Crossing the Color Line: The Effects of Racial Integration during the Korean War

Daniel Indacochea†

November 2019

Most recent version here.

Abstract

The racial integration of the US Army during the Korean War (1950-1953) is one of the largest and swiftest desegregation episodes in American history. This paper argues that racial integration during the war improved white survival rates at the expense of blacks and resulted in less anti-black prejudice among white veterans decades afterwards. Using a novel military casualty file, I quantify exogenous changes in racial integration across units to show that integrated whites were 3% more likely to survive their injuries than segregated whites, whereas integrated blacks were 2% less likely to survive their injuries than segregated blacks. To explore the long-term effects of racial integration, I link individual soldiers to post-war social security and cemetery data using an unsupervised learning algorithm. With these linked samples, I show that a standard deviation increase in wartime racial integration caused white veterans to live in more racially diverse neighborhoods and marry spouses with less white-sounding names. In aggregate, these results are rare documented examples of large-scale interracial contact reducing prejudice on a long-term basis.

JEL Codes: J12, J15, N42
Keywords: *Racial Integration, Contact Hypothesis, Korean War*

---

†University of Toronto, Department of Economics. E-mail: daniel.indacochea@mail.utoronto.ca. I thank Aloysius Siow, Shari Eli, Robert McMillan, Ismael Mourifié, James G. Mackinnon, Eli Berman, Alexander Chernoff, David Rosé, Alastair Fraser, Marc-Antoine Schmidt, and seminar participants at the University of Toronto, the 13^{th} Defense and Security Economics Workshop, and the 53rd Annual Conference of the CEA for valuable comments and feedback. Financial support from SSHRC is gratefully acknowledged. Any errors are my own.

I Introduction

The racial integration of the Eighth US Army during the Korean War (1950-1953) is one of the largest and swiftest desegregation episodes in American history (Mershon and Schlossman, 1998). At the outset of the war, blacks and whites were rigidly segregated both by unit and by occupation. All-white units were largely responsible for combat, whereas blacks were assigned menial tasks in non-combat support roles[1]. But the early years of the Korean War were one of the most disastrous military campaigns in American history, and many of the all-white combat units became dangerously depleted. In a desperate bid to bolster manpower, some of these badly depleted all-white units began to accept black reinforcements. And so began integration in Korea—not by decree or fiat—but informally with what was intended to be a temporary emergency measure. Because of the hurried pace of integration—and the soldiers' lack of experience in integrated environments generally—the Army feared that racial strife would hinder combat performance in these newly integrated units. As a precaution, the Army sent teams of social scientists to foxholes in 1951—a time when integrated and segregated military units were still operating in parallel—to interview white officers commanding both racially mixed and segregated units (Bogart, 1992). Code named "Project Clear", the primary concern of this study was to determine what effect, if any, did racial integration have on combat performance. Project Clear offered some reassurance, as white officers perceived little difference in the performance of racially-mixed rifle squads[2] under their command. Moreover, these researchers were also eager to understand the state of race relations in these newly integrated units. What they found was that—compared to their segregated peers—white officers commanding integrated platoons reported *less* anti-black prejudice.

For all of Project Clear's foresight, it nevertheless has several drawbacks: black soldiers are never interviewed, interviews only take place once, and the project relies on self-reported measures of combat performance and prejudice. What then would more objective measures reveal about the performance of newly integrated units, the treatment of blacks within these units, or any shifts in veterans' prejudicial attitudes after the war? In this paper, I will address these questions about the broad effects of racial integration by linking historical records of individual American veterans of the Korean War to their

[1] A notable exception is the all-black 24th regiment, which was the last remaining of the four all-black "Buffalo Soldier" regiments from the American Civil War.

[2] Table 1 outlines the organizational structure of the US military.

long-run, post-war outcomes—both within the army and in their civilian lives. In the absence of training data, I use an unsupervised learning algorithm to match Korean War veterans to Social Security and National Cemetery data. These matched datasets enable me to test the relationship between racial integration during the war and the revealed preferences of veterans after the war in terms of where they lived and whom they married. Using these merged datasets, I first find that, during the war, integration led to increased white survival rates at the expense of blacks. Given that blacks were initially confined to non-combat support roles, these results reflect a convergence in hazardous combat assignments. Second, I find that after the war, integrated whites lived in more racially integrated zip codes and married women with less white-sounding names.

Much of what economists know about large-scale racial integration comes from the closely related literature on the court-ordered desegregation of US public schools, which generally reports that blacks benefited from desegregation policies while the effect on whites was mixed. For example, Guryan (2004) reports that school desegregation reduced the drop-out rates of blacks but had no effect on whites. But unlike school desegregation—which was initiated expressly for the benefit of blacks (Reber, 2010)—the Army integrated for reasons of military efficiency (MacGregor, 1981). Thus this paper evaluates individual and team success in a context where social progress was incidental to war aims. Another unique feature of the Korean War is that it is one of the earliest desegregation episodes in American history. Although the landmark *Brown v. Board of Education* case ruled racial segregation of public schools to be unconstitutional in 1954, school desegregation did not begin in earnest until the mid-1960s. By contrast, formal integration in the Korean War was complete by May of 1952.

A common response to desegregation in other contexts was so-called "white flight". This phenomenon of white raises sample selectivity concerns, as highly-prejudiced individuals are more likely to self-select out of contact with blacks. Reber (2005) finds that actively busing black students into previously all-white schools *accelerated* the exodus of white families moving into the suburbs. A similar white-flight response was observed during the Second Great Migration (1940-1970), the timing of which envelopes the Korean War (1950-1953). During this time, 5 million black migrants from the rural South relocated to urban centers in the North and West, which in turn led to the post-WWII relocation of whites from city centers to suburban rings. Boustan (2007) causally estimates that from 1940-1970, 2.5 white urban residents departed for every black arrival. However, this kind selective migration is less of a concern in a war context; the vast majority of soldiers held low-ranks and therefore had no say as to which unit they served in. In my sample, 80% of soldier held the rank of corporal or lower (see Table 2 for a hierarchy of ranks). To further mitigate sample-selectivity concerns, I repeat the analysis on a sample restricted to draftees.

An expanding number of empirical studies explores the adverse effects of group heterogeneity on outcomes such as team performance or the provision of public goods. Alesina et al. (1999) develop and empirically confirm a model that shows ethnically fractionalized counties in the U.S. spend a smaller share on productive goods such as education and infrastructure. In a sample of U.S. firms, Adams and Ferreira (2009) find that directors in gender-diverse corporate boards have better attendance records but that these firms are more vulnerable to costly outside takeovers. Similar efficiency losses have been tied to group fractionalization in military contexts. During the American Civil War (1861-1865), Costa and Kahn (2003) find that socially fractionalized infantry companies endured more arrests and desertions[3]. A major difference between the Korean War and these other settings, however, is that racial integration changed both the composition *and the size* of regiments. Thus, the experiment I consider is what happens to regiment (firm) productivity—as measured by aggregate survival rates—when restrictions on where soldiers (labour) can be allocated are removed[4].

A third literature that this paper relates to is on the so-called "contact hypothesis" (Allport, 1954), which posits that inter-group contact in a cooperative environment reduces inter-group prejudice. A consistent finding from social psychology is that this type of cooperative inter-group contact does indeed reduce prejudice, but these studies are typically conducted in laboratory settings and rely on non-behavioral measures of prejudice (Pettigrew and Tropp, 2006). In the past 15 years, economists have generalized the contact hypothesis beyond laboratory settings across a variety of revealed-preference measures (e.g. classrooms: Rao, 2019; dorm rooms: Boisjoly et al., 2006,Camargo et al., 2010, Burns et al., 2015, and Air Force Academy squadrons: Carrell et al., 2016). One drawback of real-world settings is that they are often contaminated by so-called "network effects" (Camargo et al., 2010). If,

[3]By contrast, reduced enlistment terms during the Korean War kept desertion rates below 2%, compared to 10% in the Civil War.

[4]Unfortunately, I do not observe the size of regiments in my dataset. Some regimental war diaries contain data on regiment size, which I am in the process of converting into a machine-readable format.

say, a white individual were assigned a black roommate and reports having formed more black friendships afterward, from this the researcher cannot distinguish between a change in prejudicial attitudes or simply an overall increased exposure to blacks. Network effects could in theory be eliminated if black and white strangers were integrated and then isolated again. And although such a scenario may sound fanciful, the Korean War may provide a suitable approximation: whites and blacks from different counties and states were integrated in the Korean Peninsula only to be separated again upon returning to the United States. Thus, in terms of where veterans ultimately live or whom they marry, network effects are less pervasive compared to laboratory settings, and any change in behavior toward the out-group plausibly reflects a genuine change in prejudice—be it attitudes or beliefs[5]. In a closely related study, Schindler and Westcott (2015) look at the deployment of black American G.I.s in the UK during WWII and the racial attitudes of UK citizens decades later. Schindler and Westcott (2015) find that counties that experience a standard deviation increase in exposure to black G.I.s had 0.03 standard deviation fewer members of far-right nationalist parties in 2010. Remarkably, these magnitudes resemble those documented in this paper, which is notable given the distinct outcomes I consider here. This paper differs from Schindler and Westcott (2015) in that I consider a treatment group who had direct contact with blacks.

This paper makes several contributions to our understanding of large-scale racial integration, group fractionalization, and inter-group contact theory. First, I construct a novel measure of wartime integration that I use throughout my analysis. This is possible thanks to a remarkably rich casualty dataset I have obtained on all 103,000 individual Korean War Army veterans who were killed or wounded during the war which, to my knowledge, has yet to be analyzed by economists. Such casualty data make it possible to quantify the extent of integration: if black and white soldiers were physically placed in the same foxholes, then their injuries should correlate. I find that between the first and last year of the war, casualty correlations spiked from 2% to 91%.

Second, I evaluate the effect of integration on unit performance by modeling the survival rate of wounded soldiers. To do so, I develop a modified difference-in-differences strategy where the identifying assumption is that any departure from parallel trends in survival can be attributed to continuous

[5]The datasets I use in this paper are not rich enough to distinguish taste-based discrimination from statistical discrimination. For ease of exposition, I use the term "prejudice" as a stand-in for either form of discrimination.

changes in integration levels. Using this strategy, I find that a one standard deviation increase in the integration measure is associated with a 3 percent increase in white survival rates but a 2 percent decrease in black survival rates. This result in part reflects the fact that, for blacks, racial integration meant being increasingly reassigned to more dangerous combat roles that had hitherto been assigned almost exclusively to all-white units.

Third, I show that racial integration reduced anti-black prejudice after the war in terms of where veterans lived and whom they married. To my knowledge, this paper is the first to document a reduction in prejudice decades after a large-scale, natural integration experiment takes place. In lieu of survey methods, the outcomes that I consider in this paper—namely, the integration of a veteran's last zip code of residence and the "non-whiteness" of the wife's given name—are long-term behavioral outcomes. In order to match Korean War veterans to post-war civilian datasets, I use the expected-maximization (EM) algorithm advocated by Abramitzky et al. (2019) to link individuals in the casualty file to social security and national cemetery data. These merged datasets enable me to test the relationship between racial integration during the war and their revealed preferences after the war in terms of the demographics of their last zip code of residence and the race of whom soldiers married. Since I do not have direct data on the wife's race, I construct a race index based on her given name, year of birth and state of death. In terms of residential choice, I find that a standard deviation increase in regimental integration during the war increased residential integration by 0.06 standard deviations—roughly the difference in integration between present-day Central Harlem and the Upper East Side. Moreover, a one standard increase in regimental integration increased the non-white name index by 0.02 standard deviations.

The remainder of this paper is organized as follows. Section 2 details the origins of racial segregation in the US military and why, during the Korean War, this system was ultimately abandoned. Section 3 presents the casualty data, the six other datasets that veterans are matched to, and the EM algorithm used to match them. Section 4 introduces the casualty similarity index as a measure of wartime integration and outlines the empirical strategy. Section 5 investigates outcomes during the war; namely, the effects of integration on casualty survival. Section 6 shifts the focus to the post-war period by investigating where integrated veterans ultimately chose to reside and whom they married. Section 7 concludes.

II Historical Background

I The origins of racial segregation in the military

Throughout American history, blacks have served in every major conflict. Their participation began in the colonial militias and continued all through the American Revolution. But following the founding of the United States, the history of blacks in the military closely mirrors that of American civil society.

Federal and state legislators passed laws to restrict military service to whites only, though these prohibitions were ignored to meet manpower needs in time of extreme crisis. During the Civil War, 186,000 blacks served in segregated all-black units (Bowers et al., 1997). In the years following the Civil War, public interest in securing civil rights for both free blacks and newly liberated slaves ran high, and the position of blacks in the military improved considerably.

Congress passed measures in 1866 which guaranteed blacks the right to serve, and in 1869 that required the Army to maintain four all-black regiments on a permanent basis. An 1867 federal statute made blacks eligible for militia duty, and many states—all of which had previously excluded non-whites except during emergencies—thereafter admitted blacks to serve in all-black militias. These Reconstruction Era (1863-1877) reflected the new paradigm that had that swept across other public institutions including schools, transportation and hospitals: blacks had to accept a *de facto* “separate but equal” doctrine in order to gain access to public benefits. This doctrine was codified into law by the U.S. Supreme Court in the famous 1896 case *Plessy v. Ferguson*, which established the legal precedent for racial policy in both the military and civilian society for the next seventy years.

II Blacks in the military during WWI and WWII

The late-nineteenth century saw a resurgence in racial hostility, and the white public’s enthusiasm for enforcing the rights of blacks had waned. From 1890 to 1910, several states once again excluded blacks from militia service. But the looming American involvement in World War I in April 1917 rekindled the national debate on the role of blacks within the military. By this time, blacks made up 10% of the national population, and it was clear that the manpower needed to support the war effort in Europe would need to include blacks. But Army officials were vehemently opposed to placing blacks in combat roles. Those blacks who did join the American Expeditionary Forces did so in segregated units, most of which were confined to non-combat support roles. Even the four traditional all-black regiments were outright excluded from deploymen to Europe (Dalessandro et al., 2009).

