2] In 1950, 26% of African American men worked in semi-skilled jobs with this share rising to over 45% in 1970. This trend break in the occupational structure among Black men occurred both within and outside the South (see Figure 1), and related work has shown that occupational upgrading can explain much of the narrowing in Black-white economic outcomes over this period (Maloney, 1994; Margo, 1995; Collins, 2000; Aizer, Boone, Lleras-Muney and Vogel, 2020).[3] Given the large wage difference between low- and semi-skilled jobs, this upgrading meant a substantial change in the economic position of Black workers and families. This raises the question which factors can explain this occupational change at mid-century.

In this paper I seek to answer the question of how Black workers in the South were able to make such rapid and significant occupational advances during and after the Second World War, even while living under the repressive Jim Crow regime. My main hypothesis focuses on the permanent labor reductions brought by the differentially higher wartime mortality rates among semi-skilled white soldiers. I show that their disappearance from the labor market opened new opportunities for Black workers to move into semi-skilled jobs. I construct a data set of newly digitized mortality records which I combine with WWII Army Enlistment data for over 8 million soldiers as well as with county-level employment information from 1920 to 1960 for the U.S. South. Difference-in-differences results show that between two and three African American men entered semi-skilled employment for each fallen white soldier with a semi-skilled pre-war occupation. This is similar to the employment increase during World War I when firms hired an additional Black worker for each Black worker they already employed, thus doubling their African American workforce (Whatley, 1990). Given the significantly larger dimensions

[1] The average semi-skilled job in 1940 paid more than one and a half times the wage of the average low-skilled job, see Appendix Figure A1.

[2] For an example see the study of the aircraft industry by Weaver (1945).

[3] Also the Great Migration contributed to this change, however, Smith and Welch (1989) estimate that about 20% of Black-white wage convergence at this time was due to migration.

of World War II, one would expect not only one Black worker to replace a fallen semi-skilled white soldier but at least one or two additional Black workers to follow into semi-skilled work. This larger than proportional effect can be motivated with the declining descrimination on part of employers at the time which was an important driver of occupational upgrading among Black workers (Aizer et al., 2020), and race-based referral networks in hiring (Montgomery, 1991; Royster, 2003).

Since the 1940s were a period of significant socioeconomic changes, I rule out other explanations including the industrialization of the South, selective migration of Black workers within and out of the South, differences in education and incomes, skill-biased technological change in agriculture that released Black labor into the sectors with higher shares of semi-skilled employment, selection on observables and unobservables, differences in historic anti-Black sentiment as measured by past lynchings and land in cotton production, federal spending programs related to the war, soldiers' characteristics, and legacies of New Deal spending. I also show that it is only deaths among semi-skilled white workers that affect Black occupational upgrading and not deaths in other race and skill groups. The Fair Employment Practice Committee is also less of a concern in the South where it had little enforcement power (Collins, 2001).

I then test whether the mechanism behind this occupational change was a reduction in racial barriers to entry, or if it this was simply driven by improved economic opportunities that benefited everyone regardless of race. I analyze repeated cross sectional data on individual workers from the U.S. Census between 1920 to 1960 in a triple differences setting to compare the changes in the probability of occupational upgrading from low- to semi-skilled employment for Black and white workers from before to after the war in low- and high-casualty rate commuting zones.[4] War casualties among semi-skilled white workers did not improve the occupational prospects for low-skilled white men relative to the opportunities they already had before the war. Instead, the casualty induced labor shortages reduced racial barriers to entry to these jobs as firms decided to not hire less qualified white men over more qualified Black workers whom they would not have considered in the absence of such labor shortages. This is consistent with the results by Aizer et al. (2020). Results show that the average WWII casualty rate among semi-skilled white soldiers and the associated multiplier effects can explain 22.6% of the rise in semi-skilled employment among Southern Black men from 1940 to 1960.[5]

[4]Since the county of residence is not provided after 1940, I use commuting zones to assign the treatment as these can be consistently constructed between 1920 and 1960.

[5]The *multiplier effects* refer to follow-up employment after entrance of the first Black worker who replaced a

To better understand the selection of new semi-skilled workers, I build and test a simple theoretical model in which Black and white individuals work in either agriculture, which employs all low-skilled jobs, or manufacturing, which employs all semi-skilled jobs. Workers move across the two industries based on their ability in each sector. While it is preferable to work in manufacturing where wages are higher, there is a also switching cost of moving across sectors which is differentially higher for Black than for white workers due to discrimination. Once the discriminatory cost for Black workers is removed or lowered enough, the model predicts that more Black than white workers will upgrade from agricultural to manufacturing work, and that those Black workers will be positively selected compared to both other Black workers in agriculture as well as new white manufacturing entrants. The model can motivate the observed empirical patterns in post-war characteristics of Black and white semi-skilled workers with Black workers having higher education, wages, and house values.

This paper provides new empirical evidence for a causal link between wartime casualties and Black economic advancement at mid-century. I contribute to the literature on the labor market effects of the World War II draft which has mainly focused on white women.[6] Work considering Black men has examined the effect of veteran status on manufacturing employment (Collins, 2000) or on post-war education via the G.I. Bill which, however, did not help African American men in the South (Turner and Bound, 2003). Labor economists at the time also noted the large scale occupational upgrading of Black workers during and after the war (Weaver, 1945; Wolfbein, 1947). Later studies by Maloney (1994) and Margo (1995) find that Black occupational upgrading had a sizable influence on the reduction in Black-white earnings differentials during the 1940s. Also Collins (2000) shows how such upgrading improved economic mobility among Black workers in the north. A recent study by Aizer et al. (2020) shows that spending in the defense industry was an important factor in the occupational upgrading of Black workers during the war which led to a significant decline in the Black-white wage gap. They also provide evidence for intergenerational spillovers that raised Black children's educational attainment. I complement this literature on wartime occupational upgrading and Black economic progress by exploring a particular channel of permanent labor shortages brought by skill- and race-specific wartime mortality rates. Black men faced lower draft rates due to lim-

fallen semi-skilled white soldier. The argument is based on learning by employers about the type of Black workers (Whatley, 1990), as well as the importance of employee-based referrals for new hires (Montgomery, 1991), and the importance of race-based networks in hiring (Royster, 2003).

[6]See Goldin (1991); Acemoglu, Autor and Lyle (2004); Goldin and Olivetti (2013); Jaworski (2014); Shatnawi and Fishback (2018), and Bailey and Collins (2006) for Black women.

ited space in the barracks of a segregated military.[7] Racist attitudes also saw Black men as unfit for combat and therefore placed them in relatively safer support tasks in the military (see Lee, 1965; Flynn, 1993). Hence more Black men were available during and after the war to fill jobs that remained open due to war-related deaths. Unlike most female workers, African American men maintained their wartime labor market gains (Wolfbein, 1947; Collins, 2000). Female workers of either race were displaced once draftees returned to their pre-war jobs (Acemoglu et al., 2004). Considering war casualties instead can potentially explain the persistent gains for Southern Black men.

2 Historic Background

2.1 Black Economic Progress at Mid-Century

The economic situation for African Americans improved little from the first half of the 20th century to World War II.[8] Since the early 1940s, however, the economic gaps between Black and white Americans began to shrink significantly (Smith and Welch, 1989). Both Maloney (1994) and Margo (1995) document a substantial decrease in Black-white wage differentials from 1940 to 1950. Other advances include increased Black home ownership rates after 1940 (Boustan and Margo, 2013) and higher economic mobility (Collins, 2000).[9]

The majority of Black workers were employed in low-skilled jobs before 1940, many of which were in agriculture,[10] which changed with the onset of the war and the resulting labor shortages. Figure 1 shows the substantial occupational upgrading of Black men from low- to semi-skilled employment from 1940 onward. The figure also highlights that this upgrading not only occurred in the industrial areas of the north, but that also Black Southerners entered these higher paying jobs at a much increased rate.[11] Overall one million African Americans entered semi-skilled employment during the war years (Wolfbein, 1947). The share of semi-

[7] Some service branches would not even accept Black men at all. One example are the Marines which African Americans only could join in 1943 after an executive order by President Roosevelt.

[8] Myrdal (1944) provides an account of the pre-war conditions of Black Americans in the U.S.: "They own little property; even their household goods are mostly inadequate and dilapidated. Their incomes are not only low but irregular. They thus live from day to day and have a scant security for the future." (p. 205).

[9] For a review of post-war Black economic progress see Smith and Welch (1989).

[10] The 1940 Census shows that 46% of Black workers were employed in agriculture.

[11] To make the meaning of *semi-skilled* more tangible in this context, Appendix Figure A1 displays the semi-skilled occupations with the largest inflows of Black workers from 1940-50, as well as the wage ratio of each job to the average wage paid to Black men in low-skilled jobs in 1950. Most of the occupations listed are operatives jobs in the durable-goods manufacturing sector (metal, stone/cement, wood, and textiles) and transport. Occupational definitions are those used by the Census Bureau in 1950. Panel (b) shows that all of these jobs lie substantially above the average earnings received by Black workers in low-skilled jobs.

skilled Black men rose by 8 percentage points between 1870 and 1940 but increased by 11.4 percentage points from 1940 to 1950 alone. Black men therefore made more occupational progress in this one decade than for most part of the first half of the 20th century.

With 16 million Americans serving in the military during World War II, labor shortages opened up job opportunities for groups that previously had received little consideration by employers. Wolfbein (1947) describes that for Black workers between 1940 and 1944 a "significant shift occurred from the farm to the factory as well as considerable upgrading of Negro workers, many of whom received their first opportunity to perform basic factory operations in a semiskilled or skilled capacity." (p. 663).[12] The literature thus far has focused on the labor market effects of the World War II draft on women.[13] For Black men the unfilled labor demand also reduced barriers to entry to industries and jobs that were previously inaccessible to them. This particularly concerned employment in semi-skilled jobs and in manufacturing where racist attitudes against Black workers were high (Weaver, 1945; Maloney, 1994). Weaver (1945) describes how labor shortages in the aircraft industry opened job opportunities for Black workers beyond low-skilled work.[14] Government intervention via the Fair Employment Practice Committee further helped Black workers during the war years but was generally ineffective in the South (Collins, 2001), where also the post-war G.I. Bill benefits had little effect on Black veterans in terms of college access (Turner and Bound, 2003).

While the link between the war and women's labor market outcomes has been studied in previous work, its role in the occupational upgrading of Black men in the postwar period is less well understood. A recent paper by Aizer et al. (2020) has provided important evidence on the impact that wartime labor shortages and expenditures on war contracts in the private sector had on the Black occupational upgrading. They show that these gains not only persisted at least until 1970 but that they also led to increased educational attainment among Black children. Earlier work, such as Collins (2000), documents large wage gains for Black men due to the occupational upgrading from low- to semi-skilled employment, especially in war related industries. Maloney (1994) and Margo (1995) note the importance of the occupational upgrading of

[12] It should be noted though that not all farm labor was low-skilled and many semi-skilled occupations were not in classical factory settings. I thank an anonymous referee for pointing this out to me.

[13] See Goldin (1991); Acemoglu et al. (2004); Goldin and Olivetti (2013); Jaworski (2014); Shatnawi and Fishback (2018), and Bailey and Collins (2006) for Black women.

[14] This development was not always welcome. Collins (2001) cites the example of the Philadelphia Transportation Company where white workers went on strike in 1944 because they disliked working with Black Americans in the same jobs. The strike was eventually broken by the Army's threat to re-evaluate the striking workers' draft deferments.

Black workers in explaining the overall wage gains at the time. Unlike most female workers who gained employment during the war, African American men managed to maintain their wartime labor market gains even after the draft ended. Hence the draft alone cannot explain this pattern as it ended with the war.

2.2 Racial Differences in Draft and War Mortality Rates

The U.S. mainly relied on the draft and the Selective Service Act to muster their military for World War II. Of the 16 million service personnel, over 10 million were inducted via the draft. Given that the draft was enacted during peacetime,[15] it had to be significantly more just and equal than the prior drafts to pass the substantial resistance by politicians and the public. Going to college or buying out was not possible, unlike in the later drafts for the Korean and Vietnam wars. Volunteering was forbidden in 1942 due to the British experience, where overenthusiastic volunteering had negatively affected the labor supply in war production (Chambers, 1987). The end of volunteering also meant that draftees did not choose their service branch as this choice was mainly reserved for voluntarily enlisted soldiers. The broad acceptance of the draft is reflected in the fact that out of the 40 million men assessed by local draft boards only 11,896 registered as conscientious objectors (Flynn, 1993). Of the 16 million soldiers, only some 50,000 deserted compared to over 200,000 desertions among the 2.2 million Union Army soldiers during the Civil War (Glass, 2013).

Black men were drafted less frequently than white men. This was due to racism in the military, draft boards, and essentially at all institutional levels, as well as the reluctance by Black men to serve a society that so heavily discriminated against them (see Qian and Tabellini, 2020). In the first six months of the draft, not one Black man was called for service. In some service branches, such as the Marines, Black men were not allowed at all. Even in late 1943, only 5.6% (ca. 375,000) of all soldiers were Black compared to their population share of a little over 10% (Lee, 1965). The Army General Classification Test (AGCT) often placed Black draftees in the bottom two out of five categories. This led to rejection or placement in service and support units, though such results mostly reflected the poor Jim Crow education received by Black Americans rather than actual aptitude (Lee, 1965). Another reason was segregation in the military. The Army and other service branches could not build barracks fast enough

[15] The U.S. had not entered the conflict yet when the draft was enacted on September 16, 1940. This first peacetime draft in U.S. history inducted men aged 21 to 45 via a lottery. This was abandoned after the attack on Pearl Harbor when the draft was substantially scaled up and the service age was changed to include men aged 18 to 44. See Flynn (1993) or Chambers (1987) for a historical treatment of the draft.

to maintain segregation, hence Black men were inducted at significantly lower rates (Flynn, 1993). Roosevelt's aim of employing Black soldiers according to their share in the population, which was around 10%, was only achieved in the later stages of the war when their share in the military rose just above 9%.

Also mortality rates were lower among Black soldiers. Due to the racist attitudes that saw Black men as unfit for combat, they typically served in support units that tended to be outside the line of fire (Lee, 1965). At the end of 1942, the share of enlisted Army soldiers serving in combat units was 40.2% for white soldiers but only 18.4% for Black soldiers.[16] 31.7% of Black soldiers but only 12.8% of white soldiers were employed in support and service units.[17] This uneven distribution of risk led to the racial gap in casualty rates that were higher among white soldiers. Appendix Figure A2 shows WWII mortality rates by race and occupational skill group. While high-skilled soldiers were safer in both racial groups, the mortality rate among Black soldiers was about half of that of white soldiers in any skill group.

3 White War Casualties and the Black Occupational Upgrade

Can war casualties among semi-skilled white soldiers explain the persistent occupational upgrading of Black workers from low- to semi-skilled employment during and after the war? To test this hypothesis, I compute county-specific casualty rates by race and occupational skill group by matching two data sources which are the WWII Enlistment Records and the WWII Honor List of Dead and Missing for the Army and Army Air Force.[18] The Army kept meticulous records of their drafted and enlisted soldiers during the war. Upon entry, an IBM punch card would store a soldier's name, unique Army serial number, age, education, race, marital status, residence, date and place of entry, and their pre-war occupation codified in three-digit groups using the Dictionary of Occupational Titles of 1939. The National Archives and Records Administration digitized these enlistment records which have been used in studying the effect of the Rosenwald schools on Black education (Aaronson and Mazumder, 2011), compulsory schooling and English-only instruction laws (Lleras-Muney and Shertzer, 2015), and cash transfers and poverty (Aizer, Eli, Ferrie and Lleras-Muney, 2016), among others.

The data do not contain soldiers in other service branches such as the Navy, Marines, or

[16] It should also be noted that Black and white units were segregated. Some Black fighting units, such as the Tuskegee Airmen or the 452nd Anti-Aircraft Artillery Battalion achieved remarkable successes. Yet the first Medal of Honor for service in WWII was awarded to a Black veteran only in 1997 by Bill Clinton.

[17] See Appendix Figure A3 for details.

[18] The Air Force only became an independent service branch after the war in 1947.

Coast Guard. However, the 8.3 million individuals in the Army data comprise the majority of the 10 million drafted men during World War II. Due to the high manpower demands by the armed forces there was almost no scope for drafted soldiers to choose a service branch (Flynn, 1993). Volunteering provided more choice regarding the branch of service but was forbidden in 1942 to give the military more control over who entered into service (Flynn, 1993). The end of volunteering came before the largest battles and casualties were sustained but after the majority of the drafting was completed (see Appendix Figure A4). It therefore would have been difficult to form a prior as to which service branch was the least dangerous in order to enlist strategically.

I further add information on each soldier's survival outcome. I digitized data on over 310,000 fallen soldiers from the WWII Honor List of Dead and Missing for the Army and Army Air Force. The remaining 100,000 deaths were suffered by the Navy, Marines, and the Coast Guard.[19] The casualty records include the name, state and county of residence, cause of death, and the Army serial number. The unique serial number is what identifies soldiers across the two data sources.[20] More details on merging the enlistment and casualty records is provided in the data appendix.[21]

Using the information on residence, race, pre-war occupation and casualty status, the casualty rate among semi-skilled white soldiers in county c can be computed as,

$$\text{Casualty rate}_c = \frac{\text{white semi-skilled casualties}_c \times 100}{\text{white semi-skilled soldiers}_c} \tag{1}$$

which is the percentage of those who went to war and who needed a replacement at their pre-war workplace, but who did not return from the war. The denominator was chosen to be the number of serving semi-skilled white soldiers rather than the total number of semi-skilled white workers in a county. Using the latter is potentially problematic because workers in war related industries had a higher chance of receiving deferments. Without exact knowledge about the number of deferred men it is not possible to compute an accurate measure of wartime demand for alternative labor such as women of either race or Black men. In the regressions below I

[19] I provide evidence that there is almost zero correlation between casualties in the Navy, Marine Corps, and Coast Guard with those in the Army and Air Force. See Appendix Table A1.

[20] Appendix Figure A5 shows examples of the enlistment and casualty records.

[21] Summary statistics for the matched data for different sample splits comparing Black and white soldiers, enlisted and drafted, and northern with Southern soldiers are reported in Appendix Table A2. Black soldiers had a substantially lower probability of dying since there were few segregated Black fighting units and hence African American soldiers were mainly employed in relatively safer support units (Lee, 1965).

control for the draft rate. The draft rate captures how many white workers were gone during the war whereas the casualty rate among semi-skilled white soldiers captures the variation among those in this particular group that did not return.[22]

The spatial distribution of this casualty rate measure for counties in Southern states is plotted in Figure 2. The casualty rate measure can be constructed for the whole of the U.S. but the outcome variable of interest, i.e. the share of semi-skilled Black workers, can only be computed at the county-level for the mapped Southern states. The county-level Census data does not provide occupational counts by race for counties outside the South. Appendix Table A3 reports Moran's I statistic and Appendix Table A4 reports the Getis-Ord statistic for local spatial autocorrelation at different distance thresholds. While the casualty rate measure displays significant spatial correlation in both tests, once state fixed effects are included this vanishes entirely. I provide a further test for selectivity in the casualty rate measure by regressing it on pre-war county characteristics measured in 1940.[23]

Lastly, I combine the casualty and draft rates with county level data from the U.S. decennial Census from 1920 to 1960 which is available in ten years intervals. The main outcome of interest is the percentage share of semi-skilled Black male workers in county c and decade t. Following the U.S. Census Bureau's occupational classification of 1950, semi-skilled jobs are those classified in the craftsmen and operatives categories.[24] Aggregate data on the number of employed workers by skill group at the county level is available for the U.S. Census files between 1920 and 1960. These data are not available in a digitized format and were collected specifically for this project. Appendix Figure A8 shows an example of the data. Only Southern states tabulated occupational counts by race.[25] For the 16 states plus D.C. there is a total of 1,388 counties.

[22]For robustness checks, I later also use the casualty measure with the denominator being all semi-skilled white workers in 1940 (see appendix A).

[23]Appendix Figure A6 plots the resulting coefficients. The WWII casualty rate does not show strong evidence of selection along the lines of wealth, education, income, government spending, agricultural and manufacturing industries, urbanization, or unemployment. This is consistent with evidence provided by Kriner and Shen (2010) who show that there was no significant difference in casualty rates across socioeconomic groups during WWII. Only from the Korean War onwards such a gap emerged. Appendix Figure A7 shows that volunteering was not different comparing the South to the non-South (panel a), and within the South there were no differences in volunteering between above- and below-median casualty rate counties (panel b).

[24]In other classification systems, craftsmen count as skilled workers but for the purpose of this paper I consider them together. The other common feature between the two employment groups is that neither require college training, although it takes significantly longer to train craftsmen than operatives.

[25]These are Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, South Carolina, Oklahoma, Tennessee, Texas, Virginia, and West Virginia, and Washington D.C. Note that even though I refer to mentioned states as "South", this deviates from the typical definition of the South as the former Confederacy, unless stated otherwise.

3.1 Evidence from Data on Southern Counties, 1920-70

The evolution of the share of semi-skilled Black workers employed from 1920 to 1960 across casualty rate quartiles is plotted in Figure 3, panel a.[26] The figure shows how this variable followed the same trend across counties in all casualty rate quartiles before the war. After the war, the share of semi-skilled Black workers rose but it rose differentially faster in counties with the highest WWII casualty rates among semi-skilled white soldiers. The graph also shows pre-war level differences between counties across casualty quartiles. Panel b of the same figure therefore provides the parallel trends plot after partialling out the share of manufacturing employment, Black population, and the urbanization rate. These factors influenced the share of Black workers in semi-skilled positions as well as the share of semi-skilled white men who died in the war, e.g. counties with very low levels of semi-skilled employment before the war would not have sent any semi-skilled white men but also would not have the opportunity to upgrade Black workers, thus being in the bottom quartile while displaying large differences to the top quartile of the casualty rate. Once these factors are taken out, the initial level difference shrinks from 13 percentage points to less than one percentage point.[27]

The pre-war parallel trends in the share of semi-skilled Black workers makes this a suitable setting for a difference-in-differences framework which removes such differences and any other unobserved time-invariant county characteristics. The difference-in-differences regression to be estimated is,

$$\%\text{ semi-skilled Black men}_{ct} = \alpha_c + \lambda_t + \beta\text{ Casualty rate}_c \times \text{Post-war}_t + X'_{ct}\phi + \eta_{ct} \qquad (2)$$

where % semi-skilled Black men$_{ct}$ is the outcome in county c in decade t and the WWII casualty rate among semi-skilled white soldiers interacted with a post-war indicator is the main treatment of interest which allows for variable treatment intensities. Under the usual parallel trends assumption and in the absence of time-varying confounding factors, the coefficient β captures the causal effect of a one percentage point increase in the WWII casualty rate among semi-skilled white soldiers on the share of semi-skilled Black workers after the war.[28]

[26]The raw correlation between casualty rates and the share of semi-skilled Black men in the cross section of counties and across time is shown in Appendix Figure A9. Conditional scatter plots that partial out county characteristics in 1940 are shown in Appendix Figure A10.

[27]Notice that the main requirement for the empirical strategy here is that the casualty rate is as good as randomly assigned *conditional* on time and county fixed effects for which I provide visual evidence in Figure 4 as well as under different model specifications in Appendix Figure A11.

[28]With a continuous treatment, β will be a weighted average of the pairwise comparisons of the difference-in-

Time-invariant observable and unobservable determinants of the share of semi-skilled Black workers across counties, including any pre-war level differences, are absorbed by county fixed effects α_c. Time-varying shocks common to all counties are controlled for by time fixed effects λ_t. Alternative specifications include state-by-year fixed effects ρ_{st} or county-specific linear time trends $\alpha_c t$ to probe for robustness of the results with respect to treatment of the time dimension. This allows for partialling out state- or county-specific secular changes in the outcome that would have occurred in the absence of the casualty shock such as state-specific legislation, for instance.

The vector X_{ct} contains controls that account for potential changes in observables that might determine the share of semi-skilled Black workers but also correlate with the casualty rate among semi-skilled white soldiers. The draft rate inversely accounts for the remaining workforce during the war in the county and controls directly for the share of the male population under threat of being killed in the war. To account for spillover effects, I include the average casualty rate in the adjacent counties of a given county c. Also included are the log of WWII related spending per capita (Fishback and Cullen, 2013; Aizer et al., 2020),[29] the share of rural population and the share of Black men from the county-level Census, the Republican vote share from data by Clubb, Flanigan and Zingale (2006), interactions of the post-war indicator with the number of lynchings between 1900 and 1930 per 1,000 Black population, the number of slaves in 1860, the number of Rosenwald schools per 1,000 Black population to capture determinants of Black education (Aaronson and Mazumder, 2011), and the share of acres flooded by the Mississippi in 1928 (Hornbeck and Naidu, 2014).

I also control for manufacturing and agricultural variables such as the number of manufacturing establishments per capita, the average employment per firm, log value added per worker, and the share of manufacturing employment. Agricultural controls include the share of land in agriculture, the share of acres in cotton, the share of cash tenants, and the average value of machinery per farm. The latter seeks to control for technological changes in the agricultural sector. In particular, the use and quality of tractors expanded at the time, especially in the South and released labor from agriculture (see Olmstead and Rhode, 2001).

differences estimates for each value of the treatment, which deviates from the standard approach of estimating such a model. Appendix Table A5 estimates the difference-in-differences model in a more traditional sense by using an indicator for a county being in the top quartile of the WWII casualty rate among semi-skilled white soldiers interacted with a post-war dummy as main treatment variable.

[29] Data for WWII expenditure comes from the County and City Data Book 1947 published by the United States Department of Commerce (2012)

Finally, to account for the major economic changes brought by the Great Depression in the decade just prior to the war, I include measures of New Deal spending per capita from Fishback, Horrace and Kantor (2006). This includes government loans, money for public works, funds from the Agricultural Adjustment Act (AAA), and by the Federal Housing Administration (FHA), as well as the unemployment rate in 1937 interacted with decade fixed effects. Note that some of these variables are potentially endogenous to the casualty treatment and thus I consider specifications where I use the pre-war values of these variables in 1940 only and I also show that the controls do not change the results in any way. All monetary values are deflated to 2010 U.S. dollars using the CPI provided by the Bureau of Labor Statistics.[30] Summary statistics are reported in Table 1. All remaining variation in the outcome, which is not captured by the listed control variables, is absorbed in the error term η_{ct}. Standard errors are clustered at the county level to account for heteroscedasticity and autocorrelation.

3.1.1 Difference-in-Differences Results

The main results from the estimation of equation (2) are reported in Table 2 under different model specifications. The baseline result from only including time and county fixed effects in column (1) estimates that a one percentage point increase in the WWII casualty rate among semi-skilled white soldiers increased the share of semi-skilled Black workers by 0.54 percentage points. Column (2) includes all control variables mentioned in the previous section.[31] To take into account the possibility that some of these time-varying measures could be outcomes of the semi-skilled casualty rate themselves, column (3) includes all controls measured at their 1940 values and interacts them with the post-war indicator. While this reduces the risk of including potentially bad controls, it also takes out variation of these controls that might be useful but results remain stable nevertheless. Columns (4) and (5) relax the parallel trends assumption by including state-by-year fixed effects and county-specific linear time trends, respectively. Finally, column (6) estimates equation (2) using the doubly-robust selection algorithm by Belloni, Chernozhukov and Hansen (2014). Their machine learning covariate selection algorithm tests for the stability of treatment effects and potentially improves inference on such parameters.[32]

[30] An overview of all data sources used to compile the final estimation sample is given in the data appendix.

[31] Appendix Table A6 reports results from the estimation of equation (2) using also median household income per capita as control for the period from 1940 to 1960 to see whether occupational upgrading of Black workers occurred in significantly richer or poorer areas that may also relate with the local casualty rate. Since general wage income is only available in the Census since 1940, it was omitted from the main specification. Results with the income control remain unchanged.

[32] Suppose that a large set of observed controls includes the most relevant covariates to explain the relation of interest but that these variables are unknown to the econometrician. First, the outcome is regressed on the

The casualty rate coefficient is statistically significant at the 1% level in all cases and with a coefficient of 0.54 percentage points in the preferred specification in column (2).

