The Labor Market Outcomes of WWI Veterans: Positive Results Despite Limited Support

Nikolai Boboshko

September 2020

Abstract: After military service, returning troops receive substantial benefits from the government. An exception to this is WWI veterans, who are unique in that they received only limited support after the war. What benefits they gained were granted late, years after the end of the war, creating an opportunity to study the impact of military service that is not confounded by large support packages. I implement a family fixed effects model to study the labor market outcomes of these veterans. WWI military service is associated with short-run labor market penalties. These adverse effects are expected, given limited aid to veterans. However, against expectations, there is substantial veteran labor market gains in the long-run. These gains are twice as large for men from a disadvantaged background, but so are the associated costs of military service. These positive results do not imply that veterans will always perform well in the labor markets and that comprehensive support policies are not necessary. Instead, the positive veteran premium is in part due to the unique characteristics of the early 20th-century labor markets.

1 Introduction

Veterans receive substantial benefits after the end of their military service. These include free vocational training and university education, home and business loans, priority in public service jobs, free or low-cost medical services, and additional pension compensation. The importance of these benefits, such as the WWII GI Bill, cannot be overstated (Humes, 2006). The broad support that veterans enjoy has been found to increase their educational attainment (Angrist and Chen, 2011) and raise their earnings (Card and Cardoso, 2012; Greenberg et al., 2020). Yet, as most veterans are eligible for these large compensation packages it is difficult to understand the role these support policies have on veteran outcomes.

In order to find a group of veterans that did not receive generous benefits, it is necessary to look at those who served in WWI. Veterans of earlier wars were compensated with land and cash transfers, while veterans of later wars received substantial benefits modelled on the WWII GI Bill. In contrast, WWI veterans received benefits only years after the war. Therefore, they present a unique opportunity to study what happened to veterans without the presence of substantial benefit policies.

Notwithstanding the uniqueness of WWI veterans, we know little about them and their outcomes. We do not know how well they reintegrated into the economy, despite receiving limited government support. My paper fills this gap by studying the impact that WWI military service had on veteran's employment and earnings. To account for selection into the army, I implement a family fixed-effects model. It compares outcomes between matched siblings. Implementing family fixed-effects is a non-trivial task as panel datasets do not exist for this period. I create a large historical panel dataset by taking advantage of Census Data digitized by IPUMS (Ruggles et al., 2020) that contains the first and last names of every individual in the United States. With this and other identifying information, it possible to track the same individual across multiple census years. Specifically, I use the 1900 census to identify siblings while they reside in the same household. Then these siblings are identified in later census years that contain information on outcomes of interest, such as employment. This matching procedure creates a large and representative panel that contains hundreds of thousands of individuals.

A key limitation in the panel data is that the US Census did not collect individual income data until 1940. Therefore, the analysis relies on constructing an earnings proxy called an earnings score. It is the predicted earnings based on an individuals industry, occupation, region and race complied from a multiple of sources. The key results in the paper are robust to several different earning imputations and in 1940, the year for which it is possible to compare score data to actual income data, the results are identical.

Applying the family fixed effects model to the newly created panel data, I find that WWI veterans are less likely to be employed after finishing military service. These adverse employment effects appear in 1920, immediately after the war. They then persist up to 1940, 28 years after the end of WWI. Although these estimates are negative, they are small. In 1920 veterans are 2.0 percentage points (pp) less likely to be in the labor force. By 1940 the employment rate for veterans is 1.6 pp smaller than that of non-veterans.

The impact of veteran status on my measure of earnings is initially negative. In 1920 veterans earn 1.7 percent less than non-veterans. This negative result is what we would expect given that veterans missed out on civilian labor market experience, experienced wartime trauma, were paid token wages, and did not benefit from a comprehensive post-service benefit program. However, by 1930 the estimate flips. The average veteran earns 9.6 percent more than the average non-veteran. This large premium begins to fade; by 1940, the veterans' earnings premium falls to about 6.0 percent.

Without an explicit random assignment, the results may be due to endogeneity that can occur within families. I demonstrate that the result is robust to alternative empirical strategies and specifications. First, a concern with family fixed effects estimates is that siblings can be exposed to drastically different home environments. To control for different home environments, I go beyond most sibling comparison papers and limit the sample to twin pairs. Twin estimates control for all time-varying factors within the family. In the second test, the baseline model is modified to include controls for pre-treatment covariates that should not be affected by serving in WWI, such as educational attainment and earnings before the war.

Estimates are qualitatively similar among these tests. These are vigorous robustness checks and not often implemented in family fixed-effects models. For typical papers, the sample sizes are too small to implement twin estimates credibly. Many papers look at the impact of educational policies on younger children and consequently cannot control for educational attainment. Data limitations prevent them from testing balance on the lagged outcome variables.

To explore the heterogeneity of the results, I leverage the panel nature of the data and split the sample based on the father's earnings in 1900. Men whose father earned little can be considered disadvantaged and likely raised in worse childhood environments than advantaged men. Prior academic work identified positive wage impact of military service on men who are likely disadvantaged. For example, army service increased wages for those with low educational attainment (Card and Cardoso, 2012) and African-Americans (Greenberg et al., 2020). Those studies focus on peacetime service, after which veterans were provided with substantial benefits, such as funding for education. Will they hold for wartime military service or if veterans do not receive large post-war benefits? WWI veterans serve as a test case as they were exposed to risks that do not exist in peacetime and were not provided with support programs that are likely to benefit the disadvantaged, such as the GI Bill (Olson, 1974).

I find that in 1920 there are no differences in outcomes between veterans with high or low earning fathers. By 1930 and 1940, substantial differences appear. Veterans whose father earned in the bottom quartile gain a earning premium more than twice as high than those with fathers in the top quartile. This premium does not come without costs, employment penalties are higher for disadvantaged men, while non-existent for others. These results imply that wartime military service in WWI, even without support packages, benefits disadvantaged men and improves intergenerational mobility, but does so at a cost.

Examining the mechanisms behind the results, I present evidence that the 1920 earnings penalty can be accounted for by the loss of civilian labor market experience during WWI. Civilian experience during this time was especially valuable as increased wartime demand combined with depressed labor supply created opportunities for occupational advancement not available in peacetime. Veterans should not have been able to take advantage of such opportunities. I find that is the case as they are less likely to be employed in industries that experience high wage growth during WWI. Furthermore, I estimate a Mincer wage equation (Mincer, 1974) and find that the 1920 penalty is what we would expect if we combine the average length of service and estimates of returns to experience.

The positive 1930 earnings premium occurs largely due to two factors unique to the early 20th Century United States. First, the experience-earnings profile flattens quicker in the past than in the present. By 1930 differences in experience no longer matter, allowing veterans to catch up. Second, I find that veterans were more likely to engage in high return activities. They were more likely to move across states and counties, more likely to reside in urban areas, and had a higher propensity to leave low paying farming occupations. Compared to the present, where the population is mainly urban and non-farmer, the decision to abandon farming or leave rural areas were available to a substantial portion of the population. Exploring the fade out in the veteran earnings between 1930 and 1940, I find that veterans experienced identical earnings declines compared to non-veterans with similar skills. That an identical wage decrease is observed in non-veterans with similar skills implies that the veteran premium fell due to structural changes. These changes lowered the returns to skills that veterans acquired by 1930, lowering the veteran premium.

