A Farewell to Army Segregation: The Effects of Racial Integration During the Korean War

Daniel Indacochea*
Department of Economics
University of Toronto

May 2020

Most recent version here.

Abstract

The racial integration of the U.S. Army during the Korean War (1950-1953) is one of the largest, swiftest desegregation episodes in American history. Integration began in an effort to reinforce badly depleted all-white units, and went on to become Army-wide policy for reasons of military efficiency. The first part of this paper evaluates whether the Army achieved its goal of improving efficiency as measured by the survival rates of wounded soldiers. Using casualty data, I develop a novel wartime integration measure to quantify exogenous changes in racial integration over time and across regiments. Based on a two-way fixed effects model, I find that a one standard deviation increase in regimental integration improved overall casualty survival rates by 3%. The second part of the paper explores the effects of wartime racial integration on the prejudicial attitudes of veterans after the war. To do so, 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 one standard deviation increase in wartime racial integration caused white veterans to live in more racially diverse neighborhoods and marry spouses with less distinctively white names. These results provide suggestive evidence that large-scale interracial contact reduces prejudice on a long-term basis.

Keywords: *Racial Integration; Military Efficiency; Preference Formation; Residential Sorting, Korean War.*
JEL Codes: J12, J15, N42.

---

* I thank Eli Berman, Francine Blau, Alexander Chernoff, Shari Eli, David Finer, Alastair Fraser, Jingyuan Huang, Lawrence Kahn, Dena Krieger, James MacKinnon, Robert McMillan, Ismael Mourifié, Swapnika Rachapalli, David Rosé, Marc-Antoine Schmidt, Aloysius Siow, seminar participants at the University of Toronto, the 13th Defense and Security Economics Workshop, and the 53rd Annual Conference of the CEA, and veterans at the KWVA 2019 Annual Membership Meeting for valuable comments and feedback. Financial support from SSHRC is gratefully acknowledged. Any errors are my own. E-mail: daniel.indacochea@mail.utoronto.ca.

1 Introduction

The racial integration of the Eighth U.S. Army during the Korean War (1950-1953) is one of the largest, 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, while blacks were assigned primarily to non-combat support roles.[1] Following the deadly campaigns in the early years of the Korean War, many all-white combat units became severely depleted, and in a bid to bolster manpower, some of these depleted units began to accept black reinforcements as a temporary emergency measure (Blair, 1989). Thus integration began informally, well before it was Army-wide policy.

At the time, the Army feared that racial strife would hinder combat performance in these newly integrated units. So as a precaution, it 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 (Hausrath, 1954). Code-named "Project Clear," the primary concern of this study was to determine what effect, if any, racial integration had on combat performance. Project Clear offered some reassurance in this regard, as white officers perceived little difference in the performance of racially-mixed rifle squads under their command.[2] Given the circumstances, however, Project Clear had several drawbacks: black soldiers were never interviewed, interviews only took place once, and the project relied on self-reported measures of unit performance.

This raises the question what would more objective measures reveal about the performance of newly integrated units, or the treatment of blacks within those units? The first part of this paper addresses these questions by taking a novel quantitative approach. The approach requires a measure of integration that varies across units over the course of the war. In an idealized setting, the researcher would have access to data that tracked the racial

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

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

composition of units at the smallest level over the course of the war. With such granular data, the researcher could then calculate a similarity index to measure wartime integration of larger units over time.[3] Yet to the best of my knowledge, no such data are readily available. Although military archives contain wartime "Morning Reports" which tracked the daily flow of soldiers in and out of companies, those files report neither the race nor the total stock of soldiers within those companies.[4] In the absence of data on the racial composition of small units, I am nevertheless able to construct a measure of integration thanks to a remarkably rich casualty dataset covering all 103,000 individual Korean War Army veterans who were killed or wounded during the war—a file that reports each soldier's race, regiment and date of casualty.

Such casualty data make it possible to quantify the extent of integration: intuitively, if black and white soldiers were physically placed in the same foxholes, then the timing of their injuries should coincide highly. More concretely, for each regiment, I calculate the daily distribution of black and white casualties at nine different points during the war. Based on these temporal casualty distributions, I construct a regimental similarity index which I use as my measure of integration throughout the paper. I find that between the first and last three months of the war, the regimental similarity index rose from 8% to 60%. This stark contrast is not driven by unobserved trends. Using regimental war diaries, I present a case study that shows credibly that integration also varied across units within a given time period. Having variation both across regiments and over time then allows me to use panel methods to model the effects of integration on various outcomes of interest.

Considering that the immediate goal of integration was to reinforce depleted all-white combat teams, I evaluate the effect of integration on team performance by modeling the

[3] In the residential segregation literature, researchers routinely calculate dissimilarity indexes—the complement of a similarity index—based on the spatial distribution of blacks and whites within some geographic unit. For example, Cutler et al. (1999) use tract-level demographic data to calculate similarity indexes for metropolitan statistical areas across the U.S.

[4] Alternatively, a researcher could in theory reconstruct individual assignments from Official Military Personnel Files (OMPF), but 80% of Korean War era OMPFs were destroyed in a disastrous archival fire in 1973. Moreover, the files that did survive cannot be accessed by archivists (or other researchers) until 62 years after the service member's separation from the military.

survival rates of wounded soldiers. Conceptually, by accepting black reinforcements, all-white combats teams increased their manpower, which in turn made them less vulnerable to enemy attacks and better equipped to evacuate wounded soldiers. The identifying assumption is that those combat teams that did not accept blacks serve as suitable counterfactuals for those that did. Empirically, this raises the concern that the first combat teams to integrate blacks might have benefited from unobserved leadership quality. To account for this possibility, I include regiment fixed effects in each specification. Similarly, I include time fixed effects to mitigate concerns that the similarity index is merely capturing unobserved time trends.

Using this two-way linear fixed effects model, I find that a one standard deviation (SD) increase in integration caused a 4% increase in the survival rates of wounded white soldiers relative to those who remained segregated. By contrast, the results indicate that a one SD increase in integration caused a 3% *decrease* in the survival rates of wounded black soldiers. Given that blacks were initially confined to non-combat support roles, these results are best viewed as a convergence in hazardous combat assignments. What is less obvious is how racial integration affected overall survival rates. Based on a pooled sample of blacks and whites, I find that a one SD increase in integration caused a 4% increase in the *overall* survival rates of wounded soldiers. Of course, one should expect integration to be correlated with the unobserved jeopardy that combat teams were in, which means this effect size is likely an underestimate. Taking the view that black and white soldiers are close substitutes, this result is consistent with the hypothesis that segregation impeded the efficient allocation of scarce labor.[5]

This positive effect of integration on overall survival rates is perhaps surprising given the vast number of empirical studies documenting the adverse effects of group heterogeneity on various outcomes, such as team performance or the provision of public goods. Alesina et al. (1999) show that ethnically fractionalized counties in the U.S. spend a smaller share

[5] One limitation of the casualty data is that they do not include the overall size of regiments at a given point in the war. Some regimental war diaries contain data on regiment size, which I am in the process of converting into a machine-readable format.

on productive goods such as education and infrastructure. In the same vein, Adams and Ferreira (2009) find that directors in gender-diverse corporate boards in the U.S. have better attendance records but that these firms are more vulnerable to costly outside takeovers. Similar efficiency losses have also 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.[6] A major difference between the Korean War and these other settings, however, is that in addition to composition, racial integration changed the size of regiments. In a depleted all-white combat unit, the marginal productivity of an additional black soldier was relatively high compared to the over-strength support units that blacks were drawn from. Thus, the experiment I consider is what happens to regiment (firm) productivity, as measured by aggregate survival rates, when restrictions on where soldiers (labor) can be allocated are removed. My results indicate that the efficiency gained from freely allocating labor exceeded any adverse effects of racial integration.

From a longer-term perspective, the natural restrictions that war places on selective migration and sorting present a rare opportunity to investigate the effects of interracial contact on prejudice.[7] Without such restrictions, highly prejudiced individuals would be able to self-select out of contact, and standard econometric analysis would overstate the social benefit of intergroup contact. The Army's unique ability to compel its soldiers is in stark contrast to civilian settings, where large-scale desegregation in turn propelled white flight. For example, Reber (2005) finds that actively busing black students into previously all-white schools *accelerated* the exodus of white families moving into the suburbs; likewise, during the Second Great Migration (1940-1970), Boustan (2007) estimates that 2.5 white urban residents departed for every black arrival.

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

[7] 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.

Although enlistment contracts required soldiers to remain in the Korean Peninsula, one possibility is that integration led to a kind of preemptive white flight: would-be volunteers with a distaste for integration were dissuaded from joining the Army as news of racial integration reached the home front. However, I find no evidence that this took place. Between the first and second halves of the war, the fraction of wounded white volunteers who came from the Deep South—historically the most prejudiced region in the country—*increased* by 0.5 percentage points.[8] Moreover, once deployed, 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 soldiers held the rank of corporal or lower (see Table 1 for a hierarchy of ranks), and results from balance tests reported in Table 3 indicate that white Southerners were just as likely to be integrated as white Northerners.

One of the fascinating aspects of racial integration during the Korean War is that we have contemporaneous evidence regarding the effect of integration on prejudice. Researchers from Project Clear interviewed white officers about their attitudes towards blacks and found that—compared to their segregated peers—officers commanding integrated platoons reported *less* anti-black prejudice (Bogart, 1992). Given that integration appeared to reduce prejudice contemporaneously, did it also produce any lasting effects?

To address this issue, the second part of the paper studies the long-run effects of wartime integration by examining the post-war behavior of veterans toward people outside their race group. The two behavioral measures I consider are (i) where veterans chose to live and (ii) whom they married. To study veterans' residential choice, I link veterans in the casualty file to social security death data using the expectation maximization (EM) algorithm following Abramitzky et al. (2019). These linked data enable me to observe the deceased veteran's last zip code of residence, for which I in turn construct a residential similarity index using block-level Census data. Based on a sample of veterans born after 1930, I find that a one SD increase in wartime integration increased residential integration by 0.1 SD—roughly the

[8] The definition of the Deep South used here are the states of Alabama, Georgia, Louisiana, Mississippi, and South Carolina.

difference in integration between present-day Central Harlem and the Upper East Side. To further mitigate sample-selectivity concerns, I repeat the analysis on a sample restricted to draftees which yields a larger, albeit statistically insignificant, estimate of the effect size. An arguably higher bar for prejudice reduction is exogamy—marrying outside one's own race or ethnic group. To obtain data on wives, I again use the EM algorithm but this time to link soldiers to national cemetery data. Since at this stage 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 intermarriage, I find that a one SD increase in wartime integration led to a 2 percentage point (or 0.08 SD) increase in the wife's non-white name index. For both residential choice and intermarriage, I replicate this analysis for veterans born before 1930 and find no effect.

The circumstances under which integration in Korea came about conform to the conditions of the so-called "contact hypothesis" (Allport, 1954). In particular, Allport (1954) hypothesizes that intergroup contact will reduce prejudice if two groups have (i) equal status (e.g. age/rank); (ii) common goals (e.g. a common enemy); (iii) intergroup cooperation (e.g. combat teams); and (iv) support of authorities (e.g. officers). A consistent finding from social psychology is that this type of intergroup 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 the laboratory across a variety of observational settings (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., 2019).[9]

One drawback of real-world applications is that they are often contaminated by network effects (Camargo et al., 2010). If, say, a white individual were assigned a black roommate and reports having formed more black friendships afterward, in this context the researcher

[9] In a closely related study, Schindler and Westcott (2015) find that U.K. counties that experienced a one SD increase in exposure to black G.I.s during WWII had 0.03 SD fewer members of far-right nationalist parties in 2010. This paper differs from Schindler and Westcott (2015) in that I consider a treatment group who had direct contact with blacks.

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. The Korean War provides a suitable approximation to such a scenario: 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 lived or whom they married, network effects are less pervasive compared to other physical settings, and any change in behavior toward the out-group plausibly reflects a genuine change in prejudice—be it in attitudes or beliefs. Overall, these results relating to veterans' post-war behavior provide suggestive evidence that large-scale interracial contact reduces prejudice on a long-term basis.

The remainder of the paper is organized as follows. Section 2 describes the origins of racial segregation in the U.S. military and why, during the Korean War, this system was ultimately abandoned. Section 3 presents the casualty data and formally defines the regimental similarity index used to measure wartime integration. Section 4 investigates outcomes during the war—namely, the effects of integration on casualty survival. Section 5 shifts the focus to the post-war period by introducing the residential data and investigating how integration influenced where veterans ultimately chose to locate. Section 6 presents the national cemetery data that identifies wives, details the construction of the non-white name index, and analyzes the effects of integration on intermarriage. Section 7 concludes.

2 Historical Background

2.1 The origins of racial segregation in the military

From a historical perspective, blacks served in major U.S. conflicts beginning in the colonial militias and continuing through the American Revolution (Mershon and Schlossman, 1998). Following the founding of the United States, the history of blacks in the military closely mirrors that of American civil society. In 1783, federal and state legislators passed laws to restrict military service to whites only, though these prohibitions were ignored to meet manpower needs in times 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 advancing 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, requiring the Army to maintain four all-black regiments on a permanent basis. A federal statute in 1867 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) laws reflected the new paradigm apparent in 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.

2.2 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 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 concerning 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 required to support the war effort in Europe would need to include blacks. Yet 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 deployment in Europe (Dalessandro et al., 2009).

2.3 Black activism and Executive Order 9981

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,[10] and the Marine Corps continued to exclude blacks outright.

The growing activism and political influence of the black electorate began to turn the tide of racial policies within the military. In July 1948, President Harry Truman issued Executive Order 9981, which stipulated equal treatment and opportunity regardless of race or ethnicity across all branches of the military. Truman made it clear that this decree ultimately meant blacks and whites were to be integrated. Like subsequent integration directives in the context of public schools, however, 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.

2.4 The Korean War (1950-1953)

Despite the fact that 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 2). Traditional Army structure is formed in threes:

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

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 all-black battalion, but smaller units were either all-white or all-black.[11]

In addition to being segregated across units, blacks and whites 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). One of the few advances in racial policy the Army had made during the intra-war years was to remove race quotas that limited black manpower; yet because Truman's executive order prohibited the formation of new all-black units, the Army simply over-staffed those already in existence.

The early years of the Korean War were one of the most deadly campaigns in American military history, and many of the all-white combat units became severely depleted. This can be seen in Figure 2, which plots the monthly casualties throughout the war, and shows that the majority of both black and white casualties occurred early on.

In a bid to bolster manpower, some commanders of severely 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 1st 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

[11] All-black units could have either black or white officers.