The magnitude of the effect implies that each fallen white semi-skilled soldier was replaced by two to three Black workers.[33] Notice that this includes multiplier effects such as those documented by Whatley (1990) who shows that firms employed another Black worker for each Black worker they already employed during World War I. If a fallen semi-skilled white WWII soldier was replaced by one Black worker, we would expect at least one additional Black worker to join the firm. Since WWII was considerably larger and longer lasting, however, this multiplier effect might have been even stronger. An additional two Black workers who joined the Black replacement worker for a fallen soldier therefore is not an implausible magnitude. There are at least two potential contributing factors: first, the World War II period was characterized by steeper labor shortages than World War I, and second, it also significantly changed employers' views of Black workers, leading to updated beliefs among employers and hiring practices as argued by Aizer et al. (2020). Whatley (1990) discusses the changing racial views of employers once they hire minorities for the first time but also internal referrals and networks have been shown to lead to additional employment. Employers value referrals from their existing workforce due to their incentives for providing truthful information (Montgomery, 1991), hence Black workers who found jobs during the war might have provided such referrals. Additionally, Royster (2003) highlights the importance of race-based networks which might be another channel for this larger than proportional effect.

Given the stability of the results in Table 2 it seems unlikely that the estimated casualty rate effect is driven by selection on the observables or differential secular trends in the outcome. To probe for the sensitivity of the previous results with respect to the unobservable components, Table 2 reports the coefficient sensitivity test by Oster (2019) for all specifications.[34] A value of

controls, their squares, and all cross-term interactions, after which the most significant predictors are selected either via LASSO or a simple t-test from a multiple regression if the sample size permits. Here a t-test sufficed. The same is repeated for the treatment, i.e. the casualty rate in this case. In a final step, equation (2) is re-estimated using the union of controls selected in either of the previous two steps. The idea is that the regression learns the most important predictors of outcome and treatment which would be problematic omitted variables.

[33] Appendix Table A7 provides the corresponding regression result using the levels of the outcome in a first differenced regression rather than working with shares. When including interactions with variables capturing larger labor shortages, urban, and manufacturing centers, the coefficient is reduced to 1.5 workers who replace on fallen soldier.

[34] Her test considers a standard linear regression model of the form $Y = \beta X + W_1 + W_2 + \epsilon$, where $W_1 = \Psi w^o$ is a vector of observable controls and W_2 is an index of unobservables. The treatment variable X here is the casualty rate. She then defines the selection relationship as $\delta \frac{Cov(W_1,X)}{Var(W_1)} = \frac{Cov(W_2,X)}{Var(W_2)}$ and solves for δ (the degree to which selection on unobservables is less than or larger than selection on observables) which would be required to produce $\beta = 0$. This uses the coefficient and R^2 movement from the controlled and uncontrolled regressions

$\delta = 1$ implies that selection on the unobservables would have to be equally important compared to selection on the observables in order to explain away the estimated treatment effect. The typical threshold for results to be considered reasonably robust is $\delta > 1$, which is the case in almost all specifications with the exception of the linear county-specific time trends model.

To provide an estimate for the dynamics of this casualty rate effect and a test for the parallel trends assumption, I next interact the casualty rate treatment in equation (2) with Census year indicators instead of the post-war indicator only. The omitted year is 1940, which makes this last pre-treatment period the reference year for all other coefficients. Figure 4 plots the resulting coefficients.[35] The interaction of the casualty rate with the 1920 and 1930 indicators are essentially placebo treatments since the war should not have had an effect on Black workers' occupational status before there were any casualties. These two coefficients reassuringly are almost zero and far from significant. The coefficients for 1950 and 1960 estimate stable coefficients close to 0.5, showing a persistent effect of the WWII casualty rate among semi-skilled white soldiers on the post-war share of semi-skilled Black workers. Miller (2017) assesses the affirmative action policies under President Johnson in 1965 and also finds that the effect of affirmative action policies remained even after their removal.[36]

Table 3 explores effect heterogeneity by including interactions of the casualty rate with indicators for above median values for the draft rate, months of labor shortages,[37] outmigration during the war,[38] log military spending per capita, and the average machine value per farm. Counties with above median draft rates, months of labor shortages, or outmigration saw the largest increase of the casualty rate effect which raised the effect size to between .629 and .745. This is consistent with Aizer et al. (2020) who find that labor shortages were an important driver of the occupational upgrading among Black workers and it also suggests that much of the upgrading already occurred during the war. Also higher war expenditures and machine values per farm were associated with larger effect sizes as more mechanized agricultural sectors could more easily release labor to the manufacturing sector.

results in a bounding argument.

[35]The regression underlying the plot includes controls and state-by-year fixed effects. Appendix Figure A11 provides the same plot for all specifications presented in Table 2 with and without controls, with the covariates fixed at their 1940 values, and the doubly-robust selection algorithm. The linear county-specific trends specification cannot be estimated in this setting.

[36]While this is certainly not to compare affirmative action policies to World War II, the idea is that an exogenous break in racial barriers to entry into certain occupations can have persistent effects.

[37]Months of labor shortages her a measured as those in which the number of drafted soldiers exceeds 10% of the pre-war labor force.

[38]Outmigration between 1940 and 1950 is measured as the population difference between the two Census years minus deaths plus births which were taken from Bailey, Clay, Fishback, Kantor, Severnini and Wentz (2018).

A key falsification test is to consider the effect of casualty rates in other skill groups for both Black and white soldiers on the occupational upgrading of Black workers. If the claim is correct that it was the death of semi-skilled white soldiers that led to the occupational upgrade of African American men, then we should not see any effect coming from casualty rates in other skill-race groups. Results are reported in Table 4 which includes casualty rates by race and skill group in the regression while controlling for the draft rates in each group. The estimated coefficients for the semi-skilled white casualty rate are similar to those in the baseline specification. There is no detectable effect for the casualty rates in any other group. The exception is a weakly significant negative effect coming from low-skilled Black soldiers for whom a one percentage point increase in their casualty rate decreases the share of semi-skilled Black workers by 0.04 to 0.14 percentage points. This result is intuitive given that these are the workers who, had they survived, would have replaced the deceased semi-skilled white soldiers after the war.

Further robustness checks are reported in the Appendix. Section A1.1 shows robustness of the event study regressions to different model specifications. Section A1.2 provides evidence for robustness towards selections on observables with potentially mismeasured controls (Pei, Pischke and Schwandt, 2019). Section A1.3 rules out that results are driven by selective migration of Black workers. Section A1.4 shows robustness to the potential selection of soldiers. Section A1.5 alters the treatment denominator in equation (1) from white semi-skilled soldiers$_c$ to white semi-skilled workers$_c$, and rules out denominator bias by fixing the outcome denominator at 1940 levels. Section A1.6 drops states one-by-one to show that no single state is driving the results. Section A1.7 corrects standard errors for spatial autocorrelation. Section A1.8 weights by population size.

4 Racial Barriers and Selection into Semi-Skilled Work

In this section I first test whether the Black occupational upgrading was likely a result of a reduction in racial barriers to entry into semi-skilled occupations due to the war-induced labor shortages, or if this was simply due to an overall improvement in economic conditions that benefited both Black and white workers. I use repeated cross-sectional data on individual workers from the Census in a triple differences setting to show that Black workers saw an increased probability of semi-skilled employment in commuting zones with higher casualty rates among semi-skilled white soldiers after the war. This same relationship does not exist

for white workers. The casualty shock appears to have mainly acted to reduce racial entry barriers. I then build a simple theoretical model to better understand the selection of Black and white workers into semi-skilled employment in response to such reduced racial barriers to entry. Using the same triple differences setting, results show that new Black semi-skilled workers are positively selected while new white workers in such occupations are negatively selected.

There are additional advantages of using the individual level data. It not only allows for a more direct comparison of Black and white workers to differentiate between the removal of racial barrier and overall economic conditions, but it also lets me study the differential effects of white WWII casualties on Black and white occupational outcomes across industries as well as across the South and non-South.[39] I use the individual level data of the 1920 to 1960 U.S. Census files by Ruggles, Flood, Goeken, Grover, Meyer, Pacas and Sobek (2018). This includes the 1% micro Census files from 1920 to 1950, and the 5% file of 1960. County of residence is not available in the later Census years, however, commuting zones can be consistently constructed for the sample period. Overall there are 722 commuting zones which are clusters of counties that share a common labor market.[40] Appendix Figure A12 plots the WWII casualty rate among semi-skilled white soldiers at the commuting zone level.

The estimation sample includes the non-institutionalized working age (16-65) Black and white male population who were participating in the labor force at the enumeration date not enrolled in school or classified as unpaid family workers. The micro level data provide the advantage of using white workers as an additional control group. If casualties resulted in a labor supply shock only, then one would expect occupational upgrading to occur for both Black and white workers. However, if semi-skilled professions had higher barriers to entry for Black workers that were removed due to the labor shortages induced by the casualties, then only this group should see an effect on their probability to be employed in such jobs.

I compare the probability of semi-skilled employment between Black and white workers,

[39]The previous county level data for Black and white employment are only available for the South.

[40]The crosswalk used to construct the commuting zones for 1950 are available on David Dorn's website (http://www.ddorn.net/data.htm), and the crosswalk files for the other years were kindly shared by Felix König. For years when county of residence is not available in the Census, the crosswalks are merged to the state economic area (SEA) identifier in 1950 and the Mini-PUMA identifiers in 1960. To construct the commuting zone casualty rate (and associated commuting zone level controls), the commuting zone and county identifiers were combined in a many-to-many merge after which the data was collapse to a weighted average at the commuting zone level. The weight was determined by the share of land of a county belonging to a given commuting zone. Following David Dorn's instructions, the person-specific weight variable from the Census used in regressions was also multiplied with this area weight.

before and after the war, and across commuting zones with differing casualty rates using repeated cross-sectional data from the Census in a triple difference (DDD) regression,

$$\begin{aligned}\Pr\left(\text{semi-skilled} = 1\right)_{izt} &= \beta_1 \left(\text{casualty rate}_z \times \text{post-WWII}_t\right) \\ &+ \beta_2 \left(\text{casualty rate}_z \times \text{Black}_{izt} \times \text{post-WWII}_t\right) \\ &+ \alpha_z + \lambda_t + \delta \text{Black}_{izt} + X'_{izt}\gamma + \epsilon_{izt} \end{aligned} \tag{3}$$

where i, z, and t index individuals, commuting zones, and Census years, respectively. The outcome is an indicator for whether an individual is a semi-skilled worker. The coefficients of interest are β_1 for white men and the triple interaction coefficient β_2 for Black men. The overall effect on the employment probability for African American men is given by the sum of the two coefficients. Individual controls include age, marital status, year of birth, a self-employment indicator, farm status, and industry fixed effects, and α_z and λ_t are commuting zone and time fixed effects. Standard errors are clustered at the commuting zone level.

The triple differences regression seeks to eliminate potentially confounding trends in the employment probability of Black workers in semi-skilled jobs across commuting zones that are unrelated to the war casualties. It also accounts for changes in the employment probability of all workers in high-casualty commuting zones which might have happened due to other shocks that occurred at the same time such as a general improvement in economic conditions that benefited all workers. If only β_1 increases, this would indicate that WWII deaths among semi-skilled white soldiers simply implied a labor shortage in the sector that led to better opportunities for both low-skilled Black and white workers. If instead β_1 is zero while β_2 increases after the war, this would provide evidence for the casualty shock to have reduced racial barriers to entry into semi-skilled jobs for Black men. This would be consistent with the historic account provided by Weaver (1945) for the aircraft industry.

To visualize the relationship, I interact the casualty rate$_z$ and casualty rate$_z$ $\times$ Black$_{izt}$ variables with Census year fixed effects in equation (3), leaving out 1940 as baseline. The resulting coefficients for Black and white men are plotted in Figure 5. There is no significant casualty rate effect before the war for either group and remains insignificant for white workers also in the post-war period. For Black workers there is a positive post-war effect with a 1.5 to 2.5 percentage points rise in the semi-skilled employment probability for every one percentage point increase in the commuting zone WWII casualty rate among semi-skilled white soldiers. Com-

pared to the earlier coefficient plot using county-level data from the South which implied a level shift, here we also see a change in trajectory. This is explained by increasing job opportunities outside the South after the war. Once the differential treatment effect is modeled as trend term for the non-South, which is shown in panel b of the same figure, the plot replicates the pattern of the county-level result in Figure 4.

Table 5 reports results from estimating equation (3) for different model specifications. The triple difference coefficient for Black workers is positive and significant in all specifications and ranges between 2.2 to 2.7 percentage points for the full sample and between 1.1 and 1.7 percentage points for workers in the South. There is little to no effect on white workers which is consistent with the results in Collins (2000) and Aizer et al. (2020), as well as with the historic account (Wolfbein, 1947). The average commuting zone casualty rate of 3.38%, together with the associated multiplier effects discussed earlier, explains 45.4% of the increase in semi-skilled employment among Black men in the post-war sample period in the full sample.[41] For the South, the average casualty rate and associated multiplier effects can explain 22.6% of the occupational upgrading of Black workers in the post-war period. The smaller effect in the South is expected due to the smaller size of the manufacturing sector, which at the time was the main employer of semi-skilled workers, and the lesser mobility of African Americans as compared with the rest of the country.

The results show that the employment gains for Black workers not only occurred in the north or west but that also Black Southerners gained significantly in terms of the occupational upgrading. Another advantage of the micro data is that I can further deal with potential migration responses. I therefore interact an indicator for whether an individual lives outside their state of birth with time fixed effects and the Black indicator in column (4). The same interactions are applied to the education variable. The results are unchanged by this inclusion and thus do not appear to be driven by differential migration or education between Black and white workers. It should be noted that migration and education are potential outcomes of the treatment, hence results from this specification are to be taken with caution. Yet it potentially sheds light on whether the occupational upgrading effect can be explained away by differential migration or educational attainment across Black and white workers over time.

Next, I analyze whether the occupational upgrading of Black workers is concentrated in

[41] The share of semi-skilled Black men in 1940 in the sample was 20.47% which rose to 39.8% in 1960. Given the coefficient estimate, the overall increase that can be explained by the average commuting zone casualty rate and associated multiplier effects is $\frac{.026 \times 3.38}{.3983 - .2047} = .454$. For the South, this computation becomes $\frac{.3.38 \times 0.012}{.3676 - .1884} = .226$.

particular sectors. Table 6 repeats the analysis for the manufacturing sector as a whole, and for the durable and non-durable manufacturing sub-sectors, as well as for telecommunications, retail, and mining and construction. Unlike the manufacturing sectors, the jobs in the telecommunications and retail often involved direct customer contact and therefore employers sought to avoid employment of Black workers in such positions (Anderson, 1982). Given that these sectors remained segregated throughout and after the war, they should not show any occupational gains made by Black men. The results provide evidence that Black occupational upgrading was particularly pronounced in all manufacturing sectors with a 5 to 8 percentage points increase in the probability of semi-skilled employment for Black workers for a one percentage points increase in the WWII casualty rate among semi-skilled white soldiers.[42] If detectable, effects are negative in the other sectors which is not surprising for telecommunications and retail in particular. The patterns are similar across the South and the full sample.

4.1 Selection into Manufacturing and Economic Outcomes

Having established that, unlike low-skilled Black workers, low-skilled white men did not see an increase in their opportunities to obtain semi-skilled employment in response to the WWII casualty shock, one remaining question is which workers entered these semi-skilled jobs and what were their characteristics. First, to add some structure to this question consider the following model based on Borjas (1987). Assume that there are two sectors, agriculture $\mathcal{A}$ and manufacturing $\mathcal{M}$, where $\mathcal{A}$ employs only low-skilled workers and $\mathcal{M}$ employs only semi-skilled workers, for simplicity. Both sectors have roughly similar labor demands. There are two groups of workers g, Black b and white w. Further assume that workers select into sectors based on maximizing income y, where for both groups income in each sector is determined by,

$$y_{\mathcal{A}} = \mu_{\mathcal{A}} + \epsilon_{\mathcal{A}} \quad \text{and} \quad y_{\mathcal{M}} = \mu_{\mathcal{M}} + \epsilon_{\mathcal{M}} \tag{4}$$

where μ_s is the average income received from working in sector s and ϵ_s is a worker's innate productivity in sector s,[43] which are normally distributed with mean zero, variances $\sigma^2_{\mathcal{A}}, \sigma^2_{\mathcal{M}}$, and covariance $\sigma_{\mathcal{AM}}$. The errors can be interpreted as the de-meaned skill premium to a worker's income. Workers have full knowledge of their own μ, ϵ, and the cost c of switch-

[42] It is worth to note that the most important manufacturing sector for semi-skilled Black workers is sawmills instead of the typical factory jobs described by Wolfbein (1947). I thank an anonymous referee for highlighting this.

[43] I omit individual subscripts for clarity.

ing sector is common knowledge. The pre-war probability of switching from $\mathcal{A}$ to $\mathcal{M}$ for a white worker can be expressed as,

$$\begin{aligned}
\Pr(\text{switch})_w &= \Pr\left(y_{\mathcal{M}} - c > y_{\mathcal{A}}\right) \\
&= \Pr\left(\epsilon_{\mathcal{M}} - \epsilon_{\mathcal{A}} > \mu_{\mathcal{A}} - \mu_{\mathcal{M}} + c\right) \\
&= \Pr\left(\frac{\nu}{\sigma_\nu} > \frac{\mu_{\mathcal{A}} - \mu_{\mathcal{M}} + c}{\sigma_\nu}\right) \\
&= \Pr\left(\frac{\nu}{\sigma_\nu} > z_w\right) \\
&= 1 - \Phi\left(z_w\right) \qquad (5)
\end{aligned}$$

where $\nu = \epsilon_{\mathcal{M}} - \epsilon_{\mathcal{A}}$, $z_w = \frac{\mu_{\mathcal{A}} - \mu_{\mathcal{A}} + c}{\sigma_\nu}$, and $\Phi(\cdot)$ is the CDF of the standard normal distribution. A worker switches from agriculture to manufacturing if the expected individual wage in manufacturing is higher than that in agriculture after paying the switching cost. Black workers follow a similar switching process in the pre-war period with the only difference being that Black workers face increased switching costs due to discrimination d, such that

$$\Pr(\text{switch})_b = \Pr\left(y_{\mathcal{M}} - c - d > y_{\mathcal{A}}\right) = 1 - \Phi\left(z_b\right) \qquad (6)$$

and $z_b = \frac{\mu_{\mathcal{A}} - \mu_{\mathcal{A}} + c + d}{\sigma_\nu}$. Since $z_b > z_w$, we would expect relatively fewer Black than white workers in $\mathcal{M}$ before the war. To provide a visual form of comparative statics, consider the distribution of ν among Black and white workers. This distribution in the difference of sector-specific abilities can be plotted as shown in Figure 6 since the difference of two normal random variables is itself normal with $\nu \sim N(0, \sigma^2_{\mathcal{A}} + \sigma^2_{\mathcal{M}} - 2\sigma_{\mathcal{AM}})$.[44] In the absence of the race-specific switching cost d faced by Black workers, we would expect $z_b = z_w$, if the skill distributions in $\mathcal{A}$ and $\mathcal{M}$ are the same for both groups. In this case, the mass between 0 and z_b^{pre} in panel (b) of Figure 6 represents all Black workers who could have productively filled a manufacturing job but did not switch over from manufacturing due to the additional race-related cost d.

Next, assume that the World War II casualty related labor shortages reduced both c and d. Using the comparative statics provided in Figure 6, there are two results. First, a reduction in both types of switching costs should lead to a relatively larger inflow of Black than white work-

[44] The threshold points, z_g^{pre} and z_g^{post}, are motivated by the data. In 1940, when considering *only* the low- and semi-skilled population of Southern male workers, 50% of white men were in a semi-skilled as opposed to a low-skilled job. Likewise, only 14% of Black men were in semi-skilled employment at that time. These employment shares rose to 30% and 60% in 1950 for Black and white workers, respectively.

ers into $\mathcal{M}$. Second, new entries into $\mathcal{M}$ will be positively selected in terms of their relative manufacturing ability if they are Black, and negatively selected if they are white workers. The first result is due to the fact that d is only relevant for Black individuals. According to the data, the share of semi-skilled employment among Black men rose by 16 percentage points between 1940 and 1950 while that of white men rose by 10 percentage points.

The second result, given the data, is because white workers in the post-war period have $\nu < 0$, i.e. they would have been relatively more productive in agriculture than manufacturing but decided to move to manufacturing anyway due to higher income which became attractive after the reduction in switching costs. Likewise, the mass of Black workers entering $\mathcal{M}$ after the war, which is the mass between z_b^{pre} and z_b^{post} in panel (b) of the same figure, have $\nu > 0$. This also implies that even before the war Black workers had to be particularly more able in order to gain access to semi-skilled jobs relative to white workers.

4.1.1 Testing the Model Predictions Empirically

The simple theoretical model predicted that more Black than white workers should upgrade from low- to semi-skilled work in response to the WWII casualty shock among semi-skilled white soldiers, and that Black workers should be positively selected whereas white workers should be negatively selected. The first prediction was already confirmed by estimating the triple differences regression in equation (3). To test the second prediction of the model, I compare the post-war economic outcomes of Black and white men in semi-skilled jobs in the same repeated cross-section triple differences setting as before. In particular, I re-estimate equation (3),

$$\begin{aligned} y_{izt} = {} & \beta_1 \left(\text{casualty rate}_z \times \text{post-WWII}_t\right) \\ & + \beta_2 \left(\text{casualty rate}_z \times \text{Black}_{izt} \times \text{post-WWII}_t\right) \\ & + \alpha_z + \lambda_t + \delta \text{Black}_{izt} + X'_{izt}\gamma + \epsilon_{izt} \end{aligned} \tag{7}$$

where y_{izt} is the economic outcome for individual i in commuting zone z in year t. Economic outcomes include the natural log of a worker's annual wage, their years of education, an binary outcomes for internal migrant status, which equals one if they do not currently live in their state of birth, for whether a person lives in a metropolitan area, and for home ownership, as

well as the natural logarithm of house values.[45] All regressions include time and commuting zone fixed effects, the same controls as in equation (3), and standard errors are again clustered at the commuting zone level. Note that the resulting coefficients will capture a mix of selection of workers into better paying jobs as well as the economic gain resulting from the occupational upgrade.[46]

Results for the full sample and for the Southern sub-sample of semi-skilled Black and white workers are reported in panels A and B in Table 7, respectively. As predicted by the model, semi-skilled Black men who live in commuting zones with higher WWII casualty rates among semi-skilled white soldiers after the war have higher wages, levels of education, or house values. For every percentage point increase in the casualty rate measure, Black semi-skilled workers' wages are associated with an approximate 3.1% increase whereas average wages of semi-skilled white workers were not significantly affected. For an average casualty rate 3.38% at the commuting zone level, the difference in Black-white wages for semi-skilled workers declined by 10.5 percentage points in the full sample and 10.1 percentage points in the South. This implies a post-war decline in the Black-white semi-skilled wage differential of 25%.[47] Since these comparisons here are within skill group, this limits the possibility of the results being confounded by the great compression, the general narrowing of the wage distribution across and within skill and education groups which occurred at mid-century (see Goldin and Margo, 1992).

The difference in years of education between Black and white semi-skilled workers was 3.5 years in 1940 and declined to 1 year in 1950. At the average casualty rate, the post-war Black-white difference in years of education declined by 0.7 years, which accounts for one quarter of this decline. This potentially captures two reinforcing effect. The first is the positive and negative selection of new Black and white manufacturing workers, respectively, as predicted by the model. The second is that Black men might have increased their investment in educational attainment due to the increased labor market opportunities and the therefore raised returns to education (see Aizer et al., 2020).

[45] In Appendix B1 I show the same analysis for women.

[46] In Appendix B2 I replace the casualty rate in equation (7) with an indicator for semi-skilled employment as treatment to provide a more direct estimate of the benefits associated with the upgrading from low- to semi-skilled jobs. However, in the absence of a good measure of ability that can partial out the selection effect it is not possible to fully disentangle the two and for the purpose of the exercise here the casualty rate is arguably more exogenous and provides a direct reduced form link to the war.

[47] The conditional wage difference between Black and white workers was taken from the triple differences regression as estimated by the included race dummy. In the South, Black men earned 0.439 log points less than white men and 0.369 log points in the full sample.

The probability of migrating to outside their state of birth declined significantly for semi-skilled Black workers, which is intuitive given the job opportunities in their home state. The probability of urban status is positive for Black workers, while it is larger and negative for white workers which is consistent with the white flight finding by Boustan (2010). While the probability of home ownership did not increase for white individuals and actually decreased for Southern Black workers in response to the casualty shock,[48] the value of dwellings occupied by Black workers increased substantially. For white workers, this pattern is the opposite but not significant. Overall, the evidence points towards Black workers being positively selected into post-war semi-skilled employment, a pattern which is the opposite for white workers, and that this development is significantly linked to the war casualties that brought down racial barriers to entry as predicted by the theoretical model. In comparison, Black women also gained more than white women, however, their labor market successes were temporary and essentially disappear again by 1960 which is consistent with the results by Acemoglu et al. (2004) (see Appendix B1).

5 Conclusion

The 1940s marked a turning point in the economic progress of Black Americans. In this paper I provide a new explanation for the large economic gains made by Black Southerners by relating this development to World War II. Prior to the war, more than 80% of Black men worked in low-skilled jobs. These paid significantly less than semi-skilled jobs from which Black workers were barred due to racial barriers to entry. During the war, Black men saw both lower draft and mortality rates. This meant that more Black men were available to work both during and after the war. This raises the question of how much of the Black occupational upgrading at the time can be explained by the higher casualty rates among semi-skilled white soldiers.

Having digitized and compiled a novel data set on WWII casualty rates by race, occupation, and location, I provide evidence in a county-level difference-in-differences setting from 1920-70 that counties with higher mortality rates among semi-skilled white soldiers saw larger increases in the share of semi-skilled Black workers after the war. The importance of this occupational upgrading and its significance for Black Southerners is hard to overstate. The average semi-skilled job paid twice the salary of the average low-skilled job, implying a substantial increase in disposable income for Black workers who secured such employment during the

[48] This is likely due to the move to urban areas where home ownership rates tend to be lower.

war years. With incomes determining individual's standards of living, health, political power, and the fortunes of their children, this was an important step forward for Black workers in the South. This finding is robust and I can rule out alternative explanations including differential migration or education of Black and white workers, the industrialization of the South at the time, the selection of soldiers, World War II related spending, historic anti-Black sentiment proxied by the share of cotton production and the number of slaves in 1860, the legacy of New Deal spending, selection on observables and unobservables, and different model specifications and time trends.

Using repeated cross-sectional Census data at the individual level in a triple differences setting, I also provide evidence that the semi-skilled white casualty rate was not only a shock to the supply of labor of young white men but that it reduced racial barriers to entry for Black workers into semi-skilled employment. The removal of such barriers in the aircraft industry during the war were described by Weaver (1945). This reduction in racial barriers to entry is consistent both with the historical account (see Weaver, 1945; Anderson, 1982) but also with a simple model that describes the transition of Black and white workers between agriculture and manufacturing based on ability in each sector as well as differential racial transition costs.

The overall contribution of this paper is to provide a new explanation for Black economic progress in the South at mid-century and to highlight the importance of group-specific barriers to entry into certain occupations that can significantly hamper the economic progress of this group. Avenues for future work include the further exploration of how this occupational upgrading of Black Southerners and the implied increase in disposable income of Black households impacted their ability to lobby for more political rights, how it affected Black-white social relations, and whether this increase in income made Black Americans more attractive customers for white store owners who would not have commonly served them out of fear of losing their white customers.

References

Aaronson, Daniel and Bhashkar Mazumder, "The Impact of Rosenwald Schools on Black Achievement," *Journal of Political Economy*, 2011, *119* (5), 821–888.

Acemoglu, Daron, David H. Autor, and David Lyle, "Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Midcentury," *Journal of Political Economy*, 2004, *112* (3), 497–551.

Aizer, Anna, Ryan Boone, Adriana Lleras-Muney, and Jonathan Vogel, "Discrimination and Racial Disparities in Labor Market Outcomes: Evidence from WWII," *NBER Working Paper 27689*, 2020.