In summary, WWI veterans had lower employment rates and did experience an initial earnings penalty. However, a surprising result is that by 1930 veterans experienced an earnings premium. Positive results do not mean that most veterans will easily reintegrate back into civilian labor markets without any support. The positive findings should not be extrapolated to the present. WWI veterans earn higher wages in part due to unique features of the early 20th Century U.S. labor markets.

2 Military Service During WWI

2.1 WWI Military Service, Costs and Compensation

The United Stated states participated in WWI from April 1917 to November 1918. Approximately 4.7 million Americans would go on to serve in the conflict. In this section, I briefly summarize the costs that the millions of veterans incurred from military service, as well as the benefits they gained. Starting with costs, participants in a military conflict expose themselves to substantial harms that we expect to lower wages. Veterans forgo civilian labor market experience, higher civilian wages, suffer physical and mental trauma, and risk death. The average length of wartime military service likely ranged between 312 and 390 days (Herbert, 1931). While in service, the average army pay was a dollar a day. This pay is approximately equal to 47 percent of the 1920 civilian wages of men in the sample. However, the cost of living in the army was lower due to free lodging and food.

In addition to the labor market penalties, veterans also suffered physical and mental trauma. During the war, 116 thousand servicemen died, 63 thousand due to diseases and viral infections, 53 thousand due to combat. Another 204 thousand were wounded (Leland, 2020), 4.4 thousand lost all or a portion of a limb (Department of Veteran Affairs, 2007), and 76 thousand suffered from "shell shock" (Chrisinger, 2019). It is difficult to state what portion of those injuries permanently affected human capital. Some could have been temporary, while injuries such as lost limbs had permanent effects.

To compensate veterans for the costs of military service, the U.S. government initiated a targeted response aimed at veterans who suffered wounds and disabilities. These policies consisted of rehabilitation, training, and disability payments. Significant medical rehabilitation and training services were provided to disabled veterans. The goal was to reintegrate them into the civilian labor market, promote self-sufficiency, and reduce government dependence.

The training component consisted of either school-based training or a vocational apprenticeship program. This was a popular initiative, 675'000 applied, more than the total number of those officially wounded. Out of the applicant pool, 180'000 entered the program. The completion rate was high; 129,000 completed the rehabilitation and training process. Many of the employers participating in the apprenticeship component guaranteed employment by completion. About 97 percent of those who finished the program were placed in a job (Schmick, 2018).

Finally, a pension was provided to disabled veterans if their income was below a given threshold. The size and the eligibility of the pension rules were modified several times, but the overall amount was modest in size. For example, in 1939, a disabled veteran would receive a $480 per year pension payment (Schmick, 2018). This payment is equivalent to 42 percent of the average 1940 sample income.

While supporting disabled veterans was a popular policy, providing broad benefits to all was controversial. Calvin Coolidge, the 30th president of the United States, stated that “Patriotism… bought and paid for is not patriotism” before he vetoed a bill that would have provided additional compensation (Greenberg, 2006). Due to its contentiousness, the broad benefits available to all veterans have an odd structure and went through multiple changes. In 1924 U.S. congress passed the World War Adjusted Compensation Act. It provided veterans bonus payments that had the structure of a financial instrument, similar to a bond.

The value of this financial instrument was equal to one dollar for each day a veteran served in the United States and $1.25 for each day served overseas, then multiplied by 1.25. The resulting amount would accrue a yearly interest of four percent. It would mature in 1945 and then be paid out in full to the veterans. The average 1945 payment would have been almost $1,000 (Herbert, 1931). For reference, the yearly earnings of men in my sample in 1940 are close to $1,100. Therefore, this payment would have been a substantial amount, despite being received 27 years after the end of the war.

In practice, veterans did not have to wait until 1945 to receive payment. Beginning in 1924, veterans could withdraw 22.5 percent of the value of the financial instrument, with the value calculated based on the 1945 payment amount. For example, if the 1940 payment to a veteran was $1,000 in 1945, they could receive an early payment of $225. The total amount that could be withdrawn was modest, equal to approximately 29 percent of the yearly earnings of men in 1920. However, these early payments also came at a cost, as interest would not be accumulated on the withdrawn amount, lowering the lifetime size of the benefits.

The benefit structure then went through two legislative changes. In 1931 the amount that could be withdrawn was further increased to 50 percent. Finally, in 1936 the benefit program was ended. Veterans would immediately receive a payment equal to the value of the bond in 1936. The average 1936 payment was $547 (Hausman, 2016). Although this sum was substantial, in monetary terms, it was much smaller and provided much later than the benefits received by veterans of later wars. If the reader is to take away anything from the confusing benefit structure, it should be that the veterans did receive a very late payment, almost 20 years after the war, equal to half of the average income.

2.2 Previous Research on WWI Veterans

Research on WWI is limited. Mazumder (2017) finds that immigrants who served in WWI were more likely to integrate and be accepted into American society. Work by Laschever (2005) identified positive peer effects among veterans during the Great Depression. The social networks formed in the military helped veterans find employment opportunities. Schmick (2018) studied the impact of a significant rehabilitation program for disabled veterans. Completion of the program leads to increased employment but did not affect earnings.

Two papers study the impact on WWI military service on labor market outcomes. Gabriel (2016) employs linear regression to study veteran income in 1940. To account for selection, the right-hand side variables include age, education, race as well as occupation, industry, and county of residence. He finds that veterans earn four percent more on average. However, including controls such as occupation and residence is problematic. These variables themselves are likely affected by veteran status and are therefore endogenous. Due to endogenous variables, the positive results are either difficult to interpret or are a conditional estimate. Finally, using the 1940 census alone to study veterans runs into a severe data issue due to large non-reporting of veteran status that biases estimates towards zero. In contrast, I rely on the 1930 census for information on veteran status.

Tan (2020) identifies the impact of WWI military service in 1930 through a regression discontinuity design. He exploits the variation in the eligibility for the WWI draft that occurs due to fixed registration dates. To be eligible, Americans had to be within a specific age range on the draft registration day. By comparing men who just made the age cutoff to those who just missed it is possible to identify an unbiased effect of military service. Seasonal effects exist around the cutoff and are differenced out by using cohorts not affected by the policy. For most veterans, there are no statistically significant effects from military service on the earnings proxies. Unfortunately, the estimates vary with the bandwidth choice and are noisy. For example, the point estimates for young veterans range from a large eight percent earning score premium to a modest three percent penalty. The standard errors on the point estimate are also substantial, creating a broad 95 percent confidence interval. Therefore, the true estimate can be as much as +/- 18 points away from the coefficient.

My contribution to this literature is to estimate the impact of WWI military service on short-run outcomes immediately after the war as well as to provide long-term results using estimates that account for selection by comparing siblings. Providing both short-run and long-run results is different from previous papers that focus on a single point in time. It is also quantitatively important as the short-run and long-run effects on earnings differ substantially. Sibling estimates have the advantage of being precise, do not include endogenous controls, and are robust to strong tests for their validity. Thus, this is the first paper to identify precise, significant, and robust estimates for WWI veterans that have a causal interpretation. Finally, none of the papers explored heterogeneity of outcomes based on men's father's earnings or other measures of skill and childhood circumstances.