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 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, segregation 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 24$^{\text{th}}$ 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.

3 Measuring Wartime Integration

In this section, I first introduce the casualty data; second, I explain how the timing of black and white casualties conveys information about integration; and third, I formally define the casualty-based similarity index used as my measure of integration throughout the rest of the paper.

3.1 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 recorded as casualties (i.e. killed or wounded) during the Korean War. This casualty information was recorded in theater using IBM punch cards (Figure 3) by teams of specialized units called 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 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 casualty, 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 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 soldiers to their post-war outcomes (see appendix The Expectation Maximization Algorithm for details). Table 4 contains an example of a typical observation contained in the TAGOKOR casualty file.

The summary statistics reported in Table 5 highlight the circumstances under which racial integration took place. Panel A summarizes each variable in the period before formal integration was announced on October 1, 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 blacks were not gratuitously put at risk once they were integrated.

Table 5 also addresses 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 bulk of this change stems from the introduction of the draft. Of the 5 million American soldiers who served globally during the Korean War era, 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 of 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. Moreover, when we further restrict our attention to the six states that form the Deep South, the number of white volunteers actually increased by 0.6 percentage points. To further mitigate sample selectivity concerns, I repeat my analysis of residential sorting below with a sample restricted to draftees.

3.2 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, 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 nearing completion, 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 9$^{\text{th}}$ Regiment presents a useful case study. The 9$^{\text{th}}$ Regiment was composed of three battalions: the all-white 1$^{\text{st}}$ and 2$^{\text{nd}}$ Battalions and the all-black 3$^{\text{rd}}$ 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 3$^{\text{rd}}$ Battalion. The photograph in Figure 5 shows a newly-integrated fireteam from the 9$^{\text{th}}$ Regiment in a foxhole in November 1950. By this time, the other thirty regiments had not begun to integrate. Figure 6a 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 6a only plots daily black and white casualties between October 1950 to December 1950 for the newly-integrated 9$^{\text{th}}$ Regiment. The casualty pattern now bears a striking resemblance to that of the last year of the war, by which time racial integration on the Korean Peninsula was nearing completion.

3.3 A casualty-based similarity index

Throughout this paper, I quantify these plausibly exogenous changes in wartime integration using a regimental similarity index.[12] Let B_{rd} and W_{rd} be the number of black and white casualties in regiment r on day d of the war. I divide the war into nine equal periods denoted by $t \in \{1, ..., 9\}$—i.e. nine periods of about 120 days, or 37 months divided by 9. Let D_t denote the set of days in period t (e.g. $D_1 = \{1, \ldots, 120\}$). The total number of blacks/whites casualties in regiment r during period t of the war is given by

$$B_{rt}^{Total} \equiv \sum_{d \in D_t} B_{rd}, \tag{1}$$

$$W_{rt}^{Total} \equiv \sum_{d \in D_t} W_{rd}, \tag{2}$$

for $t \in \{1, ..., 9\}$. We can then define similarity index for period t as

$$s_{rt}^{BW} \equiv 1 - \frac{1}{2} \sum_{d \in D_t} \left| \frac{B_{rd}}{B_{rt}^{Total}} - \frac{W_{rd}}{W_{rt}^{Total}} \right|. \tag{3}$$

By construction, s_{rt}^{BW} can take on values between 0 and 1; its distribution is plotted in Figure 7. An s_{rt}^{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_{rt}^{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.

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

4 Wartime Outcomes

4.1 Data on casualty survival rates

As mentioned, TAGOKOR details twenty different categories of casualty. In this section, a soldier is coded as having survived if they are listed as one of three casualty types: "returned to duty" (RTD), "evacuated" (EVC), or "returned to military control" (RMC).

Figures 8a and 8b show the casualty survival rates of blacks and whites over time, respectively, and Figure 8c jointly depicts these race-specific survival rates. The black casualty survival rate begins elevated and exceeds that of whites throughout most of the war. Although the mean survival rates of wounded black and white soldiers mostly converged after six months, these trends could be masking important heterogeneity across units. For example, it could be the case that the majority of blacks remained in relatively safe support roles, but the few that were integrated into combat roles were assigned especially hazardous missions. The next subsection examines the differential effects of integration on survival, and the one after that explores how blacks were treated in highly integrated units.

4.2 Integration, manpower and casualty survival rates

During the Korean War, a wounded soldier would first be attended to at a battalion aid station. Each aid station typically contained one doctor and a mix of medics and litter bearers (i.e. soldiers who carried stretchers). The wounded would then be transported behind the front line to a larger collection station. Once the wounded had been stabilized, they would then be transported to a Mobile Army Surgical Hospital (MASH) unit or a division clearing station, depending on the type of wound, after which they were transported to an evacuation hospital. In the most serious cases, wounded soldiers would be airlifted to Japan for additional care.

The battalion aid station—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."

Early in the Korean War, many all-white battalions were severely depleted, which would have impaired evacuation efforts. Those all-white units that were willing to accept black troops and thus be returned to "safer" sizes would have been better able to evacuate wounded troops to safety. By contrast, for blacks, integration meant being pulled away from support roles and thrust into more dangerous combat roles. Overall, then, integration should have increased the probability that whites survived their injuries relative to segregated whites, as measured by the "white casualty survival rate". Likewise, integrated blacks would have been less likely to survive relative to blacks who remained in support roles.

4.3 A linear model for survival rates

I test this relationship with a linear model:

$$y_{irt} = \beta_0 + \beta_1 s_{rt}^{BW} + \boldsymbol{X}_{irt}\boldsymbol{\beta}_2 + \lambda_t + \gamma_r + \varepsilon_{irt}, \tag{4}$$

where y_{irt} 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 (4) is a difference-in-differences model in spirit, in that any departure in parallel trends is a result of integration. An implication of this model is that during periods when integration levels were steady, two regiments should have parallel survival rates. But it is usually infeasible to pinpoint periods where integration levels remain steady because the integration measure I construct moves too infrequently (i.e.

every four or six months) to demonstrate parallel trends convincingly. However, information on the exact date that particular units were integrated is available for a handful of units. For example, Figure 9 shows a screen shot of the 9th Infantry Regiment's war diary 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 September 16, 1950, after having spent the first eight weeks of the war away from combat on reserve duty. Figure 10 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 September 16, 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 October 1, 1951 (i.e. day 463 of the war), the date that formal integration began. Two features 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.

4.3.1 Estimation results

The results of estimating Equation (4) are reported in Table 6. In Panel A, column (1), we see that a one standard deviation increase in the integration measure increased white survival probability by 3.3 percentage points. Column (2) adds regiment fixed effects to control for unobserved regiment quality. To account for the fact that both integration and survival rates increased as the war progressed, column (3) includes period fixed effects. Using a baseline survival rate of 70%, the effect of integration on survival is estimated to be between 4.4%-6.0%.

Typically, combat units that experienced the most severe battle shocks were more likely to both draw on blacks for reinforcements and be rotated away from the front lines. As such, failure to account for the vulnerability of a given unit will bias the estimation upwards.

More formally, let Z_{rt} be a measure of "vulnerability" of regiment r at the start of period t. Then the true relationship between survival and integration is given by

$$y_{irt} = \beta_0 + \beta_1 s_{rt}^{BW} + \boldsymbol{X}_{irt}\boldsymbol{\beta}_2 + \lambda_t + \gamma_r + \rho Z_{rt} + \varepsilon_{irt}, \tag{5}$$

and the conditional mean of the least squares estimator[13] for Equation (4) is now given by

$$\mathbf{E}(\hat{\beta}_1) = \beta_1 + \rho \cdot \frac{\mathrm{Cov}(s_{rt}^{BW}, Z_{rt})}{\mathrm{Var}(s_{rt}^{BW})}. \tag{6}$$

Since $\rho > 0$ and $\mathrm{Cov}(s_{rt}^{BW}, Z_{rt}) > 0$, the bias is upward. To proxy for vulnerability, I set $Z_{rt} = 1 - \overline{y}_{r,t-1}$, where $\overline{y}_{r,t-1}$ is the casualty survival rate in period $t-1$, which means that $1 - \overline{y}_{r,t-1}$ is the corresponding casualty *fatality* rate.

The preferred specification is column (4), which controls for the previous period's fatality rate. As expected, the fatality rate is positively associated with survival in the next period, and controlling for it reduces the estimated effect of integration on survival. Now we see that a standard deviation change in integration increased casualty survival by 2.3 percentage points. Given that mean survival rate for whites was 70%, this represents an increase of 2.9%. This effect size is economically meaningful since it is comparable to holding a rank of private. A wounded private was 1.5 percentage points less likely to survive his injury, which reflects the fact that privates were put in more hazardous positions. Panel B repeats this exercise for blacks. Again we see that moving from column (3) to column (4) reduces the upward estimation bias when failing to account for unit vulnerability. As predicted, for blacks integration is associated with a *decrease* in black survival rates by 3.0 percentage points, or 4%.

That the direction of the signs for blacks and whites differs is to be expected given that integration implied a convergence in hazardous combat assignments. What is less obvious is the effect integration had on *overall* survival rates. I test for this directly by estimating

[13] Here I abuse notation and assume that s_{rt}^{BW} and Z_{rt} have been purged of the other covariates.

Equation (4) using a pooled sample of blacks and whites. Table 7 reports results. In each specification, integration has a positive effect on overall survival rates ranging from 1.9-2.4 percentage points, or 2.4-3.2%. Overall, these results are consistent with the hypothesis that segregation impeded the efficient allocation of scarce labor. Also of interest is how estimates of the black coefficient varies across specifications. In column (1), the coefficient on black is strongly positive, which is in part due to race being correlated with unobserved combat exposure. Column (2) adds regiment fixed effects, which attenuates the coefficient on black slightly, though it remains significantly different from zero. Recall that segregation occurred at units of battalion size or smaller, which means controlling for regiment does not completely account for differences in combat exposure by race. Column (3) adds period fixed effects, which further attenuates the black coefficient because over the course of the war, the share of blacks increased as the intensity of combat waned. Column (4) controls for private status and having a support military occupational specialty, and again the black coefficient remains positive though marginally significant.

4.4 How blacks were treated in newly integrated units

Columns (5)-(6) in Table 7 explore the relationship between integration and the safety of blacks more directly. Regardless of the extent of integration, holding a rank of private is strongly predictive of fatality, which is expected given that these soldiers were exposed to more dangerous roles and were the least experienced. Column (5) is restricted to a sample of regiment-period pairs whose integration measure belong to the bottom quartile. In these relatively segregated regiments, being black was highly predictive of survival, with a magnitude two and a half times the size of that of holding the rank of private. By contrast, column (6) is restricted to a sample of regiment-period pairs whose integration measures belong to the top quartile. Here we see that blacks were no more or less likely to survive their wounds than whites, and the null hypothesis consistent with no discrimination cannot therefore be rejected.

5 Integration and residential choice

From a longer-term perspective, the natural restrictions that war places on selective migration and sorting present a rare opportunity to investigate the effects of interracial contact on prejudice. Without such restrictions, highly prejudiced individuals would be able to self-select out of contact, and standard econometric analysis would overstate the social benefit of intergroup contact. The Army's unique ability to compel its soldiers is in stark contrast to civilian settings, where large-scale desegregation in turn propelled white flight. Therefore, the second part of the paper studies the long-run effects of wartime integration by examining the post-war behavior of veterans toward people outside their race group. The two behavioral measures I consider are (i) where veterans chose to live (this section) and (ii) whom they married (Section 6).

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

The Social Security Death Index (SSDI) 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 8 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 the SSN was issued, and dates of birth and death.

The SSDI also reports the last zip code of residence prior to death, 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_{bt} (W_{bt}) 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 and whites in zip code z in year t is given by

$$B_{zt}^{Total} \equiv \sum_{b \in D_z} B_{bt}$$
$$W_{zt}^{Total} \equiv \sum_{b \in D_z} W_{bt}.$$

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

$$ZipSim_{zt}^{BW} \equiv 1 - \frac{1}{2} \sum_{b \in D_z} \left| \frac{B_{bt}}{B_{zt}^{Total}} - \frac{W_{bt}}{W_{zt}^{Total}} \right|. \tag{7}$$

As was the case with the regimental similarity index, $ZipSim_{zt}^{BW}$ can take on values between 0 and 1 by construction. The term to the right of the minus sign in Equation (7) is the familiar dissimilarity Index first proposed by Duncan and Duncan (1955). Hence $ZipSim_{zt}^{BW}$ is the complement the dissimilarity index. A $ZipSim_{zt}^{BW}$ 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, $ZipSim_{zt}^{BW}$ 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 versus Santa Fe, New Mexico) for reasons unrelated his racial preferences—be it work opportunities, family ties, and so forth. Here I define a 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 $ZipSim^{BW}_{zt}$ within a CBSA. In the case of, say, the New York housing market, the within-CBSA normalized zip code similarity index y_{zt} is given by

$$y_{zt} = \frac{ZipSim^{BW}_{zt} - \mathbf{E}_{NY}(ZipSim^{BW}_{zt})}{\sigma_{NY}(ZipSim^{BW}_{zt})}. \tag{8}$$

The distribution of y_{zt} is shown in Figure 11. That y_{zt} is roughly symmetrical indicates that, within a housing market, the similarity index is distributed symmetrically. I also use the block-level data to 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.[14]

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), the proportion of blacks and whites in the last zip code of residence, and the extent of integration between these two groups.

5.2 A linear model for residential choice

In this subsection, I estimate the effect of wartime integration on the residential similarity of each soldier's last zip code of residence. Recall that, by assumption, the veteran's racial preferences enter into his utility function only when choosing which neighborhood (i.e. "zip code") to reside in *within* a given housing market, not across housing markets. Thus, in this subsection I model neighborhood choice conditional on having already chosen a housing

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

market.

To capture this within-housing market neighborhood choice, I estimate a linear model:

$$y^{BW}_{cirstz} = \beta_0 + \beta_1 s^{BW}_{rt} + \boldsymbol{X}_i\boldsymbol{\beta}_2 + \boldsymbol{Z}_{iz}\boldsymbol{\beta}_3 + \boldsymbol{C}_c\boldsymbol{\beta}_4 + \gamma_r + \lambda_t + \zeta_s + \varepsilon_{cirstz}, \tag{9}$$

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.

Where appropriate, I will sometimes 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."[15]

5.3 Estimation results

Table 9 reports estimation results for equation (9) for a sample restricted to white veterans. In column (1) we see that a one SD increase in treatment leads to a 0.06 SD increase in residential 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 (not reported). Clearly, house prices are endogenous since they correlate

[15] 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 (9) more closely resemble intention-to-treat (ITT) effects. Typically, ITT effects represent lower bounds on the treatment effects they approximate in terms of magnitude.