_ , Shari Eli, Joseph Ferrie, and Adriana Lleras-Muney, "The Long-Run Impact of Cash Transfers to Poor Families," *American Economic Review*, 2016, *106* (4), 935–971.

Anderson, Karen Tucker, "Last Hired, First Fired: Black Women Workers during World War II," *Journal of American History*, 1982, *69* (1), 82–97.

Bailey, Martha J. and William J. Collins, "The Wage Gains of African-American Women in the 1940s," *Journal of Economic History*, 2006, *66* (3), 737–777.

_ , Karen Clay, Price Fishback, Shawn Kantor, Edson Severnini, and Anna Wentz, "U.S. County-Level Natality and Mortality Data, 1915-2007," *Inter-university Consortium for Political and Social Research [distributor], 2018-05-02. https://doi.org/10.3886/ICPSR36603.v2*, 2018.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen, "High-Dimensional Methods and Inference on Structural and Treatment Effects," *Journal of Economics Perspectives*, 2014, *28* (2), 29–50.

Borjas, George J., "Self-Selection and the Earnings of Immigrants," *American Economic Review*, 1987, *77* (4), 531–553.

Boustan, Leah, "Was Postwar Suburbanization "White Flight"? Evidence from the Black Migration," *Quarterly Journal of Economics*, 2010, *125* (1), 417–443.

_ and Robert A. Margo, "A silver lining to white flight? White suburbanization and African-American Homeownership, 1940-1980," *Journal of Urban Economics*, 2013, *78*, 71–80.

Chambers, John W., *To Raise an Army: The Draft Comes to Modern America*, Free Press of Macmillan, New York, NY, 1987.

Clubb, Jerome M., William H. Flanigan, and Nancy H. Zingale, "Electoral Data for Counties in the United States: Presidential and Congressional Races, 1840-1972," *ICPSR08611-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2006-11-13. DOI: https://doi.org/10.3886/ICPSR08611.v1*, 2006.

Collins, William J., "African-American Economic Mobility in the 1940s: A Portrait from the Palmer Survey," *Journal of Economic History*, 2000, *60* (3), 756–781.

_ , "Race, Roosevelt, and Wartime Production: Fair Employment in World War II Labor Markets," *American Economic Review*, 2001, *91* (1), 272–286.

Fishback, Price V. and Joseph A. Cullen, "Second World War spending and local economic activity in US counties, 1939-58," *Economic History Review*, 2013, *66* (4), 975–992.

_ , William C. Horrace, and Shawn Kantor, "The impact of New Deal expenditures on mobility during the Great Depression," *Explorations in Economic History*, 2006, *43* (2), 179–222.

Flynn, George Q., *The Draft, 1940-1973*, University Press of Kansas, Lawrence, KS, 1993.

Glass, Charles, *Deserter: A Hidden Story of the Second World War*, The Penguin Press, New York, NY, 2013.

Goldin, Claudia, "The Role of World War II in the Rise of Women's Employment," *American Economic*

Review, 1991, *81* (4), 741–756.

__ **and Claudia Olivetti**, "Shocking Labor Supply: A Reassessment of the Role of World War II on Womens Labor Supply," *American Economic Review Papers and Proceedings*, 2013, *103* (3), 257–262.

__ **and Robert A. Margo**, "The Great Compression: The Wage Structure in the United States at Mid-Century," *Quarterly Journal of Economics*, 1992, *107* (1), 1–34.

Hornbeck, Richard and Suresh Naidu, "When the Levee Breaks: Black Migration and Economic Development in the American South," *American Economic Review*, 2014, *104* (3), 963–990.

Jaworski, Taylor, "You're in the Army Now: The Impact of World War II on Women's Education, Work, and Family," *Journal of Economic History*, 2014, *74* (1), 169–195.

Kriner, Douglas and Francis Shen, *The Casualty Gap - The Causes and Consequences of American Wartime Inequalities*, Oxford University Press, Oxford, UK, 2010.

Lee, Ulysses, *The Employment of Negro Troops*, in Conn, S. (eds.) United States Army in World War II, Center of Military History U.S. Army, Washington D.C., 1965.

Lleras-Muney, Adriana and Allison Shertzer, "Did the Americanization Movement Succeed? An Evaluation of the Effect of English-Only and Compulsory Schooling Laws on Immigrants," *American Economic Journal: Economic Policy*, 2015, *7* (3), 258–290.

Maloney, Thomas N., "Wage Compression and Wage Inequality Between Black and White Males in the United States, 1940-1960," *Journal of Economic History*, 1994, *54* (2), 358–381.

Margo, Robert A., "Explaining Black-White Wage Convergence, 1940-1950," *Industrial and Labor Relations Review*, 1995, *48* (3), 470–481.

Miller, Conrad, "The Persistent Effect of Temporary Affirmative Action," *American Economic Journal: Applied Economics*, 2017, *9* (3), 152–190.

Montgomery, James D., "Social Networks and Labor-Market Outcomes: Toward an Economic Analysis," *American Economic Review*, 1991, *81* (5), 1408–1418.

Myrdal, Gunnar, *An American Dilemma: The Negro Problem and Modern Democracy*, Harper and Brothers Publishers, New York, NY, 1944.

Olmstead, Alan and Paul W. Rhode, "Reshaping the Landscape: The Impact and Diffusion of the Tractor in American Agriculture, 1910-1960," *Journal of Economic History*, 2001, *61* (3), 663–698.

Oster, Emily, "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business and Economic Statistics*, 2019, *37* (2), 187–204.

Pei, Zhuan, Jörn-Steffen Pischke, and Hannes Schwandt, "Poorly Measured Confounders are More Useful on the Left than on the Right," *Journal of Business and Economic Statistics*, 2019, *37* (2), 205–216.

Qian, Nancy and Marco Tabellini, "Discrimination, Disenfranchisement and African American WWII Military Enlistment," *Harvard Business School Working Paper No. 21-005*, 2020.

Royster, Dierdre A., *Race and the Invisible Hand: How White Networks Exclude Black Men from Blue-Collar Jobs*, University of California Press, Berkeley, CA, 2003.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek, "IPUMS USA," *Version 8.0 [dataset]. Minneapolis, MN: IPUMS, 2018. https://doi.org/10.18128/D010.V8.0*, 2018.

Shatnawi, Dina and Price V. Fishback, "The Impact of World War II on the Demand for Female Workers in Manufacturing," *Journal of Economic History*, 2018, *78* (2), 539–574.

Smith, James P. and Finis R. Welch, "Black Economic Progress After Myrdal," *Journal of Economic Literature*, 1989, *27* (2), 519–564.

Turner, Sarah and John Bound, "Closing the Gap or Widening the Divide: The Effects of the G.I. Bill and World War II onthe Educational Outcomes of Black Americans," *Journal of Economic History*, 2003, *63* (1), 145–177.

United States Department of Commerce, "County and City Data Book [United States] Consolidated File: County Data, 1947-1977. ICPSR07736-v2," *Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2012-09-18. DOI: https://doi.org/10.3886/ICPSR07736.v2*, 2012.

Weaver, Robert C., "Negro Employment in the Aircraft Industry," *Quarterly Journal of Economics*, 1945, *59* (4), 597–625.

Whatley, Warren C., "Getting a Foot in the Door: Learning, State Dependence, and the Racial Integration of Firms," *Journal of Economic History*, 1990, *50* (1), 43–66.

Wolfbein, Seymour L., "Postwar trends in Negro employment," *Monthly Labor Review*, 1947, *December*, 663–665.

Tables

Table 1: County Data Summary Statistics, 1920-1960

	obs.	mean	st. dev.	min	max
Main Outcome					
% semi-skilled Black workers	6,443	14.65	14.20	0.00	87.55
% semi-skilled Black workers, 1940	1,386	12.43	12.57	0.00	67.62
Military					
WWII casualty rate (semi-skilled white soldiers)	6,443	3.16	2.20	0.00	22.22
Mean casualty rate neighboring counties	6,430	1.15	1.69	0.00	11.53
% men drafted	6,443	9.39	13.19	0.00	61.59
Log WWII spending per capita	6,443	0.45	1.36	0.00	9.13
Demographics					
Log median family income	3,663	9.48	0.61	7.76	11.14
% with high school degree	3,691	20.20	9.36	3.70	71.60
% rural population	6,443	80.85	23.13	0.00	100.00
% Republican vote share	5,938	12.02	20.54	0.00	100.00
% Black population	6,443	23.56	21.18	0.00	90.77
% Black men	6,443	23.29	20.93	0.00	89.89
Lynchings per 1,000 Black pop, 1900-30	6,342	0.43	8.66	0.00	500.00
Rosenwald schools per 1,000 Black pop	6,342	0.67	1.29	0.00	26.32
% acres flooded by the Mississippi, 1928	6,443	0.54	5.69	0.00	100.00
Number of slaves in 1,000s in 1860	6,443	1,458.60	2,163.57	0.00	17,957.00
Agriculture					
% of land in agriculture	6,440	63.99	22.97	0.06	100.00
% of lang in cotton production	6,440	7.19	10.21	0.00	74.41
% cash tenants	6,441	6.63	7.53	0.00	78.28
Mean value of machines per farm	6,440	2.99	5.08	0.00	219.46
Manufacturing					
Manufacturing firms per capita	6,034	1.19	0.94	0.00	29.73
Mean manufacturing firm size	5,916	36.13	39.00	0.00	629.00
Log value added per worker	5,406	12.46	0.97	0.00	14.79
% manufacturing employment	5,916	3.98	4.47	0.00	100.00
New Deal Controls					
New Deal loans per capita, 1933-35	6,420	5.87	20.00	0.00	573.87
Relief per capita, 1933-39	6,420	9.81	26.24	0.00	949.11
Public works per capita, 1933-39	6,420	6.27	24.07	0.00	844.37
AAA spending per capita, 1933-39	6,420	6.85	28.85	0.00	852.11
FHA loans insured per capita, 1934-49	6,420	1.45	6.55	0.00	195.79
Unemployment rate, 1937	6,438	10.90	5.71	0.26	42.29

Note: Summary statistics for 1,388 counties in Southern states between 1920 and 1960. Monetary values are deflated to 2010 dollars.

Table 2: County Level Difference-in-Differences Results, 1920-1960

	Outcome: % semi-skilled Black workers (pre-war mean = 12.433)					
	(1)	(2)	(3)	(4)	(5)	(6)
Casualty rate × Post-war	0.539***	0.543***	0.493***	0.555***	0.718***	0.443***
	(0.130)	(0.147)	(0.154)	(0.152)	(0.228)	(0.142)
Controls		Yes		Yes	Yes	Yes
1940 controls × decade			Yes			
State × decade FE				Yes		
Linear county time trends					Yes	
Lasso selection						Yes
Observations	6,443	4,903	4,723	4,903	4,903	5,864
Counties	1,388	1,317	994	1,317	1,317	1,379
Adj. R^2	0.864	0.885	0.880	0.891	0.924	0.866
Oster's δ	1.877	1.674	1.436	1.954	1.131	1.253

Note: Difference-in-differences regressions of the county-level share of semi-skilled Bblack workers on the WWII county casualty rate among semi-skilled white soldiers interacted with a post-war indicator. The estimation sample uses decennial U.S. Census data on counties in Southern states from 1920 to 1960. Controls include county and decade fixed effects, the county draft rate, average casualty rate in the neighboring counties, log WWII spending per capita, share of Black men, share of rural population, no. of manufacturing establishments per capita, average manufacturing firm size, log manufacturing value added per worker, share of employment in manufacturing, share of land in agricultural production, share of acres in cotton production, share of cash tenants, average value of machinery per farm, lynchings per 1,000 Black population between 1900 and 1930, no. of Rosenwald schools per 1,000 Black population, share of acres flooded by the Mississippi in 1928, no. of slaves in 1860, Republican vote share, New Deal spending per capita 1933-35 (loans, public works, AAA, FHA loans), and the unemployment rate in 1937. Time-invariant controls are interacted with decade fixed effects. Monetary values are deflated to 2010 U.S. dollars. The doubly-robust selection method implements the Belloni et al. (2014) machine learning covariate selection algorithm for testing the stability of treatment effects with respect to the observables. The test for selection on unobservables by Oster (2019) is reported in the final row by computing the coefficient of proportionality δ for which the coefficient on the semi-skilled casualty rate among white soldiers would equal zero. Standard errors clustered at the county level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Treatment Heterogeneity

	Outcome: % semi-skilled Black workers (pre-war mean = 12.433)				
	(1)	(2)	(3)	(4)	(5)
Casualty rate× Post-war	0.202	0.181	0.077	0.268*	0.371***
	(0.140)	(0.154)	(0.084)	(0.148)	(0.139)
Interactions of Casualty rate× Post-war with **above-median indicators** for					
draft rate	0.516***				
	(0.129)				
months of labor shortages		0.466***			
		(0.131)			
outmigration			0.663***		
			(0.135)		
log military spending per capita				0.439***	
				(0.130)	
mean machine value per farm					0.283**
					(0.130)
Observations	6,443	6,443	6,443	6,443	6,443
Counties	1,388	1,388	1,388	1,388	1,388
Adj. R^2	0.866	0.865	0.860	0.865	0.865

Note: Difference-in-differences regressions of the county-level share of semi-skilled Black workers on the WWII county casualty rate among semi-skilled white soldiers interacted with a post-war indicator. The casualty rate is further interacted with indicators for whether a given interaction variable was above its median value. The number of months of labor shortages was computed as the number of months during which the number of soldiers from a county was 10% or more of that county's male labor force in 1940. Outmigration was estimated as population growth from 1940 to 1950 minus deaths plus births using data from Bailey et al. (2018). The estimation sample uses decennial U.S. Census data on counties in Southern states from 1920 to 1960. Controls include county and decade fixed effects, the county draft rate, average casualty rate in the neighboring counties, log WWII spending per capita, share of Black men, share of rural population, no. of manufacturing establishments per capita, average manufacturing firm size, log manufacturing value added per worker, share of employment in manufacturing, share of land in agricultural production, share of acres in cotton production, share of cash tenants, average value of machinery per farm, lynchings per 1,000 Black population between 1900 and 1930, no. of Rosenwald schools per 1,000 Black population, share of acres flooded by the Mississippi in 1928, no. of slaves in 1860, Republican vote share, New Deal spending per capita 1933-35 (loans, public works, AAA, FHA loans), and the unemployment rate in 1937. Time-invariant controls are interacted with decade fixed effects. Monetary values are deflated to 2010 U.S. dollars. Standard errors clustered at the county level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: County Level Difference-in-Differences Results, 1920-1960

	Outcome: % semi-skilled Black workers (pre-war mean = 12.433)					
	(1)	(2)	(3)	(4)	(5)	(6)
Low-skilled white	-0.081	-0.094	-0.001	-0.093	-0.068	-0.089
	(0.144)	(0.227)	(0.164)	(0.212)	(0.307)	(0.135)
Semi-skilled white	0.615***	0.609***	0.462***	0.630***	0.788***	0.310**
	(0.145)	(0.168)	(0.171)	(0.172)	(0.247)	(0.147)
High-skilled white	-0.227	-0.091	0.049	-0.164	-0.275	-0.066
	(0.181)	(0.199)	(0.201)	(0.205)	(0.380)	(0.165)
Low-skilled Black	-0.091**	-0.114*	-0.075	-0.092	-0.041	-0.140***
	(0.041)	(0.062)	(0.049)	(0.067)	(0.092)	(0.048)
Semi-skilled Black	0.084	0.050	0.084	0.052	0.082	0.036
	(0.065)	(0.059)	(0.059)	(0.052)	(0.098)	(0.048)
Low-skilled Black	-0.027	-0.058	0.042	-0.049	0.031	0.042
	(0.071)	(0.068)	(0.079)	(0.071)	(0.126)	(0.079)
Controls		Yes		Yes	Yes	Yes
1940 controls × decade			Yes			
State × decade FE				Yes		
Linear county time trends					Yes	
Lasso selection						Yes
Observations	6,443	4,903	4,723	4,903	4,903	5,296
Counties	1,388	1,317	994	1,317	1,317	1,340
Adj. R^2	0.865	0.886	0.886	0.891	0.924	0.889
Oster's δ	1.764	1.533	1.114	1.827	0.502	0.749

Note: Difference-in-differences regressions of the county-level share of semi-skilled Black workers on the WWII county casualty rate among semi-skilled white soldiers interacted with a post-war indicator. The estimation sample uses decennial U.S. Census data on counties in Southern states from 1920 to 1960. Controls include county and decade fixed effects, the county draft rate, average casualty rate in the neighboring counties, log WWII spending per capita, share of Black men, share of rural population, no. of manufacturing establishments per capita, average manufacturing firm size, log manufacturing value added per worker, share of employment in manufacturing, share of land in agricultural production, share of acres in cotton production, share of cash tenants, average value of machinery per farm, lynchings per 1,000 Black population between 1900 and 1930, no. of Rosenwald schools per 1,000 Black population, share of acres flooded by the Mississippi in 1928, no. of slaves in 1860, Republican vote share, New Deal spending per capita 1933-35 (loans, public works, AAA, FHA loans), and the unemployment rate in 1937. Time-invariant controls are interacted with decade fixed effects. Monetary values are deflated to 2010 U.S. dollars. The doubly-robust selection method implements the Belloni et al. (2014) machine learning covariate selection algorithm for testing the stability of treatment effects with respect to the observables. The test for selection on unobservables by Oster (2019) is reported in the final row by computing the coefficient of proportionality δ for which the coefficient on the semi-skilled casualty rate among white soldiers would equal zero. Standard errors clustered at the county level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Micro Census Triple Differences Results, 1920-1960

	Outcome: $\Pr(\text{semi-skilled}_{izt}) = 1$					
	Panel A: All U.S.					
	(1)	(2)	(3)	(4)	(5)	(6)
Casualty rate × Post-war	-0.002	-0.002	-0.000	-0.004	-0.003	-0.005
	(0.006)	(0.006)	(0.004)	(0.004)	(0.005)	(0.005)
Casualty rate × Black × Post-war	0.026***	0.027***	0.023***	0.025***	0.022***	0.022***
	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)
Individual controls		Yes	Yes	Yes	Yes	Yes
CZ Controls			Yes	Yes	Yes	Yes
Migration and Educ				Yes		
State time trends					Yes	
Commuting zone time trends						Yes
Observations	3,534,193	3,534,193	3,528,795	2,312,296	3,528,795	3,528,795
Adj. R^2	0.043	0.053	0.055	0.089	0.056	0.057
	Panel B: South					
	(1)	(2)	(3)	(4)	(5)	(6)
Casualty rate × Post-war	-0.007	-0.009	-0.006	-0.014**	-0.011	-0.016**
	(0.008)	(0.008)	(0.007)	(0.006)	(0.007)	(0.007)
Casualty rate × Black × Post-war	0.012***	0.014***	0.012***	0.017***	0.011***	0.011***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)
Individual controls		Yes	Yes	Yes	Yes	Yes
CZ Controls			Yes	Yes	Yes	Yes
Migration and Educ				Yes		
State time trends					Yes	
Commuting zone time trends						Yes
Observations	1,083,576	1,083,576	1,082,514	724,463	1,082,514	1,082,514
Adj. R^2	0.078	0.089	0.092	0.110	0.094	0.096

Note: Difference-in-difference-in-differences regression of a semi-skilled indicator on the commuting zone WWII casualty rate among semi-skilled white soldiers interacted with a post-WWII dummy, and with a Black indicator for individuals living in 722 commuting zones in the whole U.S. and 300 commuting zones in the South. The estimation sample contains repeated cross-sectional data from the decennial U.S. micro Census from 1920-60 on non-institutionalized, working Black and white males aged 15-65 who are not currently attending school. All regressions include commuting zone and Census year fixed effects. Individual level controls include age, marital status, age and place of birth dummies. Column (4) adds cross-state migration and education controls interacted with race and time fixed effects. Commuting zone level controls are the WWII draft rate, log WWII spending per capita, share of Black men, share of rural population, no. of manufacturing establishments per capita, average manufacturing firm size, log manufacturing value added per worker, share of employment in manufacturing, share of land in agricultural production, share of acres in cotton production, share of cash tenants, average value of machinery per farm, lynchings per 1,000 Black population between 1900 and 1930, no. of Rosenwald schools per 1,000 Black population, share of acres flooded by the Mississippi in 1928, no. of slaves in 1860, Republican vote share, New Deal spending per capita 1933-35 (loans, public works, AAA, FHA loans), and the unemployment rate in 1937. Time-invariant controls are interacted with decade fixed effects. Monetary values are deflated to 2010 U.S. dollars. Observations are weighted by their individual sample line weight times the spatial weight used to construct the commuting zone level treatment variable. Standard errors clustered at the commuting zone level in parentheses. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Triple Differences by Sector

	Outcome: $\Pr(\text{semi-skilled}_{izt}) = 1$					
	Panel A: All U.S.					
	Mfg All	Mfg Durable	Mfg Non-durable	Telecom	Retail	Mining and Construction
	(1)	(2)	(3)	(4)	(5)	(6)
Casualty rate × Post-war	-0.003	-0.003	0.010	-0.006	0.002	0.013**
	(0.006)	(0.006)	(0.010)	(0.015)	(0.004)	(0.005)
Casualty rate × Black × Post-war	0.072***	0.063***	0.079***	-0.026	-0.011***	-0.006
	(0.004)	(0.004)	(0.005)	(0.018)	(0.003)	(0.004)
Observations	1,119,584	433,224	686,515	30,815	373,498	425,896
Adj. R^2	0.050	0.047	0.061	0.116	0.033	0.125
	Panel B: South					
	Mfg All	Mfg Durable	Mfg Non-durable	Telecom	Retail	Mining and Construction
	(1)	(2)	(3)	(4)	(5)	(6)
Casualty rate × Post-war	-0.004	-0.007	0.003	0.002	0.003	0.009
	(0.006)	(0.007)	(0.007)	(0.027)	(0.006)	(0.007)
Casualty rate × Black × Post-war	0.054***	0.054***	0.050***	-0.085***	-0.017***	-0.011**
	(0.003)	(0.004)	(0.004)	(0.021)	(0.003)	(0.005)
Observations	269,943	134,155	135,788	7,232	111,754	212,679
Adj. R^2	0.119	0.104	0.147	0.190	0.051	0.182

Note: Difference-in-differenece-in-differences regression of a semi-skilled indicator on the commuting zone WWII casualty rate among semi-skilled white soldiers interacted with a post-WWII dummy, and with a Black indicator. The estimation sample contains data from the decennial U.S. micro Census from 1920-60 on non-institutionalized, working Black and white males aged 15-65. Regression results for semi-skill (columns 1-3) and high-skill (columns 4-6) intensive sectors. All regressions include commuting zone and Census year fixed effects. Individual level controls include age, marital status, age and place of birth dummies. Commuting zone level controls are the WWII draft rate, log WWII spending per capita, share of Black men, share of rural population, no. of manufacturing establishments per capita, average manufacturing firm size, log manufacturing value added per worker, share of employment in manufacturing, share of land in agricultural production, share of acres in cotton production, share of cash tenants, average value of machinery per farm, lynchings per 1,000 Black population between 1900 and 1930, no. of Rosenwald schools per 1,000 Black population, share of acres flooded by the Mississippi in 1928, no. of slaves in 1860, Republican vote share, New Deal spending per capita 1933-35 (loans, public works, AAA, FHA loans), and the unemployment rate in 1937. Time-invariant controls are interacted with decade fixed effects. Monetary values are deflated to 2010 U.S. dollars. Observations are weighted by their individual sample line weight times the spatial weight used to construct the commuting zone level treatment variable. Standard errors clustered at the commuting zone level in parentheses. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

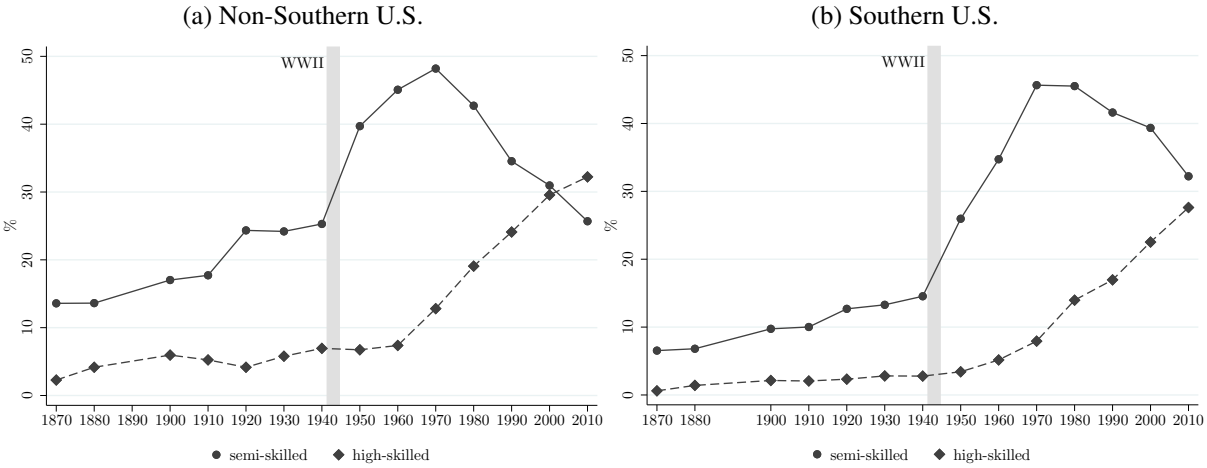
Table 7: WWII Casualties and Economic Characteristics of Semi-Skilled Workers

Outcome:	ln(wage)	Education	Migrant	Urban	Owns home	ln(house val.)
	Panel A: All U.S.					
	(1)	(2)	(3)	(4)	(5)	(6)
Casualty rate × Post-war	-0.007	-0.073**	0.004	-0.061***	-0.003	-0.030
	(0.022)	(0.032)	(0.007)	(0.020)	(0.004)	(0.020)
Casualty rate × Black × Post-war	0.031**	0.209***	-0.012**	0.006**	-0.001	0.049***
	(0.014)	(0.034)	(0.005)	(0.003)	(0.004)	(0.013)
Observations	1,114,851	1,114,851	1,587,521	1,587,521	1,523,089	522,632
Adj. R^2	0.493	0.431	0.323	0.737	0.271	0.394
	Panel B: South					
	(1)	(2)	(3)	(4)	(5)	(6)
Casualty rate × Post-war	-0.055	-0.055	-0.007	-0.104***	0.001	-0.032
	(0.035)	(0.050)	(0.008)	(0.029)	(0.007)	(0.029)
Casualty rate × Black × Post-war	0.030**	0.223***	-0.014***	0.003	-0.012***	0.049***
	(0.014)	(0.029)	(0.004)	(0.004)	(0.004)	(0.015)
Observations	346,013	346,013	465,132	465,132	439,274	156,717
Adj. R^2	0.480	0.405	0.440	0.704	0.258	0.379

Note: Difference-in-differenece-in-differences regression of economic outcomes on the commuting zone WWII casualty rate among semi-skilled white soldiers interacted with a post-WWII dummy, and with a Black indicator for individuals in semi-skilled employment living in 722 commuting zones in the whole U.S. (panel a) and the U.S. South (panel b). The estimation sample contains data from the decennial U.S. micro Census from 1920-60 on non-institutionalized, working Black and white males aged 15-65 who are not currently attending school. All regressions include commuting zone and Census year fixed effects. Urban is an indicator for whether an individual resided in a metropolitan area. Owns home is a binary outcomes for whether an individual owns their home. The log house value, log wages, and education variables are only available from 1940 onward. Log house value is also missing for 1950. Individual level controls include age, marital status, age and place of birth dummies. Commuting zone level controls are the WWII draft rate, log WWII spending per capita, share of Black men, share of rural population, no. of manufacturing establishments per capita, average manufacturing firm size, log manufacturing value added per worker, share of employment in manufacturing, share of land in agricultural production, share of acres in cotton production, share of cash tenants, average value of machinery per farm, lynchings per 1,000 Black population between 1900 and 1930, no. of Rosenwald schools per 1,000 Black population, share of acres flooded by the Mississippi in 1928, no. of slaves in 1860, Republican vote share, New Deal spending per capita 1933-35 (loans, public works, AAA, FHA loans), and the unemployment rate in 1937. Time-invariant controls are interacted with decade fixed effects. Monetary values are deflated to 2010 U.S. dollars. Observations are weighted by their individual sample line weight times the spatial weight used to construct the commuting zone level treatment variable. Standard errors clustered at the commuting zone level in parentheses. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