Between 1916 to 1940, 1.6 million blacks relocated from the rural South to urban centers in the North and hundreds of thousands more to industrial centers elsewhere in the South. These massive demographic shifts dramatically altered the American political landscape. Although race relations remained an afterthought for most whites, blacks began to grow as a political force, as those who left the South were now free from poll taxes, literacy requirements and violence. Blacks increasingly demanded a more prominent role within the military (see Figure 1), which many viewed as a frontier in the battle for civil rights. In response to mounting political pressure, the military removed barriers to black participation, with the Navy even going so far as to integrate its training programs. However, the Army only expanded black opportunities within the circumscribed limits of racial segregation[6], and the Marine Corps continued to exclude blacks outright.

III Executive order 9981

The growing activism and political influence of the black electorate began to turn the tide on racial policies within the military. In July 1948, President Harry Truman issued Executive Order 9981, which stipulated equal treatment and opportunity regardless of race of ethnicity across all branches of the military. Truman made it clear that this decree ultimately meant blacks and whites were to be integrated. But like subsequent integration directives in the context of public schools, Executive Order 9981 was too vague to have an immediate impact. Although the Air Force did begin integrate, the Army and the Marine Corps remained largely unchanged.

IV The Korean War (1950-1953)

Although the legal groundwork for integration had been laid by the outset of the Korean War in June 1950, blacks and whites nevertheless remained rigidly segregated at the battalion level down (See Table

[6]A brief exception occurred following the Battle of the Bulge (December 1944-January 1945), when 2,500 were reassigned to depleted all-white infantry companies, though these units remained segregated at the platoon level.

1). Traditional Army structure is formed in threes: three battalions to a regiment, three companies to a battalion, and so on. This meant that a regiment could contain, say, two all-white battalions and one black battalion, but smaller units were either all-white or all-black[7].

In addition to being segregated across units, blacks and white were also segregated by occupation. Most of the combat roles were assigned to all-white units, while black units were largely confined to non-combat support roles. This occupational segregation stemmed in part from official dissatisfaction with the performance of the all-black 24th Regiment early in the war (Bowers et al., 1997). These all-black units were typically over-strength. One of the few advances in racial policy the Army had made during the intra-war years was to remove race quotas the limited black manpower. However, because Truman's executive order prohibited the formation of new all-black units, the Army simply over-staffed those already in existence. But the early years of the Korean War were one of the most disastrous military campaigns in American history, and many of the all-white combat units became dangerously depleted. This can be seen emphatically in Figure 8, which plots the monthly casualties throughout the war. Evidently, the majority of both black and white casualties occurred early. In a desperate bid to bolster manpower, some commanders of badly depleted all-white units reached informal agreements with personnel officers to accept black soldiers as replacements. One of those units was the 9th Regiment. The Lieutenant Colonel in command of the 9th Regiment's 1th battalion describes his decision to accept black reinforcements:

> I was very, very low on men—less than half strength—and raised hell to get more troops. The division [personnel officer] called and, knowing that I had previously commanded a battalion of black troops [in the Twenty-fifth Infantry], said he had almost two hundred who would transfer to the infantry if they could serve with me. I agreed. In fact, I was proud to have them. [The Second Division commander, Major General] Keiser asked me if I realized what a can of worms I was opening up, to which I said, "So what? They are good fighting me. I need men."

One feature evident from this passage is that the initial decision to integrate was made by high-ranking officers. This meant that whichever units accepted blacks had more to do with the idiosyncrasies of a

[7] All-black units would have either black or white officers.

few commanders than the average characteristics of the thousands of enlisted men in those units.

Another turning point in the road to integration was the appointment of General Matthew Ridgway as the commander of the Eighth Army in January 1951. In Ridgway's view, it was "both un-American and un-Christian for free citizens to be taught to downgrade themselves in this way, as if they were unfit to associate with their fellows or to accept leadership themselves." In May 1951, Ridgway sent a telegraph to the Pentagon outlining his plans for army-wide integration. It was to begin with infantry units, followed by non-infantry combat units, and finally non-combat support units.

In October 1951, the all-black 24th Regiment was disbanded and redistributed across the Army. One of the novel features about troop replacement during the Korean War is that for the first time, reinforcements were introduced individually instead of being rotated in as entire units. Integration proceeded in both directions, with blacks being gradually introduced into previously all-white units, and whites being gradually introduced into previously all-black units.

III Data

This paper considers three main outcomes of interest—namely, casualty survival, residential sorting and intermarriage—of which the latter two require merging multiple datasets. In this section, I detail these datasets and how they are merged. For a summary, Figure 7 depicts which datasets are linked to which.

I Data on integration and survival: Korean War casualty file

The Office of the Adjutant General's Korean War (TAGOKOR) casualty file contains individual records on all 103,000 U.S. Army officers and enlisted men who were listed as casualties (i.e. killed or wounded) during the Korean War. This casualty information was recorded in theater using IBM punch cards (Figure 2) by teams of specialized units known as Mobile Machine Records Units (MRUs). These MRUs were dispatched to collect logistical information on individual units—including casualties—which ensured casualties were recorded within days of their occurrence.

TAGOKOR is unique among public-use American casualty files in that it is the only one that reports individual-level information on killed *and* living soldiers. This information on Korean War survivors is what makes it possible to study the behavior of a subset of soldiers after the war. Broadly speaking, the data in TAGOKOR fall into two categories: military and civilian. The military data details twenty different categories of casualties, including those killed, wounded, hospitalized, missing in action, and captured. Moreover, it records each soldier's rank, military occupational specialty (e.g. "cook, rifleman"), service number, as well as their race, regiment and exact date of casualty. These last three pieces of information are crucial because they are what enable me to construct an integration measure for each regiment at different points in the war. In terms of civilian data, TAGOKOR details each soldier's first and last name, middle initial and county of residence at the time of the war. Year of birth is only recorded for soldiers who were killed in action (KIA), but this information turns out to still be of use for matching surviving soldier's to their post-war outcomes (see matching appendix). Table 3 contains an example of a typical observation contained in the TAGOKOR casualty file.

The summary statistics reported in Table 6 highlight the circumstances under which racial integration took place. Panel A summarizes each variable in the period before formal integration was announced on 1-Oct-1951. Comparing Columns (1) and (2), we see that whites were 3.1 percentage points less likely to survive their injuries than blacks, reflecting the initial occupational segregation between the two groups, and the need to reinforce depleted all-white battalions early in the war. In Panel B we see this same casualty survival gap decrease to 0.5 percentage points, as blacks gradually joined whites in more hazardous combat roles. That the gap remains positive is evidence that black lives were not gratuitously put as risk once they were integrated.

Table 6 also speaks to the possible importance of white flight in the Korean War context. Although soldiers already stationed in the Korean peninsula had little control as to whom they served with, it is possible that potential volunteers with a distaste for integration were dissuaded from joining the Army as news of racial integration reached the home front. Indeed, as the war progressed, we see by comparing across Panels A and B that the number of whites soldiers originating from the South—historically the most prejudiced region in country—declined by 6.6 percentage points. However, the majority of this change stems from the introduction of the draft. Of the 5 million American soldiers who served during the Korean War, 1.5 million were draftees. These draftees were overrepresented in the Army relative to other branches of the armed forces because volunteers usually chose to enter the Navy or the Air Force. A unique feature about TAGOKOR is that one can indirectly observe the soldier's draft status. TAGOKOR reports each soldier's service number, and those with the prefix "US" were draftees. In Columns (3) and (4), I restrict the sample to soldiers who volunteered for the Army (i.e. prefix not equal to "US"). Now we see that the fraction of volunteers who came from the South only declined by 2.7 percentage points. To further mitigate sample selectivity concerns, I repeat my analysis of residential sorting with a sample restricted to draftees.

II Data on last zip code of residence: Social Security Death Index

The copy of the Social Security Death Index (SSDI) that I have obtained contains records for over 90 million Americans who died between 1940 to 2010. This file was created from the Social Security Administration's (SSA) Death Master File extract. It contains records for deceased individuals who had been assigned social security numbers (SSN) and whose deaths were reported to the SSA. From 1973 onward, the SSDI covers between 93% to 96% of deaths of individuals aged 65 or older (Hill and Rosenwaike, 2001). This high coverage rate for older cohorts implies that the vast majority of Korean War veterans appear in these records.

Table 4 presents a typical observation contained in the SSDI. The SSDI does not contain any military information that could be used to uniquely identify Korean War veterans. However, the SSDI does contain enough personally identifiable information for it to be linked to veterans in the casualty file. In particular, for each deceased individual, the SSDI reports first and last names, middle initial (if death occurred after 2000), the state where SSN was issued, and dates of birth and death. (Appendix for matching details).

The SSDI also reports the last zip code of residence, which is the geographical unit at which I model a veteran's residential choice. For this model, the outcome of interest is the extent of racial integration in the veteran's last zip code. For each zip code, I construct a residential similarity index using block-level census data. More formally, let $B_{b,t}$ ($W_{b,t}$) be the number of blacks (whites) in Census block b in year t, and let D_z be the set of blocks that belong to zip code z. The total number of blacks/whites in zip code z in year t is given by

$$B_{z,t}^{Total} \equiv \sum_{b \in D_z} B_{b,t}$$
$$W_{z,t}^{Total} \equiv \sum_{b \in D_z} W_{b,t}.$$

The residential similarity index for zip code z in year t is defined as

$$ZipSim_{z,t}^{BW} \equiv 1 - \frac{1}{2} \sum_{b \in D_z} \left| \frac{B_{b,t}}{B_{z,t}^{Total}} - \frac{W_{b,t}}{W_{z,t}^{Total}} \right| \tag{1}$$

By construction, $ZipSim_{z,t}^{BW}$ can take on values between 0 and 1. The term to the right of the minus sign in Equation 1 is the familiar dissimilarity Index first proposed by Duncan and Duncan (1955). Hence $ZipSim_{z,t}^{BW}$ is the compliment the dissimilarity index. An $ZipSim_{z,t}^{BW}$ value of 0 indicates complete segregation (i.e. none of the census blocks contain both blacks and whites), and a value of 1 indicates complete integration (i.e. the distribution of blacks and whites across census blocks is identical). Intuitively, $ZipSim_{z,t}^{BW}$ is the proportion of blacks (or whites) who would have had to have be relocated to other Census blocks in order to completely segregate the two groups.

To simplify the analysis, I assume that a veteran chooses which housing market to reside in (e.g. New York-Newark-Jersey City vs Santa Fe, New Mexico) for reasons orthogonal his racial preferences—be it work opportunities, family ties, and so forth. Here I define as housing market as a core-based statistical area (CBSA) so that the veteran's choice set is limited to all zip codes within a given CBSA. I account for the veteran's choice set by normalizing $ZipSim_{z,t}^{BW}$ within a CBSA. In the case of, say, the New York housing market, the with-CBSA normalized zip code similarity index $y_{z,t}$ is given by

$$y_{z,t} = \frac{ZipSim_{z,t}^{BW} - \mathbf{E}_{NY}(ZipSim_{z,t}^{BW})}{\sigma_{NY}(ZipSim_{z,t}^{BW})}. \tag{2}$$

The distribution of $y_{z,t}$ is shown in Figure 13. That $y_{z,t}$ is roughly symmetrical indicates that, within a housing market, the similarity index is distributed symmetrically. I also use the block-level data compute and normalize the black share of each zip code as well. To account for neighborhood quality, I matched zip codes to annual house price data published on Zillow.com[8].

In summary, I link casualty data to social security data, which in turn I link to Census and house price data. The linked dataset contains 16,081 individual Korean War veterans with data on individual characteristics (e.g. age, race), wartime characteristics (e.g. measure of integration, rank, occupational specialty, county of residence), and the proportion of blacks in the last zip code of residence.

III Data on wives: national cemeteries

The National Cemetery Administration maintains and updates data on all military veterans and their dependents who are buried in one of the 138 national cemeteries across the United States. As of November 2018, these records contain records on 250,000 unique veteran/wife pairs for a sub-sample of Korean War Army veterans. Each record details both the veteran's and his wife's first and last name, middle initial, date of birth, date of death and the veteran's military rank. Table 5 details a typical observation contained in a national cemetery record. To assign an integration measure to these veterans in national cemeteries, I link them to individuals soldiers in the casualty file using an unsupervised learning method (see Appendix X matching). A limitation of this dataset is that it does not contain information on the race of the wife, and matching directly to administrative data that contains race such as SS-5 application forms decimates the sample size. To avoid this problem, I construct a "blackness/non-whiteness" name index for each wife. Mathematically, this name index is the probability that the wife is black/non-white given her first name, year of birth and state of burial. Plots of these data are shown in Figure 12.

IV Methodology

I Variation in integration across time and regiments

The rich casualty data contained in TAGOKOR make it possible to quantify the extent of wartime racial integration; intuitively, if black and white soldiers were physically placed in the same foxholes,

[8]House price data is available at `https://www.zillow.com/research/data/`

then their injuries should correlate. This phenomenon is evident in Figure 4, which plots daily black and white casualties in the early and late stages of the war. During the first three months of combat—at a time when blacks and whites were rigidly segregated in units of battalion-size (about 800 soldiers) or smaller—there is no discernible relationship between the timing of black and white casualties ($\rho = 0.02$, see Figure 4a). But these casualty patterns change sharply in the last year of the war. By this time, integration was virtually complete, and the relationship between black and white daily casualties is practically linear ($\rho = 0.91$, see Figure 4b).

This stark contrast is not driven by some unobserved trends, as integration also varied considerably across regiments within a given time period. To illustrate this point, the story of the 9th Regiment presents a useful case study. The 9th Regiment was composed of three battalions: the all-white 1st and 2nd Battalions and the all-black 3rd Battalion. After the two all-white battalions sustained heavy losses in the first three months of the war, these units began to accept black reinforcements from the all-black 3rd Battalion. The photo in Figure 3 shows a newly-integrated fire team from the 9th Regiment in a foxhole in November 1950. By this time the other thirty regiments had not begun to integrate. Figure 5a plots daily black and white casualties between October 1950 to December 1950 for all regiments *except the 9th*. As expected, the casualty pattern is nearly identical to that of the previous three months (Figure 4a). For comparison, Figure 5a only plots daily black and white casualties between October 1950 to December 1950 for the newly-integrated 9th Regiment. The casualty pattern now bears a striking resemblance to that of the last year of the war, by which time racial integration was a *fait accompli*.