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 estimated 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, then the white veteran who was more integrated during the war is more willing to live in 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, whereas the county of origin controls lose their explanatory power.

Table 10 shows the estimation results for Equation (9) for black veterans. None of the treatment effects 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 imposing barriers to their residential choices. In the years enveloping 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 9 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.

5.4 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 U.S. public schools and the migration of millions of blacks to predominantly white regions of the U.S. 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 a form of 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 (9). Table 11 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 9—which ranges from 0.04-0.06. These larger point estimates for young veterans point to heterogeneity in treatment effects. These heterogeneous effects could stem from a number of sources. One possibility is that older veterans were already homeowners and as such were less geographically mobile. Another possibility is that younger veterans had less social status, which contact theory predicts would make them more responsive to intergroup contact. With the data currently available to me, I am unable to distinguish between these distinct mechanisms.

Table 12 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, although here again the much smaller sample size reduces statistical power.

6 Integration and intermarriage

A high bar for prejudice reduction is exogamy—an individual's willingness to marry outside of their own race or ethnicity. This section explores whether wartime integration made soldiers more exogamous upon returning home.

6.1 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 data 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 13 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 the Expectation Maximization Algorithm. 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.

The integration that soldiers experienced during the Korean War satisfied the four conditions for the contact hypothesis outlined by Allport (1954), who was interested in preference formation from a sociological perspective. 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 that these four Allportian conditions were satisfied, this raises the question to what extent, if any, did integration reduce the prejudicial attitudes of veterans?

6.2 Linear and fractional response models for intermarriage

I test the relationship between integration and exogamy with two models. The first is a linear model, which I estimate by OLS:

$$y_{irst} = \beta_0 + \beta_1 s_{rt}^{BW} + \boldsymbol{X}_i \boldsymbol{\beta}_2 + \gamma_r + \lambda_t + \zeta_s + \varepsilon_{irst} \tag{10}$$

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 of origin, and period of casualty fixed effects, respectively.

By construction, the wife's name index y_{irst} is bounded between zero and one (recall Figures 12a and 12b). In order to account for this characteristic of the left-hand-side variable, I use the quasi-maximum likelihood estimation (QMLE) framework for fractional response models developed by Papke and Wooldridge (1996). I first specify the following functional form for the expectation of the name index y_{irst}:

$$\begin{aligned}\mathbf{E}(y_{irst}|s_{rt}^{BW}, \boldsymbol{X}_i, \gamma_r, \lambda_t, \zeta_s) &= \Phi(\beta_0 + \beta_1 s_{rt}^{BW} + \boldsymbol{X}_i \boldsymbol{\beta}_2 + \gamma_r + \lambda_t + \zeta_s),\\ &\equiv \Phi(\boldsymbol{W}_{irst} \boldsymbol{\Theta} + \gamma_r + \lambda_t + \zeta_s),\end{aligned} \tag{11}$$

where Φ is the standard 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 (11) using a Bernoulli quasi-likelihood method, where the likelihood function for observation i is given by:

$$L_i(\boldsymbol{\Theta}, \gamma_r, \lambda_t, \zeta_s) = [\Phi(\boldsymbol{W}_{irst}\boldsymbol{\Theta} + \gamma_r + \lambda_t + \zeta_s)]^{y_{irst}} [1 - \Phi(\boldsymbol{W}_{irst}\boldsymbol{\Theta} + \gamma_r + \lambda_t + \zeta_s)]^{1-y_{irst}}. \tag{12}$$

This QMLE approach has two particularly attractive features: first, it yields consistent estimates of $\mathbf{\Theta}, \gamma_r, \lambda_t, \zeta_s$ provided that the conditional expectation in Equation (11) is specified correctly; this is true even if the Bernoulli likelihood in Equation (12) is misspecified. Second, the variance of y_{irst} need not be specified at all.

6.3 Estimation results

Estimation results for both the linear and fractional response model for white veterans are reported in Table 14. 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.[16] 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, or 0.8 SD. 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, or 0.04 SD). 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, although 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 11. 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; results are broadly the same.

6.4 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 15 reports estimation results for Equation (11) for a sample of men who were at least 23 years of age during the Korean

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

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 significant negative effect on the wife's name index. 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, and (ii) older veterans were more likely to already be married. Again, the data currently available to me is not sufficiently rich to distinguish between these mechanisms.[17]

7 Conclusion

The racial integration of the U.S. Army during the Korean War was one of the largest and swiftest desegregation episodes in American history. Integration began informally, as a bid to reinforce badly depleted all-white units, and by the end of the war had become Army-wide policy. Unlike the desegregation of U.S. public schools—which was initiated expressly for the benefit of blacks—the Army integrated for reasons of military efficiency.

The first part of my paper examines whether the Army managed to improve efficiency as measured by the survival rates of wounded soldiers. To this end, I developed a novel wartime integration measure to quantify exogenous changes in racial integration across time and regiments. Using a two-way fixed effects model, I found based on a pooled sample of blacks and whites that a standard deviation increase in integration caused a 3% increase in the *overall* survival rates of wounded soldiers. On the whole, these estimates indicate that the Army managed to improve efficiency once it abandoned racial segregation and could freely allocate scarce labor (soldiers). Survival rates are admittedly only a partial measure of overall military efficiency, as integration may have improved combat efficiency at the

[17] The full count 1950 Census is scheduled to be released in April 2022, which will contain valuable data on spouses. Future work could restrict the sample to soldiers who were not married as late as 1950 and identify treatment effects off of men who served early in the war.

expense of some unobserved tasks. I therefore leave it to future work to consider broader measures of efficiency such as casualty rates, unit citations, other-than-honorable (OTH) discharges, and awarded medals.

Historically, large-scale desegregation episodes have often had unintended consequences, with a common response to both school desegregation and black in-migration being so-called "white flight." The second part of this paper is therefore devoted to understanding the long-run effects of wartime integration by observing the behavior of Korean War veterans in their post-war civilian lives. In particular, by matching veterans to social security and cemetery data, I am able to observe where veterans lived and whom they married. Here I found that a standard deviation increase in wartime integration caused young veterans to live in neighborhoods that were 0.1 standard deviations more integrated—roughly the difference in integration between present-day Central Harlem and the Upper East Side. Similarly, a standard deviation increase in wartime integration caused veterans to marry spouses with names that were less distinctly white. 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. Overall, these results on veterans' post-war behavior support the view that large-scale interracial contact served to reduce 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. *Quarterly Journal of Economics*, 114(4):1243–1284.

Allman, E. S., Matias, C., Rhodes, J. A., et al. (2009). Identifiability of Parameters in Latent Structure Models with Many Observed Variables. *The Annals of Statistics*, 37(6A):3099–3132.

Allport, G. W. (1954). *The Nature of Prejudice.* Addison-Wesley.

Blair, C. (1989). *The Forgotten War: America in Korea, 1950-1953.* Anchor.

Bogart, L. (1992). *Project Clear: Social Research and the Desegregation of the United States Army.* NJ: Transaction Books.

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. *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.

Carreira-Perpinán, M. A. and Renals, S. (2000). Practical identifiability of finite mixtures of multivariate bernoulli distributions. *Neural Computation*, 12(1):141–152.

Carrell, S. E., Hoekstra, M., and West, J. E. (2019). The Impact of College Diversity on Behavior Toward Minorities. *American Economic Journal: Economic Policy*, 11(4):159–82.

Costa, D. L. and Kahn, M. E. (2003). Cowards and Heroes: Group Loyalty in the American Civil War. *Quarterly Journal of Economics*, 118(2):519–548.

Cutler, D. M., Glaeser, E. L., and Vigdor, J. L. (1999). The Rise and Decline of the American Ghetto. *Journal of political economy*, 107(3):455–506.

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.

Gyllenberg, M., Koski, T., Reilink, E., and Verlaan, M. (1994). Non-uniqueness in probabilistic numerical identification of bacteria. *Journal of Applied Probability*, 31(2):542–548.

Hausrath, A. H. (1954). Utilization of Negro Manpower in the Army. *Journal of the Operations Research Society of America*, 2(1):17–30.

Hill, M. E. and Rosenwaike, I. (2001). The Social Security Administration's Death Master File: The Completeness of Death Reporting at Older Ages. *Social Security Bulletin*, 64:45.

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.

Schindler, D. and Westcott, M. (2015). Shocking Racial Attitudes: The Cultural Legacy of Black GIs in Europe. Technical report, mimeo.

Train, K. E. (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press.

Tables

Table 1: Glossary of Army ranks

Pay Grade	Rank	Abbreviation	Classification
E-1	Private	PVT	Enlisted Soldier
E-2	Private Second Class	PV2	Enlisted Soldier
E-3	Private First Class	PFC	Enlisted Soldier
E-4	Specialist	SPC	Enlisted Soldier
E-4	Corporal	CPL	NCO
E-5	Sergeant	SGT	NCO
E-6	Staff Sergeant	SSG	NCO
E-7	Sergeant First Class	SFC	NCO
E-8	Master Sergeant	MSG	NCO
E-8	First Sergeant	1SG	NCO
E-9	Sergeant Major	SGM	NCO
E-9	Command Sergeant Major	CSM	NCO
E-9	Sergeant Major of the Army	SMA	NCO
W-1	Warrant Officer 1	WO1	Warrant Officer
W-2	Chief Warrant Officer 2	CW2	Warrant Officer
W-3	Chief Warrant Officer 3	CW3	Warrant Officer
W-4	Chief Warrant Officer 4	CW4	Warrant Officer
W-5	Chief Warrant Officer 5	CW5	Warrant Officer
O-1	Second Lieutenant	2LT	CO
O-2	First Lieutenant	1LT	CO
O-3	Captain	CPT	CO
O-4	Major	MAJ	Field Officer
O-5	Lieutenant Colonel	LTC	Field Officer
O-6	Colonel	COL	Field Officer
O-7	Brigadier General	BG	General Officer
O-8	Major General	MG	General Officer
O-9	Lieutenant General	LTG	General Officer
O-10	General	GEN	General Officer
O-10	General of the Army	GA	General Officer

Notes: NCO and CO denote "Noncommissioned Officers" and "Commissioned Officers", respectively. Source: `https://www.federalpay.org/military/army/ranks`.

Table 2: Hierarchy 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
[a]Regiment/Brigade/Legion	1,000-5,500	Colonel/Brigadier General
[b]Battalion/Cohort	300-800	Lieutenant Colonel
[b]Company	80-150	Captain/Major
[b]Platoon	15-45	Second Lieutenant/Lieutenant
[b]Squad/Section	8-14	Sergeant
[b]Fireteam	2-4	Lance Corporal/Corporal

Notes: [a]I observe which regiment individual soldiers served in.
[b]Formally segregated before Oct 1951.

Table 3: 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

Notes: Panel A. shows the balance test for the first half of the war (June 25, 1950 to December 22, 1951). Panel B. show the balance test for the second half of the war (December 23, 1951 to July 27, 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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 4: 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
Casualty Code	4	Wounded by Missile
Casualty Group Code	F	Returned to Duty
Place of Disposition		
Day of Disposition	25	25
Blank		
Month of Disposition	5	May
Year of Disposition	A	1951
Year of Birth		
[a]MOS code	4812	Heavy Weapons Infantryman
Troop Sequence Number	7003	IN DIV - 3RD
Element Sequence	60	IN DIV INF REGT
Unit	7	7 Infantry
Racial Group Code	1	White
Component	1	USA - RA (Reg Army)
Line of Duty		
Disposition of Evacuations		

Notes: [a]MOS denotes "Military Occupational Specialty". Searchable records are available at `https://aad.archives.gov/aad/series-description.jsp?s=531`.

Table 5: 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
—Deep South (%)	8.3	21.9	8.4	20.2
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
—Deep South (%)	7.0	28.6	9.0	29.4
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
	($n = 29,494$)	($n = 4,463$)	($n = 12,430$)	($n = 1,572$)

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

Table 6: The effects of integration on survival by race

Dependent variable: casualty survival dummy

	(1)	(2)	(3)	(4)
		A. Whites		
Integration (s_{BW})	0.033***	0.042**	0.031**	0.023**
	(0.007)	(0.020)	(0.014)	(0.010)
Private	-0.015	-0.012	-0.022***	-0.015***
	(0.015)	(0.012)	(0.008)	(0.006)
Lag fatality rate				0.354***
				(0.073)
Intercept	0.751	0.705	0.714	0.668
N	$63,702$	$63,702$	$63,702$	$42,725$
		B. Blacks		
Integration (s_{BW})	-0.021	-0.030	-0.013	-0.030**
	(0.021)	(0.037)	(0.016)	(0.015)
Private	-0.003	-0.016***	-0.030**	-0.036**
	(0.012)	(0.005)	(0.014)	(0.015)
Lag fatality rate				0.282***
				(0.050)
Intercept	0.776	0.690	0.749	0.506
N	$8,658$	$8,658$	$8,658$	$6,159$
Regiment FE		Y	Y	Y
4-month FE			Y	Y
#Regiments	35	35	35	35
#Periods	9	9	9	9

Notes: The dependent variable is a binary survival dummy that equals 1 if the wounded soldier died from their injury and 0 otherwise. The fatality rate is given by $1 - \overline{y}_{r,t-1}$. Panels A and B report regression results for whites and blacks, respectively. Standard errors calculated using two-way clustering by regiment and four-month period of casualty. Period 1 and 2nd Regiment are the omitted groups in the applicable columns. Only the first six periods are included since integration became formal policy by the last year of the war. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 7: The effects of integration on survival (pooled sample)

Dependent variable: casualty survival dummy

	Full Sample				Segregated	Integrated
	(1)	(2)	(3)	(4)	(5)	(6)
Integration (s_{BW})	0.020^{***}	0.024^{**}	0.019^{***}	0.019^{***}		
	(0.006)	(0.010)	(0.007)	(0.007)		
Black	0.023^{***}	0.019^{**}	0.008	0.010^{*}	0.055^{**}	0.008
	(0.007)	(0.008)	(0.006)	(0.006)	(0.023)	(0.012)
Private				-0.017^{***}	-0.020^{***}	-0.015^{***}
				(0.005)	(0.005)	(0.006)
Support role				0.009	-0.032	0.045^{**}
				(0.007)	(0.023)	(0.023)
N	$80,886$	$80,886$	$80,886$	$80,886$	$23,403$	$21,986$
Regiment FE		Y	Y	Y	Y	Y
6-month FE			Y	Y	Y	Y
#Regiments	24	24	24	24	14	14
#Periods	6	6	6	6	5	6