3 Data

My main analysis estimates the outcomes of veterans in 1920, 1930, and 1940. This is done using new datasets that consist of linked census records. In the baseline dataset, I link the 1900 census to the 1930 census. This link is necessary for two reasons. First, in the 1900 census, future veterans reside with their parents. I use this information to identify siblings and implement family fixed effects estimates. Second, only the 1930 census contains information on veteran status for the whole population. With information on siblings and veteran status, the baseline dataset is then linked to 1920 and 1940. These resulting datasets allow me to measure labor market outcomes in adulthood, test heterogeneity based on father's occupation, and track both the immediate and long-run impact WWI military service.

3.1 Census Record Linkage

To create the linked sample, I first use the full count census data from 1900. I limit the sample to men between the ages of 0 to 13 who are born in the United States. This restriction ensures that the men in the sample reside at home with their parents so that it is possible to identify sibling pairs. It also covers the vast majority of future WWI veterans. I then match these individuals to later censuses using first name, last name, place of birth, and year of birth. This match is done by implementing an automated linking procedure outlined in Feigenbaum (2016).

I will briefly summarize the linking process here; more details are available in Feigenbaum (2016). The goal of the automated linking procedure is to replicate the heuristics that a trained researcher will use to link census records manually. The first step is to draw a random sample of observations and their potential matches. Then, it is necessary to identify links in this sample manually, taking into account factors such as the similarity between first and last names, the total number of potential matches, and any potential age differences. The resulting links are coded as one, while un-matched links are coded as zero.

With all the links coded, the next step is to estimate a probit model where the dependent variable is equal to one if there is a match. The coefficients from the probit can then be applied to predict the probability that any two observations are a match for the full sample, not only for the random subset.

Given these probabilities, only the match with the highest probability among the potential matches is a candidate for being a true match. However, not all matches with the highest probabilities are likely to be true. For example, even the best possible match can have a very low probability of being true. In that case, it is better to leave observations unmatched instead of making a false link. Alternatively, there could be too many potential matches with high probabilities, making it challenging to identify the correct link among many candidates. Therefore, an additional step is necessary in order to identify which matches with the highest probability are true and which ones are false.

In this last step, it is necessary to pick two parameters. The first parameter is the probability cutoff that observations have to meet to be considered matched. For example, it could be that only observations with a probability of 0.8 are true matches. The purpose of this cutoff is to take into account cases where the best match, as in one with the highest probability, is unlikely to be true. The second parameter is the ratio between the potential match with the highest probability and the match with the second highest. This parameter removes cases where there are many potential matches with high probabilities. Setting these parameters to be higher results in more accurate matches, reducing Type I error. However, it increases the chance that true matches will not be included, increasing Type II error.

The optimal values of these two parameters are obtained using the random sample with manual links. That sample contains two essential pieces of information, the match probabilities, and the manual matches. With the probabilities, it is possible to observe the matches that would result for different values of the two parameters. These matches can then be compared to the links that were identified manually. The procedure assumes that the manual links are the true probabilities. Therefore, we now observe both true links and links obtained from the automated process. With true and automated links, I can now calculate Type I and Type II errors for different values of the two parameters. The last step is to use the information on the error rates to choose parameter values that minimize both types of errors.

The resulting matched sample is limited to observations that have non-missing values of critical characteristics. Siblings in the matched sample are identifies using the 1900 census. Observations who reside in the same household have the same last name and have the same mother and father are classified as siblings. The main sample for family fixed effects estimation is further limited to men with at least one other identified sibling.

3.2 Summary Statistics

Summary statistics for full count census data and matched samples are available in Table 1. My implementation of the Feibenbaum (2016) procedure leads to match rates comparable to the literature. In 1930, there are ten million observations in the full count data, and I was able to find matches for 39 percent of them. For comparison, other implementations of automated linking procedures result in match rates between 28 to 54 percent (Bailey et al. 2019). The matched 1930 dataset is then further linked to 1920 and 1940. I expect the match rate to decline as it is harder to track the same individual over multiple census years. For example, it is harder to find a 1900 observation in both 1930 and 1940, than only in 1930. This declining match rate is indeed the case; it falls to 20% and 21% in 1920 and 1940, respectively.

High match rates are encouraging, but are meaningless if the matched sample is not representative of the population. I test this by comparing summary statistics between the matched sample and the full count census data. Fortunately, the matched sample closely resembles the full count census data. The employment and labor force participation rates are nearly identical in the two samples. The proportion literate is nearly identical, as well as is the average age and the fraction of WWI veterans. The matched sample completed slightly more years of schooling and is about 2.0 pp more likely to work as a farmer, relative to a base of 15 to 16 percent.

One dimension that the match does not perform well on is linking African-Americans. Blacks should make up 11 percent of the population from 1920 to 1940. However, in the matched data, their proportion ranges from five to eight percent. The error is particularly egregious in 1920, where they only make up five percent of the matched sample, half of their population proportion.

Limiting the matched sample to siblings generally does not affect the resulting summary statistics. Small differences exist. The proportion of siblings who are farmers is usually two to three pp higher than in the matched sample. Siblings have also completed 0.11 fewer years of education. A major difference is the proportion of African-Americans. It further declines to between three to seven percent in the sibling sample.

The similarity in the summary characteristics between the matched sample and the population is encouraging. The matching procedure generated a representative sample, and estimates obtained using matched data are likely to be externally valid. The conclusion also holds for the sibling sample that will be used for family-fixed effects estimation. One notable exception, though, is the low match rate of African-Americans. To account for the difficulty in matching Blacks as well as other small differences, I follow best practices and weight all the estimates by the inverse probability of matching (Bailey et al., 2019). This practice ensures that groups that are underrepresented due to differential matching rates receive higher weights.

The last two columns compare veterans and non-veterans in the matched sibling sample. Veterans tend to be slightly younger than non-veterans and have marginally lower employment and labor force participation rates. However, other characteristics strongly suggest that veterans should earn more than non-veterans. Veterans are more likely to be literate, have more years of schooling, reside in more urban areas, are less likely to be Black, and are less likely to work in farming occupations. They have almost a year more of schooling. Differences in farming and education are particularly large. The percentage of veterans who are farmers, a low paying occupation, is between 10 to 14 percent, while between 22 to 24 percent of non-veterans are in farming occupations.

3.3 Earnings Score

A challenge with historical data is that the census did not record income until 1940. Even in that decade, 1940, only wage and salary income was recorded, but a significant portion of the working population did not work for a salary. A common solution in historical work is to construct earning proxies. These measures are commonly known as earning scores.

To construct an earnings score, I first calculate the median 1940 wage in each industry, occupation, race, and region cell. Each observation in the sample is then assigned the cell-specific median wage. A potential problem with this assignment is structural economic changes between 1920 and 1940. The median wage for each cell calculated using 1940 data can differ from the true median wage in 1920 or 1930. To adjust for trends over time, I use industry level annual earnings data compiled by Lebergott (1964) and calculate wage changes between 1940-1930 and 1940-1920. The values assigned to each cell in 1920 and 1930 are then deflated using the information on national industry wage changes. These adjusted cell median wages are the earnings score for wage and salary workers.