II Measuring integration

Throughout this paper, I quantify these plausibly exogenous changes in wartime integration using a regimental[9] similarity index. Let $B_{r,d}, W_{r,d}$ be the number black/white casualties in regiment r on day d of the war. I divide the war into six equal periods denoted by $t \in \{1, ..., 6\}$ (i.e. six periods of 185 days, or 37 months divided by 6). Let D_t denote the set of days in period t (e.g. $D_1 = \{1, \dots, 185\}$).

[9] I use the term "wartime integration" or "regimental similarity" to distinguish this measure from the *residential* similarity of zip codes. The former is an explanatory variable, the latter is an outcome.

The total number of blacks/whites casualties in regiment r during period t of the war is given by

$$B_{r,t}^{Total} \equiv \sum_{d \in D_t} B_{r,d} \tag{3}$$

$$W_{r,t}^{Total} \equiv \sum_{d \in D_t} W_{r,d} \tag{4}$$

for $t \in \{1, ..., 6\}$. We can then define similarity index for period t as

$$s_{r,t}^{BW} \equiv 1 - \frac{1}{2} \sum_{d \in D_t} \left| \frac{B_{r,d}}{B_{r,t}^{Total}} - \frac{W_{r,d}}{W_{r,t}^{Total}} \right| \tag{5}$$

By construction, $s_{r,t}^{BW}$ can take on values between 0 and 1; its distribution is plotted in Figure 6. The term to the right of the minus sign in Equation 5 is the familiar dissimilarity Index first proposed by Duncan and Duncan (1955). Hence $s_{r,t}^{BW}$ is the compliment the dissimilarity index. An $s_{r,t}^{BW}$ value of 0 indicates complete segregation (i.e. blacks and whites are never injured on the same day), and a value of 1 indicates complete integration (i.e. the distribution of casualties across days is identical for blacks and whites). Intuitively, $s_{r,t}^{BW}$ is the proportion of blacks (or whites) who would have had to have been injured on different days in order to completely segregate the two groups.

V Results: wartime outcomes

I Integration and casualty survival rates

In any armed conflict, a wounded soldier's chances of survival depends critically on how quickly their wounds are attended to. During the Korean War, the first line of care was organized around two groups: a battalion aid station and a separate forward collecting station. The latter contained eight men composed of a doctor, medics, and litter bearers. The wounded would then be transported to a larger collection station located behind the front line. Once the wounded had been stabilized, they would be transported to a Mobile Army Surgical Hospital (MASH) unit or a division clearing station, depending on the type of wounds. From there the wounded would be transported to an evacuation hospital. If the wounds were serious enough, the wounded would then be airlifted to a hospital in Japan.

The battalion aid station (i.e. the first line of care) was located only a few hundred yards from the front line. This proximity to the front line meant that the success of any evacuation attempt depended critically on battle conditions as medical evacuation vehicles routinely came under enemy fire. As one Army medic put it: "We were always in danger of being attacked by the enemy, overrun by the enemy, being shelled by artillery, shelled by mortar, and grenades thrown into the station". Thus, a successful evacuation depended critically on the ability of nearby combat troops to provide "suppressive fire"[10].

Early in the Korean War, however, many all-white battalions were dangerously depleted, which would have impaired evacuation efforts. Thus, those all-white units that accepted black troops and therefore returned to "safer" sizes should have been better able to evacuate wounded troops to safely. By contrast, prior to integration, black soldiers had been largely confined to non-combat support roles. As such, for blacks integration meant being pulled away from these support roles and thrust into more dangerous combat roles. Overall, then, integration should have increased the probability that whites survived their injuries (i.e. "white casualty survival rate") but decreased that of blacks. I test this relationship with a linear model:

$$y_{irt} = \beta_0 + \beta_1 s_{rt}^{BW} + \boldsymbol{X}_{irt}\boldsymbol{\beta}_2 + \lambda_t + \gamma_r + \varepsilon_{irt} \tag{6}$$

where $y_{i,r,t}$ is a dummy for whether soldier i in regiment r in period t survives their injury, s_{rt}^{BW} is the normalized casualty similarity index, $\boldsymbol{X}_{irt}$ is a vector of controls, and λ_t and γ_r are time and regiment fixed effects, respectively.

When these fixed effects are included, Equation 6 is a difference-in-differences model in spirit, only that any departure in parallel trends is a result of integration. An implication of this model is that during periods when integration levels are steady, two regiments should have parallel survival rates. It is infeasible to pinpoint periods where integration levels remain steady because the integration measure I construct moves too infrequently (i.e. every six months) to convincingly illustrate parallel trends. However, information on the exact date that particular units were integrated is available for a handful of units. For example, Figure 10 shows a screen shot of the 9th Infantry Regiment's war diary

[10]That is, combat troops would attack the enemy to prevent the medical team from being themselves attacked.

for September 1950. The text indicates that the all-black 3rd battalion was reunited with the badly depleted 1st and 2nd battalions of the 9th Infantry on 16-Sep-1950, after having spent the first eight weeks of the war away from combat on reserve duty. Figure 11 plots the time series of the survival rates of the 9th and 23rd Infantry regiments which—as members of the 2nd Infantry Division—cooperated closely in combat throughout the war. The first dashed vertical line denotes 16-Jun-1950 (i.e. day 83 of the war), the day that the three battalions of 9th Regiment were united; by contrast, the 23rd Regiment did not have any all-black units. The second vertical dashed line denotes 1-Oct-1951 (i.e. day 463 of the war), the date that formal integration began. Two things are apparent from this figure. First, after formal integration began, the two regiments exhibit parallel trends. Second, there is some evidence that survival rates improved for whites in the 9th Regiment immediately after being joined by the all-black 3rd battalion.

The results of estimating Equation 6 are reported in Table 7. In Panel A, Column (1), we see that a one standard deviation increase in the integration measure increased white survival probability by 2.2 percentage points. To account for the fact that both integration and survival rates increased as the war progressed, Column (2) includes period fixed effects. Columns (3) adds regiment fixed effects. The preferred specification is Column (4), which adds individual-level information on rank and military occupational specialty. Holding a rank of private meant that a wounded soldier was 1.5 percentage points less likely to survive their injury, which reflects the fact that privates were put in more hazardous positions. Here we see that a standard deviation change in integration increases casualty survival by 2.2 percentage points. Given that mean survival rate for whites was 75%, this represents an increase of 2.9%. Panel B repeats this exercise for blacks. We see in Column (4) that, as predicted, that integration is associated with a *decrease* in black survival rates by 1.8 percentage points—or 2.3%.

These results are not evidence of discrimination *per se*. As seen in Figure 9c, the black casualty survival rate exceeded that of whites throughout most of the war. These results likely reflect the fact that, early in the war, integrated blacks were thrust into more hazardous roles than segregated blacks. However, in the absence of discrimination, we should not expect to see a strong relationship between integration and survival because *all* units were integrated by decree, not to mitigate the risk of annihilation.

VI Results: post-war outcomes

The integration that soldiers experienced during the Korean War arguably met the four conditions for the contact hypothesis outlined by Allport (1954). In particular, black and white soldiers had i) equal status (conditional on rank) ii) common goals iii) a cooperative setting iv) support from authorities. Given the four Allportian conditions were satisfied, this raises the question to what extent, if any, did integration reduce the prejudicial attitudes of veterans?

I Integration and residential choice

I.1 A linear model for residential choice

In this section, I estimate the effect of wartime integration on the residential similarity of each soldier's last zip code of residence. To simplify the analysis, I assume that a soldier chooses which housing market to reside in (e.g. New York City vs Santa Fe, New Mexico) for reasons orthogonal his racial preferences—be it work opportunities, family, and so forth. (I assess the plausibility of this simplifying assumption later in this section.) But when choosing which neighborhood (i.e. "zip code") to reside in *within* a given housing market, the veteran's racial preferences enter into his utility function. Thus, in this section I model neighborhood choice conditional on having already chosen a housing market. To capture this within-housing market neighborhood choice, I estimate a linear model:

$$y^{BW}_{cirstz} = \beta_0 + \beta_1 s^{BW}_{rt} + \boldsymbol{X}_i\boldsymbol{\beta}_2 + \boldsymbol{Z}_{iz}\boldsymbol{\beta}_3 + \boldsymbol{C}_c\boldsymbol{\beta}_4 + \gamma_r + \lambda_t + \zeta_s + \varepsilon_{cirstz} \tag{7}$$

where y_{irstz} is the residential similarity index (within-CBSA z score) for soldier i who served in regiment r, resided in state s during the war, was wounded in period t, originated from county c, and last resided in zip code z, s^{BW}_{rt} is the normalized integration measure, $\boldsymbol{X}_i$ is a vector of individual-level controls, $\boldsymbol{Z}_{iz}$ is a vector of last zip-code level characteristics, and γ_r, ζ_s and λ_t are regiment, state-of-origin, and period-of-casualty fixed effects, respectively.

For ease of exposition, I will refer to the similarity index of the last zip code of residence as "residential similarity", and the similarity index of wartime integration as the "treatment"[11]. In Column (1) we see that a standard deviation (sd) change in regimental similarity leads to a 0.06 sd change in zip code similarity. Column (2) additionally controls for the share of casualties in a regiment that were black, which approximates the share of blacks that served in that regiment. Interestingly, the share of blacks has no explanatory power, whereas the treatment effect barely changes. Simply put, increasing the share of blacks in a regiment does nothing to influence residential sorting patterns if veterans remain segregated within that regiment. Column (3) adds additional information about the last zip code of residence: average house prices and the share of blacks—both normalized within-CBSA—as well as their respective quadratic terms. Clearly, house prices are endogenous since they correlate with unobserved neighborhood quality. The positive coefficient on house prices indicates that, conditional on racial composition, high-quality neighborhoods are more integrated. The treatment effect is attenuated to 0.04 but remains statistically significant. Column (4) adds controls for county of origin, which indicate that veterans from poorer, blacker and more rural counties are less integrated after the war. Crucially, adding these controls does not alter the treatment effect, which is further evidence that racial integration during the Korean War was quasi-random. Intuitively, this regression implies that if one were to take two white soldiers who are close in age, come from similar counties within the same state, and served during the same period of the war—the white veteran who was more integrated during the war is more willing to live in more integrated neighborhoods after the war.

One potential threat to this empirical exercise is that veterans were integrated on the basis of some unobserved factor that is also correlated with residential similarity. For example, perhaps soldiers from counties that were more integrated to begin with were prioritized for integration during the war. I explore this possibility in Column (5), which is restricted to a sub-sample of veterans who last resided in a different state than the one they were living in at the time of the war. For these veterans, the treatment effect is somewhat larger (0.042) and remains statistically significant, where as the county of origin controls lose their explanatory power.

Table 9 shows the estimation results for Equation 7 for black veterans. None of the treatment effects

[11] Strictly speaking, the wartime integration measure applies to the regiment in which soldiers served, not individual soldiers *per se*. As such, OLS estimates of Equation 7 more closely resemble intention-to-treat (ITT) effects. Typically, ITT effects represent lower-bounds of the treatment effects they approximate.

are statistically significant. That the sample sizes are about 10% that of whites means these regressions are under-powered. Moreover, the true effect size for blacks may be smaller than that of whites because blacks faced unique barriers to their residential choices. In the years following the Korean War, commercial banks practiced so called "redlining" to systematically deny mortgages to black applicants. Likewise, zoning laws such as restricting the supply of rental properties disproportionately excluded potential low-income black residents. Although the treatment effects are statistically insignificant, it is worth noting that, as Table 8 showed for whites, the treatments effects here are stable across Columns (1)-(2) and Columns (3)-(4), which indicates that blacks were similarly assigned to integration on a quasi-random basis.

I.2 Residential choice and sample selectivity

A key challenge in studying the effects of interracial contact on prejudicial attitudes is selection bias. Even if interracial contact has a negligible effect on prejudice, but highly prejudiced individuals self-select out of contact, then standard statistical analysis will yield upwardly-biased coefficient estimates. Such sample selectivity is pervasive in many large-scale desegregation episodes. As discussed earlier, a common response to both the desegregation of US public schools and the migration of millions of blacks to predominantly white regions of the US was so-called "white flight"; that is, white out-migration in response to black in-migration.

Several features of the Korean War mitigate these concerns. First, soldiers in Korea had little control as to which units they served in, particularly low-ranking combat soldiers. In my sample, 80% of soldiers held the two lowest ranks, Private or Corporal. It is nevertheless possible that integration in Korea led to preemptive white flight: as word about integration reached the general public, white civilians who might otherwise have enlisted in the army instead refused to serve alongside blacks. Two historical facts mitigate these concerns: first, the Army attempted to conceal the fact of integration. As Mershon and Schlossman (1998) put it, "The Army and the Truman administration, fearful of possible hostility to this momentous and rapid institutional change, went to great lengths to minimize publicity and maximize secrecy during the early 1950s".

Second, as the war progressed, the army came to increasingly rely on the draft to meet manpower requirements. This increased reliance on the draft is evident in my casualty sample. Between the first to last year of the war, the share of draftees in my casualty file climbed from 6% to 40%. If preemptive white flight is present in my sample, then the treatment effects for volunteers should be smaller than those of draftees. I test this possibility directly by splitting the sample between volunteers and draftees and re-estimate Equation 7. Table 10 reports results for white veterans. Since paternity and marital deferments were in use, I restrict my sample to men who were younger than 25 years of age at the time of service. Columns (1)-(3) report results for white volunteers, and Columns (4)-(6) report results for white draftees. Two things are notable: first, contrary to the white flight hypothesis, the estimated treatment effect for white draftees exceeds those of volunteers. Although the treatment effects for draftees are not statistically different from zero, a separate pooled regression indicates that the treatment effects do not differ significantly by draft status. Second, the effects sizes appear to be larger for this younger sample—ranging from 0.11-0.20 than the full sample in Table 8—which ranges from 0.04-0.06. These larger point estimates for young veterans points to heterogeneity in treatment effects. These heterogeneous effects could be for a number of reasons. One possibility is that older veterans were already homeowners and as such less geographically mobile. Another possibility is that younger veterans had less social status, which contact theory predicts would make them more responsive to inter-group contact. With the data available to me, I am unable to distinguish between these distinct mechanisms.