Notes: The dependent variable is a binary survival dummy that equals 1 if the wounded soldier survived their injury and 0 otherwise. Standard errors calculated using two-way clustering by regiment and six-month period of casualty. (Calculating integration in 4-month bins and adding 4-month fixed effects does not qualitatively change the result.) Columns (1)-(4) uses the the full sample of regiments for which the integration measure could be constructed. In columns (5)-(6), "Segregated" and "Integrated" refer to regiments that in any given period were in the bottom and top integration quartiles, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 8: Typical observation in Social Security Death Index

Field title	Value
SSN	XXXXXXXXXX
Last Name	OWENS
Suffix	
First Name	CALVIN
Middle Name	R
Year Death	2005
Year Birth	1933
State of Death	Colorado
Zip Code	80222
SSN Area	512
Day of Death Missing	0
Day of Birth Missing	0
DOD	2005-11-07
DOB	1933-01-08
Age at Death	72.8795
FIPS of Death	Colorado
SSN FIPS	Kansas

Notes: 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 9: 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

Notes: Columns (1)-(4) report results for the full sample of white veterans whom I was able to match to the social security death index. Data on house prices is only available for a subset of zip codes, hence some observations are lost in columns (3)-(4). Column (5) is restricted to white veterans whose last state of residence differed from the state they resided in during the war. All regressors except "Age at casualty" are normalized. All regressions include period of casualty, regiment and state-of-origin fixed effects. Standard errors are two-way clustered by regiment and period of casualty. The dependent variable is a zip code similarity index which is constructed using block-level demographic data from the U.S. Census Bureau. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 10: 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

Notes: Columns (1)-(4) report results for the full sample of black veterans whom I was able to match to the social security death index. Data on house prices is only available for a subset of zip codes, hence some observations are lost in columns (3)-(4). Column (5) is restricted to black veterans whose last state of residence differed from the state they resided in during the war. All regressors except "Age at casualty" are normalized. All regressions include period of casualty, regiment and state-of-origin fixed effects. Standard errors are two-way clustered by regiment and period of casualty. The dependent variable is a zip code similarity index which is constructed using block-level demographic data from the U.S. Census Bureau. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 11: 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

Notes: Columns (1)-(3) show regression results for white veterans whose service prefix indicates they were volunteers. That is, anything apart from "US". Columns (4)-(6) show regression results for white veterans whose service prefix indicates they were drafted; that is, their prefix was "US". All individuals were aged 25 or younger at the time of their casualty. The outcome is the similarity index of the soldier's last zip code of residence, normalized within CBSA. House prices are in logs. All regression include period of casualty, regiment and state-of-origin fixed effects. Standard errors are 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 U.S. Census Bureau. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 12: 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

Notes: Columns (1)-(3) show regression results for black veterans whose service prefix indicates they were volunteers; that is, anything apart from "US". Columns (4)-(6) show regression results for black veterans whose service prefix indicates they were drafted. That is, their prefix was "US". All individuals were aged 25 or younger at the time of their casualty. The outcome is the similarity index of the soldier's last zip code of residence, normalized within CBSA. House prices are in logs. All regression include period of casualty, regiment and state-of-origin fixed effects. Standard errors are 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 U.S. Census Bureau. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 13: Typical observation in national cemetery record

Field title	Veteran	Wife
First Name	Calvin	Virginia
Middle Name	R	C
Last Name	Owens	Owens
Suffix		
Birth Date	1/8/1933	2/20/1933
Death Date	11/7/2005	2/10/2015
Section ID	16	16
Row Number		
Site Number	193	193
Cemetery Name	Ft. Logan National Cemetery	Ft. Logan National Cemetery
Cemetery Address 1	4400 West Kenyon Av	4400 West Kenyon Av
Cemetery Address 2		
City	Denver	Denver
State	CO	CO
ZIP	80236	80236
Cemetery Phone	303-761-0117	303-761-0117
Relationship to Veteran	Veteran (Self)	Wife
Veteran's First Name	Calvin	Calvin
Veteran's Middle Name	R	R
Veteran's Last Name	Owens	Owens
Veteran's Suffix		
Branch	US ARMY	US ARMY
Rank	CPL	CPL
War	KOREA	KOREA

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

Table 14: 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	Y	Y	Y	Y
Rank FE		Y	Y	Y
State of res FE			Y	Y
N	664	664	664	664

Notes: 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 are 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 15: 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	Y	Y	Y	Y
Rank FE		Y	Y	Y
State of res FE			Y	Y
N	901	901	901	901

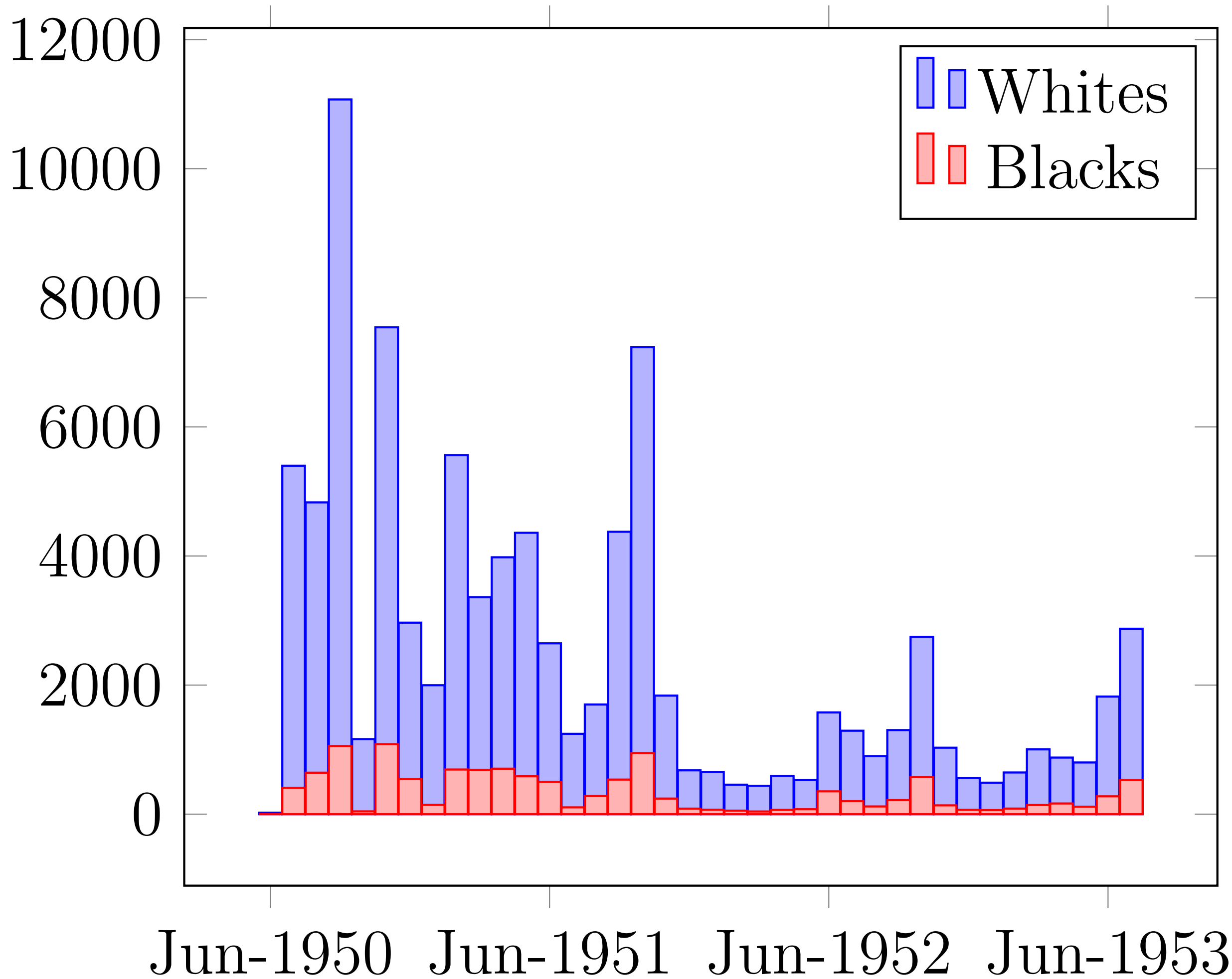
Notes: Panel A regresses the wife's non-white name index on the similarity measure. Panel B regresses the wife's black name index on the similarity measure. All regressors except "Age at casualty" or normalized. Standard errors are two-way clustered by state of origin and period of casualty. Columns (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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figures

Figure 1: Racial discrimination and black protest

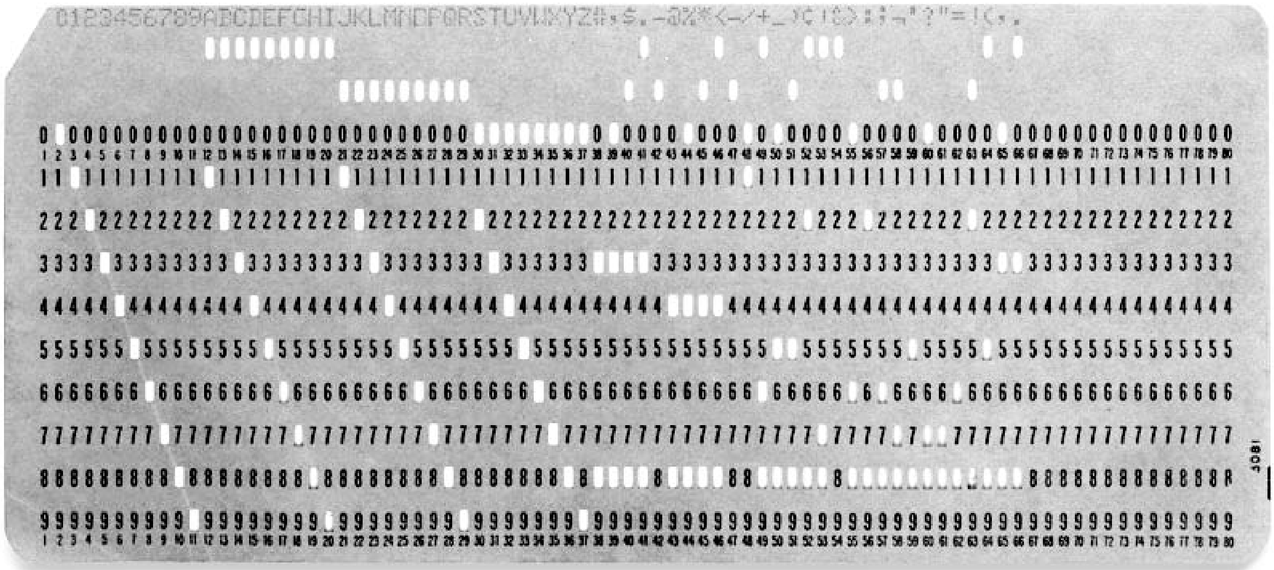
Notes: Black custodians protesting discrimination by defense contractors. The "double V" symbol—one "V" formed by the men, the second by the brooms—became the rallying cry for civil rights during the Second World War. The first "V" signifies the struggle for victory over the Axis powers, while the second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

Figure 2: Raw number of casualties in each month of the war



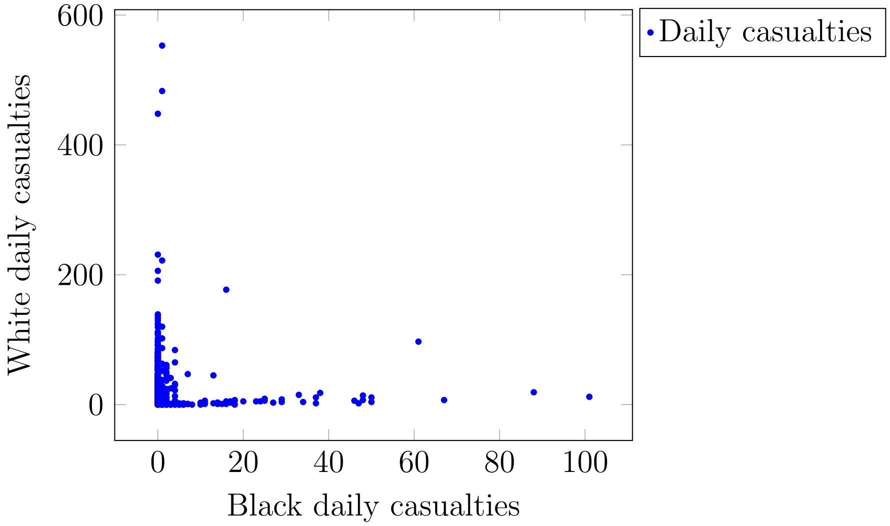
Notes: Each bar represents a month in the war. The total number of white and black casualties were 92,705 and 12,675, respectively.

Figure 3: IBM casualty punch card

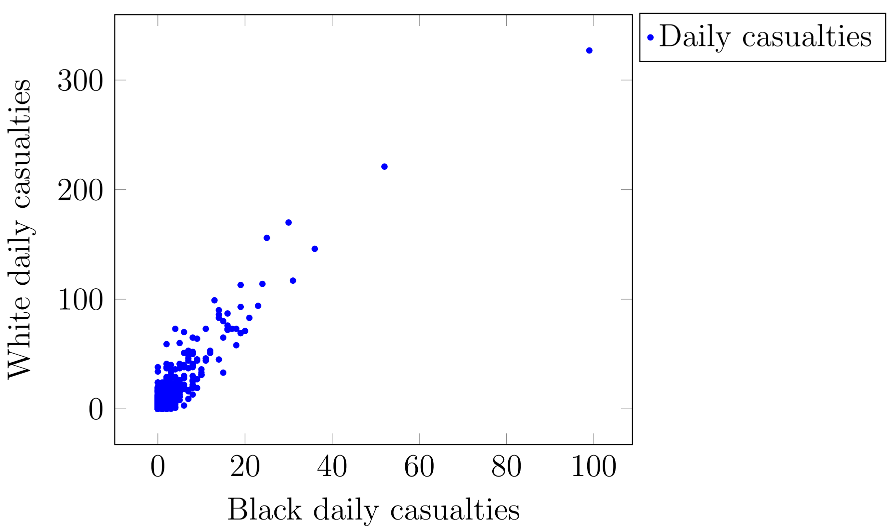


Note: IBM punch card used to document casualties during the war.