Relying on only 1940 income data would lead to the exclusion of the self-employed and drop a large portion of the population, farmers, from the analysis. To predict the income of farmers, I follow the approach in Collins and Wanamaker (2017). This procedure relies on combining two sources of data. The first is information on the income of farm laborers; it is available at the national level for every year going back to 1900 (Lebergott, 1964). The second source is the 1960 census, the earliest census with a large sample that contains information on total farm income. To combine these two sources, Collins and Wanamker (2017) assume that the ratio of farmer earnings to farm laborer earnings in 1960 is constant over time. Then this ratio can be applied to annual data on the income of farm laborers to extrapolate farmer income in different years.

To summarize, the earnings score for salary workers is the median wage in each occupation, industry, race, and region cell. It is then adjusted for national industry wage trends that have occurred over time. The earnings score for farmers is the product of a fixed 1960 income ratio between farmers and farm-laborers multiplied by the income of farm laborers in a given year.

The last four rows in Table 1 provides summary statistics on earnings measures. Comparing the matched and full sample on earnings measures reveals close similarities. The one noticeable difference is in 1940, where the matched sample has an earnings score that is $81 higher. Limiting the sample to siblings does not lead to any new major differences compared to full count data. When comparing veterans and non-veterans, those who served in WWI earn more every year by every measure. The wage income in 1940 of veterans is more than $200 higher. The earnings score measure suggests that veterans earned $52 more in 1920, $217 more in 1930, and $203 in 1940.

4 Empirical Strategy

Because selection into military service is non-random, mean comparisons between veterans and non-veterans are likely to be biased. To counteract this bias without explicit random assignment, I use family fixed effects. By comparing siblings, all-important time-invariant family characteristics, such as parent's education and permanent income, are differenced out. The key assumption with this approach is that selection into veteran status within each family is uncorrelated with unobserved variables that also determine key outcomes.

$$Y_{ij} = \alpha + \beta \text{Veteran}_{ij} + \delta X_{ij} + \gamma_j + \epsilon_i \tag{1}$$

Formally, I estimate equation (1), where i indexed individuals and j indexes families, X_ij is a vector of controls that vary within family and γ_j are family fixed-effects. The main effect of interest is β, which is the estimated effect of veteran status on an outcome, Y_ij. Standard errors are clustered at the family level. The controls in X_ij include fixed effects for age, fixed effects for the month of birth, place of birth, and indicator variables for African-American and literacy status.

A concern to identification is the individual specific error term, ϵ_i. A priori, it is not clear in what direction this would bias results. The estimates could be upward biased because the military screened recruits on basic mental and physical faculties. Due to this, the average veteran before the war was likely healthier than the average non-veteran, and there are positive returns to health status in the labor markets. Alternatively, there is potential for downward bias. The military paid a small wage and covered all living expenses. A small wage would result in men with lower earning potential paying a small opportunity cost to serve in the army. They would then be more likely to enlist.

There are also reasons to believe that the bias is zero due to random sources of variation. About 2/3rd of the veterans were drafted. Eligible men had to register with a local draft board and were recruited based on random draws from a national lottery. By including controls into a regression model, the variation that is used for identification is increasingly due the draft. Adding family fixed effects can help further exploit draft variation if preferences for military service between siblings are correlated, perhaps due to shared levels of patriotism.

Even so, I can never be sure that the variation is entirely exogenous. Therefore, later I test the validity of family fixed effect estimates by adding powerful pre-treatment controls, such as earnings before WWI and by limiting the sample to pairs of twins. This allows me to test for some, but not all, potential sources of bias.

5 Results

5.1 Baseline Estimates

Table 2 estimates the impact of WWI military service on the log earnings score and employment. Due to data limitations, in 1920, the employment results measure labor force participation. It is not possible to identify unemployed individuals in that year. For each dependent variable, the first column estimates the mean difference. The second column adds controls that are indicator variable for the month of birth, year of birth, race, father's occupation, and literacy status. The last column includes family fixed effects.

Across the different specifications and census years, there is clear evidence of a negative impact of military service on employment. Surprisingly, the results are qualitatively similar across the different models, even though mean estimates are expected to be biased. Due to this similarity, I focus the discussion on the preferred family fixed effect estimates (column 3). In 1920, veterans are 2.0 pp less likely to be in the labor force. The employment penalty in 1930 is essentially zero, only -0.3 pp, but by 1940 it increases to -1.6 pp. All the estimates are statistically significant, military service does lower the employment rate, but the penalty is qualitatively small.

The impact of military service on the earnings score, presented in columns (4) to (6), is significantly different from the employment estimates. First, evidence of non-random assignment of veteran status is evident. At the mean, without controls (column 4), veterans earn more than non-veterans. Including controls in column (5) causes the estimates to fall. Positive estimates that fall with controls imply that men are positively selected into veteran status. For example, in 1920, the average veteran earns 6.6 percent more than a non-veteran. This positive premium is unlikely to be causal, 1920 is close to demobilization, and veterans should be re-adjusting to civilian life. The positive premium must be due to the veteran's having higher earnings potential before the war. Indeed, that is the case, as including controls in column (5) causes the 1920 earnings score estimate to fall to zero.

Given that selection into veteran status is endogenous, I focus on family fixed effects estimates. The results are available in column (6). Exploiting variation within siblings, I find that serving in WWI leads to a short-run reduction in earnings score, but a long-run earnings premium. In 1920 (Panel A), veterans earn 1.7 percent less than non-veterans. This negative result is encouraging for the validity of exploiting within sibling variation. A strong prior is that the short-run effect should be negative, and that is the case only when family fixed effects are included. By 1930, the negative earning impact disappears. Instead, veterans earn a substantial premium of 9.6 percent. This large positive effect dropped somewhat by 1940. Near the end of the Great Depression, the veteran premium falls to 6.0 percent.

5.2 Heterogeneity by Father's Earning Score

Figure 1 reports estimates of the heterogeneous impact of military service by childhood economic circumstances. The sample is split into quartiles base on the father's 1900 earnings score. For example, men in the first quartile are those whose father's earning score was equal to or below the 25th percentile in 1900. This sample split allows me to observe differences in outcomes by how disadvantaged (or advantaged) the men were growing up.

The estimates for the earnings score are available in Panel A. In 1920, there are no differences between the estimates for the four groups. In the long-run, though, the estimates diverge, and the veteran premium is significantly higher for men whose father had lower earnings score. In 1930, the earnings premium is 12.6 percent for men in the lowest quartile. It then declines monotonically as father's earnings score increases. In the fourth quartile, the veteran premium is only 5.7 percent, less than half the size of the first quartile premium. Higher estimates for men with disadvantaged childhoods persists in 1940. The difference between the first and four quartiles is persistent. The premium in the fourth quartile is again half the size of that in the first, 4 percent vs. 8.2 percent.