Table 10 reports analogous estimation results for black veterans. Comparing Columns (1)-(3) to columns (4)-(6), we again see no evidence that volunteers have larger treatment effects than draftees, though here again the decimated sample size severely hinders statistical power.

II Integration and intermarriage

II.1 Linear and fractional response models for choosing a spouse

A high bar for prejudice reduction is exogamy—an individual's willingness to marry outside of their own race or ethnicity. I test the relationship between integration and exogamy with two models.

The first is a linear model that I estimate by OLS:

$$y_{irst} = \beta_0 + \beta_1 s_{rt}^{BW} + \boldsymbol{X}_i \boldsymbol{\beta}_2 + \gamma_r + \lambda_t + \zeta_s + \varepsilon_{irst} \tag{8}$$

where y_{irst} is the wife's non-white (or black) name index for soldier i who served in regiment r, resided in state s during the war, and was wounded in period t, s_{rt}^{BW} is the normalized integration measure, $\boldsymbol{X}_i$ is a vector of individual-level controls, and γ_r, ζ_s and λ_t are regiment, state or origin, and period of casualty fixed effects, respectively.

By construction, the wife's name index y_{irst} is bounded between zero and one (see Figures 12b and 12a). In order to account for this characteristic of the left-hand-side variable, I specify the following functional form for the expectation of the name index y_{irst}:

$$\begin{aligned} \mathbf{E}(y_{irst}|s_{rt}^{BW}, \boldsymbol{X}_i, \gamma_r, \lambda_t, \zeta_s) &= \Phi(\beta_0 + \beta_1 s_{rt}^{BW} + \boldsymbol{X}_i \boldsymbol{\beta}_2 + \gamma_r + \lambda_t + \zeta_s), \\ &\equiv \Phi(\boldsymbol{W}_{irst} \boldsymbol{\Theta} + \gamma_r + \lambda_t + \zeta_s), \end{aligned} \tag{9}$$

where Φ is a the normal cumulative distribution function, $\boldsymbol{W}_{irst}$ denotes the row vector $[1 \; s_{rt}^{BW} \; \boldsymbol{X}_i]$, and $\boldsymbol{\Theta}^\top$ denotes the parameter vector $[\beta_0 \; \beta_1 \; \boldsymbol{\beta}_2^\top]$. I estimate the parameters in Equation 9 using quasi-maximum likelihood where the likelihood for an observation is specified as the Bernoulli likelihood:

$$L_i(\boldsymbol{\Theta}, \gamma_r, \lambda_t, \zeta_s) = \left[\Phi(\boldsymbol{W}_{irst}\boldsymbol{\Theta} + \gamma_r + \lambda_t + \zeta_s)\right]^{y_{irst}} \left[1 - \Phi(\boldsymbol{W}_{irst}\boldsymbol{\Theta} + \gamma_r + \lambda_t + \zeta_s)\right]^{1-y_{irst}}. \tag{10}$$

The quasi-maximum likelihood estimates (QMLE) of $\boldsymbol{\Theta}, \gamma_r, \lambda_t, \zeta_s$ are consistent as long as the conditional expectation 9 is correctly specified even if the Bernoulli specification 10 is incorrect[12]. The asymptotic variance-covariance matrix of the QMLE estimates is estimated maintaining only first-moment assumptions without any additional second moment assumptions.

Estimation results for both the linear and fractional response model for white veterans are reported in Table 12. The sample is restricted to men who were 22 years of age or younger during the war. This age cut off is based on the median age of first marriage in 1950, which for men was 22.8 years[13].

[12]See Papke and Wooldridge (1996) for more details regarding this approach

[13]See https://www.thespruce.com/estimated-median-age-marriage-2303878

In Panel A, the dependent variable is the wife's non-white name index. In Column (1), we see that a one standard deviation change in integration corresponds to a 1.8 percentage point increase in the name index. This effect size exceeds that from a one standard deviation change in the black share of the county of origin (0.9 of a percentage point). The age of the soldier does not have a statistically significant effect, which is to be expected because the sample is restricted by age. Column (2) adds rank fixed effects, which barely change the estimated effect sizes. Column (3) includes state-of-residence fixed effects. The effect size of integration again barely changes, though the effect of the black share of their county of origin becomes negative and statistically significant. Column (4) estimates the fraction response model in Equation 9. The marginal effect of 3.1 percentage points is comparable to the OLS results. Panel B repeats this analysis using the wife's black name index as the dependent variable. The results are broadly the same.

II.2 Placebo test results

If racial integration indeed made veterans more exogamous, then one should not observe any effect on veterans who were already married. Since I do not observe marital status before the war, I use age as a proxy for marital status. Table 13 reports estimation results for Equation 9 for a sample of men who were at least 23 years of age during the Korean War. Now we see that across specifications, the treatment effect is no longer statistically different from zero, even though the sample size is 36% larger. Notably, age at casualty now has a strong negative effect on the wife's name indices. As with residential sorting, the fact that the treatment effect is statistically significant for the younger sample but not the older sample—and that age is strongly negative within the older sample—has at least two mutually-compatible explanations: i) interracial contact is more effective at shaping the tastes and/or beliefs of younger individuals ii) older veterans were more likely to already be married. The data available to me is not sufficiently rich to distinguish between these mechanisms[14].

[14]The full count 1950 Census is scheduled to be released in 1950, which will contain valuable spouse data. Future work could restrict the sample to soldier who were not married as late as 1950 and identify treatment effects off of men who served early in the war.

VII Conclusion

The racial integration of the US Army during the Korean War is one of the largest and swiftest desegregation episodes in American history. Integration began informally, as a desperate bid to reinforce badly depleted all-white units, but by the end of the war had become army-wide policy. Thus, unlike other large-scale desegregation episodes such as school desegregation, the army integrated not for the benefit of blacks, but for reasons of military efficiency. The first part of this paper examines whether the army managed to improve efficiency as measured by the survival rates of wounded soldiers. To this end, I develop a novel wartime integration measure to quantify exogenous changes in racial integration across time and regiments. Using a modified difference-in-differences empirical strategy, I find that a standard deviation increase in regimental integration increased white casualty survival rates by 3% relative to whites who remained segregated. By contrast, a standard deviation increased in blacks decreased black survival rates by 2% compared to blacks who remained segregated. On the whole, these figures suggest that one white life was saved for every 10 black reinforcements. These results are best viewed as a convergence in occupational hazard, as wounded black and white soldiers survived at similar rates toward the end of the war. These survival rates are admittedly only a partial measure of overall military efficiency, as integration may have improved combat efficiency at the expense of some unobserved tasks. There future work ought to consider broader measures of efficiency such as unit citations and desertion rates.

Historically, large-scale desegregation episodes have often had unintended consequences. A common response to both school desegregation and black in-migration was so-called "white flight". The second part of this paper is therefore devoted to understanding the aftermath of racial integration in the military by observing the behavior of Korean War veterans after the war. In particular, by matching veterans to social security and cemetery data, I am able to observe where veterans lived and whom they married. I find that a standard deviation increase in wartime integration caused veterans to live in neighborhoods that were 0.1 standard deviations more integrated. Similarly, a standard deviation increase in wartime integration caused veterans to marry spouses with names that were 0.02 percentage points less "white-sounding". In both cases, I observe substantial effect heterogeneity by age, though it remains unclear whether these patterns reflect age-specific changes in prejudice or the fact that older veterans were more likely to have already made residential and marital decisions prior to the war. In aggregate, these results offer some of the first and only documented examples of large-scale interracial contact reducing prejudice on a long-term basis.

References

Abramitzky, R., Mill, R., and Pérez, S. (2019). Linking individuals across historical sources: A fully automated approach. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, pages 1–18.

Adams, R. B. and Ferreira, D. (2009). Women in the boardroom and their impact on governance and performance. *Journal of financial economics*, 94(2):291–309.

Alesina, A., Baqir, R., and Easterly, W. (1999). Public goods and ethnic divisions. *The Quarterly Journal of Economics*, 114(4):1243–1284.

Allport, G. W. (1954). The nature of prejudice.

Bogart, L. (1992). Project clear: Social research and the desegregation of the u? iited states army. new brunswick.

Boisjoly, J., Duncan, G. J., Kremer, M., Levy, D. M., and Eccles, J. (2006). Empathy or antipathy? the impact of diversity. *American Economic Review*, 96(5):1890–1905.

Boustan, L. P. (2007). Black migration, white flight: the effect of black migration on northern cities and labor markets. *The Journal of Economic History*, 67(2):484–488.

Bowers, W. T., Hammond, W. M., and MacGarrigle, G. L. (1997). *Black Soldier, White Army: The 24th Infantry Regiment in Korea*. Diane Publishing.

Burns, J., Corno, L., and La Ferrara, E. (2015). Interaction, prejudice and performance. evidence from south africa. Technical report, Working paper.

Camargo, B., Stinebrickner, R., and Stinebrickner, T. (2010). Interracial friendships in college. *Journal of Labor Economics*, 28(4):861–892.

Carrell, S. E., Hoekstra, M., and West, J. E. (2016). The impact of college diversity on behavior toward minorities. *NBER working paper. No. w*, 20940.

Costa, D. L. and Kahn, M. E. (2003). Cowards and heroes: Group loyalty in the american civil war. *The Quarterly Journal of Economics*, 118(2):519–548.

Dalessandro, R. J., Torrence, G., and Knapp, M. G. (2009). *Willing Patriots: Men of Color in the First World War*. Schiffer Pub Limited.

Duncan, O. D. and Duncan, B. (1955). A methodological analysis of segregation indexes. *American sociological review*, 20(2):210–217.

Guryan, J. (2004). Desegregation and black dropout rates. *American Economic Review*, 94(4):919–943.

Hill, M. E. and Rosenwaike, I. (2001). The social security administration's death master file: the completeness of death reporting at older ages. *Soc. Sec. Bull.*, 64:45.

MacGregor, M. J. (1981). *Integration of the armed forces, 1940-1965*, volume 1. Government Printing Office.

Mershon, S. and Schlossman, S. L. (1998). *Foxholes and color lines*. Johns Hopkins University Press.

Papke, L. E. and Wooldridge, J. M. (1996). Econometric methods for fractional response variables with an application to 401 (k) plan participation rates. *Journal of applied econometrics*, 11(6):619–632.

Pettigrew, T. F. and Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of personality and social psychology*, 90(5):751.

Rao, G. (2019). Familiarity does not breed contempt: Generosity, discrimination, and diversity in delhi schools. *American Economic Review*, 109(3):774–809.

Reber, S. J. (2005). Court-ordered desegregation successes and failures integrating american schools since brown versus board of education. *Journal of Human resources*, 40(3):559–590.

Reber, S. J. (2010). School desegregation and educational attainment for blacks. *Journal of Human resources*, 45(4):893–914.

Schindler, D. and Westcott, M. (2015). Shocking racial attitudes: The cultural legacy of black gis in europe. Technical report, mimeo.

Steinmetz, E. and Iceland, J. (2003). The effects of using newly-defined metropolitan area boundaries when examining residential housing patterns. *Washington, DC: US Census Bureau working paper*.

VIII Tables

Table 1: Glossary of Army units

Units	Strength	Typical commander
region/theater	4+ army groups	six-star rank
army group front	2+ field armies	five-star general
field army	100,000-300,000	general
corps	30,000-50,000	lieutenant general
division	10,000-25,000	major general
†regiment/brigade/legion	1,000-5,500	colonel/brigadier general
*battalion/cohort	300-800	lieutenant colonel
*company	80-150	captain/major
*platoon	15-45	second lieutenant/lieutenant
*squad/section	8-14	sergeant
*fireteam	2-4	lance corporal/corporal

Note: *Formally segregated before Oct 1951. †I observe which regiment individual soldiers served in.

Table 2: Glossary of Army ranks

Pay Grade	Rank	Abbreviation	Classification	2017 Pay Range
E-1	Private	PVT	Enlisted Soldier	$1,600 (monthly)
E-2	Private Second Class	PV2	Enlisted Soldier	$1,793 (monthly)
E-3	Private First Class	PFC	Enlisted Soldier	$22,630 - $25,510
E-4	Specialist	SPC	Enlisted Soldier	$25,067 - $30,427
E-4	Corporal	CPL	Noncommissioned Officer	$25,067 - $30,427
E-5	Sergeant	SGT	Noncommissioned Officer	$27,338 - $38,794
E-6	Staff Sergeant	SSG	Noncommissioned Officer	$29,840 - $46,220
E-7	Sergeant First Class	SFC	Noncommissioned Officer	$34,502 - $62,010
E-8	Master Sergeant	MSG	Noncommissioned Officer	$49,633 - $70,787
E-8	First Sergeant	1SG	Noncommissioned Officer	$49,633 - $70,787
E-9	Sergeant Major	SGM	Noncommissioned Officer	$60,631 - $94,136
E-9	Command Sergeant Major	CSM	Noncommissioned Officer	$60,631 - $94,136
E-9	Sergeant Major of the Army	SMA	Noncommissioned Officer	$60,631 - $94,136
W-1	Warrant Officer 1	WO1	Warrant Officer	$35,597 - $61,510
W-2	Chief Warrant Officer 2	CW2	Warrant Officer	$40,554 - $67,687
W-3	Chief Warrant Officer 3	CW3	Warrant Officer	$45,828 - $80,392
W-4	Chief Warrant Officer 4	CW4	Warrant Officer	$50,184 - $93,478
W-5	Chief Warrant Officer 5	CW5	Warrant Officer	$89,233 - $116,770
O-1	Second Lieutenant	2LT	Commissioned Officer	$36,418 - $45,824
O-2	First Lieutenant	1LT	Commissioned Officer	$41,958 - $58,068
O-3	Captain	CPT	Commissioned Officer	$48,560 - $79,002
O-4	Major	MAJ	Field Officer	$55,231 - $92,218
O-5	Lieutenant Colonel	LTC	Field Officer	$64,012 - $108,752
O-6	Colonel	COL	Field Officer	$76,784 - $135,936
O-7	Brigadier General	BG	General Officer	$101,257 - $151,283
O-8	Major General	MG	General Officer	$121,860 - $175,676
O-9	Lieutenant General	LTG	General Officer	$172,224 - $186,998
O-10	General	GEN	General Officer	$15,583 (monthly)
O-10	General of the Army	GA	General Officer	$15,583 (monthly)

Source: `https://www.federalpay.org/military/army/ranks`. Pay ranges are annual unless otherwise specified.