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



(a) First three months of war



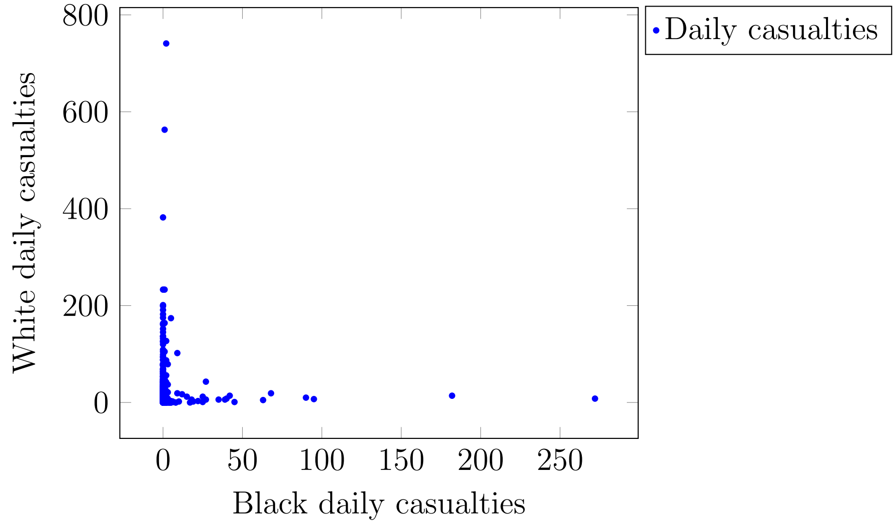
(b) Last year of war

Note: Each dot represents a day in the war for a given regiment.

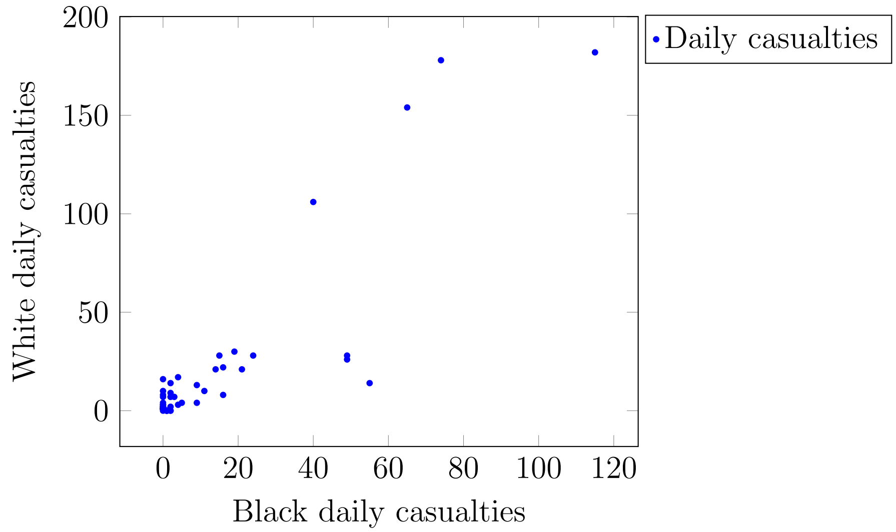
Figure 5: Example of early integration: 9^{th} Infantry Regiment (November 1950)

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

Figure 6: 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 for a given regiment.

Figure 7: Histogram for regimental similarity index

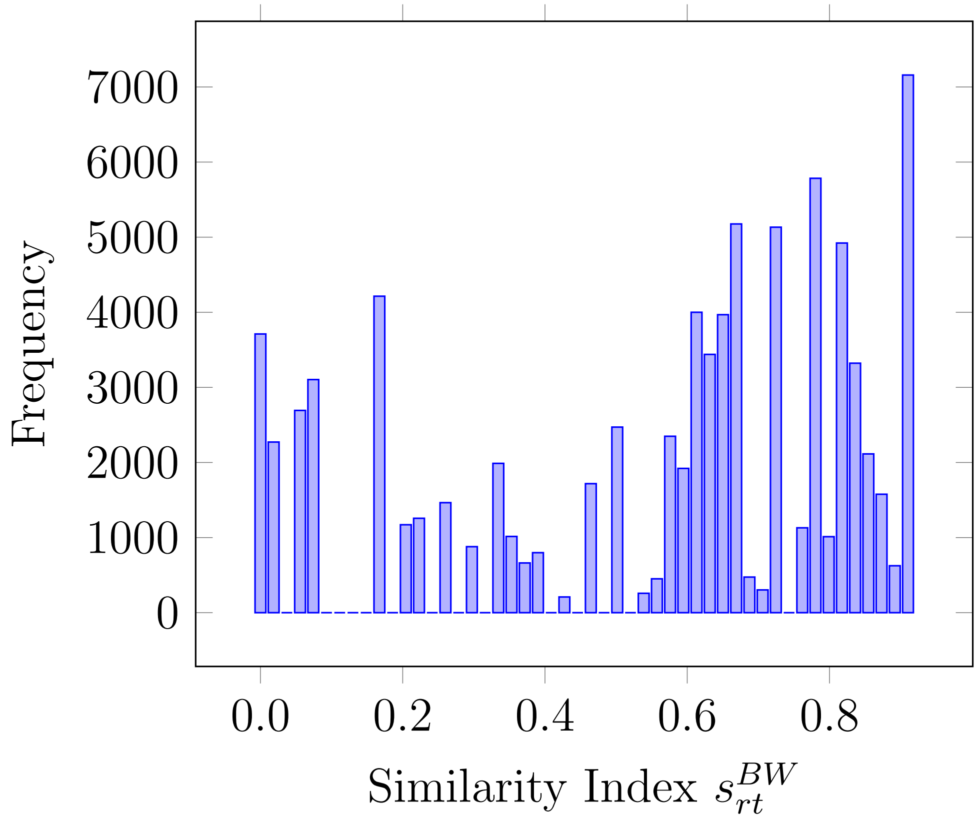
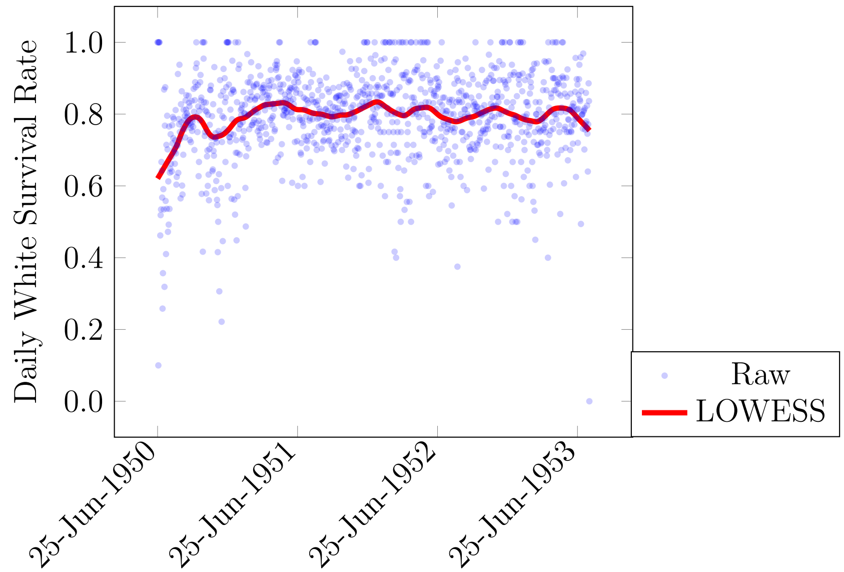
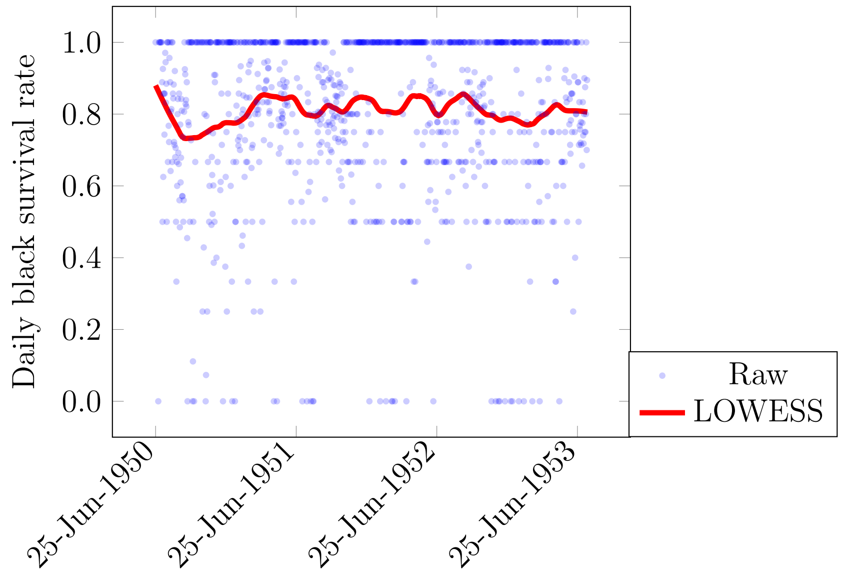


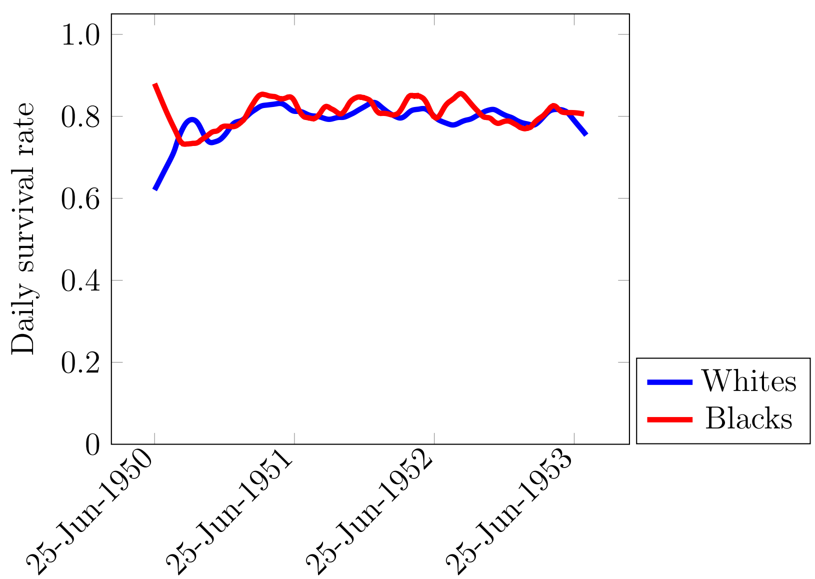
Figure 8: Daily survival rates throughout the war by race



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

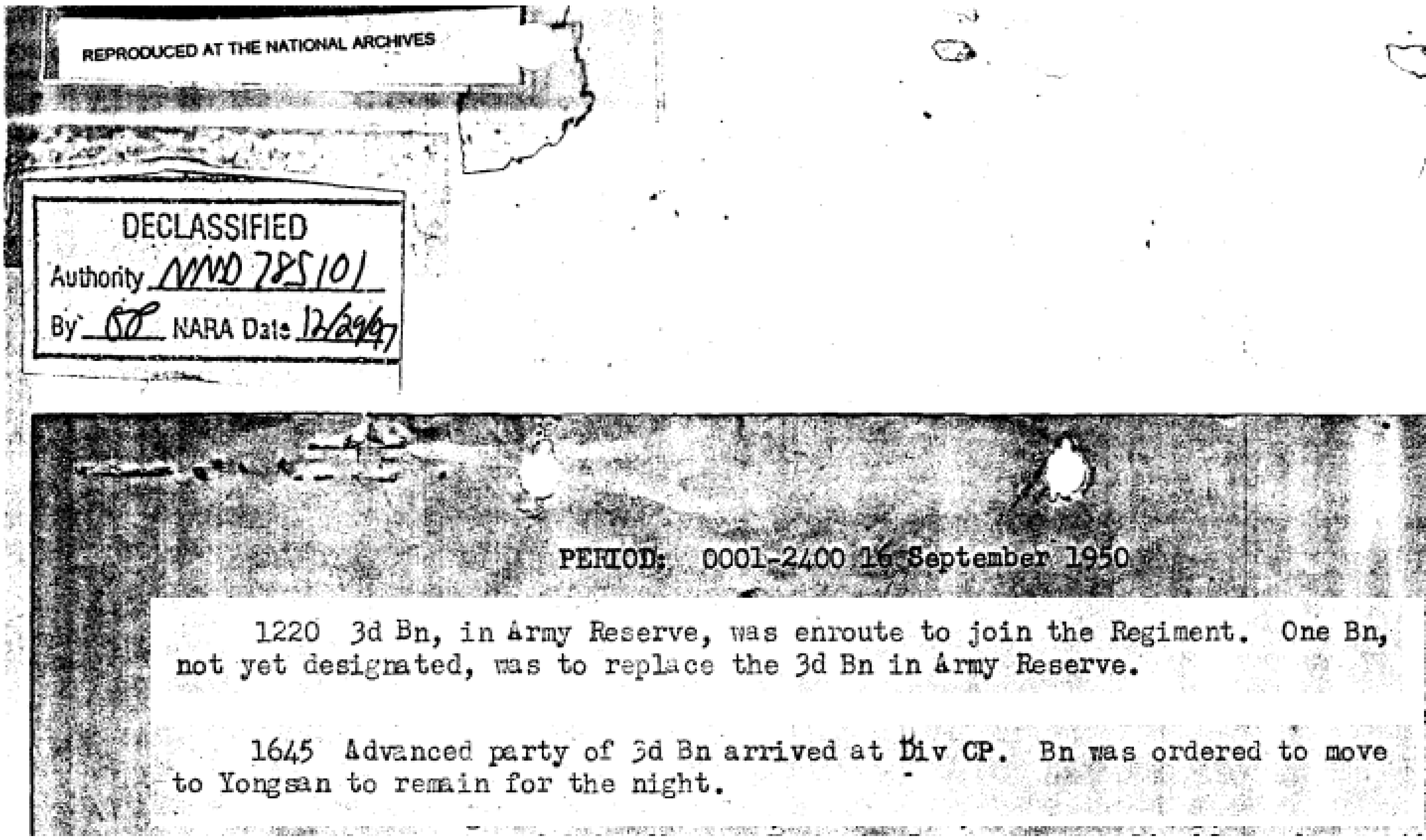


(b) Note: Each dot represents the average survival rate of wounded black soldiers on a given day in the war.



(c) Note: Figures 8a and 8b combined.