Results for the labor force participation rate (1920) and employment (1930 and 1940) are in Panel B. In 1920, childhood background has no impact on employment. By 1930 and 1940, this is no longer the case. By 1930 there is a minor penalty for the most disadvantaged men that fades out with increases in father's earning score. In 1940 there is no penalty for men from the top quarter. In contrast, the employment penalty for the lower three quartiles ranges from -2.3 pp to -1.6 pp and decreases marginally with father's earnings score.

The heterogeneity of employment and earnings score estimates imply that the positive and the negative effect of WWI military service were stronger for the disadvantaged. The earnings score premium falls as father's 1900 earnings score rises, while the employment penalty falls with father's earnings. Unfortunately, this benefit came at a cost; disadvantaged veterans were less likely to be employed. This also implies that intergenerational mobility improves for veterans after WWI.

6 Robustness

The positive long-run earnings score estimates are surprising and merit additional scrutiny. Endogenous selection into veteran status is a concern, and family fixed effects might not account for it. In this section, I present tests to show that the estimates are robust to likely sources of bias. First, in order to control for time-varying factors within a family, the sample is limited to twins. Second, to test the impact of any pre-treatment differences between veterans and non-veterans, I augment the baseline model with new controls. Specifically, the estimates are robust to including educational attainment and the earnings score before military service. The previous two checks test for non-random selection into veteran status. However, the employment rate is lower among veterans than non-veterans. If men with low earnings potential exit the labor force, the earnings estimates for veterans are biased upward. Therefore, the section concludes with estimates computed with alternative earnings score measures, including ones that assign an earnings score measure to the non-employed. The results are qualitatively similar across them.

6.1 Twin Fixed Effects

A concern with family fixed effects models is that time-varying factors within the family could be affecting the estimates. For example, an increase in the father's income with age could mean that an older sibling experienced poverty that the younger sibling did not. Differential exposure withing families is a concern for many papers that exploit within-family variation. The standard robustness check is to control for observable time-varying family characteristics that can vary over time, such as divorce and income. This information is not available in the linked data. Instead, I perform a much stronger test by limiting the sample to twins. Twin fixed effects estimates will account for both observed and unobserved factors that vary within a family over time.

Information in the 1900 census is used to identify sibling pairs who are twins. Twin pairs must reside in the same household, belong to the same family unit, have the same father and mother, have the same last name, and be born in the same year. Limiting the sample to twin pairs significantly limits the sample size and makes identifying results for 1920 and 1940 difficult. The new estimates can also differ from the baseline sibling estimates due to differences between the two samples. Therefore, the twin sample is then reweighted so that its characteristics match that of the sibling sample.

The results for the twin sample are available in Table 3, Column 1. The dependent variable in the table is the log earnings score. In 1920, contradicting the sibling comparison results, there is no score penalty for twins. However, the standard error is large, and the two estimates are not significantly different from each other. By 1930 twin veterans earn a score premium of 6.7 percent, a smaller premium than the sibling estimate of 9.6 percent. Again, the two coefficients are not significantly different from each other. By 1940 the veteran premium further increases in size to 9.3 percent but is only significant at the 10 percent level. An increasing premium between 1930 and 1940 is the opposite of the decline that occurs in the sibling sample. There the estimate falls to 6.0 percent.

It is encouraging that the most surprising results that veterans earn more than non-veterans in 1930 and 1940 hold in the twin sample. However, two differences do exist. First, in 1920 employment and earning score estimates are less negative in the twin sample. Second, the trend between 1930 and 1940 is different. In the sibling sample, veteran premium declines, while it increases in the twin sample.

Although the estimates do not align perfectly, I would like to emphasize the broad similarities that occurred despite the large reduction in sample size. In 1930 more than 589 thousand observations are used to estimate earnings impact in the sibling sample. In the twin sample, this falls close to nine thousand. Given these reductions, we expect point estimates to differ. Importantly, the result of a large positive premium in 1930 and 1940 holds. Finally, none of the differences between the two samples are statistically significant.

6.2 Sensitivity to Inclusion of Pre-Treatment Controls

Differences between siblings can also exist on observable characteristics before they enter into military service. I test the robustness of the family fixed effects model by controlling for their earnings score in 1910 and their educational attainment in 1940. If family fixed effect estimates properly control for selection, then the earnings of siblings in 1910 should not be correlated with future military service. The same could be true for K-12 educational attainment. However, military service could causally affect years of schooling as the military and the U.S. government did provide some educational support for veterans. Disabled veterans had the option to enroll in paid schooling. Education programs were also offered to veterans while they awaited demobilization in Europe.

To obtain information on the 1910 earnings score, I link the 1900-1930 matched dataset to the 1910 census. I do not then link this resulting dataset to the matched 1920, and 1940 data as performing too many links tends to limit sample size and reduce representativeness drastically. The 1930 estimates are of special interest as the veteran score premium during that decade is unexpectedly high.

The sample is limited to men who were at least 16 years old in the 1910 census and employed. The age cutoff of 16 is not arbitrary, as that is the age individuals were generally no longer considered to be child laborers. A large portion of them was also employed. This group would be at least 23 years old in 1917, the year of U.S. entry into WWI. A concern is that this group is not representative of the population. However, the loss in external validity is compensated for by the opportunity to test the validity of the family fixed effects estimates.

Table 4 column (2) provides results from the baseline specification on the merged sample. The coefficient on veteran status is 8.1 percent. It is smaller than the baseline Table 2 result of 9.6 percent. Column (3) includes a quartic in the 1910 earnings score as an additional control. Including past score leaves the coefficient unchanged at 8.1 percent. The constant coefficient implies that human capital differences that affect 1910 earnings score are not a significant source of bias between siblings.

Another potential source of pre-treatment bias are differences in educational attainment between siblings. Only the 1940 census contains information on years of schooling. Therefore, I merge the linked 1920 and the linked 1930 data to the 1940 census. Now it is possible to test the robustness of the baseline estimates to educational controls in all census years.

One concern with controlling for education is that it is potentially endogenous for two reasons. First, the veteran status might causally impact educational attainment. The military did provide veterans with educational programs while they awaited demobilization in France (Faulkner, 2017). The government also implemented comprehensive and popular educational and vocational training programs for disabled veterans (Schmick, 2018). Second, while age restrictions on military service ensured that veterans had completed their K-12 education before military service, this is not true for higher education. Therefore, years of schooling can vary post-treatment. How this endogeneity will affect the estimates is not clear.

The results for the impact of educational attainment are available in columns (4) and (5). Column (4) replicates the baseline specification on the new sample, while column (5) controls for educational attainment. Specifically, these new controls include a linear term for years of schooling and indicator variables for dropping out of middle school, graduating middle school, dropping out of high school, completing high school, and obtaining at least one year of college.

In 1920 (Panel A), if we compare column (4) to column (5), we see that the new controls have increased the size of the veteran penalty, from -2.2 percent to -2.5 percent. By 1930 (Panel B), controlling for education reduces the estimates slightly. The veteran premium drops by less than one percent, from 9.8 percent to 9.0 percent. Only in 1940 is there an important change to the veteran premium. Controlling for education lowers it from 5.9 percent to 4.0 percent.