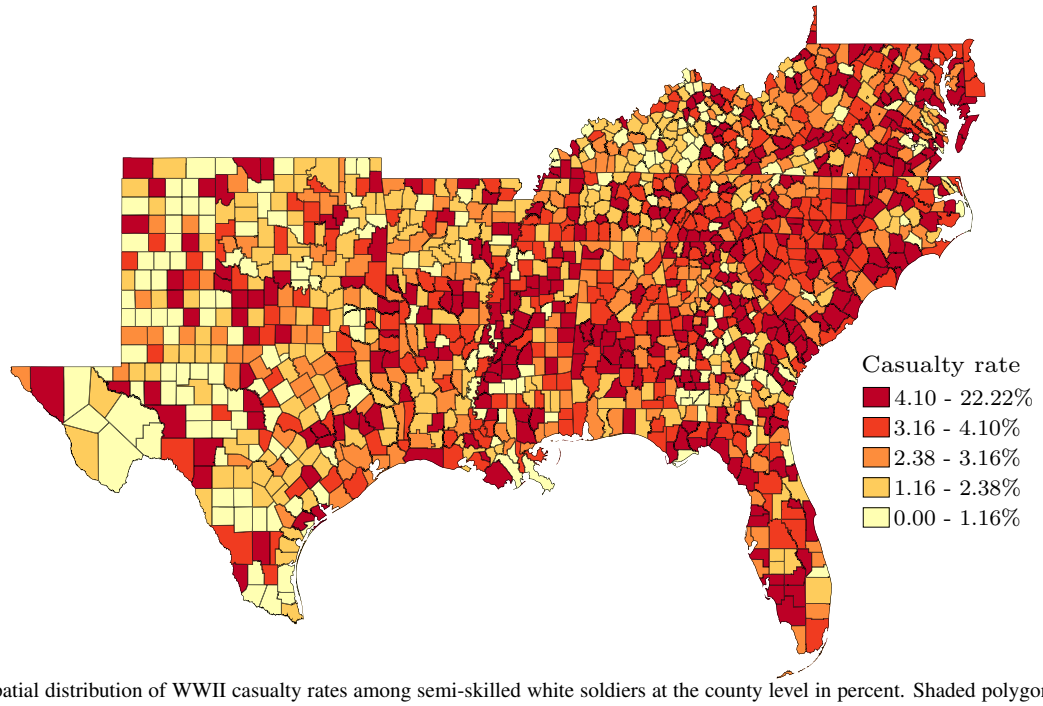
Figures

Figure 1: Share of Semi- and High-Skilled Employment among Black Men, 1870 to 2010



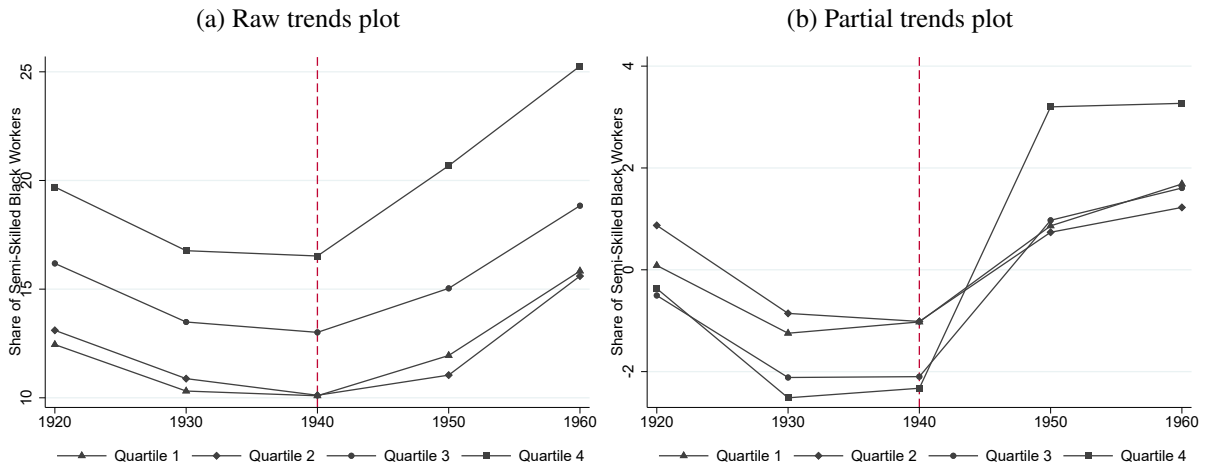
Note: Graphs are based on the public use micro data files of the 1870-2010 Decennial U.S. Censuses by Ruggles et al. (2018). The sample includes Black males aged 16 to 65 of the non-institutionalized population who are not attending school at the enumeration date. Semi-skilled jobs (dots) are operatives and craftsmen, and high-skilled jobs (diamonds) are clerks, professionals, and managers. Occupations are defined according to the 1950 Census Bureau occupational classification scheme. The years of U.S. involvement in World War II are marked with light gray background shading. Data for the South includes individuals living in the states of the former Confederacy, as well as Delaware, DC, Kentucky, Maryland, Oklahoma, and West Virginia.

Figure 2: WWII Casualty Rates among Semi-Skilled White Soldiers in the U.S. South



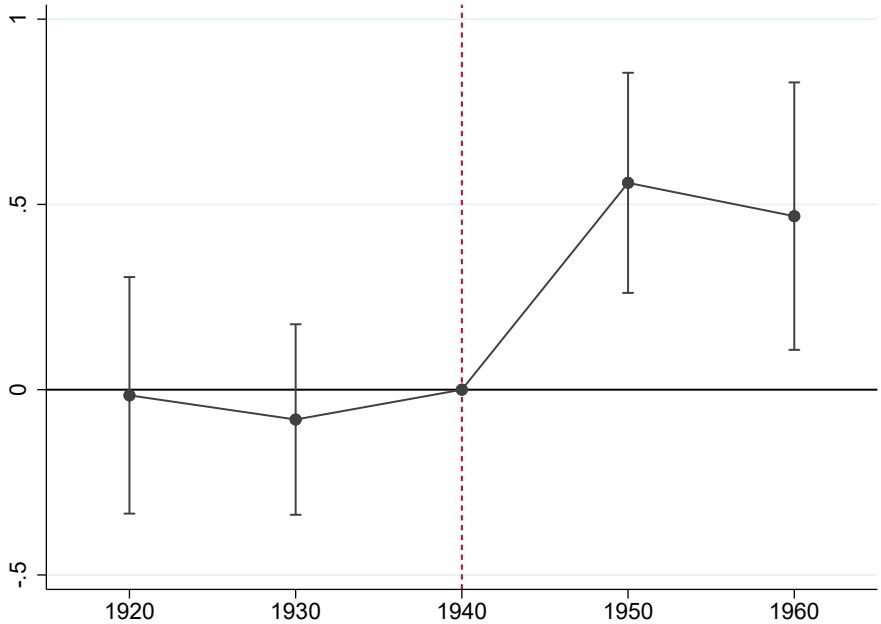
Note: Spatial distribution of WWII casualty rates among semi-skilled white soldiers at the county level in percent. Shaded polygons display the quintiles of the casualty rate distribution with ranges being shown in the legend on the side. Southern states included here are Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, South Carolina, Oklahoma, Tennessee, Texas, Virginia, and West Virginia.

Figure 3: Share of Semi-Skilled Black Workers by Casualty Rate Quartile



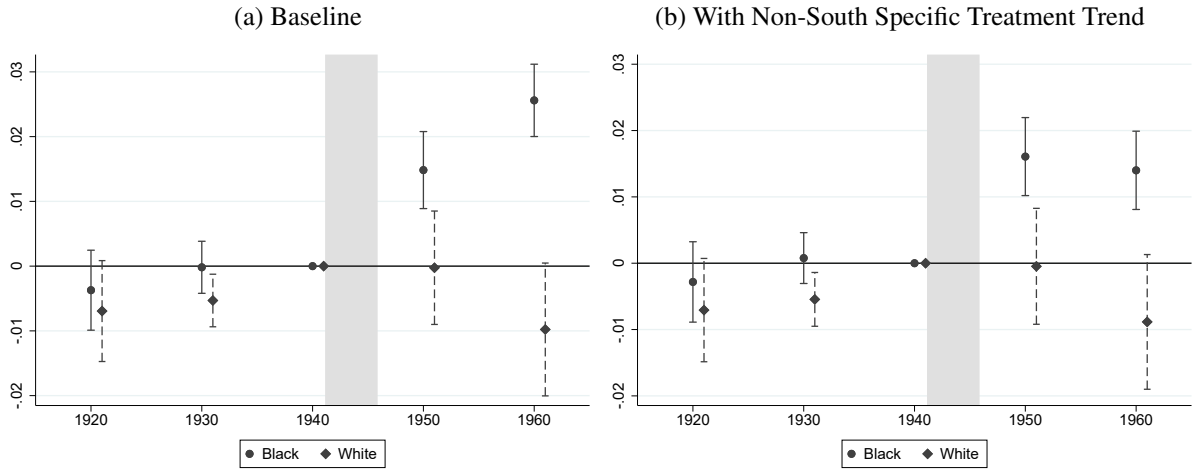
Note: The figure plots the raw outcome data for the share of semi-skilled Black workers for counties in Southern states by quartiles of the WWII casualty rate among semi-skilled white soldiers over time in panel a. Panel b shows the same figure after partialling out the share of manufacturing employment, Black population, and the urbanization rate to explain the pre-war differences in the share of semi-skilled Black workers across quartiles of the WWII casualty rate distribution among semi-skilled white soldiers. The dashed red line marks the last pre-war decade.

Figure 4: Difference-in-Differences Coefficient Plot



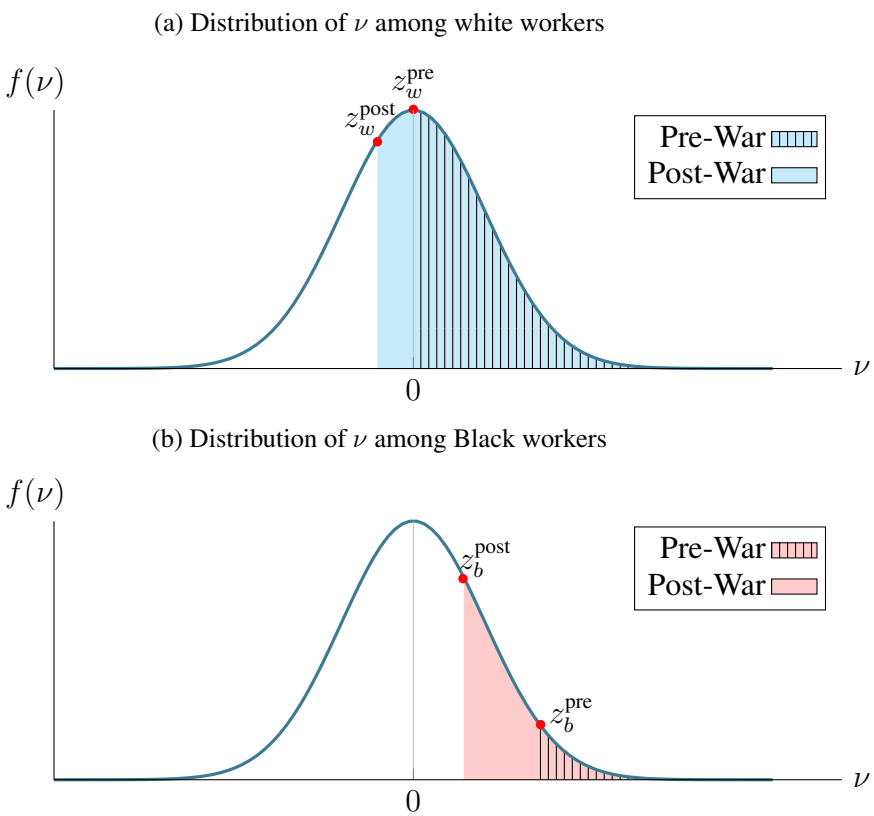
Note: Difference-in-differences regressions of the county-level share of semi-skilled Black workers on the WWII county casualty rate among semi-skilled white soldiers interacted with decade fixed effects. The omitted baseline decade is 1940 which is marked by the dashed line. This is the last pre-treatment period. The estimation sample contains counties in Southern states from 1920 to 1960. Controls include county fixed effects and state-by-decade fixed effects, the county draft rate, average casualty rate in the neighboring counties, log WWII spending per capita, share of Black men, share of rural population, no. of manufacturing establishments per capita, average manufacturing firm size, log manufacturing value added per worker, share of employment in manufacturing, share of land in agricultural production, share of acres in cotton production, share of cash tenants, average value of machinery per farm, lynchings per 1,000 Black population between 1900 and 1930, no. of Rosenwald schools per 1,000 Black population, share of acres flooded by the Mississippi in 1928, no. of slaves in 1860, Republican vote share, New Deal spending per capita 1933-35 (loans, public works, AAA, FHA loans), and the unemployment rate in 1937. Time-invariant controls are interacted with decade fixed effects. Monetary values are deflated to 2010 U.S. dollars. Standard errors clustered at the county level. Error bars show 95% confidence intervals around each coefficient estimate.

Figure 5: Triple Differences Coefficients Plot



Note: Coefficients plot from a difference-in-difference-in-differences regression of a semi-skilled indicator on the commuting zone WWII casualty rate among semi-skilled white soldiers interacted with decade dummies, and with a Black indicator. In panel b, the treatment-by-postwar interaction is further interacted with a indicator for non-Southern states to model the treatment specific trend term outside the South. White workers' coefficients represent the interaction of the casualty rate with decade dummies, plotted Black workers' coefficients are for the casualty rate interacted with decade dummies and a Black indicator. The estimation sample contains repeated cross-sectional data on individuals from the decennial U.S. micro Census from 1920-60 on non-institutionalized, working Black and white males aged 15-65. All regressions include commuting zone and Census year fixed effects. Controls include age, marital status, year of birth, a self-employment indicator, farm status, and industry fixed effects. The gray shaded area marks the war years. Observations are weighted by their individual sample line weight times the spatial weight used to construct the commuting zone level treatment variable. Standard errors clustered at the commuting zone level. Error bars show 95% confidence intervals around each coefficient estimate.

Figure 6: Distributions of Relative Talent in Manufacturing among white and Black Workers



Note: Distributions of the differences in manufacturing and agricultural talent ν among white and Black workers before and after World War II. The vertical lines show the pre-war distribution, the shaded area shows the post-war distribution. The threshold at which migration costs from agriculture to manufacturing become larger than the gains are the border of the lined and shaded areas before the war, and the border between the shaded and blank areas in the post-war period.

Online Appendix for
"World War II and Black Economic Progress"
by Andreas Ferrara

Not for Publication

Online Appendix Table of Contents

Continued on the next page

Online Appendix Table of Contents - Continued

Tables

Table A1: Correlation Between Navy/Marine Corps and Army/Air Force Casualty Rates

	Army Casualty Rate		Army Casualty Rate (semi-skilled whites)	
	(1)	(2)	(3)	(4)
Navy casualty rate	0.03235	-0.01089	0.00002	0.00063
	(0.04192)	(0.03813)	(0.00599)	(0.00764)
Controls		Yes		Yes
Observations	1,273	942	1,273	942
R^2	0.001	0.113	0.000	0.357
Outcome mean	0.332	0.332	0.043	0.044

Note: County-level cross-sectional regression of Army and Army Air Force casualty rates in general and for semi-skilled white soldiers on the casualty rate in the Navy, Marine Corps, and Coast Guard. Since the number of drafted and enlisted soldiers in the Navy, Marines and Coast Guard by county is not available, all casualty rates are expressed relative to each county's male population. If included, controls are measured in 1940 and contain the share of Black men, share of rural population, no. of manufacturing establishments per capita, average manufacturing firm size, log manufacturing value added per worker, share of employment in manufacturing, share of land in agricultural production, share of acres in cotton production, share of cash tenants, average value of machinery per farm, lynchings per 1,000 Black population between 1900 and 1930, no. of Rosenwald schools per 1,000 Black population, share of acres flooded by the Mississippi in 1928, no. of slaves in 1860, Republican vote share, New Deal spending per capita 1933-35 (loans, public works, AAA, FHA loans), and the unemployment rate in 1937. Also included is the log WWII spending per capita during the war. Monetary values are deflated to 2010 U.S. dollars. Robust standard errors in parentheses. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A2: Summary Statistics - WWII Enlistment Records

Panel A

	Black (n = 807,116)				White (n = 7,228,570)			
	mean	st. dev.	min.	max.	mean	st. dev.	min.	max.
Age	25.03	5.80	18	49	24.59	5.69	18	49
Education	9.29	1.86	8	18	10.68	2.24	8	18
AGCT	70.19	19.54	40	187	100.46	22.17	40	199
Married	0.23	0.42	0	1	0.23	0.42	0	1
Height (in.)	68.21	3.51	59	82	68.49	3.25	59	82
Weight (lbs.)	148.42	17.90	94	249	149.59	19.97	88	257
Died	0.019	0.139	0	1	0.029	0.169	0	1

Panel B

	Enlisted (n = 1,670,352)				Drafted (n = 6,622,454)			
	mean	st. dev.	min.	max.	mean	st. dev.	min.	max.
Age	22.859	5.155	18	48	25.156	5.809	18	49
Education	11.456	2.148	8	20	10.306	2.244	8	20
AGCT	133.181	27.585	1	199	95.777	22.773	1	199
Married	0.121	0.326	0	1	0.256	0.436	0	1
Height (in.)	68.821	2.839	59	82	68.328	3.414	59	82
Weight (lbs.)	149.056	19.256	90	257	149.311	20.066	88	257
Died	0.027	0.162	0	1	0.029	0.167	0	1

Panel C

	South (n = 2,249,203)				Non-South (n = 6,043,984)			
	mean	st. dev.	min.	max.	mean	st. dev.	min.	max.
Age	22.288	5.570	18	46	24.844	5.819	18	49
Education	10.157	2.207	8	20	10.680	2.280	8	20
AGCT	90.722	25.958	1	199	99.825	22.727	1	199
Married	0.252	0.434	0	1	0.220	0.414	0	1
Height (in.)	68.658	2.308	59	82	68.364	3.293	59	82
Weight (lbs.)	148.076	19.501	90	256	149.657	19.989	88	257
Died	0.028	0.166	0	1	0.028	0.166	0	1

Note: Summary statistics for data from drafted soldiers in the Army or Army Air Force between 1940 and 1946. AGCT is the Army General Classification Test, an ability test administered during the draft examinations. This measure is only available for a subset of men drafted in 1943. The similarities in the minimum values for the AGCT, education levels, and height across groups are due to the minimum requirements imposed by the Army on the draft. The indicator for a soldier's death equals one for those who were killed in combat or who died due to all other reasons such as battle and non-battle injuries, accidents, self-inflicted wounds or diseases.

Table A3: Spatial Independence Test of WWII Casualty Rates

	Distance threshold					
	200km	400km	600km	200km	400km	600km
	(1)	(2)	(3)	(4)	(5)	(6)
Moran's I	0.078***	0.064***	0.049***	-0.008	-0.005*	-0.003
	[16.473]	[26.595]	[31.875]	[-1.557]	[-1.775]	[-1.235]
Observations	1,387	1,387	1,387	1,387	1,387	1,387
State FE				Yes	Yes	Yes

Note: Moran's I for testing spatial independence of the WWII casualty rate among semi-skilled whites. For each I, the z-score is reported in squared brackets using a binary spatial weight matrix. Each county is identified by the latitude and longitude of its centroid. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: Testing for Hot and Cold Spots of WWII Casualty Rates

	Distance threshold					
Getis-Ord $G_i^*(d)$	200km	400km	600km	200km	400km	600km
z-score interval	(1)	(2)	(3)	(4)	(5)	(6)
z $\leq$ -2.58	232	347	347	0	0	0
-2.58 $<$ z $\leq$ -1.96	133	49	33	8	2	0
-1.96 $<$ z $<$ 1.96	613	371	262	1,370	1,378	1,386
1.96 $\leq$ z $<$ 2.58	130	80	59	8	7	1
2.58 $\leq$ z	279	540	686	1	0	0
Observations	1,387	1,387	1,387	1,387	1,387	1,387
State FE				Yes	Yes	Yes

Note: Getis-Ord $G_i^*(d)$ test for testing local spatial independence of the WWII casualty rate among semi-skilled whites. Local spatial independence is given when the z-score on the corresponding test statistic lies within -1.96 $<$ z $<$ 1.96. Unusually low casualty rate clusters (cold spots) are found for counties with z-scores of z $\leq$ -1.96. Conversely, unusually high casualty rate clusters (hot spots) are found for counties with z-scores of 1.96 $\leq$ z. The number of counties in each z-score bin is provided in the rows of the table. Each county is identified by the latitude and longitude of its centroid.

Table A5: Difference-in-Differences with Binary Treatment Variable

	Outcome: % semi-skilled Black workers (pre-war mean = 12.433)					
	(1)	(2)	(3)	(4)	(5)	(6)
Q4(Casualty rate) × Post-war	2.129***	1.776***	1.238**	1.612***	2.505***	1.276**
	(0.566)	(0.585)	(0.567)	(0.574)	(0.971)	(0.583)
Controls		Yes		Yes	Yes	Yes
1940 controls × decade			Yes			
State × decade FE				Yes		
Linear county time trends					Yes	
Lasso selection						Yes
Observations	6,443	4,903	4,723	4,903	4,903	5,372
Counties	1,388	1,317	994	1,317	1,317	1,363
Adj. R^2	0.864	0.885	0.879	0.890	0.924	0.869
Oster's δ	1.634	1.390	0.810	1.318	1.033	0.865

Note: Difference-in-differences regressions of the county-level share of semi-skilled Black workers on an indicator for whether a county was in the top quartile of the WWII county casualty rate among semi-skilled white soldiers interacted with a post-war indicator. The estimation sample uses decennial U.S. Census data on counties in Southern states from 1920 to 1960. Controls include county and decade fixed effects, the county draft rate, average casualty rate in the neighboring counties, log WWII spending per capita, share of Black men, share of rural population, no. of manufacturing establishments per capita, average manufacturing firm size, log manufacturing value added per worker, share of employment in manufacturing, share of land in agricultural production, share of acres in cotton production, share of cash tenants, average value of machinery per farm, lynchings per 1,000 Black population between 1900 and 1930, no. of Rosenwald schools per 1,000 Black population, share of acres flooded by the Mississippi in 1928, no. of slaves in 1860, Republican vote share, New Deal spending per capita 1933-35 (loans, public works, AAA, FHA loans), and the unemployment rate in 1937. Time-invariant controls are interacted with decade fixed effects. Monetary values are deflated to 2010 U.S. dollars. The doubly-robust selection method implements the Belloni, Chernozhukov and Hansen (2014) machine learning covariate selection algorithm for testing the stability of treatment effects with respect to the observables. The test for selection on unobservables by Oster (2019) is reported in the final row by computing the coefficient of proportionality δ for which the coefficient on the semi-skilled casualty rate among white soldiers would equal zero. Standard errors clustered at the county level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A6: County Level Difference-in-Differences Results with Income Controls

	Outcome: % semi-skilled Black workers (pre-war mean = 12.433)					
	(1)	(2)	(3)	(4)	(5)	(6)
Casualty rate × Post-war	0.552***	0.593***	0.346**	0.595***	0.969***	0.538***
	(0.132)	(0.172)	(0.148)	(0.177)	(0.371)	(0.132)
Controls		Yes		Yes	Yes	Yes
1940 controls × decade			Yes			
State × decade FE				Yes		
Linear county time trends					Yes	
Lasso selection						Yes
Observations	3,691	2,831	2,750	2,831	2,831	3,666
Counties	1,388	1,219	994	1,219	1,219	1,387
Adj. R^2	0.890	0.909	0.909	0.915	0.941	0.891
Oster's δ	1.648	1.506	0.884	1.735	1.736	1.586

Note: Difference-in-differences regressions of the county-level share of semi-skilled Black workers on the WWII county casualty rate among semi-skilled white soldiers interacted with a post-war indicator. The estimation sample contains decennial U.S. Census data on counties in Southern states from 1940 to 1960. Controls include county and decade fixed effects, the county draft rate, average casualty rate in the neighboring counties, log WWII spending per capita, share of Black men, share of rural population, log median family income, share of pop. with high school degree, no. of manufacturing establishments per capita, average manufacturing firm size, log manufacturing value added per worker, share of employment in manufacturing, share of land in agricultural production, share of acres in cotton production, share of cash tenants, average value of machinery per farm, lynchings per 1,000 Black population between 1900 and 1930, no. of Rosenwald schools per 1,000 Black population, share of acres flooded by the Mississippi in 1928, no. of slaves in 1860, Republican vote share, New Deal spending per capita 1933-35 (loans, public works, AAA, FHA loans), and the unemployment rate in 1937. Time-invariant controls are interacted with decade fixed effects. Monetary values are deflated to 2010 U.S. dollars. The doubly-robust selection method implements the Belloni et al. (2014) machine learning covariate selection algorithm for testing the stability of treatment effects with respect to the observables. The test for selection on unobservables by Oster (2019) is reported in the final row by computing the coefficient of proportionality δ for which the coefficient on the semi-skilled casualty rate among white soldiers would equal zero. Standard errors clustered at the county level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A7: Difference-in-Differences Results with First Differenced Outcome

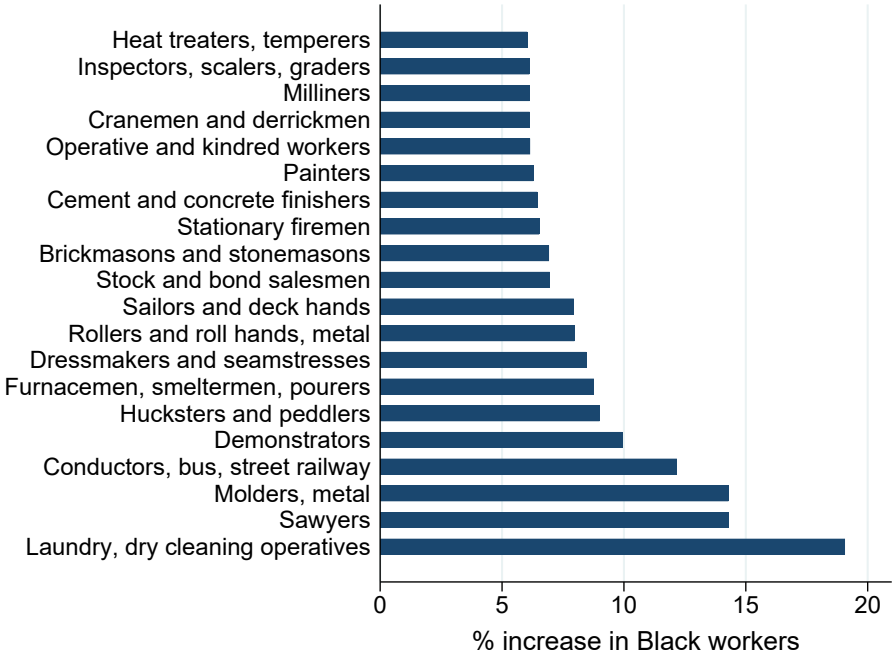
	Outcome: Δ No. of semi-skilled Black workers				
	(1)	(2)	(3)	(4)	(5)
Casualties × Post-war	3.148***	1.963	1.952*	1.532	1.520
	(0.844)	(1.398)	(0.997)	(1.704)	(1.916)
Casualties × drafted		0.000			-0.000
		(0.000)			(0.000)
Casualties × manufacturing			0.003**		0.003**
			(0.001)		(0.001)
Casualties × urbanization				0.025	0.015
				(0.023)	(0.025)
Joint F test		7.530	9.063	7.562	4.641
p-value		0.001	0.000	0.001	0.001
Observations	5,075	5,075	4,792	5,075	4,792
Adj. R^2	0.472	0.472	0.472	0.472	0.472