Table 3: Typical observation in TAGOKOR casualty file

Field title	Value	Meaning
NAME OF CASUALTY	OWENS CALVIN R	OWENS CALVIN R
SERVICE PREFIX AND NUMBER	RA17280815	RA17280815
GRADE	PFC	Private First Class
GRADE CODE	6	Private First Class
BRANCH	IN	Infantry
PLACE OF CASUALTY	L5	South Korea
DAY OF CASUALTY	31	31
MONTH OF CASUALTY	3	MARCH
YEAR OF CASUALTY	A	1951
STATE OF RESIDENCE	73	Kansas
COUNTY OF RESIDENCE	209	Wyandotte
TYPE OF CASUALTY	RTD	Returned to Duty (FECOM)
CASUALTY CODE	4	Wounded by missile
CASUALTY GROUP CODE	F	Returned to Duty (FECOM)
PLACE OF DISPOSITION		
DAY OF DISPOSITION	25	25
BLANK		
MONTH OF DISPOSITION	5	May
YEAR OF DISPOSITION	A	1951
YEAR OF BIRTH		
MILITARY OCCUPATIONAL SPECIALTY CODE	4812	Heavy Weapons Infantryman
TROOP PROGRAM SEQUENCE NUMBER	7003	IN DIV - 3RD
ELEMENT SEQUENCE	60	IN DIV INF REGT
UNIT	7	7 Infantry
RACE CODE OR RACIAL GROUP CODE	1	White
COMPONENT	1	USA - RA (Reg Army)
LINE OF DUTY		
DISPOSITION OF EVACUATIONS		

Note: Searchable records are available at `https://aad.archives.gov/aad/series-description.jsp?s=531`.

Table 4: Typical observation in Social Security Death Index

Field title	Value
ssn	XXXXXXXXXX
lastname	OWENS
suffix	
firstname	CALVIN
middlename	R
yeardeath	2005
yearbirth	1933
stateofdeath	Colorado
zipcode	80222
ssnarea	512
daydeathmiss	0
daybirthmiss	0
dod	2005-11-07
dob	1933-01-08
ageatdeath	72.8795
fipsofdeath	Colorado
ssnfips	Kansas

Note: This table depicts one of over 90 million individual records contained in the Social Security Death Index (SSDI). Although the SSDI reports social security numbers, here I have redacted it for privacy reasons. The last zip code of residence is linked to demographic data to study residential choice. The remaining fields are used to link the SSDI to the casualty file.

Table 5: Typical observation in national cemetery record

Field title	Veteran	Wife
d_first_name	Calvin	Virginia
d_mid_name	R	C
d_last_name	Owens	Owens
d_suffix		
d_birth_date	1/8/1933	2/20/1933
d_death_date	11/7/2005	2/10/2015
section_id	16	16
row_num		
site_num	193	193
cem_name	FT. LOGAN NATIONAL CEMETERY	FT. LOGAN NATIONAL CEMETERY
cem_addr_one	4400 WEST KENYON AVENUE	4400 WEST KENYON AVENUE
cem_addr_two		
city	DENVER	DENVER
state	CO	CO
zip	80236	80236
cem_phone	303-761-0117	303-761-0117
relationship	Veteran (Self)	Wife
v_first_name	Calvin	Calvin
v_mid_name	R	R
v_last_name	Owens	Owens
v_suffix		
branch	US ARMY	US ARMY
rank	CPL	CPL
war	KOREA	KOREA

Note: Searchable records are available at `https://catalog.data.gov/dataset?publisher=Department+of+Veterans+Affairs&tags=cemeteries`.

Table 6: Summary statistics for TAGOKOR

	Full Sample		Volunteers	
	White (1)	Black (2)	White (3)	Black (4)
A. Before 1-Oct-1951				
Casualty survival (%)	69.0	72.1	68.2	71.1
North (%)	51.7	35.8	51.1	37.1
South (%)	34.2	57.6	35.5	55.9
Private (%)	57.3	64.1	54.0	60.0
Officer (%)	5.6	1.7	6.2	1.9
Draftee (%)	10.6	12.9	—	—
Year of birth†	1927.5	1927.3	1927.5	1927.2
	($n = 56,223$)	($n = 7,227$)	($n = 50,244$)	($n = 6,295$)
B. After 1-Oct-1951				
Casualty survival (%)	79.2	79.7	79.3	80.1
North (%)	55.2	29.6	52.1	34.0
South (%)	27.6	62.6	32.8	40.6
Private (%)	63.2	79.0	48.9	71.2
Officer (%)	5.7	2.1	13.6	6.0
Draftee (%)	57.9	64.7	—	—
Year of birth†	1929.5	1929.8	1929.5	1930.5
	($n = 29,494$)	($n = 4,463$)	($n = 12,430$)	($n = 1,572$)

Note: Formal integration began on 1-Oct-1951. †Year of birth is reported for fatalities only.

Table 7: The effects of integration on survival by race

Dependent variable: casualty survival dummy

	(1)	(2)	(3)	(4)
		A. Whites		
Integration (ρ_{BW})	0.022***	0.017**	0.022***	0.022***
	(0.006)	(0.008)	(0.008)	(0.009)
Private				-0.015***
				(0.006)
Support role				0.007
				(0.008)
Intercept	0.745	0.679	0.653	0.655
N	$71,184$	$71,184$	$71,184$	$71,184$
R^2	0.003	0.016	0.024	0.024
		B. Blacks		
Integration (ρ_{BW})	0.006	-0.029**	-0.019*	-0.018*
	(0.019)	(0.013)	(0.011)	(0.010)
Private				-0.028**
				(0.012)
Support role				0.009
				(0.035)
Intercept	0.771	0.701	0.598	0.609
N	$9,702$	$9,702$	$9,702$	$9,702$
R^2	0.000	0.013	0.017	0.017
6-month FE	No	Yes	Yes	Yes
Regiment FE	No	No	Yes	Yes
#Regiments	24	24	24	24
#Periods	6	6	6	6

Note: Standard errors calculated using two-way clustering by regiment and six-month period of casualty. Period 1 and 2nd Regiment are the omitted groups in the applicable columns. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 8: Effect of integration on residential sorting (whites)

Dependent variable: similarity index of last zip code of residence

	Stayers and movers				Movers
	(1)	(2)	(3)	(4)	(5)
Integration (s_{BW})	0.057^{*}	0.060^{*}	0.036^{**}	0.036^{**}	0.042^{**}
	(0.024)	(0.030)	(0.013)	(0.014)	(0.015)
Age at casualty	-0.001	-0.001	-0.002^{*}	-0.002	-0.002
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)
Black share by regiment/period		-0.011	-0.031	-0.030	-0.085
		(0.035)	(0.025)	(0.027)	(0.060)
Last zip mean house price			0.213^{***}	0.211^{***}	0.176^{***}
			(0.013)	(0.013)	(0.028)
Last zip black share			0.444^{***}	0.441^{***}	0.401^{***}
			(0.011)	(0.010)	(0.016)
Origin county mean wage				-0.049^{***}	-0.009
				(0.007)	(0.012)
Origin county high school share				0.011	-0.219
				(0.152)	(0.217)
Origin county black share				-0.317^{**}	0.013
				(0.110)	(0.183)
Origin county rural share				-0.261^{***}	-0.000
				(0.053)	(0.073)
R^2_{Adj}	0.036	0.036	0.303	0.305	0.227
N	14,545	14,545	11,063	11,063	4,124

Note: Panel A. shows OLS results for white veterans. Panel B. show OLS results for black veterans. All regressors except "Age at casualty" are normalized. All regression include period of casualty, regiment and state-of-origin fixed effects. Standard errors two-way clustered by regiment and period of casualty. The dependent variable is a zip code similarity index is constructed using block-level demographic data from the US Census Bureau. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 9: Effect of integration on residential sorting (blacks)

Dependent variable: similarity index of last zip code of residence

	Stayers and movers				Movers
	(1)	(2)	(3)	(4)	(5)
Integration (s_{BW})	0.028	0.027	0.067	0.055	0.084
	(0.042)	(0.040)	(0.063)	(0.065)	(0.119)
Age at casualty	0.000	-0.000	0.001	0.001	-0.009
	(0.004)	(0.005)	(0.010)	(0.012)	(0.012)
Black share by regiment/period		0.008	-0.048	-0.052	0.060
		(0.071)	(0.093)	(0.101)	(0.264)
Last zip mean house price			0.104^{***}	0.102^{***}	-0.062
			(0.030)	(0.026)	(0.069)
Last zip black share			0.080^{***}	0.080^{***}	0.026
			(0.012)	(0.016)	(0.031)
Origin county mean wage				-0.046	0.065
				(0.041)	(0.043)
Origin county high school share				0.739	0.576
				(0.413)	(1.105)
Origin county black share				0.040	0.002
				(0.365)	(0.333)
Origin county rural share				0.069	0.624
				(0.282)	0.728
R^2_{Adj}	0.074	0.074	0.093	0.092	0.031
N	1,536	1,536	932	932	416

Note: Panel A. shows OLS results for white veterans. Panel B. show OLS results for black veterans. All regressors except "Age at casualty" are normalized. All regression include period of casualty, regiment and state-of-origin fixed effects. Standard errors two-way clustered by regiment and period of casualty. The dependent variable is a zip code similarity index is constructed using block-level demographic data from the US Census Bureau. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 10: Effect of integration on residential sorting (whites)

Dependent variable: similarity index of last zip code of residence

	Volunteers			Draftees		
	(1)	(2)	(3)	(4)	(5)	(6)
Integration (s_{BW})	0.107^{**}	0.085^{**}	0.082^{**}	0.196	0.157	0.158
	(0.037)	(0.026)	(0.023)	(0.117)	(0.115)	(0.118)
Black share by regiment/period	0.057	0.024	0.025	-0.290	-0.184	-0.190
	(0.077)	(0.041)	(0.042)	(0.241)	(0.229)	(0.228)
Last zip mean house price		0.244^{***}	0.244^{***}		0.193^{***}	0.196^{***}
		(0.009)	(0.012)		(0.038)	(0.040)
Last zip black share		0.454^{***}	0.453^{***}		0.478^{***}	0.478
		(0.024)	(0.026)		(0.031)	(0.041)
Origin county mean wage			-0.026			0.009
			(0.017)			(0.018)
Origin county high school share			-0.297^{**}			-0.224
			(0.110)			(0.157)
Origin county black share			-0.257^{*}			-0.213
			(0.102)			(0.674)
Origin county rural share			-0.240^{*}			0.114
			(0.094)			(0.235)
R^2_{Adj}	0.032	0.305	0.306	0.032	0.319	0.317
N	5,019	3,821	3,821	1,026	787	787

Note: Panel A. shows OLS results for white veterans. Panel B. show OLS results for black veterans. All regressors except "Age at casualty" are normalized. All regression include period of casualty, regiment and state-of-origin fixed effects. Standard errors two-way clustered by regiment and period of casualty. The dependent variable is a zip code similarity index is constructed using block-level demographic data from the US Census Bureau. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 11: Effect of integration on residential sorting (blacks)

Dependent variable: similarity index of last zip code of residence

	Volunteers			Draftees		
	(1)	(2)	(3)	(4)	(5)	(6)
Integration (s_{BW})	-0.060	0.127	0.110	-0.058	0.249	0.221
	(0.125)	(0.120)	(0.132)	(0.732)	(0.301)	(0.258)
Black share by regiment/period	0.058	-0.088	-0.099	0.551	0.628*	0.568
	(0.113)	(0.113)	(0.103)	(0.234)	(0.266)	(0.374)
Last zip mean house price		0.264***	0.263***		0.062	0.077
		(0.082)	(0.078)		(0.112)	(0.159)
Last zip black share		0.146***	0.141***		0.048	0.038
		(0.036)	(0.035)		(0.010)	(0.128)
Origin county mean wage			−0.050			-0.151
			(0.069)			(0.017)
Origin county high school share			0.643			-2.583
			(0.669)			(2.616)
Origin county black share			0.265			-0.297
			(0.574)			(1.919)
Origin county rural share			0.112			-0.881
			(0.552)			(1.448)
R^2_{Adj}	0.080	0.129	0.116	0.163	0.079	0.039
N	482	280	280	134	77	77

Note: Panel A. shows OLS results for white veterans. Panel B. show OLS results for black veterans. All regressors except "Age at casualty" are normalized. All regression include period of casualty, regiment and state-of-origin fixed effects. Standard errors two-way clustered by regiment and period of casualty. The dependent variable is a zip code similarity index is constructed using block-level demographic data from the US Census Bureau. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 12: Effect of integration on intermarriage