Overall, the results in 1920 and 1930 are robust to the inclusion of detailed education controls. The 1940 estimates are less so. The change in the 1940 estimates could be due to selection bias not captured by family fixed effects. Alternatively, the education variables themselves could be endogenous. The small educational programs offered to the veterans could have resulted in differences in educational attainment. Overall though, the pattern of results for veterans is unchanged. The estimates for 1920 are still negative, the 1930 premium is unexpectedly large, and the decline of the premium by 1940 still exists.

6.3 Selection into Employment

The previous two robustness tests measured whether there is endogenous selection into WWI military service. Given that there is an employment penalty for veterans, endogenous selection into employment is also a concern. If veterans with low earnings potential are more likely to exit employment than non-veterans, then the earnings score results are upward biased. To account for this, I calculate to alternative earnings score measures that impute an earnings score to those who are not employed.

The first measure, based on Carruthers and Wannamaker (2020), changes the cells used to impute the earnings score, in order to take employment status into account. Previously, a step in constructing the score was calculating the median income in each industry, occupation, race, and region cell. In this test, the cells now additionally include employment status. This measure accounts for differences in total earnings that can occur due to hours worked. Unfortunately, individuals who are not in the labor force do not report occupation and industry and are consequently not included in the new sample. To include them, I create a second measure that assigns those not in the labor force an earnings score of zero. This zero imputation is an extreme assumption. Therefore, we should think of is as the upper bound on the impact that selection into employment has on the estimates.

There are two small caveats with these measures. By including zeroes, it is no longer possible to estimate the baseline model with the log-transformed earnings score. Instead, the dependent variable in this robustness test will be the raw earnings score. To facilitate comparison, I will report the marginal effect by dividing the coefficient on veteran status by the average earnings score. Fortunately, this change in functional form has a negligible impact. The second caveat is that in 1920 it is not possible to calculate differential earnings scores based on employment, as that variable is not available.

Estimates that attempt to account for selection into employment are available in Figure 2. Imputing earnings score for those who are not employed, but are in the labor force reduces the size of the veteran earnings premium. The decline in 1930 is negligible, but in 1940 taking selection into account causes the premium to fall from 6.1 percent to 4.9 percent. This decline is a small but noticeable change. Overall, imputing wages for those who are in the labor force does not affect the large 1930 veteran premium and marginally affects the 1940 premium.

Assigning zeroes to those, not in the labor force leads to more drastic changes. The 1920 veteran penalty sharply increases, from a -0.9 percent to -4.0 percent, while the 1940 premium falls to 3.6 percent. The 1930 estimate is again unaffected.

These large changes have two interpretations. One is that not accounting for selection into the labor force significantly understates the short-run earnings penalty and overstates the 1940 veteran premium. While serving in the military can raise the earnings score, once the costs, such as labor force exit due to injuries, are taken into account, the net benefits look worse. Another interpretation is that the large changes are the product of assigning an earning score of zero to those in the labor force and should be interpreted as a worst-case bound.

6.4 Alternative Earnings Score Measures

The final robustness check measures the sensitivity of the results to alternative calculations of the earnings score. My baseline score for wage workers uses income information in each occupation-industry-region-race cell in 1940. It then adjusts that income based on changes in industry wages that occurred between a given year and 1940. For farmers, who are not wage workers, the score is the product of two numbers. One is the earnings ratio between farm owners and farmworkers in 1960. The other is the income of farmworkers in 1920, 1930 or 1940. This imputation assumes that the 1960 earnings ratio between farmers and farmworkers stays relatively constant over time.

The above calculation of the earnings score is my preferred measure for two reasons. First, it captures detailed information on earnings in each occupation-industry-region-race cell. Second, it adjusts those earnings as best as possible to take into account wage trends between 1940 and 1920 or 1930. Nevertheless, in this section, I experiment with additional measures to test robustness.

The sensitivity of the estimates to alternative earnings score calculations is available in Table 6. The first column in the table replicates the main results for ease of comparison. In column (2), I modify the baseline score. It now no longer adjusts the earnings score to take into account wage trends between 1940 and 1920 or 1930. The only notable difference between the two measures occurs in 1920 (Panel A). In the baseline results, veterans suffer a 1.7 percent earning penalty. Changing to the modified score causes the penalty to disappear; the estimated effect is now close to zero. Such differences between these two measures are likely to occur in 1920 as industry wage changes are larger between 1920-1940 than 1930-1940.

For the next score, I calculate an earnings measure using data from the 1950 census. It is simply the median earnings in each occupation and is similar to the earnings score measure provided in IPUMS (Ruggels et al. 2020). The advantage of the 1950 census is that it is the first to ask about earnings from all sources. With it, it is possible to calculate an earnings score for both salary and non-salary workers using the same census data. A disadvantage of the 1950 Census is the small sample size. Earning information is available for only 0.2 percent of the U.S. population; this leads to multiple industry-occupation-region-race cells with few observations. That is why this score I based only on occupations and not the more detailed cells previously used. A further disadvantage of this score is that it ford not adjust for wage trends over time.

The results for the earnings score based on the 1950 census are available in column (3). Compared to column (1), the overall pattern of results is the same, but some differences do exist. Using the 1950 earnings score, the estimates for 1920 are close to zero and no longer significant. It appears that, similar to the discussion above, it is likely necessary to adjust for industry wage trends to obtain negative 1920 estimates. Another small difference is that in 1930 the veteran premium falls to 7.1 percent (from 9.6 percent).

Column (4) calculates an earnings score entirely based on industry wage data. The advantage of this measure is that the average income in each industry is available in 1920, 1930, and 1940. It is no longer necessary to adjust 1940 income levels to more closely resemble any given year. Unfortunately, this measure is very broad as there are only 25 industries. Despite this, the only significant difference between it and the baseline results is in 1920. If only industry data is used to calculate the earnings score, in 1920, veterans earn 4.8 percent more than non-veterans. That this is the only difference is surprising given how different the two earnings score measures are, especially given that the cells used to calculate the baseline score are very detailed. This difference implies that in 1920 veterans were employed in higher-wage industries, but within those industries were concentrated in lower-paying occupations. After 1920, the long-run estimates are nearly identical.

The final measure is not an earnings score, but actual salary information. The 1940 census asked about income from salaried occupations. The clear advantage of using direct salary income is that it is no longer necessary to impute likely income. An important disadvantaged is that earnings from other sources, such as self-employment or farm income, were not recorded. Therefore, the sample is limited to brothers who are all wage workers. The point estimates for 1940 drop, from 6.0 percent to 4.1 percent.

Despite some differences, the overall pattern of results is consistent across the different measures. Veterans earn the less in 1920, by 1930 there is a high premium, but it declines in 1940. A key difference between the measure is that the 1920 veteran penalty is only evident using the baseline earnings score. As we expect the short-run impact of military service to be negative, this shows the importance of adjusting earnings score to more closely resemble the earnings in a given year.

7 Mechanisms

In this section, I provide additional evidence that explores the mechanisms behind the veteran earning score results.