Note: Difference-in-differences regressions of the change in the number of semi-skilled Black workers on the number of WWII county casualties among semi-skilled white soldiers. The number of fallen white soldiers is also interacted with the number of total men drafted, the number of manufacturing establishments, and the urbanization rate. Interaction variables are standardized to have mean zero and variance one such that they can be interpreted in terms of a standard deviation increase in the given variable. The estimation sample uses decennial U.S. Census data on counties in Southern states from 1920 to 1960. Controls include the number of drafted Black and white soldiers, the number of Black men in a county, and decade fixed effects. To account for different population sizes across counties, observations are weighted by their total population. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

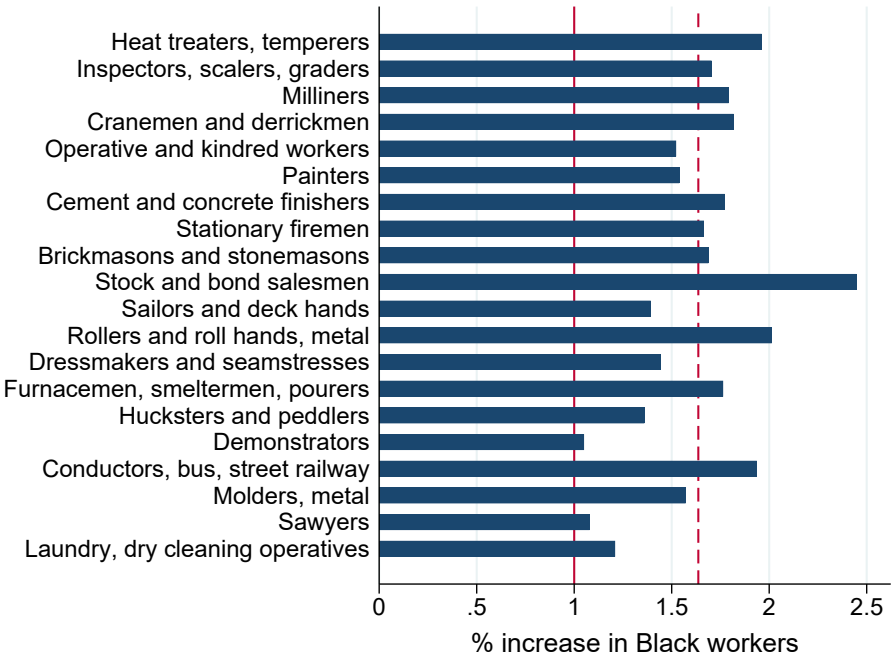
Figures

Figure A1: Inflows of Black Workers to Semi-Skilled Occupations and Relative Wage Gains

(a) %-points inflow of Black workers into semi-skilled occupations

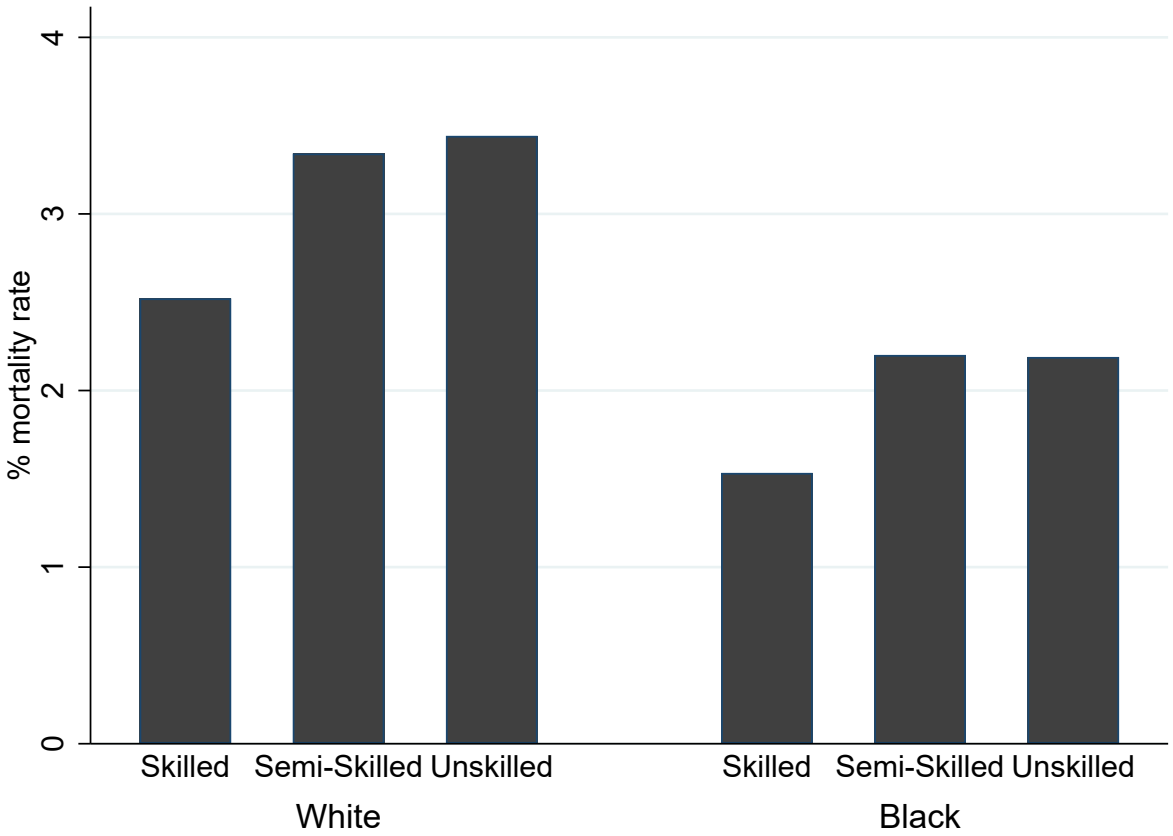


(b) Semi- to low-skilled wage ratio by semi-skilled job



Note: Graphs are based on the public use microdata files of the 1940 and 1950 Decennial U.S. Censuses (Ruggles, Flood, Goeken, Grover, Meyer, Pacas and Sobek, 2018). The sample includes males aged 16 to 65 of the non-institutionalized population who are not attending school at the enumeration date. Panel (a) shows the change in the share of Black workers in the listed semi-skilled occupations from 1940 to 1950. Occupations are defined based on the Census Bureau's 1950 occupational classification scheme. Panel (b) plots the ratio of wages paid in each occupation over the average low-skilled wage in 1950. The red line marks the average low-skilled job's wage in 1950 as baseline and the dashed red line at 1.6 marks the average wage relative to that baseline wage in low-skilled jobs.

Figure A2: World War II Mortality Rates by Race and Skill Group



Note: WWII mortality rates for white and Black soldiers by occupational skill group. Mortality rates are computed as the total number of fallen soldiers divided by the total number of soldiers within each race and skill group.

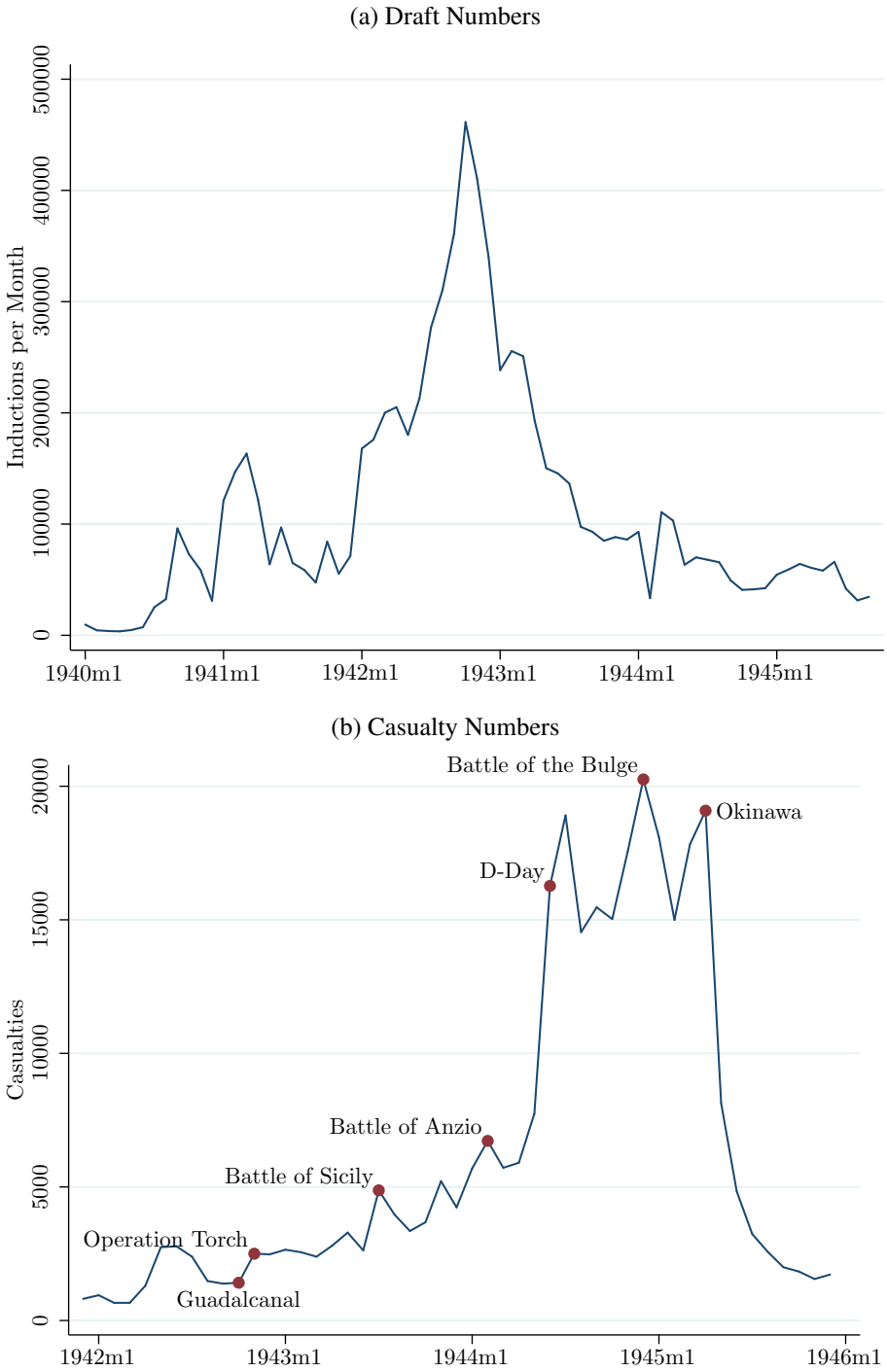
Figure A3: Racial Distribution of Troops in the Army (Enlisted Soldiers)

Units and Centers	Mobilized as of 31 Dec 42		1943 Augmentation		Total by 31 Dec 43	
	White	Negro	White	Negro	White	Negro
Total	4,532,117	467,883	2,246,233	286,767	6,778,350	754,650
Combat units	1,820,254	86,294	842,911	64,873	2,663,165	151,167
Service units[a]	578,262	148,370	263,300	90,991	841,562	239,361
AAF and services	1,190,363	109,637	810,000	90,000	2,000,363	199,637
Overhead[b]	363,820	65,880	64,155	9,145	427,975	75,025
RTC's	238,500	27,500	44,000	6,000	282,500	33,500
OCS's	72,200	800	0	0	72,200	800
Unassigned	268,718	29,402	221,867	25,758	490,585	55,160

[a] Includes AGF services but excludes AAF services.
[b] Includes men in hospitals 60 days or longer, men in replacement depots, men assigned to headquarters, station complements, and installation staffs, and men on detached lists.

Note: Racial distribution of enlisted soldiers in the U.S. Army and Army Air Force (AAF). Acronyms: AGF (Army Ground Force), RTC (Replacement Training Center), OCS (Officer Candidate School). Source: Lee (1965)

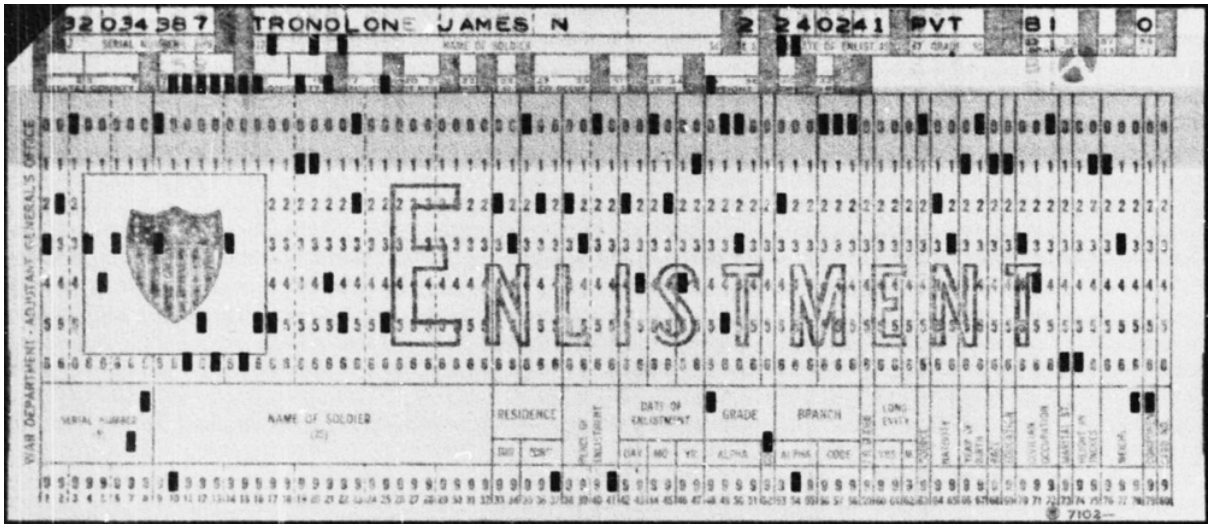
Figure A4: Number of Drafted and Fallen Soldiers by Month and Year



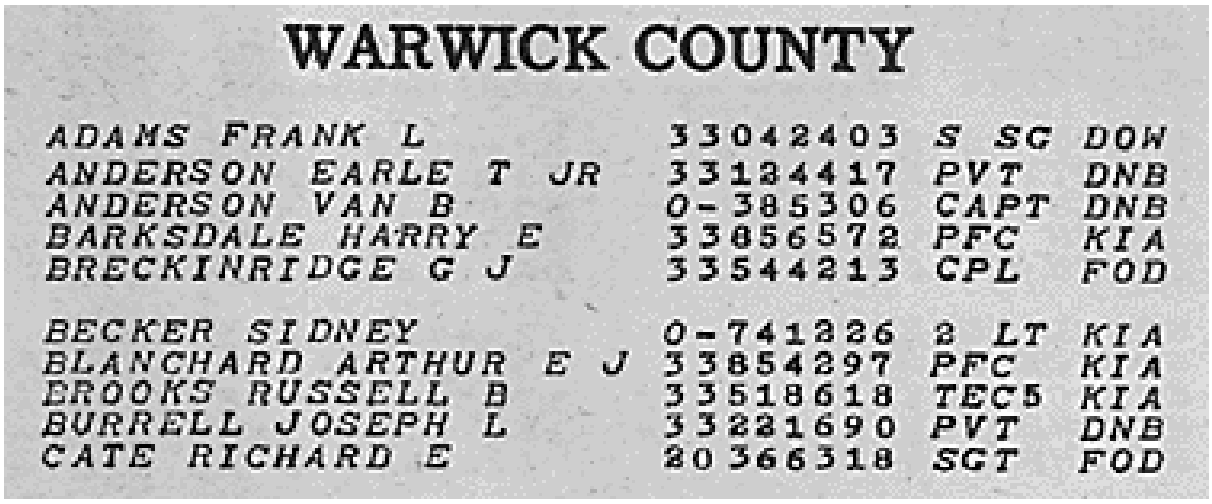
Note: Draft numbers (inductions) also include those who enlisted voluntarily prior to when voluntary enlistment was forbidden in 1942. Both draft and casualty figures are for the Army and Army Air Force only. Panel (b) shows the number of fallen soldiers per month together with major battles and operations involving U.S. Army and Army Air Force personnel. Casualties here refer to all combat and non-combat related deaths. The draft series begins with the enactment of the WWII draft in 1940 whereas the casualty series begins with the attack on Pearl Harbor. Monthly casualty counts come from the Office of the Adjutant General (1946) "Army Battle Casualties and Nonbattle Deaths in World War II - Final Report".

Figure A5: Draft and Casualty Records Example

(a) IBM Draft Punch Card



(b) WWII Honor List of Dead and Missing

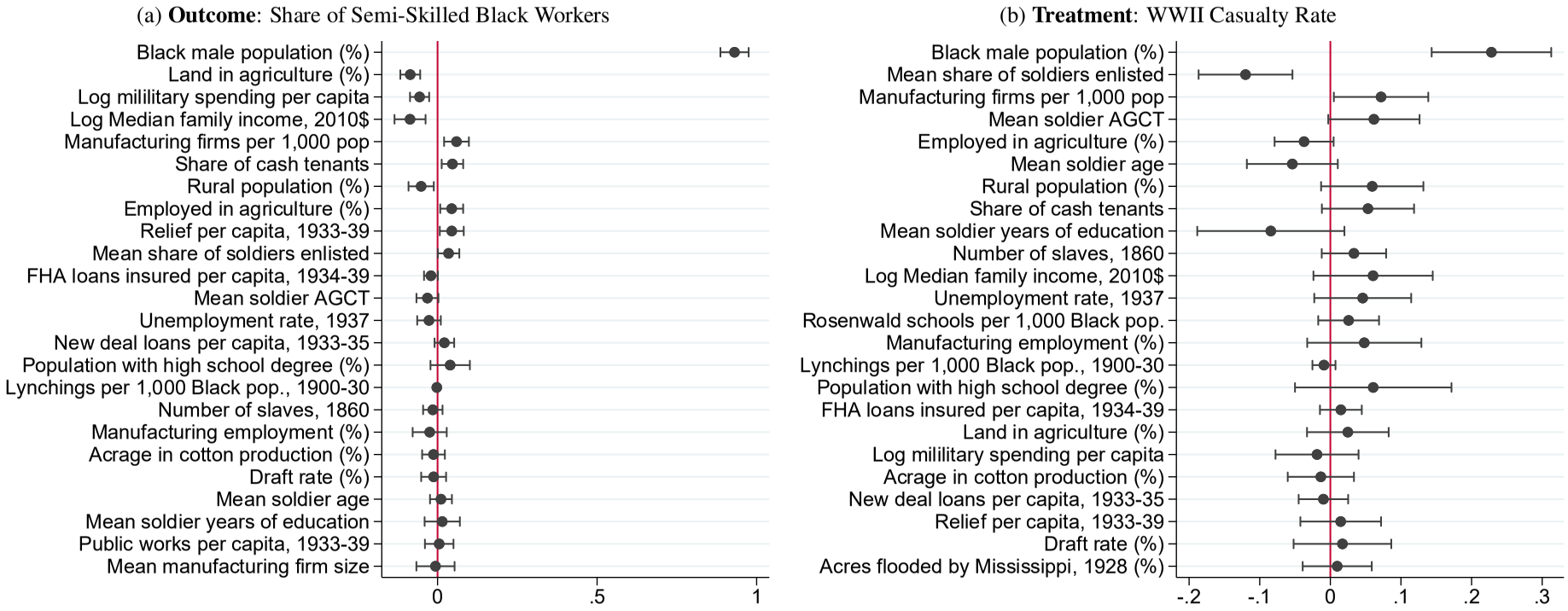


WARWICK COUNTY

ADAMS FRANK L	33042403	S SG	DOW
ANDERSON EARLE T JR	33124417	PVT	DNB
ANDERSON VAN B	0-385306	CAPT	DNB
BARKSDALE HARRY E	33856572	PFC	KIA
BRECKINRIDGE G J	33544213	CPL	FOD
BECKER SIDNEY	0-741226	2 LT	KIA
BLANCHARD ARTHUR E J	33854297	PFC	KIA
BROOKS RUSSELL B	33518618	TEC5	KIA
BURRELL JOSEPH L	33221690	PVT	DNB
CATE RICHARD E	20366318	SGT	FOD

Note: Panel a) shows the enlistment punch card for James Tronolone from Erie, New York, born in 1910. His Army serial number is shown on the top left corner of the card, his rank, date of enlistment, and service branch, among other, on the top right. Panel b) shows an excerpt from the WWII Honor List of Dead and Missing for Warwick County, Virginia. The table displays a soldier's name, their Army serial number, rank, and cause of death. Source: National Archives and Records Administration, Record Group 407: Records of the Adjutant General's Office, 1917- [AGO].

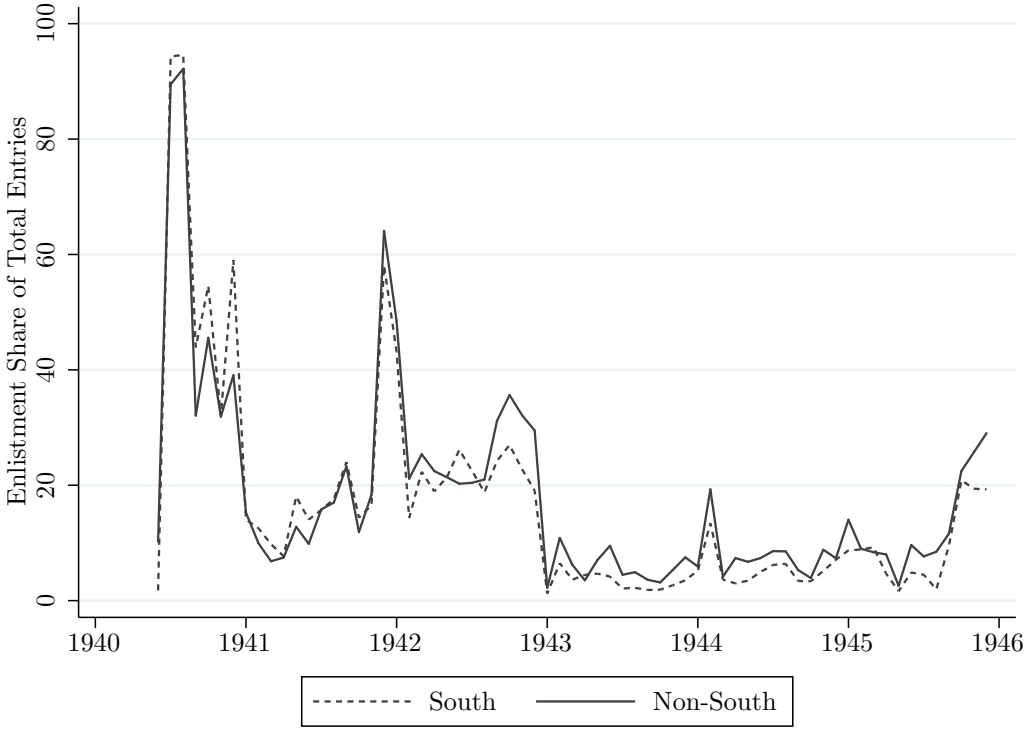
Figure A6: Observable Determinants of Outcome and Treatment



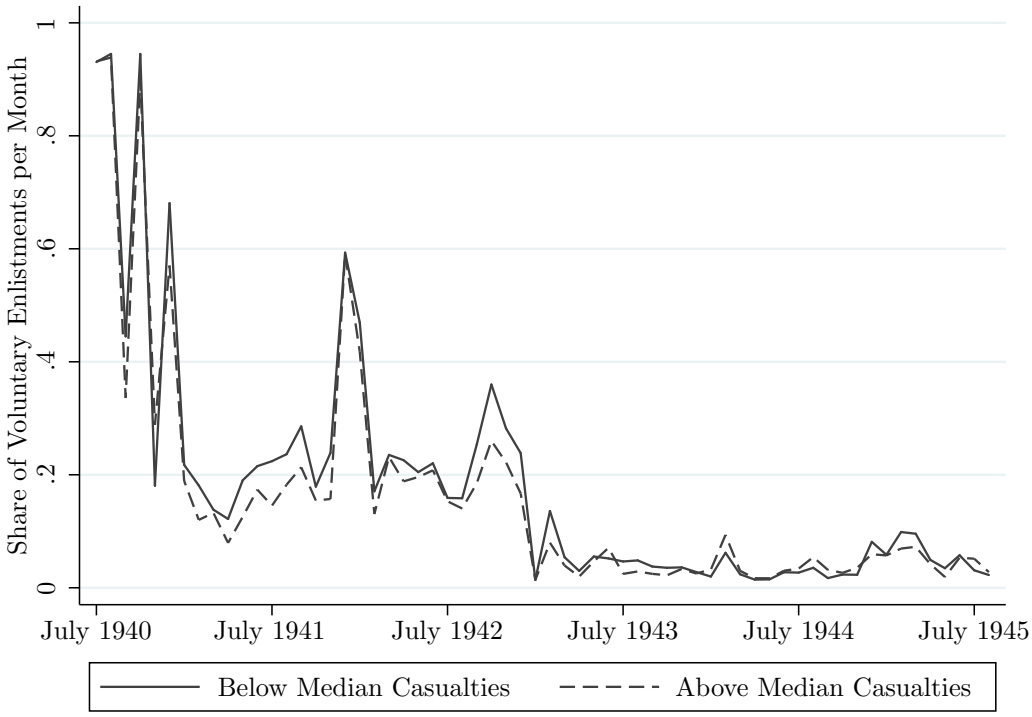
Note: Cross-sectional correlation ranking of pre-war controls from 1940 with the post-war outcome (share of semi-skilled Black workers) and treatment (WWII casualty rate among semi-skilled white soldiers) variables in 1950. All variables are de-meaned and standardized to have unit variance. Beta coefficients are ranked by the absolute value of their t-statistic to show the most important correlates from top to bottom. All regressions include state fixed effects for which coefficients have been dropped for this plot. Error bars show 95% confidence intervals.

Figure A7: Voluntary Enlistment Rates

(a) South vs. Non-South



(b) Within South



Note: Share of voluntary enlistments out of total new entries into the Army and Army Air Force by month. The drop at the end of 1942 is because voluntary enlistment was forbidden to avoid hurting the war economy due to overenthusiastic enlistments as was the case in the United Kingdom. After December 1942 only men aged 38 or older were allowed to volunteer if they demonstrated their physical and mental fitness for service.