Dependent variable: wife's non-white/black name index

	OLS			Fractional
	(1)	(2)	(3)	(4)
	A. Wife's non-white name index			
Integration (s_{BW})	0.018^{**}	0.020^{**}	0.018^{*}	0.335^{***}
	(0.005)	(0.007)	(0.008)	(0.128)
				$[0.031]^{*}$
Black share origin county	0.009	0.011	-0.011^{***}	-0.045
	(0.006)	(0.006)	(0.003)	(0.039)
Age at casualty	-0.001	-0.001	-0.001	-0.009
	(0.012)	(0.012)	(0.012)	(0.022)
R^2_{Adj}	0.003	0.000	0.006	0.048
	B. Wife's black name index			
Integration (s_{BW})	0.022^{*}	0.024^{*}	0.024^{**}	0.510^{***}
	(0.009)	(0.011)	(0.008)	(0.169)
				$[0.038]^{***}$
Black share origin county	0.006	0.007	-0.015	-0.061
	(0.008)	(0.009)	(0.009)	(0.046)
Age at casualty	0.005	0.004	0.003	0.004
	(0.008)	(0.009)	(0.008)	(0.027)
R^2_{Adj}	0.001	-0.003	0.001	0.065
Period FE	X	X	X	X
Rank FE		X	X	X
State of res FE			X	X
N	664	664	664	664

Note: Panel A. regresses the wife's non-white name index on the similarity measure. Panel B. regresses the wife's black name index on the similarity measure. All regressors except "Age at casualty" or normalized. Standard errors two-way clustered by state of origin and period of casualty. Columns (1)-(3) reports OLS estimates. Column (4) estimates the parameters in the fractional response model using pseudo-MLE. The marginal effect is shown in square brackets. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 13: Effect of integration on intermarriage

Dependent variable: wife's non-white/black name index

	OLS			Fractional
	(1)	(2)	(3)	(4)
		A. Wife's non-white name index		
Integration (s_{BW})	0.002	0.002	0.007	-0.058
	(0.006)	(0.007)	(0.008)	(0.110)
				[0.008]
Black share origin county	0.003	0.003	-0.007	-0.025
	(0.005)	(0.007)	(0.013)	(0.040)
Age at casualty	-0.007***	-0.007***	-0.007***	-0.029***
	(0.002)	(0.002)	(0.002)	(0.007)
R^2	0.011	0.015	0.026	0.039
		B. Wife's black name index		
Integration (s_{BW})	0.005	0.004	0.008	-0.012
	(0.004)	(0.007)	(0.007)	(0.131)
				[0.007]
Black share origin county	-0.002	-0.003	-0.019	-0.095**
	(0.004)	(0.005)	(0.011)	(0.046)
Age at casualty	-0.005**	-0.005**	-0.005**	-0.030***
	(0.002)	(0.002)	(0.002)	(0.009)
R^2	0.011	0.012	0.026	0.062
Period FE	X	X	X	X
Rank FE		X	X	X
State of res FE			X	X
N	901	901	901	901

Note: Panel A. regresses the wife's non-white name index on the similarity measure. Panel B. regresses the wife's black name index on the similarity measure. All regressors except "Age at casualty" or normalized. Standard errors two-way clustered by state of origin and period of casualty. Columns (1)-(3) reports OLS estimates and the adjusted R^2. Column (4) estimates the parameters in the fractional response model using pseudo-MLE and reports the pseudo-R^2. The marginal effect is shown in square brackets. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 14: Balance test for Similarity Index

Dependent variable: Similarity Index				
	q_1 (1)	$q_2 - q_1$ (2)	$q_3 - q_1$ (3)	$q_4 - q_1$ (4)
		A. First half of war		
From north	0.4559	0.0139	0.0264	0.0414
		(0.009)	(0.008)	(0.009)
From south	0.3105	0.0794	0.0152	−0.0061
		(0.008)	(0.008)	(0.008)
Age at casualty	23.281	−0.274	−0.869	−0.858
		(0.14)	(0.14)	(0.15)
N	25,978	25,978	25,978	25,978
		B. Second half of war		
From north	0.5313	−0.0049	−0.0260	−0.0073
		(0.011)	(0.011)	(0.012)
From south	0.3011	0.0034	0.0363	0.0055
		(0.010)	(0.010)	(0.011)
Age at casualty	22.457	0.080	−0.072	−0.257
		(0.14)	(0.13)	(0.14)
N	15,058	15,058	15,058	15,058

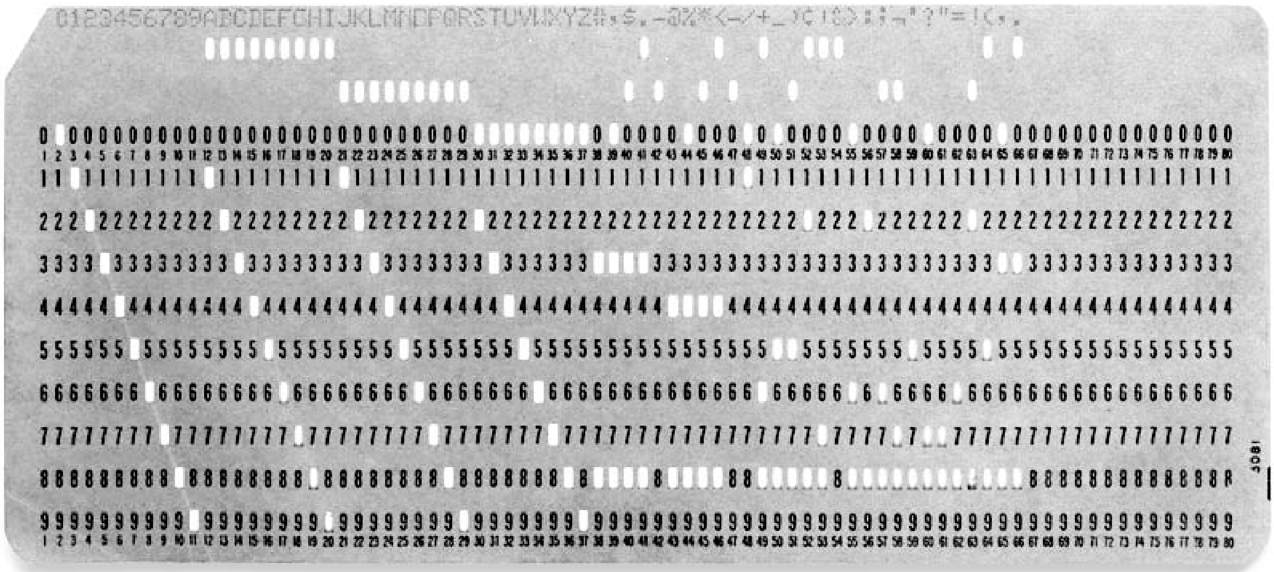
Note: Panel A. shows the balance test for the first half of the war (25-Jun-1950 to 1951-12-22). Panel B. show the balance test for the second half of the war (1951-12-23 to 27-Jul-1953). Column (1) shows the average of the regiments whose similarity fall within the first quartile (i.e. least integrated). Columns (2)-(4) show the difference in mean of the top three quartiles relative to the bottom quartile. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

IX Figures

Figure 1: Racial discrimination and black protest

Note: Black custodians protesting discrimination by defense contractors. The "double V" symbol—one "V" formed by the men, the second by the brooms—became the rallying cry for civil rights during the Second World War. The first "V" signifies the struggle for victory over the Axis powers, while the second "V" signifies victory of discrimination at home. RG-208-NP-1KK-1. Office of War Information, National Archives.

Figure 2: IBM casualty punch card

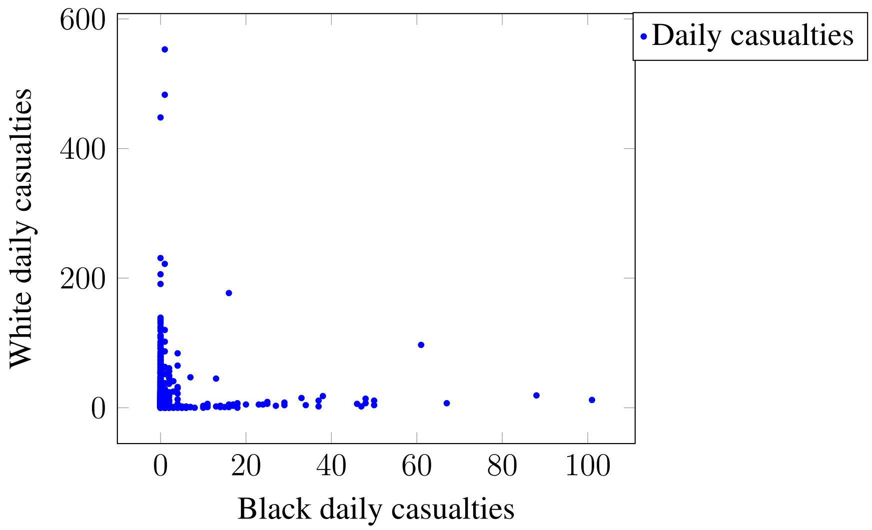


Note: IBM punch card used to document casualties during the war.

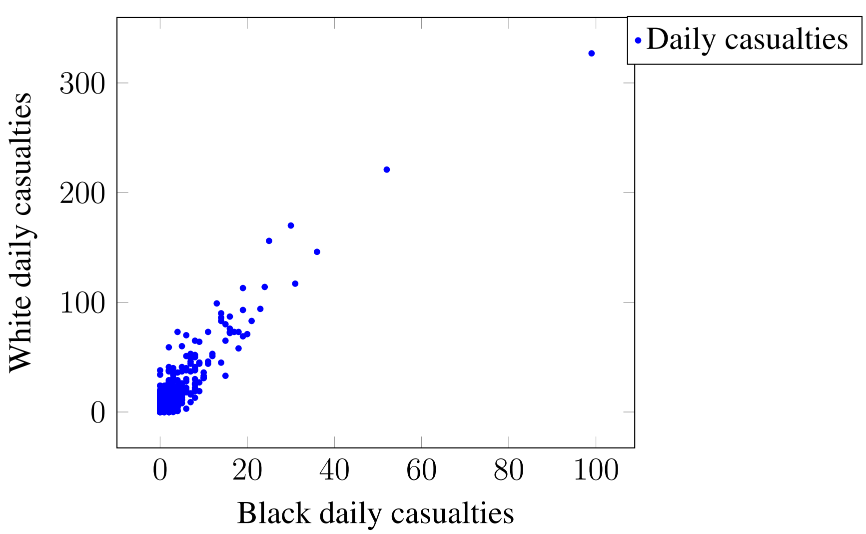
Figure 3: Example of early integration: 9th Infantry Regiment (November 1950)

Note: Early in the Korean War, the all-white 1st and 2nd battalions of the 9th Infantry Regiment, like many other all-white combat units, became severely depleted. In desperation, the commanders of the 9th Infantry willingly integrated black replacement troops. This photo, taken sometime in November 1950, shows a recently integrated fire team in a foxhole.

Figure 4: Black white casualty correlation early and late in the Korean War



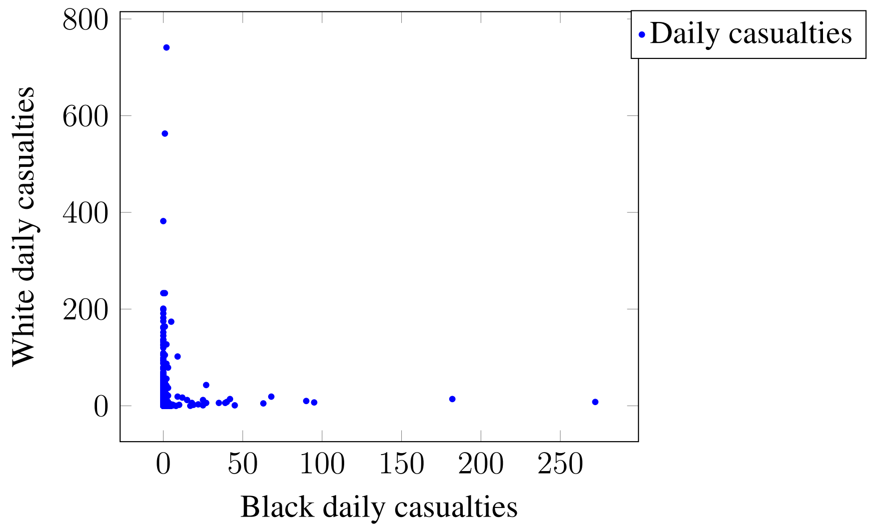
(a) First three months of war



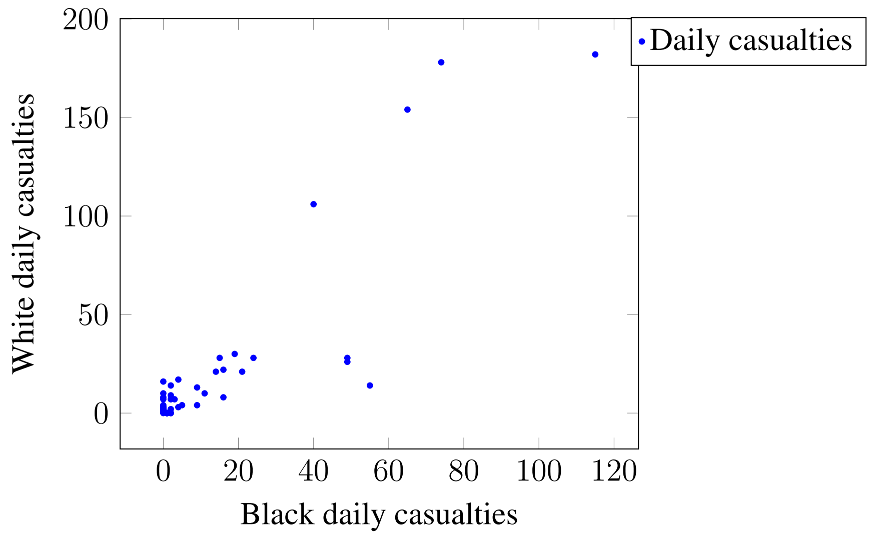
(b) Last year of war

Note: Each dot represents a day in the war.

Figure 5: Black white casualty correlation across regiments in the same time period



(a) Months 4-6 excluding 9th Regiment



(b) Months 4-6, 9th Regiment only

Note: Each dot represents a day in the war.