7.1 Lost Civilian Labor Market Experience

I began by assessing the impact of lost civilian labor market experience. Previous papers that studied the impact of military service on civilian outcomes found it to be an important factor (Angrist, 1990; Angrist and Chen, 2011). Figure 3 characterizes the relationship between the earnings score and age for men in the sample in 1920, 1930, and 1940. To ease comparison across different census years and to take into account differences in educational attainment across cohorts, the earnings score is normalized. The normalization is the following; the average score for the youngest men in each census year for each category of educational attainment is set to one. Then the score for older men, with comparable educational attainment, is calculated in relation to the youngest.

We see in Figure 3 that in 1920 for men between 20 and 34, the relationship between age and the normalized earnings score is steep. The oldest men in 1920 make about 20 percent more than the youngest. By 1930 and 1940, the relationship is essentially flat; older men do not earn more than younger men. One exception is a small but sharp increase in earnings between the youngest men and the second youngest, which appears to be either a cohort effect or an artifact of the match.

The above pattern implies that lost civilian labor market experience can account for the 1920 earnings penalty. Furthermore, due to the rapid flattening of the age-score relationship by 1930, any negative effects of lost experience disappear, contributing to the earnings premium. To quantitatively assess the importance of these patterns, I perform a back of the envelope calculation. The baseline family fixed effects model is augmented with variables that are standard in a Mincer wage regression, a quadratic in potential experience and educational attainment. Then, given that the average veteran served for approximately a year, I back out the decline in earnings score this would cause.

The results are available in Appendix Table 2 and support the conclusions one would draw from Figure 3. We see that in 1920 (Column 1), a loss of a year of civilian experience would lower earnings by 1.7 percent. This is identical to the 1920 earnings score penalty of 1.7 percent. So, the lost experience can explain all of the decline in earnings score. The negative impact of lost experience also fades out quickly due to the flattening of the experience-score relationship, by 1930 the expected loss is zero. The lack of any negative penalty due to lost experience contributes to the positive veteran premium that appears by 1930.

One dimension that the above estimates do not capture is that experience during WWI could have been particularly valuable. The combination of increased labor demand due to the war and lower labor supply due to conscription presented opportunities for occupational advancement and wage growth. These opportunities were not available to military veterans, and likely lowered their wages. The census data does not allow one to measure the direct impact of this on wages. However, I present indirect evidence; veterans were persistently less likely to be employed in industries that experience high wage growth during WWI.

The results are presented in Table 4, column (1). The dependent variable is the percent change in industry growth during WWI. In 1920 the average veteran was employed in an industry that had wage growth that was 1.429 percent lower than non-veterans. A statistically significant result that does not fade away. In 1930 veterans were employed in industries where WWI wage growth was 2.283 percent lower. By 1940 the effect declines to 1.622 percent, but it is still larger than the 1920 result.

These estimates are consistent with contemporaneous work that found short-run economic conditions can affect long-run economic outcomes. For example, university students who enter the labor market during a recession experience both short-run and long-run wage penalties relative to those who enter at other times (Kahn, 2010).

7.2 Farming Occupations and Urban Residence

This section presents evidence of decisions that veterans made to improve their earnings scores. Given that there are multiple such choices that veterans could have made, I focus on only two, the decision to reside in urban areas and to leave farming occupations. The reason for choosing these two is that, as discussed further below, they are associated with large increases in earnings score and are available to a large portion of the population.

I begin by testing for differences in the propensity to work in farming occupations. Farming jobs were low paid, but many men engaged in them. Therefore, leaving them was an option for many veterans and would lead to large increases in earnings score. In the sample, in 1920, the average farmer earned a slightly more than average, but by 1930 engaging in farming occupations carries a significant penalty. Farmers earn 55 percent and 62 percent less than non-farmers in 1930 and 1940, respectively (Table 4, Panels B and C). Therefore, leaving a farming occupation should be associated with a large increase in the earnings score. Furthermore, exiting farming occupations is a choice available to many individuals, as approximately 20 percent of the observations in any given year are engaged in farm work (see Table 1).

Results for the impact of military service on farming are available in Table 4, column (2). In 1920, veterans are 9.4 pp less likely to engage in farm work. This substantial effect begins to decline. By 1930 and 1940, veterans are 7.3 pp and 4.5 pp, respectively, less likely to be farmers.

Next, I explore the location choices of veterans, focusing on the decision to reside in urban places. In the early 20th Century, estimates from the baseline family fixed effects model imply that living in an urban area substantially raised the earnings score. In 1920 the earnings score of siblings in urban areas was 36 percent higher (Table 4, Column 3). By 1930 and 1940, the score premium in urban areas rose to over 50 percent. The decision to move to an urban place was also available to a large portion of the population, as about half resided in rural areas (Table 1).

Estimates in Table 4, column (3) show that serving in the army during WWI increased urban residence. The dependent variable in that column is an indicator equal to one if the individual resides in an urban area. Veterans are more likely to reside in an urban area relative to non-veterans by 2.7 pp in 1920, 6.4 pp in 1930, and 3.6 pp in 1940.

To summarize the above analysis, veterans were significantly less likely to work in farming occupations and more likely to reside in urban areas. These two decisions likely lead to large improvements in the earnings score. In order to test the cumulative impact of these decisions on veteran earnings, I augment the family fixed effects model with state-by-urban residence fixed effects and a farmer dummy variable. The difference in earnings score between siblings in this specification can be interpreted as the difference in their relative earnings with farmers/non-farmers and state-urban locations. However, as farming status and area of residence are endogenous to veteran status, these estimates should be interpreted with caution.

The results for the augmented model are available in column (4). Column (5) replicates the original results from Table 2, without the controls for state-by-urban status and farmer status. The difference between these two columns is due to veteran choices.

In 1920, the veteran estimate with the additional farmer and state-by-urban controls is -1.3 percent. In comparison, this is marginally lower than the baseline penalty of -1.7 percent estimated in Table 2. Therefore, the actions taken by veterans in 1920 did not significantly affect their earnings score. By 1930 and 1940, though, the choices made by veterans matter significantly. The additional controls lower the veteran earnings premium significantly to 4.1 percent in 1930 and 2.9 percent in 1940. In comparison, the baseline estimate in column (5) is a premium of 9.6 percent and 6.0 percent in 1930 and 1940, respectively. More than half of the veteran premium in these two years is due to veterans exiting farming occupations and residing in urban areas.

7.3 Structural Economic Changes and the Decline of the Veteran Premium

Between 1930 and 1940, WWI veterans received a substantial cash payment as late compensation for their service. A goal of military benefits is to aid reintegration into civilian labor markets. However, during that same period, the veteran premium declined between 1930 and 1940, from 9.6 percent to 6.0 percent. The veteran premium could decline for two reasons. First, the skills learned in the army could have degraded. Second, they could have become less valuable due to structural economic changes.

To test these two hypotheses, I match veterans with non-veterans based on observed skills in 1930. By measuring the outcomes of veterans and non-veterans in 1940, while holding 1930 skills constant, it is possible to identify why, by 1940, the veteran earnings score premium declined. Given that the two groups have comparable skills, they should both be equally affected by structural shocks that occur during that period. Therefore, if the decline in the veteran premium is entirely due to structural changes, then there should be no difference between veterans and non-veterans in 1940. However, if the outcomes between the two groups differ in 1940, then that would imply that there is a veteran-specific component. It would mean that the impact of military service on human capital is continuously evolving, even for men in their 40s. I find that once matched on skills in 1930, veterans and non-veteran earning score outcomes track perfectly. Therefore, most of the decline in the veteran premium is due to changes in the returns to skills.