Figure A8: Data Source for Semi-Skilled Employment of Black Workers

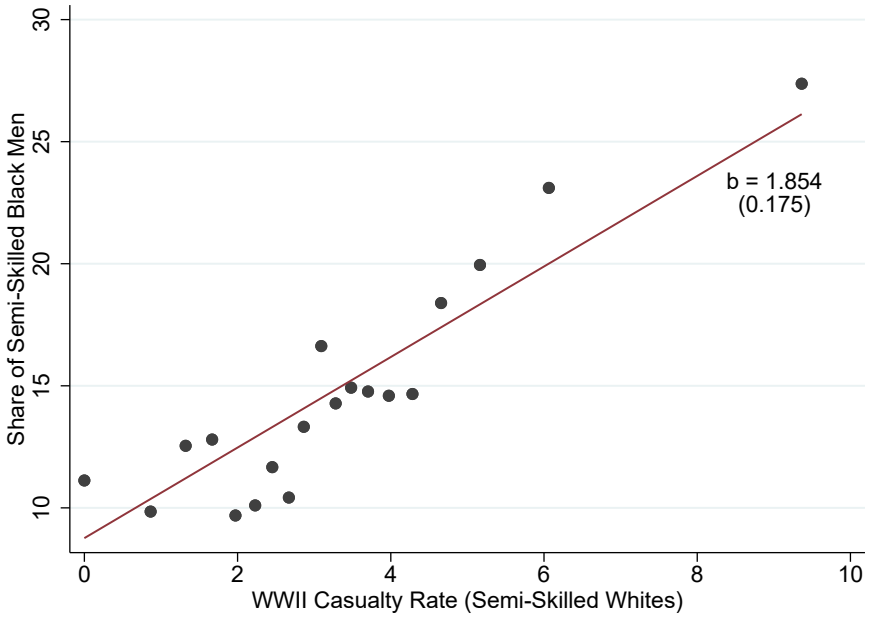
Table 23a.—NONWHITE EMPLOYED WORKERS 14 YEARS OLD AND OVER, BY MAJOR OCCUPATION GROUP AND SEX, BY COUNTIES: 1940

COUNTY AND SEX	Total employed (except on public emergency work)	Professional workers	Semiprofessional workers	Farmers and farm managers	Proprietors, managers, and officials, exc. farm	Clerical, sales, and kindred workers	Craftsmen, foremen, and kindred workers	Operatives and kindred workers	Domestic service workers	Service workers, except domestic	Farm laborers (wage workers) and foremen	Farm laborers, unpaid family workers	Laborers, except farm	Occupation not reported
Appling.....Male....	646	8	-	105	3	3	17	59	7	27	93	23	300	1
Female..	252	22	-	12	1	-	-	2	160	20	7	5	-	3
Atkinson.....Male....	504	3	-	44	3	-	3	19	1	7	54	12	355	3
Female..	136	8	-	2	-	-	-	1	106	8	2	5	1	3
Bacon.....Male....	288	1	-	48	3	11	6	16	2	4	23	5	168	1
Female..	92	4	-	3	-	-	-	1	57	7	3	4	2	1
Baker.....Male....	1,151	1	-	543	-	-	3	6	4	1	319	203	69	2
Female..	360	15	-	45	-	-	-	1	75	3	40	179	-	2
Baldwin.....Male....	1,984	23	3	392	14	13	141	147	50	248	306	124	515	8
Female..	1,484	67	-	33	2	3	2	48	923	241	41	119	2	8
Banks.....Male....	150	-	-	90	-	-	-	1	3	3	21	28	3	1
Female..	25	1	-	5	-	-	-	-	7	-	-	12	-	-
Barrow.....Male....	585	2	-	202	1	-	9	26	10	33	196	64	39	3
Female..	373	9	-	4	4	-	-	1	250	13	21	71	-	-
Bartow.....Male....	947	11	-	213	4	3	19	151	43	70	170	59	195	9
Female..	631	28	-	9	1	2	-	5	496	38	12	36	1	3
Ben Hill.....Male....	1,117	15	3	204	16	10	47	153	6	59	270	65	265	4
Female..	608	26	-	13	14	2	-	10	489	33	11	4	2	4
Berrien.....Male....	683	4	-	74	5	5	17	49	9	9	149	23	337	2
Female..	274	3	-	4	2	-	-	2	202	12	33	15	-	1
Bibb.....Male....	7,379	138	14	218	115	170	640	1,924	263	947	547	53	2,311	39
Female..	6,625	231	4	18	51	58	19	443	4,858	754	29	37	101	23

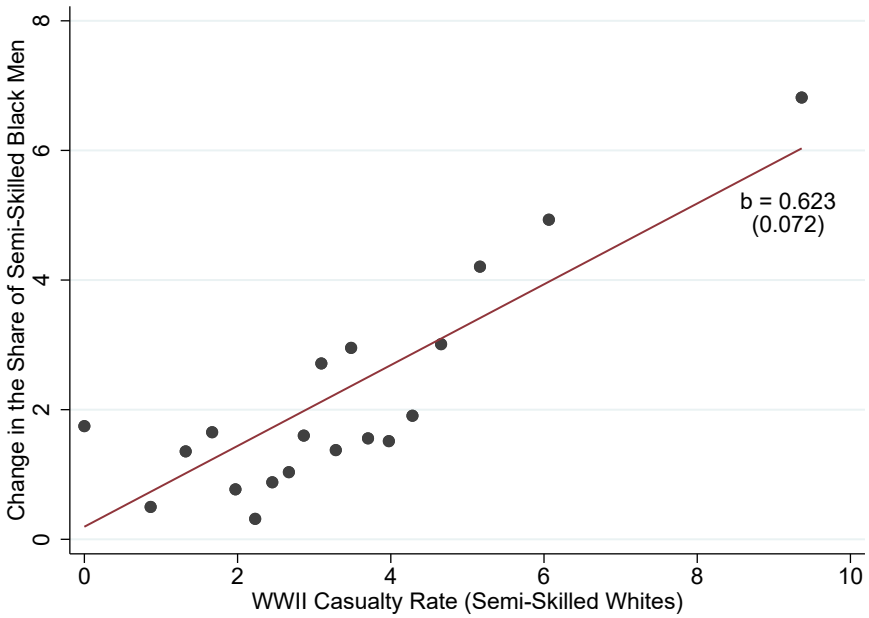
Note: Raw data source from the 1940 Census of Population and Housing for the state of Georgia (p. 278). Occupational information is reported for each skill group by county.

Figure A9: Correlation of the Black Occupational Upgrading and WWII Casualty Rates

(a) Cross-sectional correlation in 1950



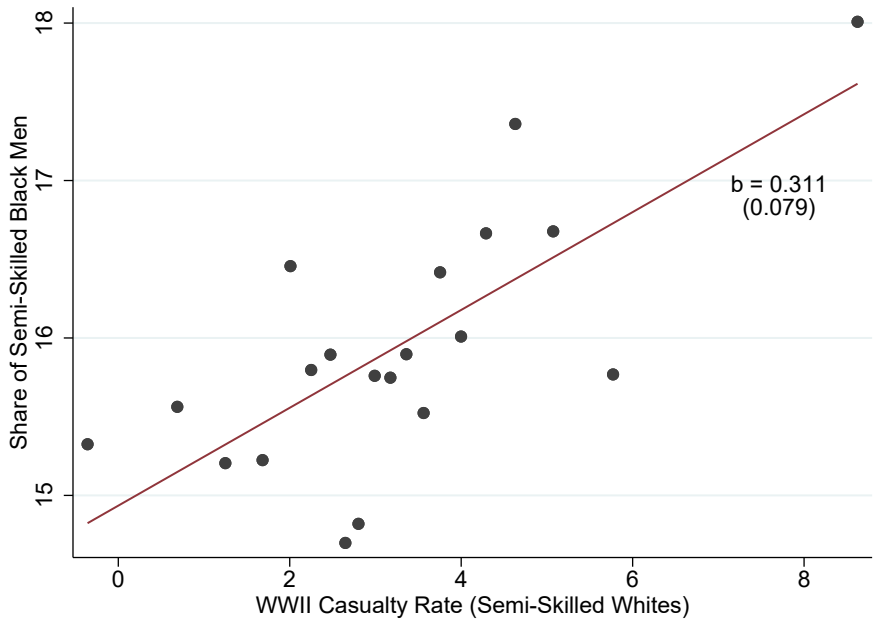
(b) Time correlation from 1940-50



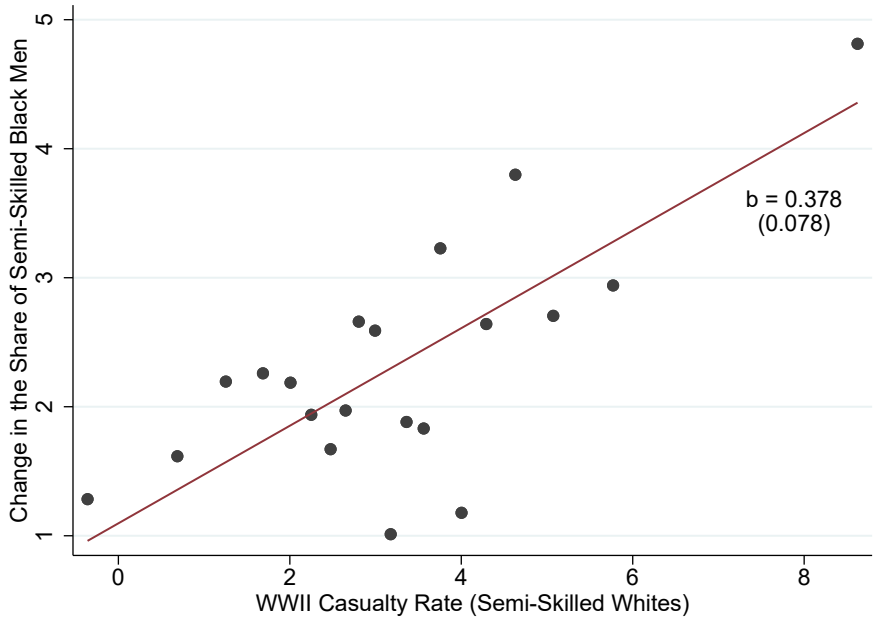
Note: Binned scatter plots showing the relation between the share of semi-skilled Black workers in 1950 (panel a) and the change in the share of semi-skilled Black workers from 1940 to 1950 (panel b) with the WWII casualty rate among semi-skilled white soldiers across Southern counties.

Figure A10: Cross-Sectional and Time Correlation of the Black Occupational Upgrading and WWII Casualty Rates

(a) Cross-sectional correlation in 1950

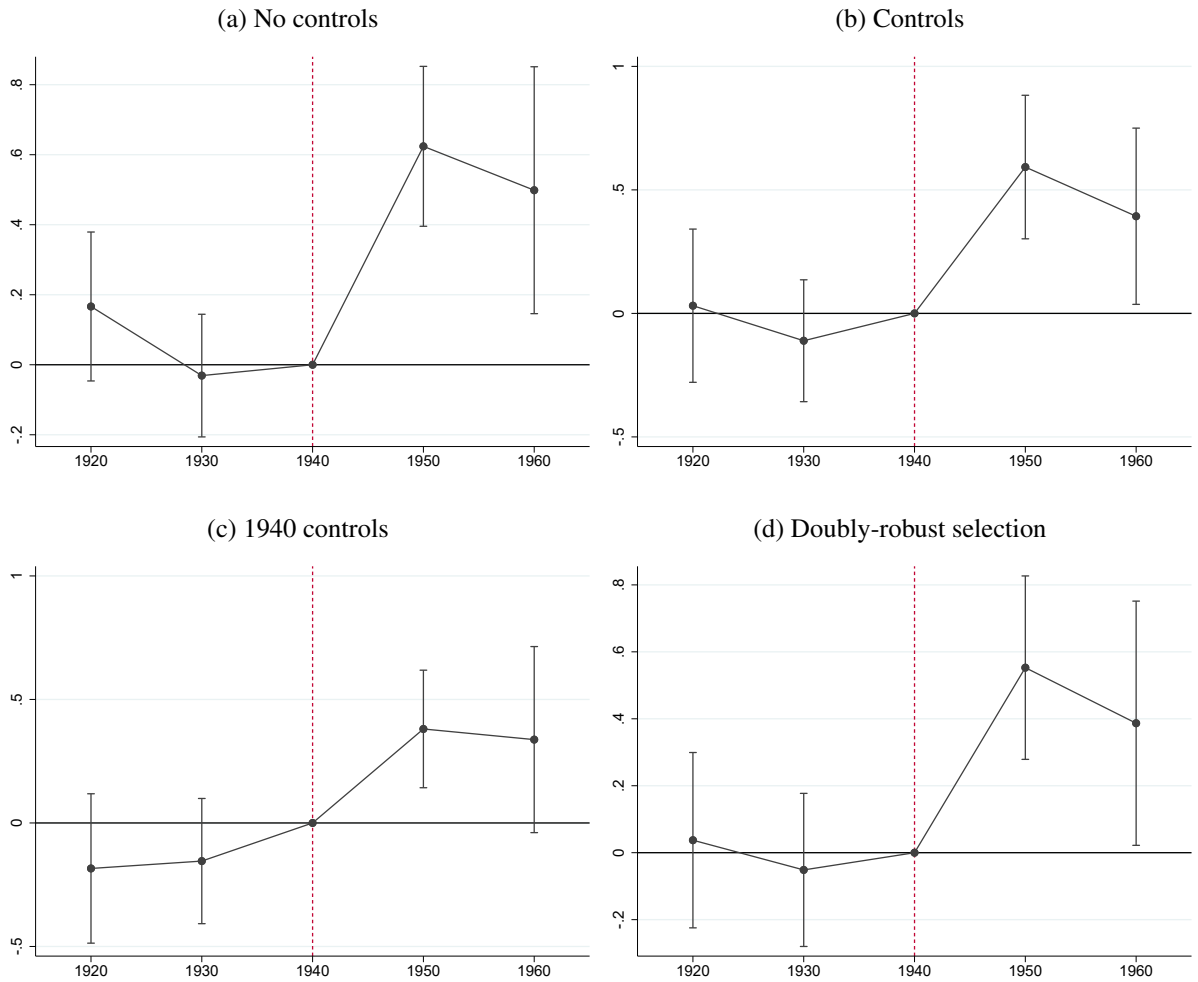


(b) Time correlation from 1940-50



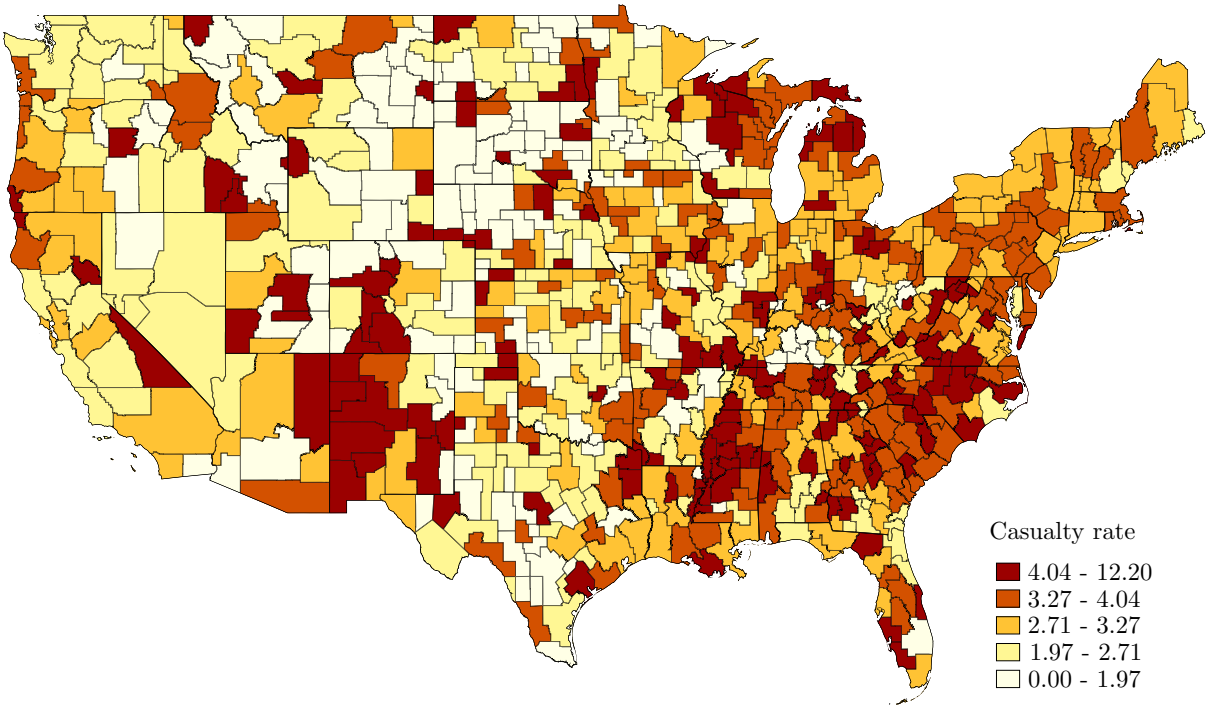
Note: Binned scatter plots showing the relation between the share of semi-skilled Black workers in 1950 (panel a) and the change in the share of semi-skilled Black workers from 1940 to 1950 (panel b) with the WWII casualty rate among semi-skilled white soldiers across Southern counties after partialling out pre-war county-level controls in 1940 which include: the county population, share of Black men, and the shares of agricultural and manufacturing employment.

Figure A11: Difference-in-Differences Coefficient Plots using Alternative Specifications



Note: Difference-in-differences regressions of the county-level share of semi-skilled Black workers on the WWII county casualty rate among semi-skilled white soldiers interacted with decade fixed effects. The omitted baseline decade is 1940 which is marked by the dashed line. This is the last pre-treatment period. The estimation sample contains counties in Southern states from 1920 to 1960. Coefficients show the effect of a one standard deviation increase in the casualty rate on the outcome in terms of percentage points. All regressions include county and decade fixed effects unless stated otherwise. If used by a given specification, controls include the county draft rate, average casualty rate in the neighboring counties, log WWII spending per capita, share of Black men, share of rural population, no. of manufacturing establishments per capita, average manufacturing firm size, log manufacturing value added per worker, share of employment in manufacturing, share of land in agricultural production, share of acres in cotton production, share of cash tenants, average value of machinery per farm, lynchings per 1,000 Black population between 1900 and 1930, no. of Rosenwald schools per 1,000 Black population, share of acres flooded by the Mississippi in 1928, no. of slaves in 1860, Republican vote share, New Deal spending per capita 1933-35 (loans, public works, AAA, FHA loans), and the unemployment rate in 1937. Time-invariant controls are interacted with decade fixed effects. The 1940 controls plot fixes all controls at their level in that year and interacts them with decade fixed effects. The doubly-robust selection method implements the Belloni et al. (2014) machine learning covariate selection algorithm to select the most relevant controls. Monetary values are deflated to 2010 U.S. dollars. Standard errors clustered at the county level. Error bars show 95% confidence intervals around each coefficient estimate.

Figure A12: Spatial Distribution of WWII Casualty Rates among Semi-Skilled Whites



Note: Spatial distribution of WWII casualty rates among semi-skilled white men at the commuting zone level in percent. Shaded polygons display the quintiles of the casualty rate distribution with ranges being shown in the legend on the side.

A Additional Analyses - County Level Data

A1.1: Parallel Trends Assumptions

In addition to the lags and leads of the casualty treatment and their effects on the share of semi-skilled Black workers in Figure 4, Figure A11 provides the same plot under different model specifications. This includes the model without covariates (i.e. the raw data less time and county fixed effects), with controls, with controls fixed at their 1940 values and interacted with time dummies, and controls selected by the Belloni et al. (2014) algorithm. The insignificance of the pre-trends and the post-war treatment effect do not hinge on any particular model specification but are indistinguishable from the coefficients plot presented in the main section.

A1.2: Selection on Observables

Table A8 estimates the DiD model in equation (2) and gradually expands the covariate set. Observing the movement of the coefficient of interest shows that the casualty rate coefficient stabilizes at around 0.5 percentage points There is no one particular control which significantly alters the results after being included. The typical argument is that the treatment effect remains stable with respect to the inclusion of observed factors, it would remain stable also with respect to unobserved factors. However, as discussed in the main section with reference to the test by Oster (2019), this is not necessarily true if, for instance, observables and unobservables are unrelated to each other but separately affect the relationship between treatment and outcome.

A downside of the coefficient stability test is that invariance of the top-row coefficient might be due to measurement error in the controls. Following Pei, Pischke and Schwandt (2019), a more powerful alternative is to take the added control to the left-hand side of the equation and test for imbalances with respect to the treatment variable. This is equivalent to running regressions with and without the added control and comparing both estimates via a SUR regression. This is a generalized Hausman test. The corresponding χ^2 test statistics and p-values are reported in the bottom two rows of Table A8. The test reveals no significant imbalances in the controls which are related to the casualty rate.

A1.3: Selective Migration of Black Workers

Even though the casualty rate need not be random in this estimation framework, a potential threat to identification are time-varying confounding factors or systematic manipulation of individuals' treatment status. With the war period being a major episode of migration for Black Southerners (Boustan, 2016), a plausible issue could arise if Black workers migrated from low- to high-casualty counties to find semi-skilled employment. In this case, the casualty rate effect picks up an additional migratory response.

To test for this possibility, I construct state and county level measures of migration and include them as controls in the main specification. The state-level outmigration information come from the 1940 backward looking migration question.[1] For each Southern state, I compute

[1] I thank an anonymous referee for this suggestion.

the share of Black Southerners who live Outside the South but who reported to have lived in the given Southern state 5 years ago. The county-level migration from 1930 to 1940, and from 1940 to 1950, are constructed as the change in population between Census years minus deaths plus births which were obtained from Bailey, Clay, Fishback, Kantor, Severnini and Wentz (2018). This is total migration and not race-specific but the idea is to also test for migration leading to amplified overall labor shortages that can explain the Black occupational upgrading. The migration measures are standardized to have mean zero and variance one and are interacted with the post-war indicator. Results are reported in Table A9 and show a significant effect of migration on the increase in Black workers in semi-skilled position, however, none of them change the main coefficient of interest.

As an additional test for the selective migration story, I re-estimate equation (2) using the share of Black population and the share of Black men in a given county as dependent variable and interacting the WWII casualty rate among semi-skilled whites with decade fixed effects, leaving the interaction with the 1940 indicator out as baseline. The results for this cross-county migration test are shown in Figure A13. None of the estimated coefficients are significant, neither statistically nor economically. This finding is consistent with the findings in Aizer, Boone, Lleras-Muney and Vogel (2020) and the previous balancing test by Pei et al. (2019) in Table A8 for the share of Black men. The result also suggests that if Black workers gained semi-skilled employment due to the war-induced lack of white workers in this skill-group, then they must have done so in their current counties of residence.

Even if the 1950 interaction in Figure A13 was significantly different from zero, it would imply that the share of Black workers in a given county increased by 0.05 percentage points for a one percentage point increase in the casualty rate. Relative to a pre-war average of 22.4%, such an increase would not be considered an economically significant migratory response. The result for the share of Black men is the same. This is not to say that African Americans were not migrating during this period. They just did not do so differentially across high- and low-casualty rate counties. Appendix B uses data from the micro Census to provide further evidence that the findings here are not driven by migration patterns by Black workers. The findings here are consistent with Aizer et al. (2020) who also do not find a significant impact of migration on Black occupational upgrading that related to labor shortages during the war.

A1.4: Selection of Soldiers

Table A10 reports DiD results of equation (2) including average soldier characteristics by county interacted with a post-war indicator. These characteristics include the average age, years of education, AGCT score (an aptitude test which is the predecessor of the AFQT), share of married, and share of voluntarily enlisted soldiers. This is to preclude the possibility that soldiers from particularly patriotic counties volunteer and die, but that these are also the types of counties where people become more attached to each other and less prejudiced on racial grounds in times of hardship.

The results are unchanged by including these variables. In addition, Figure A7 shows that there are no marked differences in voluntary enlistments between a) the South and the rest of the country and b) above and below median casualty rate counties within the South. While soldiers are certainly selected (e.g. illiterates were service ineligible), the selection into the military and into death does not appear to affect the relationship between the WWII casualty rates among semi-skilled white soldiers and the share of Black workers in this skill group.

A1.5: Alternative Treatment Denominators and Denominator Bias

In this section I consider an alternative definition of the treatment variable as compared to equation (1) which used the number of semi-skilled white soldiers as denominator. The rational was to account for unobservable draft deferments. Results using as denominator all semi-skilled white workers,

$$\text{Casualty rate}_c = \frac{\text{Number of fallen semi-skilled white soldiers}_c}{\text{Number of semi-skilled white workers}_c} \times 100 \tag{1}$$

are reported in Table A11. This casualty variable has a mean of 0.55, standard deviation of 1.39, minimum of zero, and maximum of 25.54. In all specifications the casualty rate effect is positive and significant at the 1% level. Compared to the baseline specification the coefficients are larger and slightly more volatile with respect to their magnitude when county-specific linear time trends are included. The corresponding coefficients plot for the lags and leads of this treatment variable is shown in Figure A14.

Another concern is that there might be a spurious relationship between the share of semi-skilled Black workers and the casualty rate among semi-skilled white soldiers due to a correlation between the denominators which is driving the estimated change. To account for this, I fix the outcome denominator in equation (1) at its pre-war level in 1940. This will result in shares that are not necessarily bound in the $[0, 1]$ interval but are indicative for whether results are sensitive with respect to changes in the denominator. Table A12 reports the estimation results. All but the last column show a positive effect which is significant at the 5% level or less.

A1.6: Sensitivity of Results by State

To test whether results are driven by any given state, I re-estimate the DiD specification in equation 2 using the sample with counties from the $S-1$ states. The results from this jackknife-type leave-one-out procedure are shown in Figure A15. The figure plots the estimated WWII casualty rate DiD coefficient for each iteration with the left-out state in a given regression being displayed on the vertical axis. The resulting coefficients are indistinguishable from each other as well as from the main result in Table 2.

A1.7: Spatial Clustering of Casualty Rates

U.S. military units were raised locally during WWII, a practice that was abandoned after D-Day. This policy as well as the patterns observed in the map in Figure 2 may hint towards spatial dependencies in the outcome. Such spatial correlation would pose problems for inference whereby standard errors are underestimated. To test for such spatial autocorrelation, I compute the I statistic by Moran (1950) for global spatial correlation and the Getis-Ord $G_i^*(d)$ statistic (Getis and Ord, 1992) to test for local spatial correlation. Moran's I is computed as

$$I = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} C_i C_j}{\sum_{i=1}^{n} C_i^2} \tag{2}$$

where i indexes counties with a total number of n counties, j indexes all other counties with $i \neq j$, C is the WWII casualty rate among semi-skilled whites, and w is a spatial weight matrix. Like the standard correlation coefficient, Moran's I lies in $[-1, 1]$. The z score for the corresponding test statistic is given by:

$$z(I) = \frac{I - E(I)}{\sqrt{Var(I)}}$$

Results from this test are reported in Table A3 for distance thresholds of 200, 400, and 600km. Columns (1) to (3) show the casualty rate has a small but statistically significant positive spatial autocorrelation at the 1% level across counties. Moran's I ranges between 0.049 and 0.078. However, once the casualty rate is demeaned by its state-specific averages, Moran's I drops to between -0.003 and -0.008 and becomes insignificant except for the 400km distance threshold where it is marginally significant at the 10% level. This implies that once state fixed effects are controlled for, the casualty rate measure is as good as randomly assigned across geographic space. In the main DiD specifications, these fixed effects would be absorbed by the county fixed effects.

Spatial correlation, however, may exist at a more concentrated level. To test for more local correlations, I provide estimates of the Getis-Ord $G_i^*(d)$ statistic:

$$G_i^*(d) = \frac{\sum_{j=1}^{n} w_{ij}(d) C_j}{\sum_{i=1}^{n} C_j} \tag{3}$$

where the notation is as before except that now the spatial weight matrix depends on a certain radius d within which the statistic is computed.[2] Clusters of counties with significantly higher casualty rates are referred to as hot spots. Conversely, those with significantly lower casualty rates are called cold spots.

[2]For both Moran's I and the Getis-Ord $G_i^*(d)$ binary spatial weights matrices were used. Changing these to exponential or power function type spatial weight matrices does not alter the results. Additional results with alternative spatial weight matrices are not reported here but are available on request. The Stata routine `getisord` by Kondo (2016) was used to compute this test.

Table A4 reports the results from the Getis-Ord test for the same 200, 400, and 600km distance bands as before. The table reports the number of counties within a given z-score interval. Casualty rates show local spatial independence if the z-score of $G_i^*(d)$ falls within -1.96 and 1.96. Lower z-scores than the lower bound of -1.96 indicate cold spots while higher values than 1.96 indicate hot spots. Again, columns (1) to (3) indicate local spatial correlation with a significant number of counties displaying cold spots (365 counties) and 409 counties having hot spots, out of a total of 1,387 counties. Once state fixed effects are partialled out, almost all counties lose this local spatial autocorrelation as is shown in columns (4) to (6).

Although spatial correlation appears to be accounted for by geographic fixed effects, I replicate the main findings in Table 2 and compute Conley (1999) standard errors to correct for spatial dependence.[3] Table A13 reports the results and shows that the significance of previous results is not driven by spatial autocorrelation.

A1.8: Weighting by Population Size

Even though the main regression results control for total population size, either as a time-varying right-hand side variable or fixed at its 1940 level interacted with time fixed effects, one worry might be that small counties receive too high weights in the estimation. Very small and large counties are essentially treated the same up to the extend to which controlling for population size and share variables, which was to original motivation for not working with the levels, can account for such size differences. To probe for this more formally, I repeat the main regression exercise with population size weights based on the 1940 county population. The results are reported in Table A14 which show that the results are almost identical to the main specification in terms of sign and significance. The only exception is the specification that includes the linear county-level time trends where the coefficient reduces, accompanied by a much larger increase in the standard error.