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



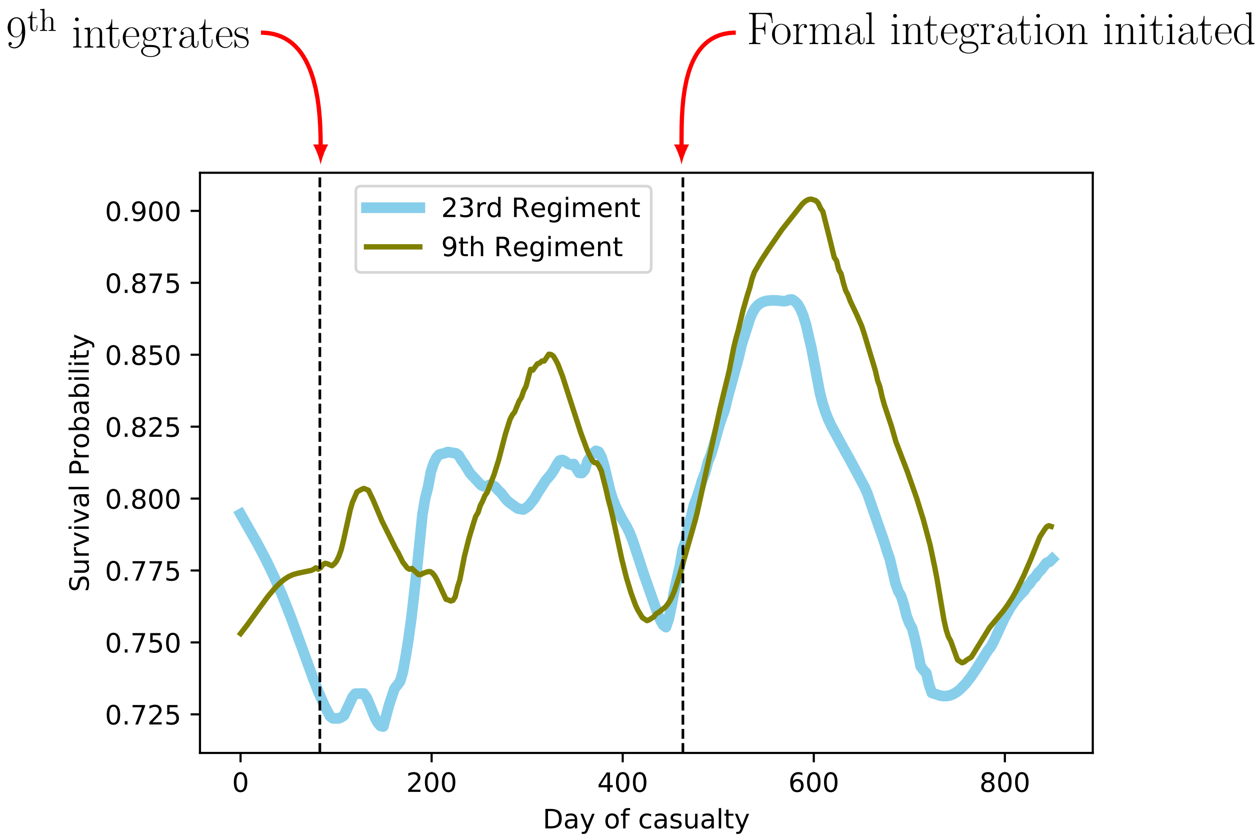
PERIOD: 0001-2400 16 September 1950

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 9^{th} Regiment

Figure 10: Casualty survival rates for 9th and 23rd Regiments



Note: The plot depicts the respective casualty survival rates of the 9$^{\text{th}}$ and 23$^{\text{rd}}$ Regiments, both of which belonged to the 2$^{\text{nd}}$ Infantry Division. The survival rates were estimated from daily casualty data using LOWESS. The first vertical line denotes September 15, 1950, and the second vertical line denotes October 1, 1951.

Figure 11: Similarity index for all U.S. zip codes (2010, standardized)

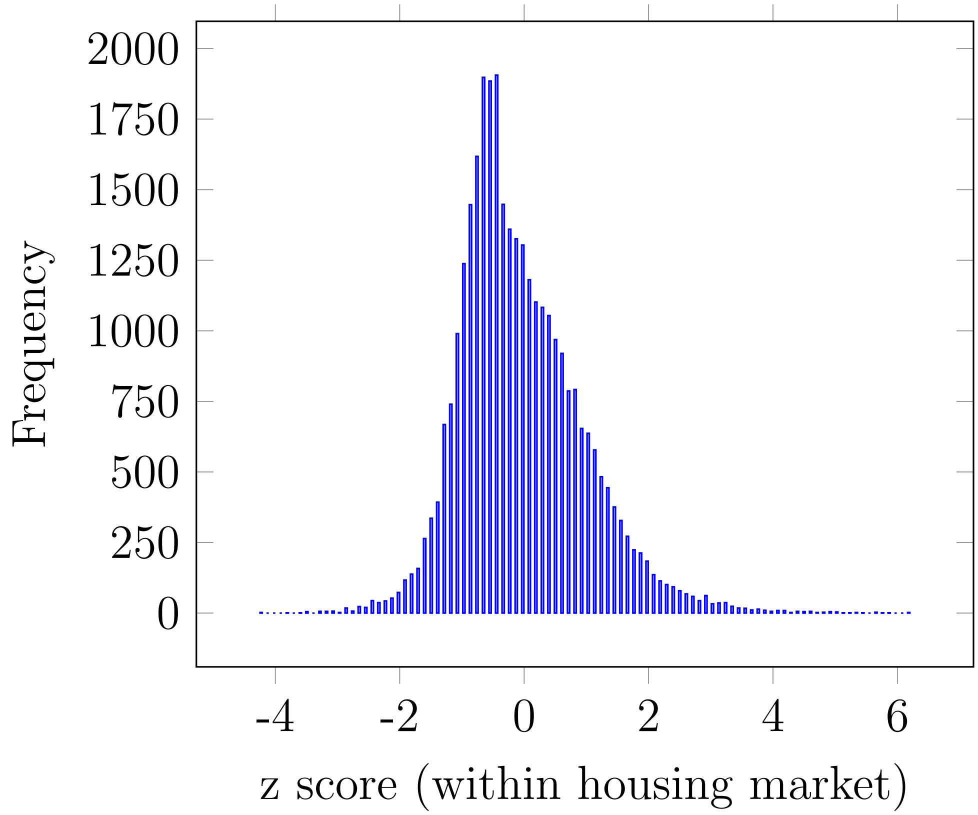
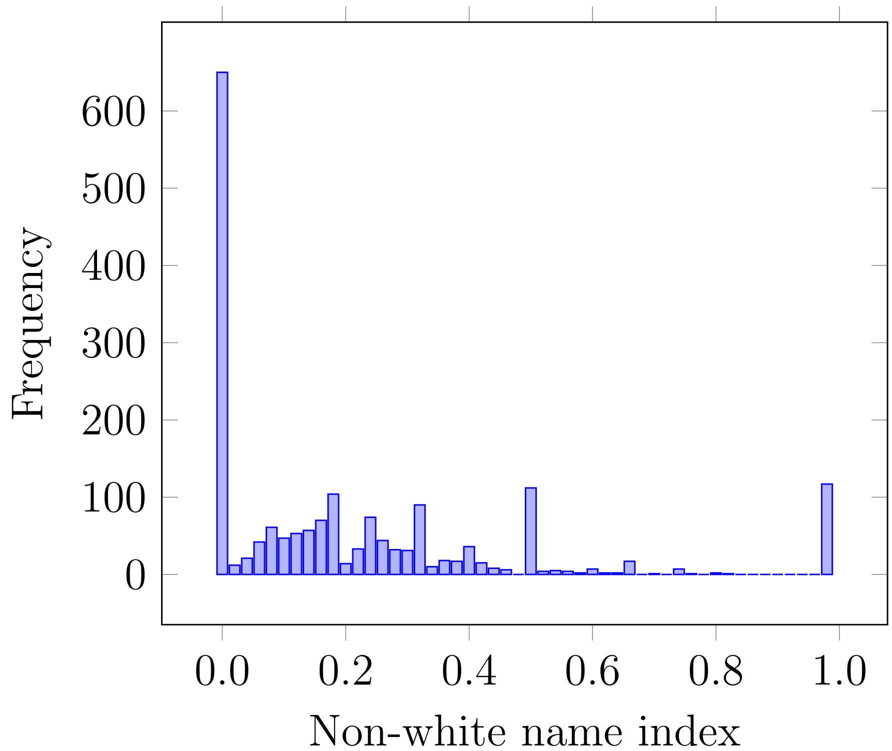
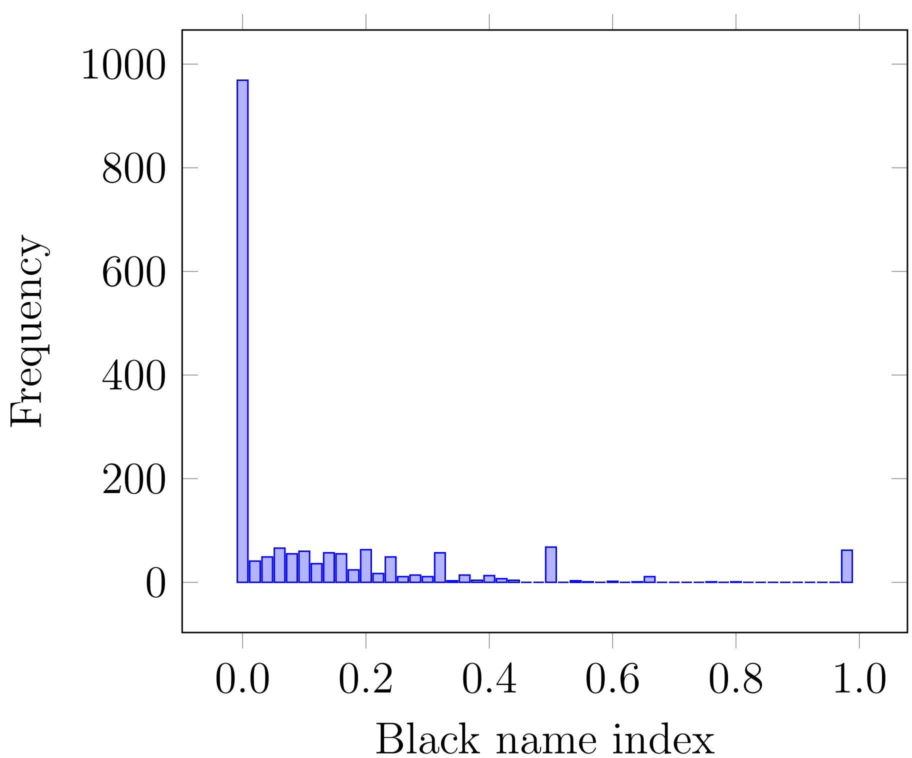


Figure 12: Racial content of wife's name



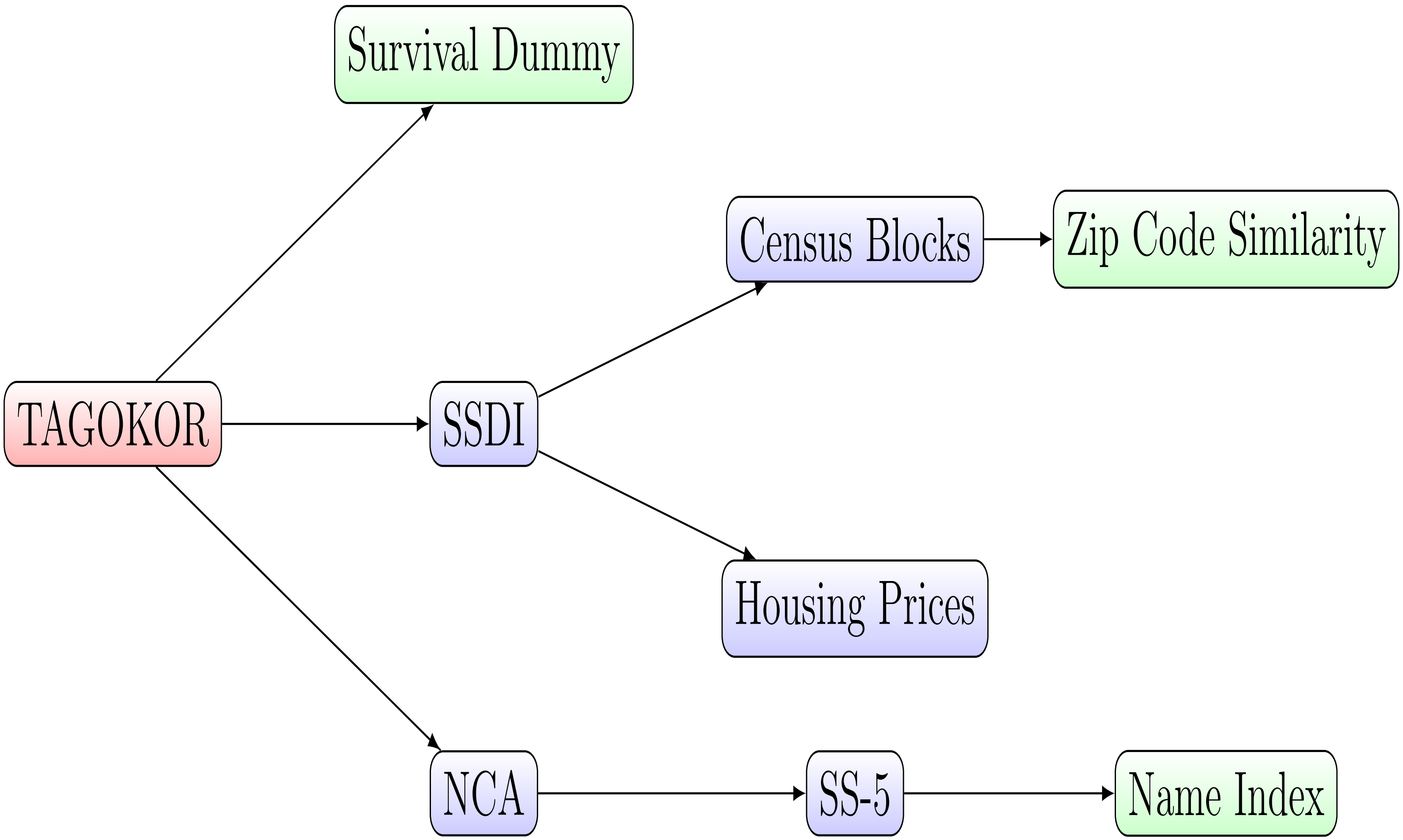
(a) Wife's black name index



(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: Summary of merged datasets



Notes: The Adjutant General's Office Korean (TAGOKOR) War casualty file contains records for all 103,000 American soldiers who were killed or wounded during the Korean War. I use this data to construct a similarity index to measure integration and model the survival rates of wounded soldiers. To examine post-war outcomes of veterans who survived the war, I link them to the social security death index (SSDI) which reports their last zip code of residence. For data on wives, I link veterans to national cemetery data, which is maintained and updated by the National Cemetery Administration. I construct a non-white name index for wives using race, state of residence and year of birth information from the SS-5 social security applications forms.