Figure 6: Histogram for regimental similarity index

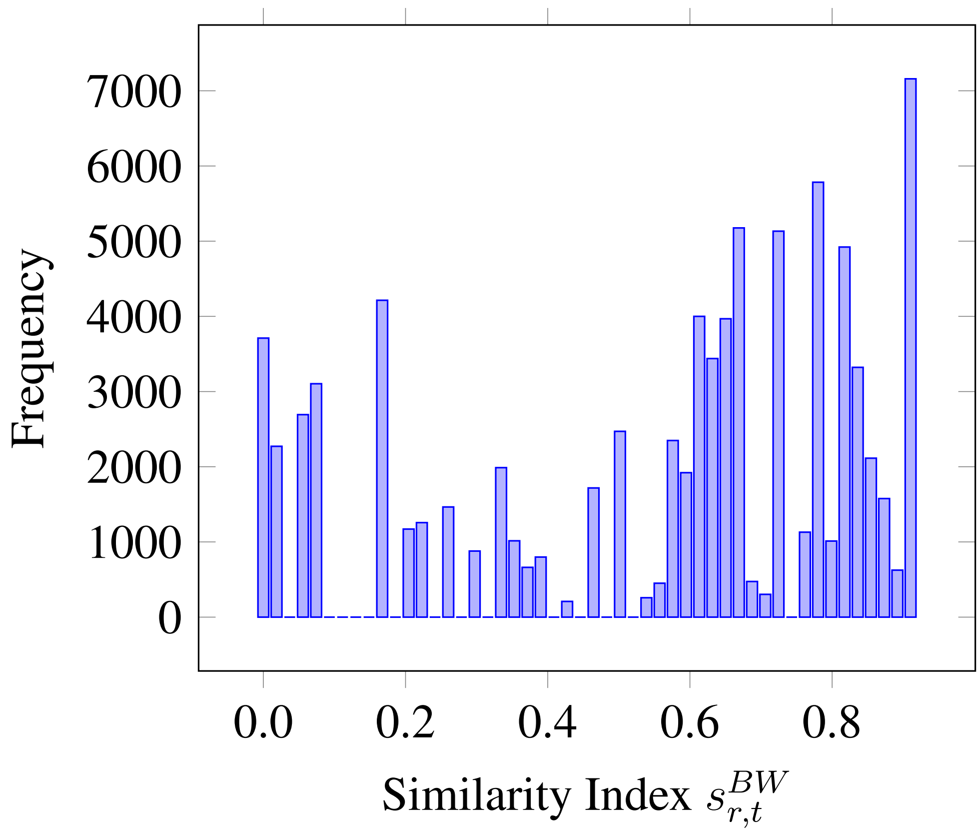


Figure 7: Summary of merged datasets

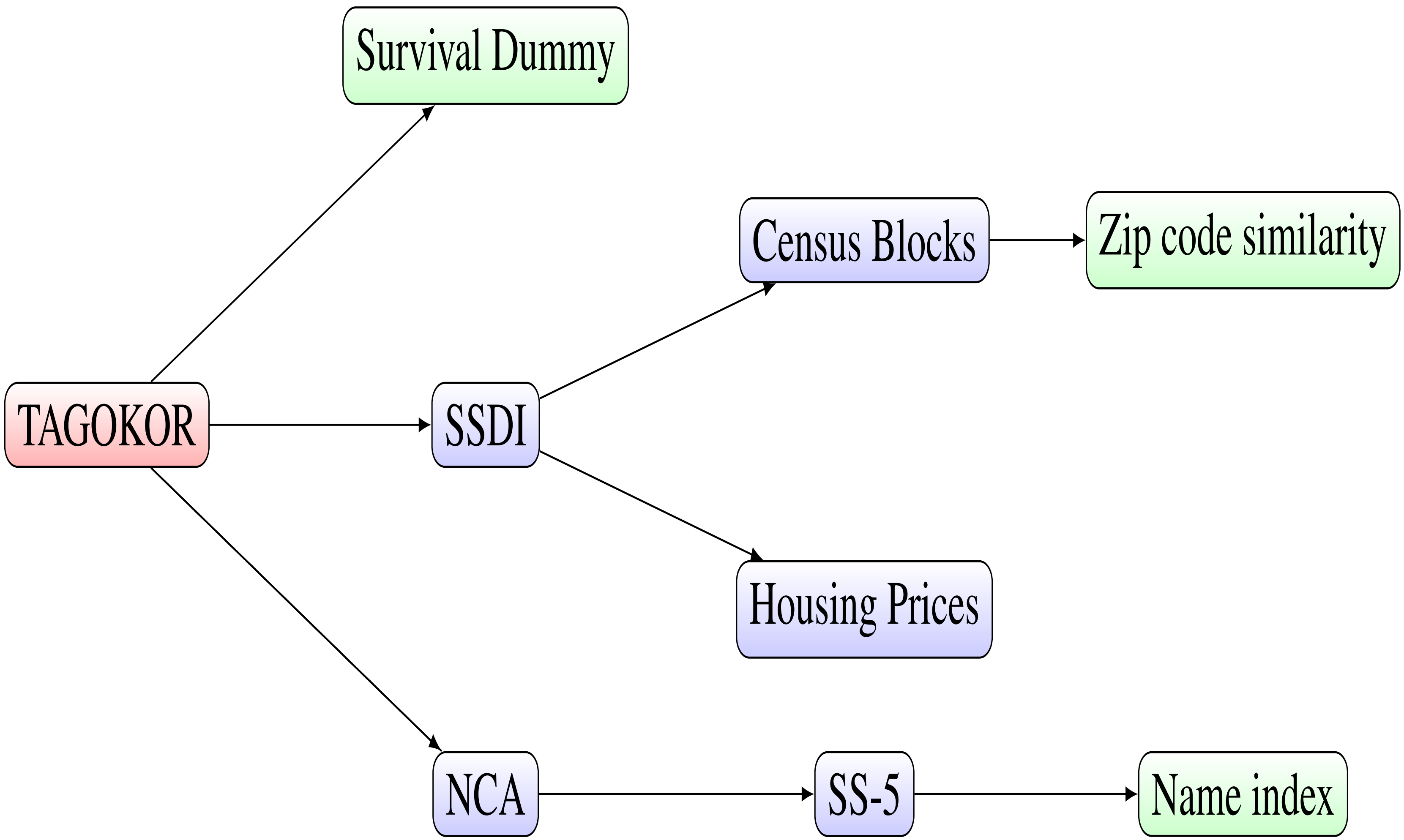
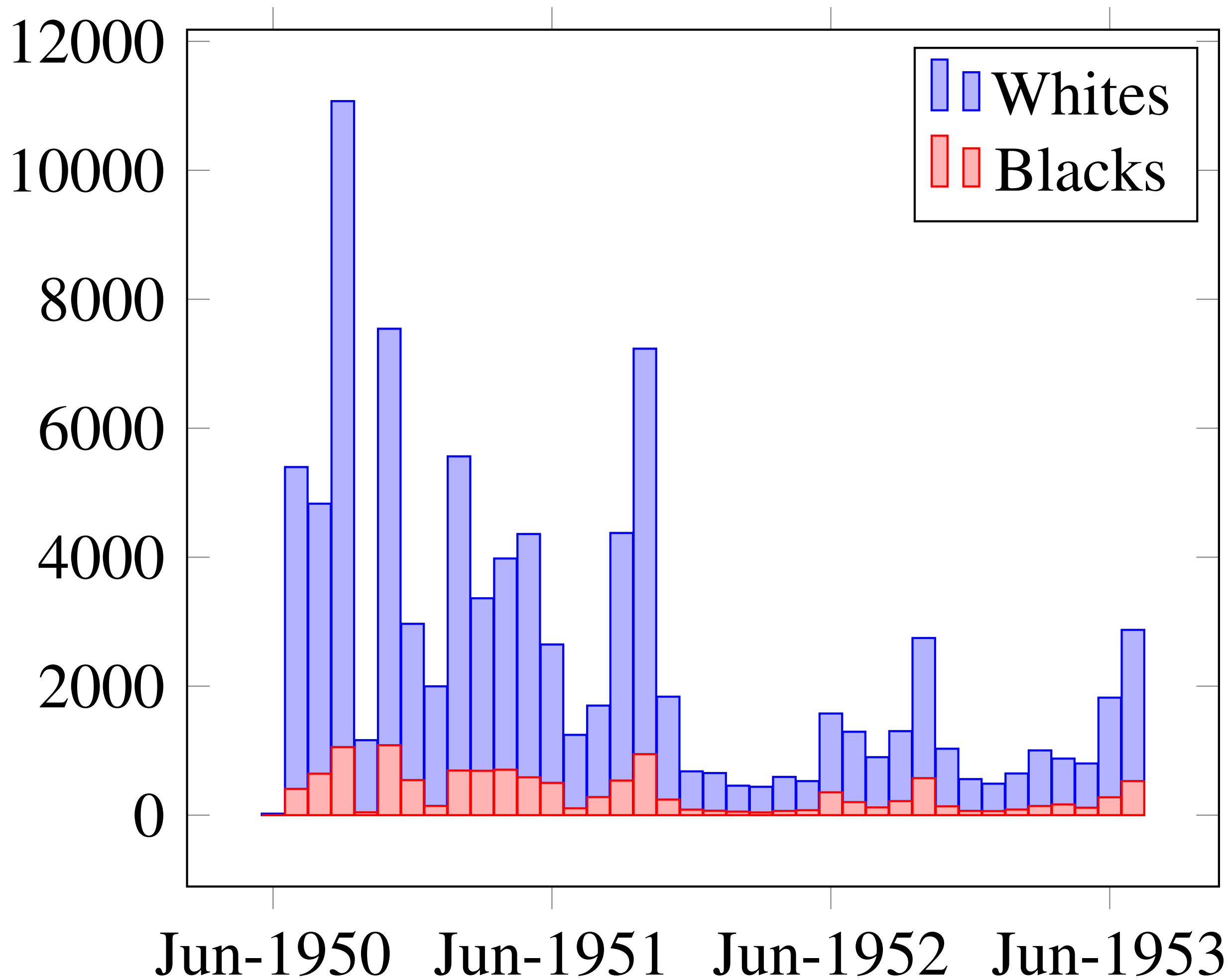
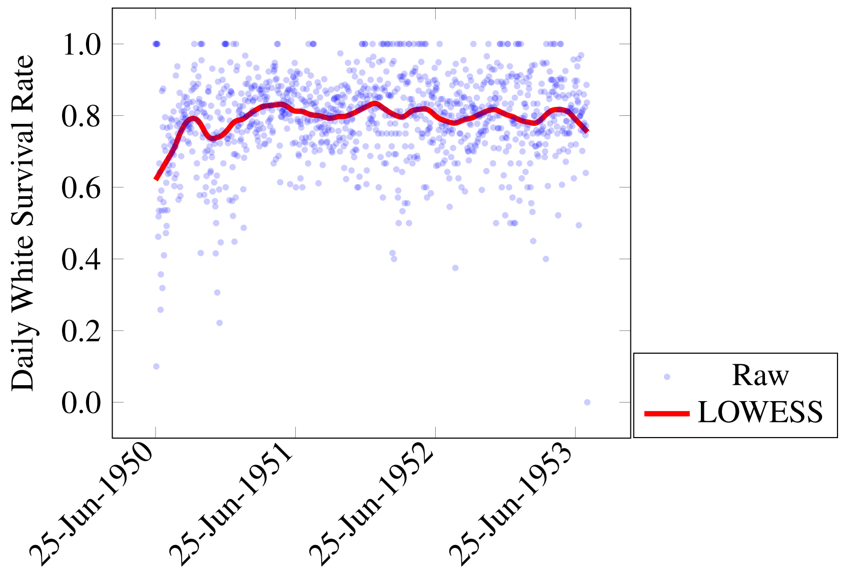


Figure 8: Raw number of casualties in each month of the war

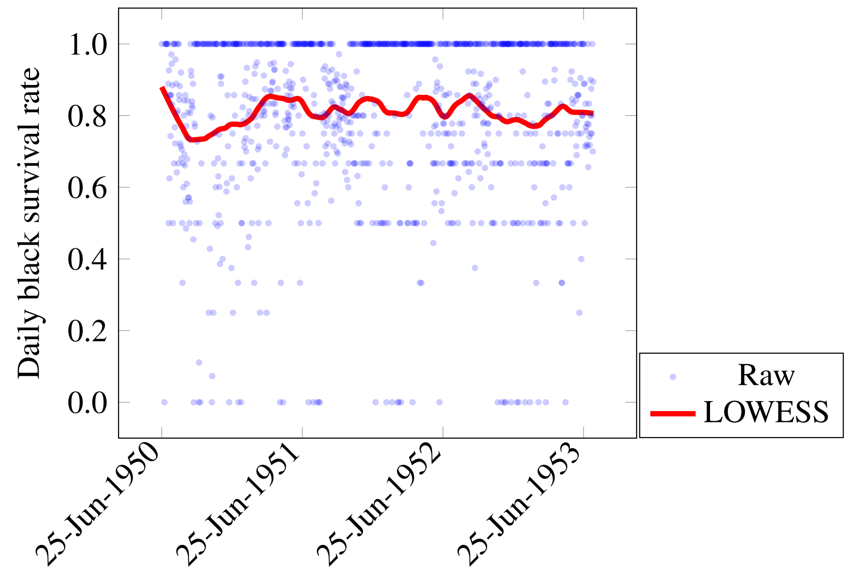


Note: Text (forthcoming)

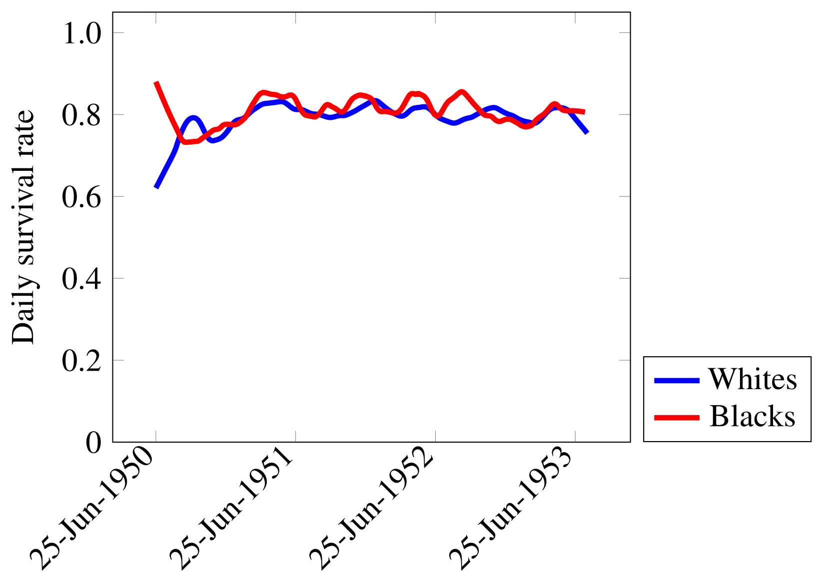
Figure 9: Daily survival rates through out the war by race



(a) Note: Each dot represents the average survival rate of wounded white soldiers on a given day in the war.