[3]Thiemo Fetzer's `reg2hdfespatial` Stata routine was used to run these regressions.

Table A8: Sensitivity Analysis Using Observable County Characteristics

Outcome: % semi-skilled Black workers	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Casualty rate × Post-war	0.486***	0.488***	0.498***	0.496***	0.518***	0.506***	0.508***	0.508***	0.508***	0.506***	0.506***
	(0.132)	(0.132)	(0.131)	(0.121)	(0.134)	(0.134)	(0.135)	(0.135)	(0.136)	(0.135)	(0.135)
Draft rate	-0.130**	-0.115**	-0.117**	-0.135**	-0.125**	-0.117**	-0.118**	-0.118**	-0.119**	-0.114**	-0.114**
	(0.052)	(0.052)	(0.055)	(0.053)	(0.050)	(0.049)	(0.049)	(0.049)	(0.051)	(0.051)	(0.051)
Log mil spending		-0.296***	-0.293***	-0.133*	-0.163**	-0.161**	-0.158**	-0.169**	-0.164**	-0.168**	-0.168**
		(0.073)	(0.072)	(0.068)	(0.071)	(0.070)	(0.070)	(0.071)	(0.070)	(0.071)	(0.071)
Casualty rate (neighbors)			0.917***	1.068***	1.079***	1.044***	1.046***	1.047***	1.046***	1.049***	1.049***
			(0.228)	(0.219)	(0.226)	(0.223)	(0.223)	(0.224)	(0.224)	(0.224)	(0.224)
% Black men				0.469***	0.454***	0.478***	0.475***	0.473***	0.474***	0.477***	0.477***
				(0.048)	(0.048)	(0.048)	(0.048)	(0.048)	(0.049)	(0.049)	(0.049)
Mfg firm size					-0.006	-0.006*	-0.006*	-0.007*	-0.006*	-0.006*	-0.006*
					(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
% cotton						-0.120***	-0.116***	-0.116***	-0.116***	-0.118***	-0.118***
						(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
% cash tenants							0.029	0.030	0.030	0.027	0.027
							(0.036)	(0.036)	(0.037)	(0.037)	(0.037)
Rosenwald schools								-0.185	-0.182	-0.166	-0.166
								(0.215)	(0.215)	(0.217)	(0.217)
New Deal relief									0.001	-0.003	-0.003
									(0.003)	(0.005)	(0.005)
1937 unempl rate										0.081**	0.081**
										(0.031)	(0.031)
Historical lynchings											-0.000
											(0.005)
Observations	6,443	6,443	6,430	6,430	5,905	5,905	5,905	5,864	5,842	5,842	5,842
Counties	1,388	1,388	1,388	1,388	1,387	1,387	1,387	1,379	1,379	1,379	1,379
Adj. R^2	0.876	0.876	0.877	0.886	0.884	0.885	0.885	0.885	0.884	0.884	0.884
Balancing Test	0.042	0.815	1.090	0.019	0.260	0.000	0.098	0.076	0.628	2.322	0.004
Balancing Test p-val	0.837	0.367	0.297	0.890	0.610	0.990	0.755	0.783	0.428	0.128	0.950

Note: Difference-in-differences regressions of the county-level share of semi-skilled Black workers on the WWII county casualty rate among semi-skilled white soldiers interacted with a post-war indicator. The estimation sample uses decennial U.S. Census data on counties in Southern states from 1920 to 1960. All regressions include county and decade fixed effects. The covariate balancing test by Pei et al. (2019) is reported in the bottom two rows of the table where the null hypothesis is that a new added control does not vary systematically across high- and low-casualty rate counties. The variables on WWII military spending, WWII casualties in neighboring counties, New Deal Relief per capita, and the unemployment rate in 1937 are interacted with a post-war indicator. Standard errors clustered at the county level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A9: Sensitivity to Outmigration

	Outcome: % semi-skilled Black workers (pre-war mean = 12.433)					
	(1)	(2)	(3)	(4)	(5)	(6)
Casualty rate × Post-war	0.556***	0.528***	0.564***	0.569***	0.549***	0.554***
	(0.146)	(0.147)	(0.147)	(0.146)	(0.147)	(0.147)
Outmigration 1930-40	0.807***			0.565**		0.339
	(0.232)			(0.257)		(0.264)
Outmigration 1940-50		0.963***			0.883***	0.789***
		(0.290)			(0.294)	(0.304)
State lvl outmigration 1935-40			0.844***	0.677**	0.774***	0.682**
			(0.258)	(0.281)	(0.261)	(0.278)
Observations	4,903	4,903	4,903	4,903	4,903	4,903
Counties	1,317	1,317	1,317	1,317	1,317	1,317
Adj. R^2	0.886	0.886	0.886	0.887	0.887	0.887

Note: Difference-in-differences regressions of the county-level share of semi-skilled Black workers on the WWII county casualty rate among semi-skilled white soldiers interacted with a post-war indicator. The estimation sample uses decennial U.S. Census data on counties in Southern states from 1920 to 1960. The regressions control for different outmigration measures interacted with a post-war indicator to test whether past and contemporaneous outmigration can explain away the casualty effect by capturing labor shortages during the war. County-level outmigration is measured as the change in population between two Census years plus births minus deaths over that period which uses the historic mortality and natality data by Bailey et al. (2018). The state-level outflow is measured using the 1940 Census backward-looking migration question and counts the number of Black Americans who live outside the South in 1940 and who report a Southern state of residence five years ago relative to the state's population. All migration measures are standardized to have mean zero and variance one. Controls include county and decade fixed effects, the county draft rate, average casualty rate in the neighboring counties, log WWII spending per capita, share of Black men, share of rural population, no. of manufacturing establishments per capita, average manufacturing firm size, log manufacturing value added per worker, share of employment in manufacturing, share of land in agricultural production, share of acres in cotton production, share of cash tenants, average value of machinery per farm, lynchings per 1,000 Black population between 1900 and 1930, no. of Rosenwald schools per 1,000 Black population, share of acres flooded by the Mississippi in 1928, no. of slaves in 1860, Republican vote share, New Deal spending per capita 1933-35 (loans, public works, AAA, FHA loans), and the unemployment rate in 1937. Time-invariant controls are interacted with decade fixed effects. Monetary values are deflated to 2010 U.S. dollars. Standard errors clustered at the county level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A10: Difference-in-Differences Results with Average Soldier Characteristics

	Outcome: % semi-skilled Black workers (pre-war mean = 12.433)					
	(1)	(2)	(3)	(4)	(5)	(6)
Casualty rate × Post-war	0.539***	0.533***	0.493***	0.562***	0.741***	0.529***
	(0.130)	(0.168)	(0.154)	(0.171)	(0.235)	(0.129)
Controls		Yes		Yes	Yes	Yes
1940 controls × decade			Yes			
State × decade FE				Yes		
Linear county time trends					Yes	
Lasso selection						Yes
Observations	6,443	4,903	4,723	4,903	4,903	6,317
Counties	1,388	1,317	994	1,317	1,317	1,384
Adj. R^2	0.864	0.874	0.880	0.881	0.921	0.864
Oster's δ	1.877	1.484	1.436	1.804	1.117	1.812

Note: Difference-in-differences regressions of the county-level share of semi-skilled Black workers on the WWII county casualty rate among semi-skilled white soldiers interacted with a post-war indicator. The estimation sample uses decennial U.S. Census data on counties in Southern states from 1920 to 1960. Controls include county and decade fixed effects, the county draft rate, average casualty rate in the neighboring counties, log WWII spending per capita, share of Black men, share of rural population, no. of manufacturing establishments per capita, average manufacturing firm size, log manufacturing value added per worker, share of employment in manufacturing, share of land in agricultural production, share of acres in cotton production, share of cash tenants, average value of machinery per farm, lynchings per 1,000 Black population between 1900 and 1930, no. of Rosenwald schools per 1,000 Black population, share of acres flooded by the Mississippi in 1928, no. of slaves in 1860, Republican vote share, New Deal spending per capita 1933-35 (loans, public works, AAA, FHA loans), and the unemployment rate in 1937, as well as the average soldier characteristics in each county including age, education, AGCT score, share of married, and share of voluntarily enlisted. Time-invariant controls are interacted with decade fixed effects. Monetary values are deflated to 2010 U.S. dollars. The doubly-robust selection method implements the Belloni et al. (2014) machine learning covariate selection algorithm for testing the stability of treatment effects with respect to the observables. The test for selection on unobservables by Oster (2019) is reported in the final row by computing the coefficient of proportionality δ for which the coefficient on the semi-skilled casualty rate among white soldiers would equal zero. Standard errors clustered at the county level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A11: Difference-in-Differences Results with Alternative Treatment Denominator

	Outcome: % semi-skilled Black workers (pre-war mean = 12.433)					
	(1)	(2)	(3)	(4)	(5)	(6)
Casualty rate × Post-war	1.286***	2.029***	1.842***	2.127***	2.712***	2.035***
	(0.328)	(0.553)	(0.298)	(0.546)	(0.605)	(0.508)
Controls		Yes		Yes	Yes	Yes
1940 controls × decade			Yes			
State × decade FE				Yes		
Linear county time trends					Yes	
Lasso selection						Yes
Observations	6,443	4,903	4,723	4,903	4,903	5,405
Counties	1,388	1,317	994	1,317	1,317	1,342
Adj. R^2	0.867	0.878	0.882	0.885	0.922	0.870
Oster's δ	3.265	2.286	1.451	2.260	1.297	1.923

Note: Difference-in-differences regressions of the county-level share of semi-skilled Black workers on the WWII county casualty rate among semi-skilled white soldiers interacted with a post-war indicator. The casualty rate in county c here is one hundred times the total number of killed semi-skilled white soldiers over the number of total semi-skilled white workers in 1940. The estimation sample uses decennial U.S. Census data on counties in Southern states from 1920 to 1960. Coefficients are expressed in terms of a one standard deviation increase in the casualty rate. Controls include county and decade fixed effects, the county draft rate, average casualty rate in the neighboring counties, log WWII spending per capita, share of Black men, share of rural population, no. of manufacturing establishments per capita, average manufacturing firm size, log manufacturing value added per worker, share of employment in manufacturing, share of land in agricultural production, share of acres in cotton production, share of cash tenants, average value of machinery per farm, lynchings per 1,000 Black population between 1900 and 1930, no. of Rosenwald schools per 1,000 Black population, share of acres flooded by the Mississippi in 1928, no. of slaves in 1860, Republican vote share, New Deal spending per capita 1933-35 (loans, public works, AAA, FHA loans), and the unemployment rate in 1937. Time-invariant controls are interacted with decade fixed effects. Monetary values are deflated to 2010 U.S. dollars. The doubly-robust selection method implements the Belloni et al. (2014) machine learning covariate selection algorithm for testing the stability of treatment effects with respect to the observables. The test for selection on unobservables by Oster (2019) is reported in the final row by computing the coefficient of proportionality δ for which the coefficient on the semi-skilled casualty rate among white soldiers would equal zero. Standard errors clustered at the county level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A12: Difference-in-Differences Results with Fixed Outcome Denominator

	Outcome: % semi-skilled Black workers (pre-war mean = 12.433)					
	(1)	(2)	(3)	(4)	(5)	(6)
Casualty rate × Post-war	0.704**	1.074***	0.820***	0.880**	1.094*	0.784**
	(0.286)	(0.373)	(0.317)	(0.369)	(0.594)	(0.356)
Controls		Yes		Yes	Yes	Yes
1940 controls × decade			Yes			
State × decade				Yes		
Linear county time trends					Yes	
Lasso selection						Yes
Observations	6,919	5,158	4,956	5,158	5,158	5,158
Counties	1,388	1,331	994	1,331	1,331	1,331
Adj. R^2	0.557	0.617	0.702	0.650	0.746	0.667
Oster's δ	0.198	0.334	0.349	0.391	0.205	0.281

Note: Difference-in-differences regressions of the county-level share of semi-skilled Black workers on the WWII county casualty rate among semi-skilled white soldiers interacted with a post-war indicator. The casualty rate in county c here is one hundred times the total number of killed semi-skilled whites over the number of total semi-skilled whites in 1940. The estimation sample uses decennial U.S. Census data on counties in Southern states from 1920 to 1960. The denominator of the outcome (number of semi-skilled workers) is fixed at 1940 values to reduce denominator bias. Controls include county and decade fixed effects, the county draft rate, average casualty rate in the neighboring counties, log WWII spending per capita, share of Black men, share of rural population, no. of manufacturing establishments per capita, average manufacturing firm size, log manufacturing value added per worker, share of employment in manufacturing, share of land in agricultural production, share of acres in cotton production, share of cash tenants, average value of machinery per farm, lynchings per 1,000 Black population between 1900 and 1930, no. of Rosenwald schools per 1,000 Black population, share of acres flooded by the Mississippi in 1928, no. of slaves in 1860, Republican vote share, New Deal spending per capita 1933-35 (loans, public works, AAA, FHA loans), and the unemployment rate in 1937. Time-invariant controls are interacted with decade fixed effects. Monetary values are deflated to 2010 U.S. dollars. The doubly-robust selection method implements the Belloni et al. (2014) machine learning covariate selection algorithm for testing the stability of treatment effects with respect to the observables. The test for selection on unobservables by Oster (2019) is reported in the final row by computing the coefficient of proportionality δ for which the coefficient on the semi-skilled casualty rate among white soldiers would equal zero. Standard errors clustered at the county level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A13: County Level Difference-in-Differences Results with Conley Standard Errors

	Outcome: % semi-skilled Black workers (pre-war mean = 12.433)					
	(1)	(2)	(3)	(4)	(5)	(6)
Casualty rate × Post-war	0.539***	0.543***	0.493***	0.555***	0.718***	0.456***
	(0.091)	(0.087)	(0.094)	(0.090)	(0.101)	(0.108)
Conley s.e. (200km)	0.084	0.085	0.092	0.088	0.108	0.093
Conley s.e. (400km)	0.090	0.084	0.092	0.087	0.112	0.102
Conley s.e. (600km)	0.091	0.087	0.094	0.090	0.101	0.108
Controls		Yes		Yes	Yes	Yes
1940 controls × time			Yes			
State × year FE				Yes		
Linear county time trends					Yes	
Doubly-robust selection						Yes
Observations	6,438	4,900	4,723	4,900	4,903	5,861

Note: Difference-in-differences regressions of the county-level share of semi-skilled Black workers on the WWII county casualty rate among semi-skilled white soldiers interacted with a post-war indicator. The estimation sample uses decennial U.S. Census data on counties in Southern states from 1920 to 1960. Controls include county and decade fixed effects, the county draft rate, average casualty rate in the neighboring counties, log WWII spending per capita, share of Black men, share of rural population, no. of manufacturing establishments per capita, average manufacturing firm size, log manufacturing value added per worker, share of employment in manufacturing, share of land in agricultural production, share of acres in cotton production, share of cash tenants, average value of machinery per farm, lynchings per 1,000 Black population between 1900 and 1930, no. of Rosenwald schools per 1,000 Black, share of acres flooded by the Mississippi in 1928, no. of slaves in 1860, Republican vote share, New Deal spending per capita 1933-35 (loans, public works, AAA, FHA loans), and the unemployment rate in 1937. Time-invariant controls are interacted with decade fixed effects. Monetary values are deflated to 2010 U.S. dollars. The doubly-robust selection method implements the Belloni et al. (2014) machine learning covariate selection algorithm for testing the stability of treatment effects with respect to the observables. Standard errors adjusted for spatial correlation using Conley (1999) standard errors with a distance threshold of 200, 400, and 600km.

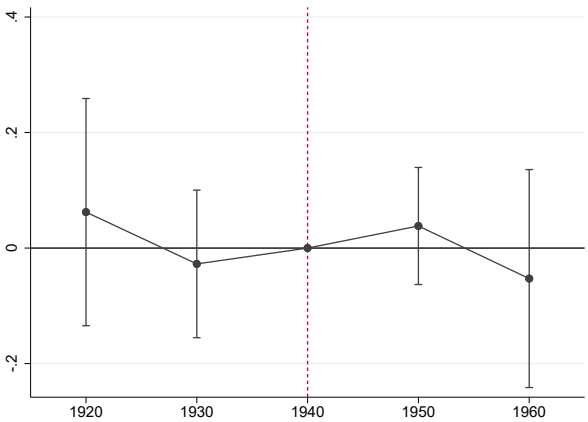
Table A14: County Level Difference-in-Differences Results Population Weighted

	Outcome: % semi-skilled Black workers (pre-war mean = 12.433)					
	(1)	(2)	(3)	(4)	(5)	(6)
Casualty rate × Post-war	0.523***	0.543***	0.488***	0.611***	0.828***	0.463***
	(0.165)	(0.161)	(0.187)	(0.178)	(0.253)	(0.175)
Controls		Yes		Yes	Yes	Yes
1940 controls × decade			Yes			
State × decade FE				Yes		
Linear county time trends					Yes	
Lasso selection						Yes
Observations	6,443	4,903	4,723	4,903	4,903	5,421
Counties	1,388	1,317	994	1,317	1,317	1,371
Adj. R^2	0.857	0.878	0.878	0.885	0.920	0.861
Oster's δ	1.208	1.116	1.057	1.324	0.903	0.933

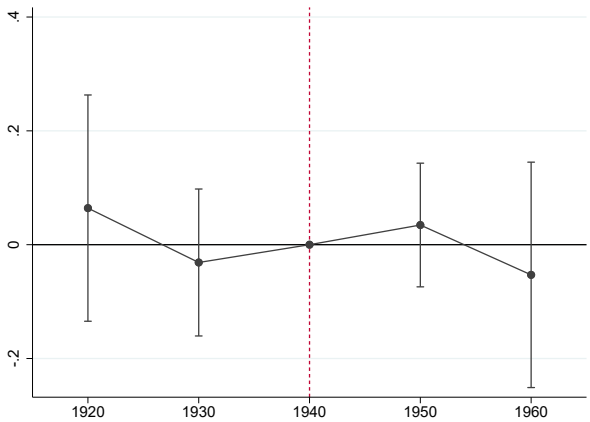
Note: Difference-in-differences regressions of the county-level share of semi-skilled Black workers on the WWII county casualty rate among semi-skilled white soldiers interacted with a post-war indicator. Regressions are weighted by 1940 county population size. The estimation sample uses decennial U.S. Census data on counties in Southern states from 1920 to 1960. Controls include county and decade fixed effects, the county draft rate, average casualty rate in the neighboring counties, log WWII spending per capita, share of Black men, share of rural population, no. of manufacturing establishments per capita, average manufacturing firm size, log manufacturing value added per worker, share of employment in manufacturing, share of land in agricultural production, share of acres in cotton production, share of cash tenants, average value of machinery per farm, lynchings per 1,000 Black population between 1900 and 1930, no. of Rosenwald schools per 1,000 Black population, share of acres flooded by the Mississippi in 1928, no. of slaves in 1860, Republican vote share, New Deal spending per capita 1933-35 (loans, public works, AAA, FHA loans), and the unemployment rate in 1937. Time-invariant controls are interacted with decade fixed effects. Monetary values are deflated to 2010 U.S. dollars. The doubly-robust selection method implements the Belloni et al. (2014) machine learning covariate selection algorithm for testing the stability of treatment effects with respect to the observables. The test for selection on unobservables by Oster (2019) is reported in the final row by computing the coefficient of proportionality δ for which the coefficient on the semi-skilled casualty rate among white soldiers would equal zero. Standard errors clustered at the county level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A13: Difference-in-Differences Cross-County Migration Test

(a) Share of Black population

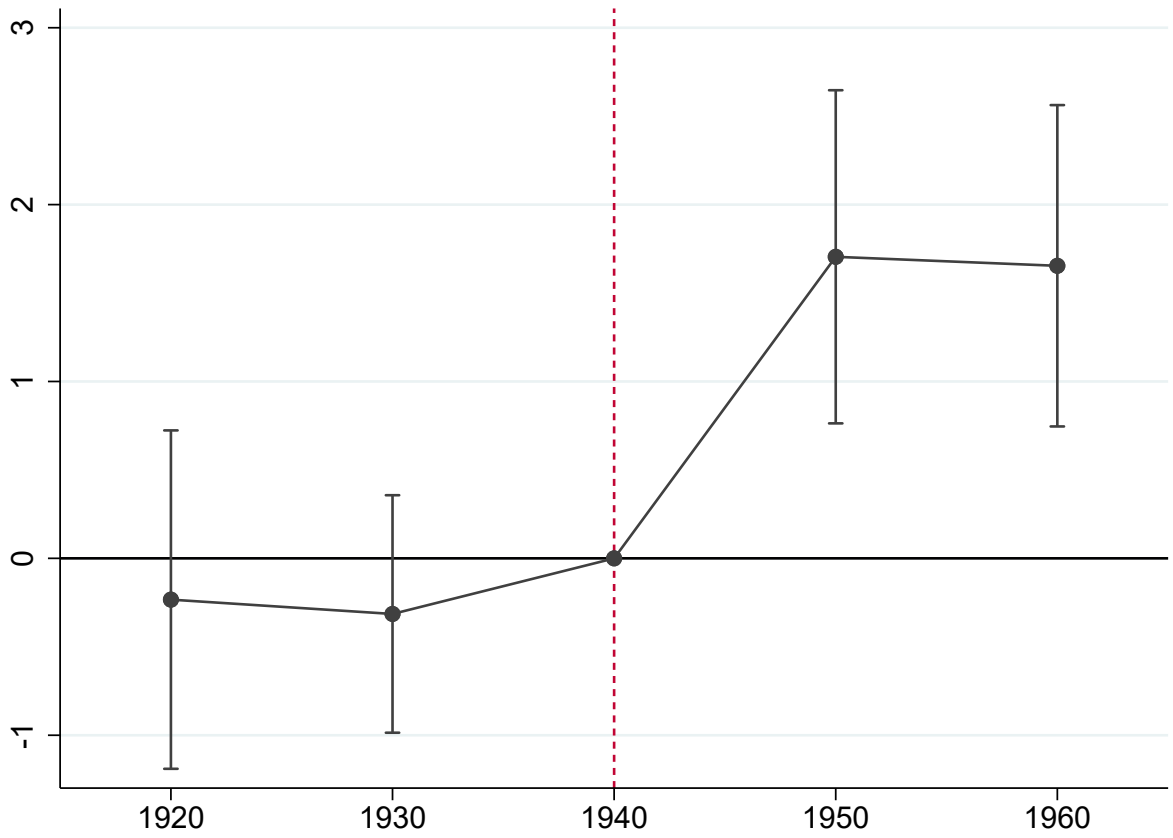


(b) Share of Black men



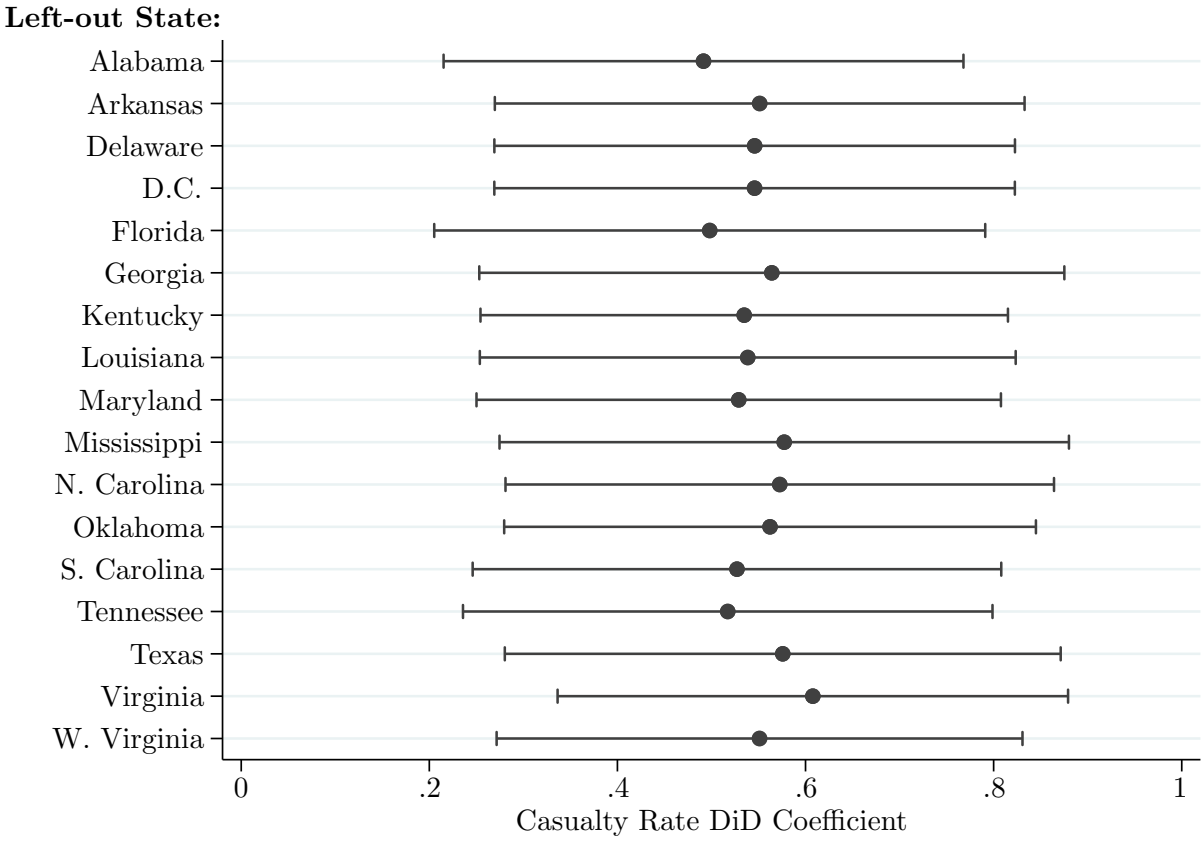
Note: Difference-in-differences regressions of the county-level share of Black population and the share of Black men in percent on the WWII county casualty rate among semi-skilled white soldiers interacted with decade fixed effects. The omitted baseline decade is 1940 which is marked by the dashed line. This is the last pre-treatment period. The estimation sample contains decennial U.S. Census data on counties in Southern states from 1920 to 1960. Controls include county fixed effects, the county draft rate, average casualty rate in the neighboring counties, log WWII spending per capita, share of rural population, no. of manufacturing establishments per capita, average manufacturing firm size, log manufacturing value added per worker, share of employment in manufacturing, share of land in agricultural production, share of acres in cotton production, share of cash tenants, average value of machinery per farm, lynchings per 1,000 Black population between 1900 and 1930, no. of Rosenwald schools per 1,000 Black population, share of acres flooded by the Mississippi in 1928, no. of slaves in 1860, Republican vote share, New Deal spending per capita 1933-35 (loans, public works, AAA, FHA loans), and the unemployment rate in 1937. Time-invariant controls are interacted with decade fixed effects. Monetary values are deflated to 2010 U.S. dollars. Standard errors clustered at the county level. Error bars show 95% confidence intervals around each coefficient estimate.

Figure A14: Difference-in-Differences Coefficient Plot with Alternative Treatment



Note: Difference-in-differences regressions of the county-level share of semi-skilled Black workers on the WWII county casualty rate among semi-skilled white soldiers interacted with decade fixed effects. The denominator in the computation of the casualty rate here is the number of all semi-skilled whites in 1940 in county c. The omitted baseline decade is 1940 which is marked by the dashed line. This is the last pre-treatment period. The estimation sample contains counties in Southern states from 1920 to 1960. Coefficients show the effect of a one standard deviation increase in the casualty rate on the outcome in terms of percentage points. Controls include county and decade fixed effects, the county draft rate, average casualty rate in the neighboring counties, log WWII spending per capita, share of Black men, share of rural population, no. of manufacturing establishments per capita, average manufacturing firm size, log manufacturing value added per worker, share of employment in manufacturing, share of land in agricultural production, share of acres in cotton production, share of cash tenants, average value of machinery per farm, lynchings per 1,000 Black population between 1900 and 1930, no. of Rosenwald schools per 1,000 Black population, share of acres flooded by the Mississippi in 1928, no. of slaves in 1860, Republican vote share, New Deal spending per capita 1933-35 (loans, public works, AAA, FHA loans), and the unemployment rate in 1937. Time-invariant controls are interacted with decade fixed effects. Monetary values are deflated to 2010 U.S. dollars. Standard errors clustered at the county level. Error bars show 95% confidence intervals around each coefficient estimate.