8 Appendix: The Expectation Maximization Algorithm

8.1 Record linkage as a missing data problem

Suppose we want to evaluate whether a given entry from historical record A refers to the same person in historical record B. Let the set of all matches be $M = \{(a, b) : a = b \ \forall a \in A, b \in B\}$, and let $z_i = 1$ if $i \in M$ and zero otherwise. Similarly, the set of non-matches (i.e. "unmatched") is given by $U = \{(a, b) : a \neq b, \ \forall a \in A, b \in B\}$. Consider the set of candidate matches given by the Cartesian product $\Omega = A \times B$. For each candidate match $i \in \Omega$, we observe the Jaro-Winkler similarity of the first names, which we denote be x_i^{JW}. For example, suppose record A contains "DANIEL", and record "B" contains four candidate matches {"DANIEL","DAN","DAVID","JOSH"}. Then x_i^{JW} takes on the values

$$
\begin{aligned}
JW_{Sim}(\text{"}DANIEL\text{"}, \text{"}DANIEL\text{"}) &= 1.00, \\
JW_{Sim}(\text{"}DANIEL\text{"}, \text{"}DAN\text{"}) &= 0.83, \\
JW_{Sim}(\text{"}DANIEL\text{"}, \text{"}DAVID\text{"}) &= 0.76, \\
JW_{Sim}(\text{"}DANIEL\text{"}, \text{"}JOSH\text{"}) &= 0.00.
\end{aligned}
$$

For true matches, the x_i^{JW} will be concentrated around 1, whereas non-matches will be distributed more evenly over the support $[0, 1]$.[18] The researcher must specify the two distributions for x_i^{JW}: one for when x_i^{JW} is drawn from a pair of true matches, and one for when x_i^{JW} is drawn from a pair of non-matches. Following Abramitzky et al. (2019), I specify a

[18] Strictly speaking, there will be spikes at values corresponding to common names like "JOHN" and "WILLIAM", which have a JW similarity relative to "DANIEL" of 0.47 and 0.54, respectively.

categorical distribution by partitioning the unit interval into four sets:

$$\begin{aligned} x_{i1} &= \mathbb{1}\left(x_i^{JW} \in [0, 0.8)\right), \\ x_{i2} &= \mathbb{1}(x_i^{JW} \in [0.8, 0.9)), \\ x_{i3} &= \mathbb{1}(x_i^{JW} \in [0.9, 1.0)), \\ x_{i4} &= \mathbb{1}(x_i^{JW} \in \{1.0\}), \end{aligned}$$

where $\mathbb{1}(\cdot)$ denotes the indicator function. The probability mass function (PMF) for $(x_{i1}, x_{i2}, x_{i3}, x_{i4})$ drawn from the matched population is given by:

$$\mathbf{P}\left((x_{i1}, x_{i2}, x_{i3}, x_{i4})|z_i = 1\right) = p_{M_1}^{x_{i1}} p_{M_2}^{x_{i2}} p_{M_3}^{x_{i3}} p_{M_4}^{x_{i4}}$$

where $p_{M_1}+p_{M_2}+p_{M_3}+p_{M_4} = 1$ with $p_k \in [0, 1]$, and $x_{i1}+x_{i2}+x_{i3}+x_{i4} = 1$ with $x_{ik} \in \{0, 1\}$ for all $k \in \{1, 2, 3, 4\}$. The categorical PMF can be thought of as a "four-sided die" (formally, a tetrahedron), where p_{M_k} the probability of landing on face k. The unmatched population is defined analogously:

$$\mathbf{P}\left((x_{i1}, x_{i2}, x_{i3}, x_{i4})|z_i = 0\right) = p_{U_1}^{x_{i1}} p_{U_2}^{x_{i2}} p_{U_3}^{x_{i3}} p_{U_4}^{x_{i4}}$$

Suppose for a moment we observed z_i and we would like to classify a new candidate record pair j. We would like to know the probability two records are a match given $(x_{j1}, x_{j2}, x_{j3}, x_{j4})$; that is, $\omega_j = \mathbf{P}(j \in M|x_{j1}, x_{j2}, x_{j3}, x_{j4})$. w_j is sometimes referred to as the "responsibility" of the matched distribution for observation j. If we had access to training data, meaning we observed z_i for all $i \neq j$, we could compute ω_j using Bayes' rule:

$$\begin{aligned}
\omega_j &= \mathbf{P}(z_j = 1 | x_j; \boldsymbol{\theta}) \\
&= \frac{\mathbf{P}(x_j | z_j = 1; \boldsymbol{\theta}) \pi_M}{\mathbf{P}(x_j | z_j = 1; \boldsymbol{\theta}) \pi_M + \mathbf{P}(x_j | z_j = 0; \boldsymbol{\theta})(1 - \pi_M)} \\
&= \frac{p_{M_1}^{x_{j1}} p_{M_2}^{x_{j2}} p_{M_3}^{x_{j3}} p_{M_4}^{x_{j4}} \pi_M}{p_{M_1}^{x_{j1}} p_{M_2}^{x_{j2}} p_{M_3}^{x_{j3}} p_{M_4}^{x_{j4}} \pi_M + p_{U_1}^{x_{j1}} p_{U_2}^{x_{j2}} p_{U_3}^{x_{j3}} p_{U_4}^{x_{j4}} (1 - \pi_M)}
\end{aligned} \tag{13}$$

where $\boldsymbol{\theta} = (\boldsymbol{p_M}, \boldsymbol{p_U}, \pi_M)$, with $\pi_M = \mathbf{P}(z_j = 1)$—that is, the proportion of records that are true matches—and $\boldsymbol{p_M} \equiv (p_{M_1}, p_{M_2}, p_{M_3}, p_{M_4})$ and $\boldsymbol{p_U} \equiv (p_{U_1}, p_{U_2}, p_{U_3}, p_{U_4})$ are the parameters for the categorical distributions of matched and unmatched observations, respectively. If we observed all the z_i, we could estimate $\boldsymbol{\theta} = (\boldsymbol{p_M}, \boldsymbol{p_U}, \pi_M)$ by splitting the sample between $z_i = 1$ and $z_i = 0$. The problem, of course, is that z_i is not observed, which means $\boldsymbol{p_M}$, $\boldsymbol{p_U}$, and π_M cannot be estimated directly by partitioning the sample as proposed.

8.2 Maximum likelihood approach

Alternatively, one could calculate $(\boldsymbol{p_M}, \boldsymbol{p_U}, \pi_M)$ using maximum likelihood. Let $\boldsymbol{X}$ denote the $n \times 4$ matrix of categorical data, which has a typical row $\boldsymbol{x}_i = (x_{i1}, x_{i2}, x_{i3}, x_{i4})$. The log likelihood equation is given by

$$\begin{aligned}
LL(\boldsymbol{X}; \boldsymbol{p_M}, \boldsymbol{p_U}, \pi_M) &= \sum_{i=1}^{n} \log \left(\mathbf{P}(\boldsymbol{x}_i | \boldsymbol{p_M}) \pi_M + \mathbf{P}(\boldsymbol{x}_i | \boldsymbol{p_U})(1 - \pi_M) \right) \\
&= \sum_{i=1}^{n} \log \left(p_{M_1}^{x_{i1}} p_{M_2}^{x_{i2}} p_{M_3}^{x_{i3}} p_{M_4}^{x_{i4}} \pi_M + p_{U_1}^{x_{i1}} p_{U_2}^{x_{i2}} p_{U_3}^{x_{i3}} p_{U_4}^{x_{i4}} (1 - \pi_M) \right)
\end{aligned} \tag{14}$$

Maximizing Equation 14 is difficult because the summation of the two probability mass functions appears inside the logarithm. As such, taking first-order conditions of Equation 14 will not yield closed-form estimators for $(\boldsymbol{p_M}, \boldsymbol{p_U}, \pi_M)$.

8.3 EM approach

The Expectation Maximization (EM) approach to this problem is to "pull the sum out of the logarithm". This can be achieved by maximizing the equation

$$\xi(\boldsymbol{\theta}|\boldsymbol{X};\boldsymbol{\theta^{(t)}}) := \sum_{i=1}^{n} \omega_i^{(t)} \log\left(\mathbf{P}(\boldsymbol{x}_i|\boldsymbol{p}_{\boldsymbol{M}}^{\boldsymbol{(t)}})\pi_M^{(t)}\right) + (1-\omega_i^{(t)})\log\left(\mathbf{P}(\boldsymbol{x}_i|\boldsymbol{p}_{\boldsymbol{U}}^{\boldsymbol{(t)}})(1-\pi_M^{(t)})\right) \tag{15}$$

where $\omega_i^{(t)} = \mathbf{P}(z_i = 1|\boldsymbol{x}_i, \boldsymbol{\theta}^{(t)})$ is the contribution defined in Equation 13 for some choice of $\boldsymbol{\theta}^{(t)}$. Since $\xi(\boldsymbol{\theta}|\boldsymbol{X};\boldsymbol{\theta}^{(t)})$ is an expectation of logs, it is sometimes referred to as the *expected complete log likelihood* (ECLL) function. It turns out that maximizing this ECLL function is tantamount to increasing along the gradient of $LL(\boldsymbol{X};\boldsymbol{\theta})$ at $\boldsymbol{\theta} = \boldsymbol{\theta}^{(t)}$.[19]

After choosing some initial value for $\boldsymbol{\theta}^{(t)}$ for $t = 0$, the EM algorithm proceeds in two alternating steps:

1. Expectation step:

$$\hat{\omega}_i^{(t)} = \frac{p_{M_1}^{(t)x_{i1}} p_{M_2}^{(t)x_{i2}} p_{M_3}^{(t)x_{i3}} p_{M_4}^{(t)x_{i4}} \pi_M^{(t)}}{p_{M_1}^{(t)x_{i1}} p_{M_2}^{(t)x_{i2}} p_{M_3}^{(t)x_{i3}} p_{M_4}^{(t)x_{i4}} \pi_M^{(t)} + p_{U_1}^{(t)x_{j1}} p_{U_2}^{(t)x_{j2}} p_{U_3}^{(t)x_{j3}} p_{U_4}^{(t)x_{j4}} (1-\pi_M^{(t)})} \tag{16}$$

2. Maximization step:

$$\hat{\boldsymbol{\theta}}_l^{(t+1)} = \underset{\boldsymbol{\theta}_l}{\operatorname{argmax}} \left(\sum_{i=1}^{n} \hat{\omega}_i^{(t)} \log\left(\mathbf{P}(\boldsymbol{x}_i|\boldsymbol{p}_{\boldsymbol{M}})\pi_M\right) + (1-\hat{\omega}_i^{(t)})\log\left(\mathbf{P}(\boldsymbol{x}_i|\boldsymbol{p}_{\boldsymbol{U}})(1-\pi_M)\right)\right)$$

$$\text{for } l = 1, \ldots, 9. \tag{17}$$

[19] For an intuitive proof, see Train (2009), pages 349 to 353.

In the case of the categorical PMFs, the $\hat{\boldsymbol{\theta}}^{(t+1)}$ is given by

$$
\begin{aligned}
p_{M_k}^{(t+1)} &= \underset{\substack{p_{M,k} \\ s.t. \sum_{k=1}^{4} p_{M_k}=1}}{\operatorname{argmax}} \left(\sum_{i=1}^{n} \hat{\omega}_i^{(t)} \log\left(p_{M_1}^{x_{i1}} p_{M_2}^{x_{i2}} p_{M_3}^{x_{i3}} p_{M_4}^{x_{i4}} \pi_M\right) + (1-\hat{\omega}_i^{(t)}) \log\left(p_{U_1}^{x_{i1}} p_{U_2}^{x_{i2}} p_{U_3}^{x_{i3}} p_{U_4}^{x_{i4}} (1-\pi_M)\right) \right) \\
&= \frac{\sum_{i=1}^{n} \hat{\omega}_i^{(t)} x_{ik}}{\sum_{k=1}^{4} \sum_{i=1}^{n} \hat{\omega}_i^{(t)} x_{ik}} \text{ for } k = 1, \ldots, 4 \\
p_{U_k}^{(t+1)} &= \underset{\substack{p_{U_k} \\ s.t. \sum_{k=1}^{4} p_{U_k}=1}}{\operatorname{argmax}} \left(\sum_{i=1}^{n} \hat{\omega}_i^{(t)} \log\left(p_{M_1}^{x_{i1}} p_{M_2}^{x_{i2}} p_{M_3}^{x_{i3}} p_{M_4}^{x_{i4}} \pi_M\right) + (1-\hat{\omega}_i^{(t)}) \log\left(p_{U_1}^{x_{i1}} p_{U_2}^{x_{i2}} p_{U_3}^{x_{i3}} p_{U_4}^{x_{i4}} (1-\pi_M)\right) \right) \\
&= \frac{\sum_{i=1}^{n} (1-\hat{\omega}_i^{(t)}) x_{ik}}{\sum_{k=1}^{4} \sum_{i=1}^{n} (1-\hat{\omega}_i^{(t)}) x_{ik}} \text{ for } k = 1, \ldots, 4 \\
\pi_M^{(t+1)} &= \underset{\pi_M}{\operatorname{argmax}} \left(\sum_{i=1}^{n} \hat{\omega}_i^{(t)} \log\left(p_{M_1}^{x_{i1}} p_{M_2}^{x_{i2}} p_{M_3}^{x_{i3}} p_{M_4}^{x_{i4}} \pi_M\right) + (1-\hat{\omega}_i^{(t)}) \log\left(p_{U_1}^{x_{i1}} p_{U_2}^{x_{i2}} p_{U_3}^{x_{i3}} p_{U_4}^{x_{i4}} (1-\pi_M)\right) \right) \\
&= \frac{1}{n} \sum_{i=1}^{n} \hat{\omega}_i^{(t)}
\end{aligned}
$$

8.4 Simulation Experiment

Simulation experiments can be used to test the performance of the EM algorithm. The experiment used here is designed to mimic the historical and social security data that I match in this paper. Table 16 lists the four variables used for matching, and the probability distribution used to model their similarity. Using first names, for example, I construct a Jaro-Winkler similarity measure which I in turn partition into four categories following Abramitzky et al. (2019). Recall that in this paper the categories are given by the sets $I_{11} = [0, 0.8)$, $I_{12} = [0.8, 0.9)$, $I_{13} = [0.9, 1)$, $I_{14} = \{1\}$. For true matches, we expect first names to be identical in both records for the overwhelming number of cases, which means $p_{M_{14}}$ should be close to 1.[20] To reflect this fact, I set $p_{M_{14}} = 0.97$ in each simulation. By contrast, if two records are not matches, then the probability of the first names are identical, $p_{U_{14}}$, will be much lower than $p_{M_{14}}$. It should be noted that because I block candidate matches on last names and first initial, the number of true non-matches who by sheer

[20] However, $p_{M_{14}}$ will not be exactly one because of transcription errors, nicknames, name changes, etc.

chance nevertheless have identical first names will be exaggerated. For this reason, I set $p_{U_{14}} = 0.3$—a relatively high value for true non-matches—throughout the simulation. Since the casualty file has about 85,000 survivors, and the SSDI has 90 million observations, the share of true matches π_M will not exceed 0.001. In the simulations performed here, I use values of $\pi_M = 0.01$ and $\pi_M = 0.001$.