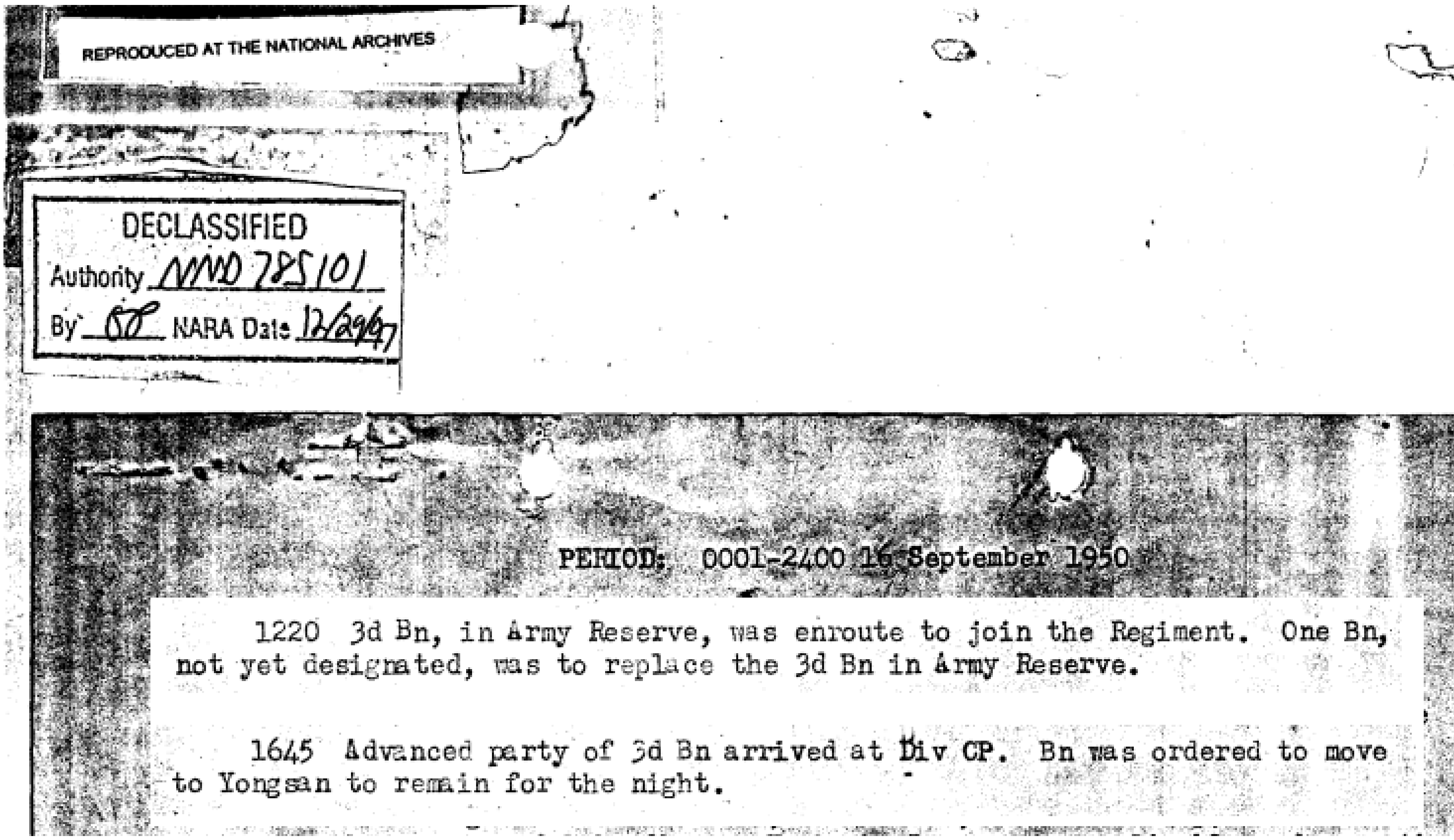


(b) Note: Each dot represents the average survival rate of wounded black soldiers on a given day in the war.



(c) Note: Figures 9a and 9b combined.

Figure 10: 9th Infantry Regimental War Diary (Sep 1950)



REPRODUCED AT THE NATIONAL ARCHIVES

DECLASSIFIED
Authority NND 785101
By OP NARA Date 12/29/97

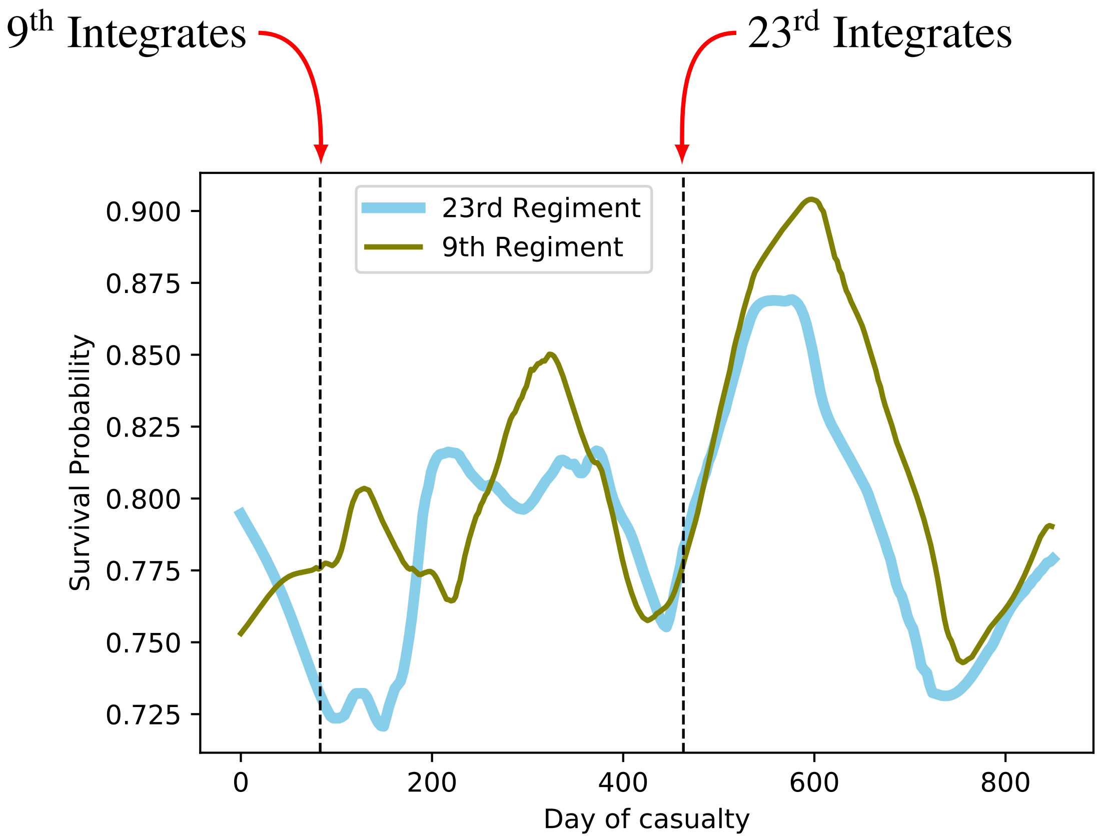
PERIOD: 0001-2400 16 September 1950

1220 3d Bn, in Army Reserve, was enroute to join the Regiment. One Bn, not yet designated, was to replace the 3d Bn in Army Reserve.

1645 Advanced party of 3d Bn arrived at Div CP. Bn was ordered to move to Yongsan to remain for the night.

Note: Regimental war diary for the 9th Regiment

Figure 11: Casualty survival rates for 9th and 23rd regiments

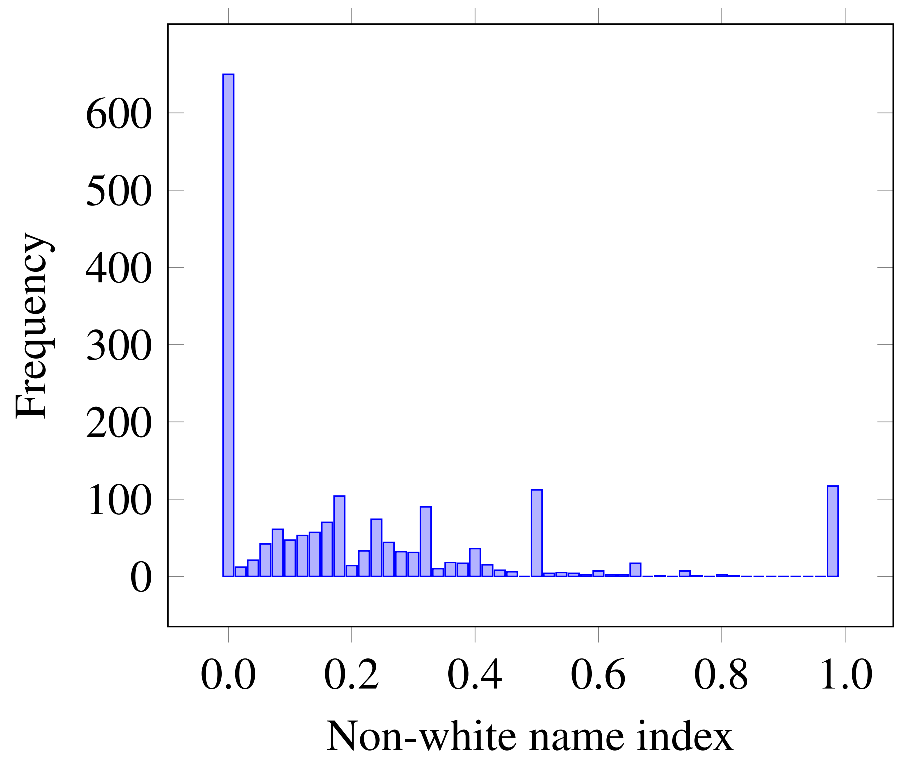


Note: The plot depicts the respective casualty survival rates of the 9^{th} and 23^{rd} Regiments, both of which belonged to the 2^{nd} Infantry Division. The survival rates were estimated from daily casualty data using LOWESS. The first vertical line denotes 16-Sep-1950, and the second vertical line denotes 1-Oct-1951.

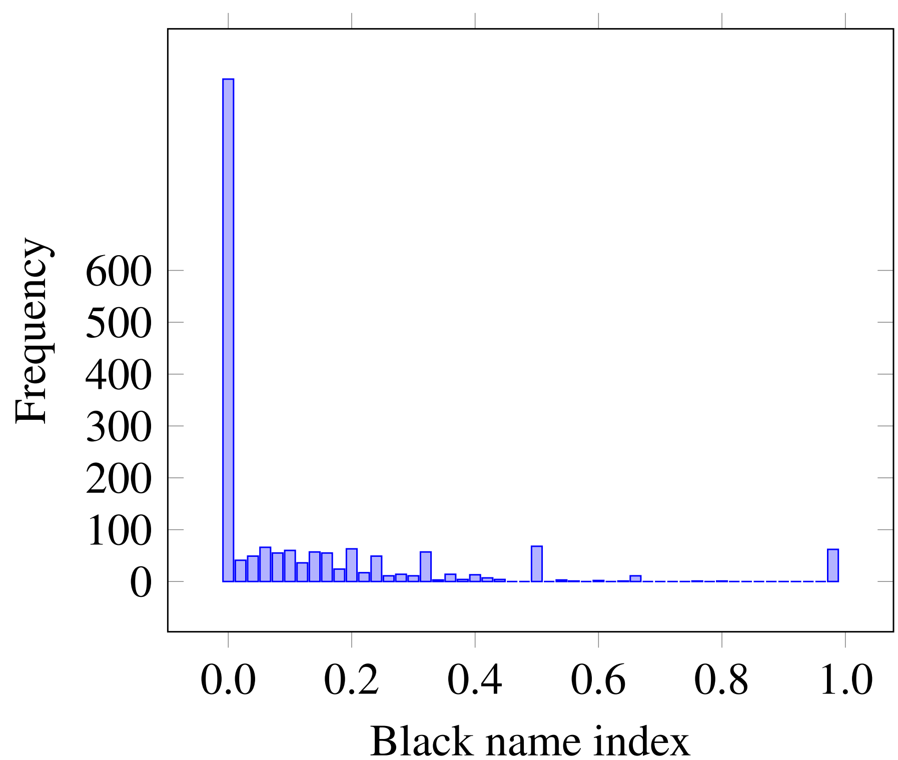
Appendix I: Matching casualty data to Social Security Death Index (forthcoming)

Appendix II: Matching casualty data to National Cemetery Administration Data (forthcoming)

Figure 12: Racial content of wife's name



(a) Wife's black name index



(b) Wife's non-white name index

Note: Name indexes for wives of white veterans who were 22 or younger during Korean War

Figure 13: Similarity index for all US zip codes (2010, standardized)