Figure A15: Leave-One Out DiD Sensitivity Check



Note: Difference-in-differences regressions of the county-level share of semi-skilled Black workers on the WWII county casualty rate among semi-skilled white soldiers interacted with a post-war indicator. The estimation sample uses decennial U.S. Census data on counties in Southern states from 1920 to 1960. Each regression leaves out all counties from a specific state at a time to assess whether results are driven by any one single state. The omitted state is listed on the left. Each regression includes county and decade fixed effects, the county draft rate, average casualty rate in the neighboring counties, log WWII spending per capita, share of Black men, share of rural population, no. of manufacturing establishments per capita, average manufacturing firm size, log manufacturing value added per worker, share of employment in manufacturing, share of land in agricultural production, share of acres in cotton production, share of cash tenants, average value of machinery per farm, lynchings per 1,000 Black population between 1900 and 1930, no. of Rosenwald schools per 1,000 Black population, share of acres flooded by the Mississippi in 1928, no. of slaves in 1860, Republican vote share, New Deal spending per capita 1933-35 (loans, public works, AAA, FHA loans), and the unemployment rate in 1937. Time-invariant controls are interacted with decade fixed effects. Monetary values are deflated to 2010 U.S. dollars. Standard errors are clustered by county. Error bars show 95% confidence intervals.

B Additional Analyses - Individual Level Data

B1.1: Effect on Women

Given the substantial literature on the effect of the World War II draft on female labor supply (Goldin, 1991; Acemoglu, Autor and Lyle, 2004; Goldin and Olivetti, 2013; Jaworski, 2014), I estimate the triple differences setting from equation (7) again for the sample of women using the same Census years between 1920 to 1960. This is to test whether war casualties among semi-skilled white men also had an effect on women's labor market outcomes of both races. The added complication is that, unlike for men, there is also an extensive margin response (whether to work or not) in addition to potential intensive margin effects (the type of work).

Table A15 reports the results from this exercise. Column 1 shows that the WWII casualty rate among semi-skilled white men did not have any substantial labor force participation (LFP) impacts on white women. However, Black women were significantly more like to work with a 1.7 percentage points increase in their LFP probability for each percentage points increase in the casualty rate. Black women were also more likely to work in semi-skilled jobs, especially in manufacturing, and less likely to work in in the service sector. Interestingly, the pattern is reversed for white women. In terms of wages, Black women saw significant gains in terms of their annual wages with approximately over 9% higher wages in response to the casualty shock. For white women no such effect exists which may be driven by the fact that Black women entered the better paying manufacturing jobs with a higher probability. These gains appear to be temporary though. Column 6 shows that Black women were again less likely to work in 1960. The coefficient negates the increase in labor force participation in this group that was estimated in column 1. This is consistent with previous work by Acemoglu et al. (2004) who show that once the war ended, most women who gained jobs during the war years due to the draft tended to lose them again with the exception of highly education women.

Table A15: Micro Census Triple Differences Results for Women, 1920-1960

	LFP	Semi-skilled	Services	Mfg.	ln(wage)	LFP 1960
	(1)	(2)	(3)	(4)	(5)	(6)
Casualty rate × Post-war	-0.004**	-0.006**	0.016***	-0.002	-0.011	0.004**
	(0.002)	(0.003)	(0.004)	(0.003)	(0.008)	(0.002)
Casualty rate × Black × Post-war	0.018***	0.027***	-0.010	0.013***	0.036***	-0.018***
	(0.002)	(0.003)	(0.007)	(0.004)	(0.008)	(0.002)
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
CZ Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,146,392	2,146,392	2,146,392	2,146,392	1,100,238	2,146,392
Adj. R^2	0.271	0.058	0.090	0.089	0.121	0.271

Note: Difference-in-differenece-in-differences regression of a semi-skilled indicator on the commuting zone WWII casualty rate among semi-skilled whites interacted with a post-WWII dummy, and with a Black indicator for women living in 722 commuting zones in the whole U.S. The estimation sample contains data from the decennial U.S. micro Census from 1920-60 on non-institutionalized, working Black and white males aged 15-65 who are not currently attending school. All regressions include commuting zone and Census year fixed effects. Individual level controls include age, marital status, age and place of birth dummies. Commuting zone level controls are the WWII draft rate, log WWII spending per capita, share of Black men, share of rural population, no. of manufacturing establishments per capita, average manufacturing firm size, log manufacturing value added per worker, share of employment in manufacturing, share of land in agricultural production, share of acres in cotton production, share of cash tenants, average value of machinery per farm, lynchings per 1,000 Black population between 1900 and 1930, no. of Rosenwald schools per 1,000 Black population, share of acres flooded by the Mississippi in 1928, no. of slaves in 1860, Republican vote share, New Deal spending per capita 1933-35 (loans, public works, AAA, FHA loans), and the unemployment rate in 1937. Time-invariant controls are interacted with decade fixed effects. Monetary values are deflated to 2010 U.S. dollars. Observations are weighted by their individual sample line weight times the spatial weight used to construct the commuting zone level treatment variable. Standard errors clustered at the commuting zone level in parentheses. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B1.2: Economic Gains from Occupational Upgrading by Race

While the casualty rate is arguably the more exogenous shock, it is still instructive to examine the effect of upgrading into semi-skilled employment on economic outcomes by race from before and after the war. A first test amounts to running the following difference-in-difference-in-differences (DDD) regression:

$$\begin{aligned} y_{izt} = & \beta_1 \left(\text{semi-skilled}_{izt}\text{post-WWII}_t\right) \\ & + \beta_2 \left(\text{semi-skilled}_{izt} \times \text{Black}_{izt} \times \text{post-WWII}_t\right) \\ & + \alpha_z + \lambda_t + \delta\text{Black}_{izt} + X'_{(i)zt}\gamma + \epsilon_{izt} \end{aligned} \tag{4}$$

where y_{izt} is the given economic outcome for individual i in commuting zone z in decade t. The regression includes fixed effects for race Black_{izt}, commuting zone α_z, and census year λ_t, as well as individual- and commuting zone-level controls $X'_{(i)zt}$. Individual level controls include dummies for age, marital status, and place of birth. Commuting zone controls include all the controls used also in section 3 which are aggregated to the county- to the commuting zone-level. Standard errors are clustered by commuting zone. Table A16 reports the results from this regression for six outcomes. The first three are indicators for urban and cross-state migration status, and home ownership. A cross-state migrant here is a person who does not reside in their state of birth.

Consistent with the model predictions from section 4.1, the average quality of semi-skilled Black workers increased while that of semi-skilled white workers decreased after the war. This interpretation holds true if we accept a worker's wage as indicative of their productivity as signal of the worker's quality. A Black worker, who upgraded from low- to semi-skilled work after the war, saw an increase in their wage of 28.3% in the full sample and an 29.2% when considering the South alone.[4] Considering the difference in wages between low-skilled agricultural jobs and semi-skilled manufacturing jobs shown in panel b of Appendix Figure A1, this is a reasonable estimate. Whites on the other hand did not see such a wage increase and actually experienced a reduction in wages in this skill group. As predicted by the model, the average quality of new white manufacturing workers should have declined, hence also average wages would decrease as a result. Another explanation that I cannot rule out is that increased competition reduced white wages as the threat of hiring a Black worker as substitute now became real. A third potential explanation is that the best semi-skilled whites left semi-skilled employment as Black workers entered these jobs and pushed into high-skilled employment as type of occupational white flight that has been shown to exist for housing (Boustan, 2010).

Moving into semi-skilled employment is also associated with increased education for Black workers and decreased education for white workers. Again, this could reflect increased investment in education by Black workers due to the newly available jobs and a reduction for whites'

[4]For this statement the coefficients were transformed from their approximate value to their exact value by computing $100 \times (e^{\beta_2} - 1)$.

education due to selection as more highly educated whites would enter high-skilled employment instead. This finding is consistent in both the full and the Southern sample. The skill-upgrade is only marginally significantly related with the probability of home ownership in the full sample with a 1.2 percentage points rise. However, when African Americans owned their home, this is now of substantially higher value for those who experience the skill-upgrade. For whites there is a negative effect on house values which might be due to outmigration of wealthier whites driving down home values (Boustan and Margo, 2013) or a decline in housing segregation that reduces prices for homes of whites (Logan and Parman, 2017).

Table A16: Effect of Occupational Upgrading on Economic Outcomes

Outcome:	ln(wage)	Education	Migrant	Urban	Owns home	ln(house val.)
	Panel A: All U.S.					
Semi-skilled × Post-war	-0.056***	-1.448***	-0.004	-0.014***	0.004*	-0.247***
	(0.007)	(0.041)	(0.006)	(0.003)	(0.002)	(0.005)
Semi-skilled × Black × Post-war	0.249***	1.755***	0.022***	0.020***	0.029***	0.315***
	(0.009)	(0.054)	(0.008)	(0.005)	(0.008)	(0.019)
Observations	1,926,990	2,312,296	3,528,795	3,528,795	3,422,362	1,084,868
Adj. R^2	0.452	0.410	0.308	0.604	0.246	0.441
	Panel B: South Only					
Semi-skilled × Post-war	-0.077***	-1.505***	0.023***	-0.018***	-0.004	-0.314***
	(0.014)	(0.066)	(0.008)	(0.005)	(0.003)	(0.010)
Semi-skilled × Black × Post-war	0.256***	1.809***	-0.014**	0.029***	0.022***	0.367***
	(0.014)	(0.065)	(0.007)	(0.007)	(0.008)	(0.020)
Observations	590,194	724,463	1,082,514	1,082,514	1,047,141	329,455
Adj. R^2	0.453	0.389	0.446	0.627	0.224	0.450

Note: Difference-in-differenece-in-differences regression of economic outcomes on the commuting zone WWII casualty rate among semi-skilled whites interacted with a post-WWII dummy, and with a Black indicator for individuals living in 722 commuting zones in the whole U.S. The estimation sample contains data from the decennial U.S. micro Census from 1920-70 on non-institutionalized, working Black and white males aged 15-65 who are not currently attending school. All regressions include commuting zone and Census year fixed effects. Owns home is a binary outcome for whether an individual owns their home. The log house value, log wages, and education variables are only available from 1940 onward. Log house value is also missing for 1950. Individual level controls include age, marital status, age and place of birth dummies. Commuting zone level controls are the WWII draft rate, log WWII spending per capita, share of Black men, share of rural population, no. of manufacturing establishments per capita, average manufacturing firm size, log manufacturing value added per worker, share of employment in manufacturing, share of land in agricultural production, share of acres in cotton production, share of cash tenants, average value of machinery per farm, lynchings per 1,000 Black population between 1900 and 1930, no. of Rosenwald schools per 1,000 Black population, share of acres flooded by the Mississippi in 1928, no. of slaves in 1860, Republican vote share, New Deal spending per capita 1933-35 (loans, public works, AAA, FHA loans), and the unemployment rate in 1937. Time-invariant controls are interacted with decade fixed effects. Monetary values are deflated to 2010 U.S. dollars. Observations are weighted by their individual sample line weight times the spatial weight used to construct the commuting zone level treatment variable. Standard errors clustered at the commuting zone level in parentheses. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B1.3: Further Robustness Checks for Migration Responses

The commuting zone casualty rate effects estimated in section 4.1 are unlikely to be driven by migration responses. While the Census lacks good measures of recent migration events for an individual, it is still possible to test for whether a person has left their state of birth. Tables A17 and A18 repeat the DDD analysis for the sub-samples of those who did not reside in their state of birth at the Census enumeration date and birth-state stayers in the country as a whole and in the South only, respectively.

Table A17: Movers vs. Birth-State Stayers, all U.S.

Outcome:	ln(wage)	Education	Urban	Owns home	ln(house val.)
	Panel A: Movers				
Casualty rate × Post-war	-0.020*	-0.058	-0.014	0.002	-0.048**
	(0.010)	(0.037)	(0.017)	(0.004)	(0.023)
Casualty rate × Black × Post-war	0.044***	0.315***	0.003	0.011***	0.054***
	(0.005)	(0.026)	(0.002)	(0.004)	(0.018)
Observations	700,329	816,365	1,215,203	1,138,945	355,222
Adj. R^2	0.411	0.381	0.632	0.248	0.407
	Panel B: Birth-State Stayers				
Casualty rate × Post-war	-0.012	-0.043	-0.007	-0.008**	-0.037*
	(0.008)	(0.028)	(0.010)	(0.003)	(0.021)
Casualty rate × Black × Post-war	-0.001	0.326***	0.000	-0.004	0.080***
	(0.006)	(0.022)	(0.003)	(0.003)	(0.009)
Observations	1,226,661	1,495,931	2,313,592	2,283,417	729,646
Adj. R^2	0.471	0.420	0.602	0.252	0.443

Note: Difference-in-differenece-in-differences regression of economic outcomes on the commuting zone WWII casualty rate among semi-skilled whites interacted with a post-WWII dummy, and with a Black indicator for individuals living in 722 commuting zones in the whole U.S. The estimation sample contains data from the decennial U.S. micro Census from 1920-70 on non-institutionalized, working Black and white males aged 15-65 who are not currently attending school. All regressions include commuting zone and Census year fixed effects. Owns home is a binary outcome for whether an individual owns their home. The log house value, log wages, and education variables are only available from 1940 onward. Log house value is also missing for 1950. Individual level controls include age, marital status, age and place of birth dummies. Commuting zone level controls are the WWII draft rate, log WWII spending per capita, share of Black men, share of rural population, no. of manufacturing establishments per capita, average manufacturing firm size, log manufacturing value added per worker, share of employment in manufacturing, share of land in agricultural production, share of acres in cotton production, share of cash tenants, average value of machinery per farm, lynchings per 1,000 Black population between 1900 and 1930, no. of Rosenwald schools per 1,000 Black population, share of acres flooded by the Mississippi in 1928, no. of slaves in 1860, Republican vote share, New Deal spending per capita 1933-35 (loans, public works, AAA, FHA loans), and the unemployment rate in 1937. Time-invariant controls are interacted with decade fixed effects. Monetary values are deflated to 2010 U.S. dollars. Observations are weighted by their individual sample line weight times the spatial weight used to construct the commuting zone level treatment variable. Standard errors clustered at the commuting zone level in parentheses. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A18: Movers vs. Birth-State Stayers, South only

Outcome:	ln(wage)	Education	Urban	Owns home	ln(house val.)
	Panel A: Movers				
Casualty rate × Post-war	-0.025*	-0.001	-0.037	0.008	-0.059*
	(0.015)	(0.069)	(0.025)	(0.006)	(0.033)
Casualty rate × Black × Post-war	0.049***	0.383***	0.001	0.004	0.097***
	(0.007)	(0.034)	(0.003)	(0.006)	(0.016)
Observations	195,623	228,198	331,025	304,065	97,363
Adj. R^2	0.435	0.393	0.672	0.247	0.418
	Panel B: Birth-State Stayers				
Casualty rate × Post-war	-0.034***	-0.106***	-0.036***	-0.003	-0.055**
	(0.012)	(0.039)	(0.013)	(0.005)	(0.024)
Casualty rate × Black × Post-war	0.011*	0.337***	0.003	-0.001	0.090***
	(0.006)	(0.022)	(0.002)	(0.003)	(0.011)
Observations	590,194	724,463	1,082,514	1,047,141	329,455
Adj. R^2	0.451	0.378	0.627	0.224	0.441

Note: Difference-in-differenece-in-differences regression of economic outcomes on the commuting zone WWII casualty rate among semi-skilled whites interacted with a post-WWII dummy, and with a Black indicator for individuals living in 300 commuting zones in the U.S. South. The estimation sample contains data from the decennial U.S. micro Census from 1920-70 on non-institutionalized, working Black and white males aged 15-65 who are not currently attending school. All regressions include commuting zone and Census year fixed effects. Owns home is a binary outcome for whether an individual owns their home. The log house value, log wages, and education variables are only available from 1940 onward. Log house value is also missing for 1950. Individual level controls include age, marital status, age and place of birth dummies. Commuting zone level controls are the WWII draft rate, log WWII spending per capita, share of Black men, share of rural population, no. of manufacturing establishments per capita, average manufacturing firm size, log manufacturing value added per worker, share of employment in manufacturing, share of land in agricultural production, share of acres in cotton production, share of cash tenants, average value of machinery per farm, lynchings per 1,000 Black population between 1900 and 1930, no. of Rosenwald schools per 1,000 Black population, share of acres flooded by the Mississippi in 1928, no. of slaves in 1860, Republican vote share, New Deal spending per capita 1933-35 (loans, public works, AAA, FHA loans), and the unemployment rate in 1937. Time-invariant controls are interacted with decade fixed effects. Monetary values are deflated to 2010 U.S. dollars. Observations are weighted by their individual sample line weight times the spatial weight used to construct the commuting zone level treatment variable. Standard errors clustered at the commuting zone level in parentheses. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Data Appendix

Merging Enlistment and Casualty Records

Merging the 8.3 million observations from the WWII Army enlistment records with the casualty records based on the Army serial number matches 78% of all casualties. These are observations which found a unique match across both data sets. For robustness I computed the soundex string distance of first- and surname and kept those matches for which it was sufficiently small in order to be sure that the match was correct. Less than one percent of these initial matches were returned to the pool of unmatched observations because of significant differences in the names that indicated a clear mismatch despite a perfect match on the serial number. The match rate is not perfect because of mistakes in the serial number made by the Optical Character Recognition (OCR) software on part of the casualty tables for which the scans are of less than ideal quality.

The remaining casualties were matched via the probabilistic string matching algorithms provided by Wasi and Flaaen (2015). A one-to-one match was used to link each casualty with a potential enlistment record based on name and serial number stratified by state of residence. Names are matched via a tokenization and serial numbers via a bigram algorithm. The match with the highest combined matching score was kept. This results in a final match rate of 94%. From a random sample of 1,000 matches the error rate was 0.6% as judged by correctness of the name, serial number, and residence. The OCR quality of the remaining 6% of casualty observations was too poor in order to clearly identify whether a given match was correct. These cases were dropped.

Sources of the U.S. Census County Data, 1920-1960

The main data source are the county aggregates of the U.S. Decennial Census of Population and Housing from 1940 to 1960 and the 100% full count micro data of the Census. For the years 1940 to 1960, the Census publishes occupational counts at the county level where Southern states report them separated for Black and white workers. For instance, see Table 23a on page 278 of the 1940 Census for Georgia shown in Figure A8 which are the raw data from which I digitized the employment information at the county level for Black workers by county and skill group. Occupations are defined according to the harmonized 1950 definition by the U.S. Census Bureau. The categories include professional, semi-professional, farmers, proprietors and managers, clerical and sales, craftsmen and foremen, operatives, domestic services, farm laborers, and laborers. Semi-skilled occupations here are taken to be the groups of craftsmen and operatives. These definitions change considerably with the 1980 Census which makes it impossible to keep a consistent measurement of the outcome variable.

Before 1940 the county level aggregates do not report these statistics. However, it is possible to construct them from the 100% full count micro data of the Census for 1920, 1930, and 1940. Before 1920 there is no reliable employment status data. This information is important

to construct the correct county aggregates. For each county, these are the sum of all currently employed workers in a given occupational group. The emphasis lies on currently employed. Given the overlap of the full count Census and the county level aggregates in 1940, this is the only definition of workers which gives a complete overlap between the two data sources with respect to the constructed and the actual county level data.

The Census data also contain information on each county's population but also on the local economies. This includes information on the number of manufacturing establishments, number of manufacturing workers, and value added. From the I compute the following controls:

Manufacturing firms per 1,000 pop $= \frac{\text{No. manufacturing establishments}_{ct}}{\text{Total population}_{ct}/1{,}000}$

Av. manufacturing firm size $= \frac{\text{Total manufacturing workers}_{ct}}{\text{No. manufacturing establishments}_{ct}}$

Manufact. value added per worker $= \ln\left(1 + \frac{\text{Total manufacturing value added}_{ct}}{\text{Total manufacturing workers}_{ct}}\right)$

Share of manufacturing workers $= \frac{\text{Total manufacturing workers}_{ct} \times 100}{\text{Total population}_{ct}}$

Share of Black men $= \frac{\text{Total no. of Black men}_{ct} \times 100}{\text{Total no. of men}_{ct}}$

Share of Black population $= \frac{\text{Total no. of Black population}_{ct} \times 100}{\text{Total population}_{ct}}$

Data on the number of slaves in 1860 by county come from the 1860 U.S. Decennial Census of Population and Housing. Additionally, information on median family income was taken from the Census files. For 1940, the median family income was computed from the 1940 100% Census micro data. Whenever information on manufacturing or income variables was not available or incomplete in the Census, these were supplemented with information from the County and City Data Books from 1947 to 1972 published by the U.S. Census Bureau.

Control Variables

Agricultural Controls

Information on agricultural variables at the county level for each decade was taken from the U.S. Agricultural Census prepared by:

- Haines, M., Fishback, P.V., and Rhode, P. (2016) "United States Agriculture Data, 1840 - 2012", Study No. ICPSR35206-v3, Inter-university Consortium for Political and Social Research 2016-06-29, Ann Arbor, MI

Constructed variables from this data set are:

acres in farm land $= \frac{\text{farm acres}_{ct} \times 100}{\text{land acres}_{ct}}$

average value of machinery per farm $= \frac{\text{value of farm machinery}_{ct} \times \text{CPI}_t}{\text{No. farms}_{ct}}$

share of cash tenants$_{ct}$ $= \frac{\text{No. cash tenants}_{ct} \times 100}{\text{Total no. tenant farmers}_{ct}}$

share of cotton in agriculture$_{ct}$ $= \frac{\text{No. acres in cotton production}_{ct} \times 100}{\text{Acres in farm land}_{ct}}$

Lynchings

Data on the number of lynchings for a given county between 1900 and 1930 come from Project

HAL: Historical American Lynching. Their definition of a lynching follows the conditions outlined by the National Association for the Advancement of Colored People (NAACP). The conditions for a murder to qualify as lynching are that there must be evidence that someone was killed; the killing must have occurred illegally; three or more persons must have taken part in the killing; and the murderers must have claimed to serve tradition or justice. The lynchings variable here is defined as: $\frac{\text{No. lynchings 1900-1930}_c}{\text{No. of Black pop}_{ct}/1{,}000}$. The data are freely available at:

- http://people.uncw.edu/hinese/HAL/HAL%20Web%20Page.htm
  (retrieved on November 2nd, 2017)

Mississippi Flooded Acres, 1928

This data comes from the data deposit by Hornbeck and Naidu (2014) at the American Economic Review website. The variable used here is defined as: $\frac{\text{flooded acres}_{c,1928} \times 100}{\text{total acres}_{c,1930}}$. The data can be accessed at:

- https://www.aeaweb.org/aer/data/10403/20120980_data.zip
  (retrieved on November 3rd, 2017)

Party Vote Shares

Data on the Republican vote share come from:

- Clubb, J.M., Flanigan, W.H., and Zingale, N.H. (2006) "Electoral Data for Counties in the United States: Presidential and Congressional Races, 1840-1972", ICPSR08611-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2006-11-13. https://doi.org/10.3886/ICPSR08611.v1

The data report congressional and presidential vote share by party for each election between 1840 and 1972. The Republican vote share here is taken to be the share of votes obtained by the Republican party in congressional elections in a Census year. If there was no election in given Census year, the nearest election was assigned.

Rosenwald Schools

The Rosenwald School variable here is defined as: $\frac{\text{No. Rosenwald Schools}_c}{\text{No. of Black pop}_{ct}/1{,}000}$.
The number of Rosenwald Schools per county was obtained from:

- http://rosenwald.fisk.edu/index.php
  (retrieved on November 2nd, 2017)

WWII Related Spending

War related spending during World War II was taken from the 1947 County and City Data Book. A digital version is provided by:

- United States Department of Commerce. Bureau of the Census. "County and City Data Book [United States] Consolidated File: County Data, 1947-1977. ICPSR07736-v2". Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2012-09-18. https://doi.org/10.3886/ICPSR07736.v2

The war related spending per capita variable here is computed as:

$$\text{Log mil. spending per capita} = \ln\left(1 + \frac{\left(\text{\$ combat equip.+\$ other equip.+\$ ind. facilities+\$ milfacilities}\right)_{c,1940}}{\text{Total population}_{c,1940}}\right)$$

References

Acemoglu, Daron, David H. Autor, and David Lyle, "Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Midcentury," *Journal of Political Economy*, 2004, *112* (3), 497–551.

Aizer, Anna, Ryan Boone, Adriana Lleras-Muney, and Jonathan Vogel, "Discrimination and Racial Disparities in Labor Market Outcomes: Evidence from WWII," *NBER Working Paper 27689*, 2020.

Bailey, Martha J., Karen Clay, Price Fishback, Shawn Kantor, Edson Severnini, and Anna Wentz, "U.S. County-Level Natality and Mortality Data, 1915-2007," *Inter-university Consortium for Political and Social Research [distributor], 2018-05-02. https://doi.org/10.3886/ICPSR36603.v2*, 2018.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen, "High-Dimensional Methods and Inference on Structural and Treatment Effects," *Journal of Economics Perspectives*, 2014, *28* (2), 29–50.

Boustan, Leah, "Was Postwar Suburbanization "White Flight"? Evidence from the Black Migration," *Quarterly Journal of Economics*, 2010, *125* (1), 417–443.

__ , "Competition in the Promised Land: Black Migrants in Northern Cities and Labor Markets," *Princeton University Press, Princeton, NJ*, 2016.

__ **and Robert A. Margo**, "A silver lining to white flight? White suburbanization and African-American Homeownership, 1940-1980," *Journal of Urban Economics*, 2013, *78*, 71–80.

Conley, Timothy G., "GMM estimation with cross sectional dependence," *Journal of Econometrics*, 1999, *92* (1), 1–45.

Getis, Arthur and J. Keith Ord, "The Analysis of Spatial Association by Use of Distance Statistics," *Geographical Analysis*, 1992, *24* (3), 189–206.

Goldin, Claudia, "The Role of World War II in the Rise of Women's Employment," *American Economic Review*, 1991, *81* (4), 741–756.

__ **and Claudia Olivetti**, "Shocking Labor Supply: A Reassessment of the Role of World War II on Womens Labor Supply," *American Economic Review Papers and Proceedings*, 2013, *103* (3), 257–262.

Hornbeck, Richard and Suresh Naidu, "When the Levee Breaks: Black Migration and Economic Development in the American South," *American Economic Review*, 2014, *104* (3), 963–990.

Jaworski, Taylor, "You're in the Army Now: The Impact of World War II on Women's Education, Work, and Family," *Journal of Economic History*, 2014, *74* (1), 169–195.

Kondo, Keisuke, "Hot and cold spot analysis using Stata," *Stata Journal*, 2016, *16* (3), 612–631.

Lee, Ulysses, *The Employment of Negro Troops*, in Conn, S. (eds.) United States Army in World War II, Center of Military History U.S. Army, Washington D.C., 1965.

Logan, Trevon D. and John M. Parman, "Segregation and Homeownership in the Early Twentieth Century," *American Economic Review Papers and Proceedings*, 2017, *107* (5), 410–414.

Moran, Patrick A.P., "Notes on Continuous Stochastic Phenomena," *Biometrika*, 1950, *37* (1/2), 17–23.

Oster, Emily, "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business and Economic Statistics*, 2019, *37* (2), 187–204.

Pei, Zhuan, Jörn-Steffen Pischke, and Hannes Schwandt, "Poorly Measured Confounders are More Useful on the Left than on the Right," *Journal of Business and Economic Statistics*, 2019, *37* (2), 205–216.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas,

and Matthew Sobek, "IPUMS USA," *Version 8.0 [dataset]. Minneapolis, MN: IPUMS, 2018. https://doi.org/10.18128/D010.V8.0*, 2018.

Wasi, Nada and Aaron Flaaen, "Record Linkage Using Stata: Pre-processing, Linking, and Reviewing Utilities," *Stata Journal*, 2015, *15* (3), 672–697.