Table 16: EM parameter estimates at various sample sizes

Casualty File	SSDI	Similarity Measure	Likelihood functions
First Name	First Name	Jaro-Winkler	$p_{M_{11}}^{x_{i1}} p_{M_{12}}^{x_{i2}} p_{M_{13}}^{x_{i3}} p_{M_{14}}^{x_{i4}}$, $p_{U_{11}}^{x_{i1}} p_{U_{12}}^{x_{i2}} p_{U_{13}}^{x_{i3}} p_{U_{14}}^{x_{i4}}$
Middle Initial	Middle Name	$\mathbb{1}$(If same)	$\theta_{M_1}^{x_{i5}}(1-\theta_{M_1})^{(1-x_{i5})}$ $\theta_{U_1}^{x_{i5}}(1-\theta_{U_1})^{(1-x_{i5})}$
YOB Distribution	DOB	YOB p-value	$p_{M_{21}}^{x_{i6}} p_{M_{22}}^{x_{i7}} p_{M_{23}}^{x_{i8}} p_{M_{24}}^{x_{i9}}$, $p_{U_{21}}^{x_{i6}} p_{U_{22}}^{x_{i7}} p_{U_{23}}^{x_{i8}} p_{U_{24}}^{x_{i9}}$
State of Residence	State SSN Issued	$\mathbb{1}$(If same)	$\theta_{M_2}^{x_{i10}}(1-\theta_{M_2})^{(1-x_{i10})}$ $\theta_{U_2}^{x_{i10}}(1-\theta_{U_2})^{(1-x_{i10})}$

Notes: Variables used for matching. The YOB distribution is conditional on rank and is constructed using the birth cohort of fatalities.

8.5 Simulation Results

8.5.1 $\pi_M = 0.01$

Table 17 reports simulation results for all 21 parameters, using $\pi_M = 0.01$ in the data-generating process (DGP) with 10,000 iterations. The first sample contains 100 draws from the (truly) matched joint distribution, whose parameters are given in the row labeled "Actual" in Panel A, and 9900 draws from the (truly) unmatched joint distribution, whose parameters are given in the row labeled "Actual" in Panel B. Consider Panel A, Sample 1.

With only 100 draws, the EM algorithm returns $\hat{\pi}_M = 0.04$, which is closer to the true value of 0.01 than to the initial value of 0.1. However, for many parameters the algorithm moves in the wrong direction. For example, even though true matches reside in the same state 80% of the time (i.e. $\theta_{M_1} = 0.8$), EM returns $\hat{\theta}_{M_1} = 0.5$, which is a worse approximation than the initial guess of 0.6. The situation is not helped by doubling the number of iterations to $20,000$, as seen in Table 18, Panel A, Sample 1. By contrast, the estimation results in Panel B are nearly identical to the actual values used to simulate the data, which reflects the fact that there are 99 times as many draws from the unmatched distribution in a given sample by design.

Sample 2 increases the number of draws from each joint distribution by a factor of 10. Since the parameter values for the unmatched joint distribution appear to have already converged in Sample 1, I will only discuss Panel A, which reports results for the matched distribution. Sample 2 now contains 1000 draws from the matched joint distribution, and the parameter estimates are markedly better. EM now returns a value of $\hat{\theta}_{M_1} = 0.74$, which is much closer to the true value of 0.8 than the initial guess of 0.6. Virtually all the parameter estimates are improved. One exception is $\hat{p}_{M_{14}} = 0.99$, which is further from the actual value than it was in Sample 1, but the other parameters from the categorical distribution, $\hat{p}_{M_{11}}, \hat{p}_{M_{12}}$ and $\hat{p}_{M_{13}}$, are all improved. In Sample 3, the parameter estimates further improve three of the four marginal distributions (Coin 1, Coin 2, Die 2). And in Sample 4, all of the distributions improve again. Die 1 presents a challenge for EM, and the situations is not resolved when the number of iterations is doubled, as shown in Table 18. This appears to be caused by the fact that the true value of $p_{M_{11}} = 0$. Other simulations results (not shown) suggest that when a non-zero mass is chosen for all $p_{M_{jk}}$ in the DGP, EM converges as expected. Thus, for the experiment considered here, EM yields a relatively high number of false positives.

8.5.2 $\pi_M = 0.001$

Preliminary results suggest that EM performs worse when π_M is small, even controlling for N_M. This may be due to the fact that $\pi_M^{(0)} = 0.1$ is now an even worse first guess. A future draft of this paper will explore this finding further.

8.5.3 Identification of mixture models and EM

Mixing models sometimes produce distributions whose parameters are not identified. For example, suppose the only information we have about a person is their state of residence, which can be modeled as a Bernoulli mixture $(\theta_M, \theta_U, \pi_M)$. Suppose further that we observe $x_i = 1$ for 20% of the sample. One possibility is that this sample was generated by the triple $(\theta_M, \theta_U, \pi_M) = (1, 0, 0.2)$. Another possibility is that the sample was generated by $(\theta_M, \theta_U, \pi_M) = (0.8, \frac{2}{15}, 0.1)$. Unlike the one class case (i.e. π_M is known to be one), the the triple $(\theta_M, \theta_U, \pi_M)$ cannot be identified. However, when joined with other distributions such as the categorical distribution used in the simulations, $(\theta_M, \theta_U, \pi_M)$ can be identified along with the categorical parameters, and EM will converge. Even with two categorical variables, the EM algorithm will not converge if two Bernoullis and two categorical variables are used, which is why the second Bernoulli (e.g. "Coin 2") is not used to compute the responsibility ω_i in the E step. The general conditions under which mixture models are identified varies by choice of marginal distribution (see Allman et al., 2009, Gyllenberg et al., 1994, Carreira-Perpinán and Renals, 2000 for details). For this reason, practitioners are encouraged to examine the properties of the EM algorithm on a well-designed simulation experiment before applying it to real data.

Table 17: EM parameter estimates at various sample sizes (10,000 iterations, $\pi_M^{True} = 0.01$)

A. Matched joint distribution

		Coin 1	Coin 2	Die 1				Die 2				
Sample	N_M	θ_{M_1}	θ_{M_2}	$p_{M_{11}}$	$p_{M_{12}}$	$p_{M_{13}}$	$p_{M_{14}}$	$p_{M_{21}}$	$p_{M_{22}}$	$p_{M_{23}}$	$p_{M_{24}}$	π_M
1	1×10^2	0.50	0.37	0.03	0.00	0.00	0.97	0.52	0.15	0.23	0.10	0.04
2	1×10^3	0.74	0.54	0.00	0.00	0.01	0.99	0.32	0.27	0.20	0.20	0.01
3	1×10^4	0.85	0.62	0.07	0.01	0.01	0.91	0.31	0.26	0.25	0.18	0.01
4	1×10^5	0.78	0.57	0.04	0.03	0.03	0.90	0.36	0.24	0.24	0.16	0.01
	True	0.80	0.70	0.00	0.01	0.02	0.97	0.36	0.24	0.24	0.16	0.01
	Initial	0.60	0.30	0.55	0.20	0.15	0.10	0.55	0.20	0.15	0.10	0.10

B. Unmatched joint distribution

		Coin 1	Coin 2	Die 1				Die 2				
Sample	N_U	θ_{U_1}	θ_{U_2}	$p_{U_{11}}$	$p_{U_{12}}$	$p_{U_{13}}$	$p_{U_{14}}$	$p_{U_{21}}$	$p_{U_{22}}$	$p_{U_{23}}$	$p_{U_{24}}$	$1 - \pi_M$
1	99×10^2	0.29	0.19	0.76	0.13	0.07	0.04	0.64	0.16	0.16	0.04	0.96
2	99×10^3	0.30	0.20	0.73	0.13	0.07	0.07	0.64	0.16	0.16	0.04	0.99
3	99×10^4	0.30	0.20	0.73	0.13	0.07	0.07	0.64	0.16	0.16	0.04	0.99
4	99×10^5	0.30	0.20	0.73	0.13	0.07	0.07	0.64	0.16	0.16	0.04	0.99
	True	0.30	0.20	0.73	0.13	0.07	0.07	0.64	0.16	0.16	0.04	0.99
	Initial	0.45	0.10	0.70	0.20	0.09	0.01	0.70	0.20	0.09	0.01	0.90

Notes: N_M (N_U) denotes the number of observations drawn from the matched (unmatched) joint distributions. Coin 2 is not used to compute the responsibility weights in the E step because otherwise EM does not converge. Pseudo-random data was generated using Python's Numpy package with the seed 1234. The maximum number of iterations is 10^4, with a tolerance of 10^{-5} on changes to the log-likelihood between iterations.

Table 18: EM parameter estimates at various sample sizes (20,000 iterations, $\pi_M^{True} = 0.01$)

		A. Matched joint distribution										
		Coin 1	Coin 2	Die 1				Die 2				
Sample	N_M	θ_{M_1}	θ_{M_2}	$p_{M_{11}}$	$p_{M_{12}}$	$p_{M_{13}}$	$p_{M_{14}}$	$p_{M_{21}}$	$p_{M_{22}}$	$p_{M_{23}}$	$p_{M_{24}}$	π_M
1	1×10^2	0.50	0.37	0.01	0.00	0.00	0.99	0.51	0.15	0.23	0.10	0.04
2	1×10^3	0.74	0.54	0.00	0.00	0.01	0.99	0.32	0.27	0.20	0.20	0.01
3	1×10^4	0.85	0.62	0.07	0.01	0.01	0.91	0.31	0.26	0.25	0.18	0.01
4	1×10^5	0.78	0.57	0.04	0.03	0.03	0.90	0.36	0.24	0.24	0.16	0.01
	True	0.80	0.70	0.00	0.01	0.02	0.97	0.36	0.24	0.24	0.16	0.01
	Initial	0.60	0.30	0.55	0.20	0.15	0.10	0.55	0.20	0.15	0.10	0.10

		B. Unmatched joint distribution										
		Coin 1	Coin 2	Die 1				Die 2				
Sample	N_U	θ_{U_1}	θ_{U_2}	$p_{U_{11}}$	$p_{U_{12}}$	$p_{U_{13}}$	$p_{U_{14}}$	$p_{U_{21}}$	$p_{U_{22}}$	$p_{U_{23}}$	$p_{U_{24}}$	$1 - \pi_M$
1	99×10^2	0.29	0.19	0.76	0.13	0.07	0.04	0.64	0.16	0.16	0.04	0.96
2	99×10^3	0.30	0.20	0.73	0.13	0.07	0.07	0.64	0.16	0.16	0.04	0.99
3	99×10^4	0.30	0.20	0.73	0.13	0.07	0.07	0.64	0.16	0.16	0.04	0.99
4	99×10^5	0.30	0.20	0.73	0.13	0.07	0.07	0.64	0.16	0.16	0.04	0.99
	True	0.30	0.20	0.73	0.13	0.07	0.07	0.64	0.16	0.16	0.04	0.99
	Initial	0.45	0.10	0.70	0.20	0.09	0.01	0.70	0.20	0.09	0.01	0.90

Notes: N_M (N_U) denotes the number of observations drawn from the matched (unmatched) joint distributions. Coin 2 is not used to compute the responsibility weights in the E step. Pseudo-random data was generated using Python's Numpy package with the seed 1234. The maximum number of iterations is 2×10^4, with a tolerance of 10^{-5} on changes to the log-likelihood between iterations. Sample 4 is in progress.

Table 19: EM parameter estimates at various sample sizes (10,000 iterations, $\pi_M^{True} = 0.001$)

A. Matched joint distribution

		Coin 1	Coin 2	Die 1				Die 2				
Sample	N_M	θ_{M_1}	θ_{M_2}	$p_{M_{11}}$	$p_{M_{12}}$	$p_{M_{13}}$	$p_{M_{14}}$	$p_{M_{21}}$	$p_{M_{22}}$	$p_{M_{23}}$	$p_{M_{24}}$	π_M
1	1×10^2	0.28	0.19	0.76	0.11	0.00	0.03	0.18	0.48	0.27	0.07	0.028
2	1×10^3	0.50	0.34	0.46	0.09	0.03	0.41	0.54	0.22	0.18	0.06	0.008
3	1×10^4	0.67	0.45	0.44	0.08	0.05	0.43	0.48	0.17	0.23	0.12	0.004
	True	0.80	0.70	0.00	0.01	0.02	0.97	0.36	0.24	0.24	0.16	0.001
	Initial	0.60	0.30	0.55	0.20	0.15	0.10	0.55	0.20	0.15	0.10	0.01

B. Unmatched joint distribution

		Coin 1	Coin 2	Die 1				Die 2				
Sample	N_U	θ_{U_1}	θ_{U_2}	$p_{U_{11}}$	$p_{U_{12}}$	$p_{U_{13}}$	$p_{U_{14}}$	$p_{U_{21}}$	$p_{U_{22}}$	$p_{U_{23}}$	$p_{U_{24}}$	$1 - \pi_M$
1	99×10^3	0.30	0.20	0.73	0.13	0.07	0.07	0.65	0.15	0.16	0.04	0.972
2	99×10^4	0.30	0.20	0.73	0.13	0.07	0.07	0.64	0.16	0.16	0.04	0.992
3	99×10^5	0.30	0.20	0.73	0.13	0.07	0.07	0.64	0.16	0.16	0.04	0.996
	True	0.30	0.20	0.73	0.13	0.07	0.07	0.64	0.16	0.16	0.04	0.999
	Initial	0.45	0.10	0.70	0.20	0.09	0.01	0.70	0.20	0.09	0.01	0.99

Notes: N_M (N_U) denotes the number of observations drawn from the matched (unmatched) joint distributions. Coin 2 is not used to compute the responsibility weights in the E step because otherwise EM does not converge. Pseudo-random data was generated using Python's Numpy package with the seed 1234. The maximum number of iterations is 10^4, with a tolerance of 10^{-5} on changes to the log-likelihood between iterations. Sample 4 is pending.