To identify individuals of comparable skill levels, I use matched 1930-1940 census data. Veterans are matched to non-veterans with comparable skills by creating skill groups using information from the 1930 census. These groups are the interaction of 1930 occupation, industry, state of residence, and urban status. They are then included as fixed effects in a 1940 regression. With these fixed effects, the 1940 regression tests whether the earnings score between veterans and non-veterans evolved differently within these skill groups. Note that this regression does not include family fixed effects, and the data is no longer limited to the family sample.

Table 5 displays estimates for the log earnings score in 1940. Results with skill group fixed effects are displayed in column (1). Within these groups, by 1940, veterans earn 0.9 percent more than non-veterans. However, even within these narrow cells, veterans could have differed from non-veterans. Column (2) includes the standard set of controls used in this paper. The coefficient on veteran status drops to 0.3 percent, but it still significant. Thus, veterans who have comparable skills to non-veterans in 1930 experienced nearly identical trends in their earnings score.

Given that there is no difference within skill cells in 1940, it means that the decline in the veteran premium is due to structural changes. This result is not unexpected; during this time, the United States was experiencing the Great Depression. Real output contracted by nearly 30 percent, and unemployment reached a height of 25 percent in 1933. Due to its size, we expect substantial changes in the labor markets and large differences in the returns to skill over time.

As an aside, the extent that veterans and non-veterans track each other within the 1930 skill groups has implications for cash payments as a form of compensation. Between 1930 and 1940, veterans received substantial cash bonuses, equal to about half of the average income of men in the sample. Despite these payments, there are no differences in earnings score between veterans and comparable non-veterans. One interpretation of the lack of differences is that the cash payments had no effect. They were received late in life and likely not invested in improving human capital. In standard life cycle models, the incentive to invest in human capital weakens with age due to smaller increases in lifetime earnings. Another reasonable interpretation of the zero results is that without the additional cash, veterans would have been worse off. Unfortunately, given the available Census data, it is not possible to identify variation in the size of the cash bonus across veterans. Therefore, it is not possible to identify the impact of these late cash payments.

8 Conclusion

This paper identifies the impact of WWI military service on labor market outcomes using within-family variation. Consistent with veterans receiving initially limited government support, they had a lower-earning score in 1920 and suffered a persistent employment penalty. However, by 1930 and 1940, WWI military service increased veteran earnings by 9.6 percent and 6.0 percent, respectively. This premium was more than twice as high for veterans who grew up in the most disadvantaged circumstances compared to those who are most advantaged. Unfortunately, this came at a cost as the employment penalty is higher for the disadvantaged too.

Exploring the mechanisms behind the earnings pattern, I find that the negative 1920 result can be fully explained by the loss of civilian labor market experience during an advantageous labor market. Half of the large 1930 premium is due to veterans abandoning low paying farming occupations and moving to better (i.e., urban) areas. The decline in the (still positive) earnings premium between 1930 and 1940 is largely due to structural change that occurred during the Great Depression.

Overall, WWI veterans did well even before most of them received substantial government support. Does this mean that policies motivated by their experiences, such as the GI Bill, are unnecessary? I argue that this is not the case, as the estimates for the WWI veterans are likely due three to factors unique to the war and early 20th-century labor markets. First, the war was relatively short, and most soldiers saw limited combat exposure. The short combat service of limited intensity lowered earning losses that occurred due to less civilian labor market experience as well as disabilities. Second, the relationship between experience and earnings is much weaker and flattens early than in the latter half of the 20th Century. Finally, the option to abandon farming occupations and move to higher-paying urban areas was available to a larger portion of the population than in the present. As currently, most of the population resides in urban areas and does not participate in farming occupations.

9 Works Cited

Angrist, J.D., 1990. Lifetime earnings and the Vietnam era draft lottery: evidence from social security administrative records. *The American Economic Review*, pp.313-336.

Angrist, J.D. and Chen, S.H., 2011. Schooling and the Vietnam-era GI Bill: Evidence from the draft lottery. *American Economic Journal: Applied Economics*, *3*(2), pp.96-118.

Bailey, M., Cole, C., Henderson, M. and Massey, C., 2019. How Well Do Automated Methods Perform in Historical Samples? Evidence from New Ground Truth (No. w24019). National Bureau of Economic Research.

Card, D. and Cardoso, A.R., 2012. Can compulsory military service raise civilian wages? Evidence from the peacetime draft in Portugal. *American Economic Journal: Applied Economics*, *4*(4), pp.57-93.

Cáceres-Delpiano, J., 2019. The Impact of Mandatory Military Service. Evidence from Spain.

Chetty, R. and Hendren, N., 2018. The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects. *The Quarterly Journal of Economics*, *133*(3), pp.1107-1162.

Chrisinger, David. 2019. The Army's Message to Returns World War I Troops? Behave Yourselves. *The New York Times.* [Accessed 12 June 2020]. Available from: https://www.nytimes.com/2019/07/31/magazine/world-war-i-veterans-treatment.html

Collins, W.J. and Wanamaker, M.H., 2020. *African American Intergenerational Economic Mobility Since 1880* (No. w23395). National Bureau of Economic Research.

Department of Veterans Affairs, 2007. V.A./DoD clinical practice guideline for rehabilitation of lower limb amputation.

Faulkner, R.S., 2017. Pershing's Crusaders: The American Soldier in World War I. University Press of Kansas.

Feigenbaum, J.J., 2016. Automated census record linking: A machine learning approach.

Gabriel, P.E., 2016. The doughboy premium: an empirical assessment of the relative wages of American veterans of World War I. Applied Economics Letters, 23(2), pp.93-96.

General Accounting Office. (1992). *Vocational Rehabilitation: Better VA Management Needed to Help Disabled Veterans Find Jobs.* Available at: https://www.gao.gov/products/GAO/HRD-92-100 (Accessed: 7 June 2020).

Greenberg, D., 2006. *Calvin Coolidge: The American Presidents Series: The 30th President, 1923-1929.* Macmillan.

Greenberg, K., Gudgeon, M., Isen, A., Miller, C., Patterson, R. 2020. *Army Service in the All-Volunteer Era.* NBER Summer Institute, 20 July, Online.

Hausman, J.K., 2016. Fiscal policy and economic recovery: The case of the 1936 veterans' bonus. *American Economic Review*, *106*(4), pp.1100-1143.

Herbert, H. 1931. Veto Messages Regarding Emergency Adjusted Compensation Act. 26 February. Washington, DC [Accessed 12 June 2020]. Available from: https://millercenter.org/the-presidency/presidential-speeches/february-26-1931-veto-messages-regarding-emergency-adjusted

Humes, E., 2006. Over here: How the GI Bill transformed the American dream. Houghton Mifflin Harcourt.

Kahn, L.B., 2010. The long-term labor market consequences of graduating from college in a bad economy. *Labour economics*, *17*(2), pp.303-316.

Laschever, R., 2005. The doughboys network: social interactions and labor market outcomes of World War I veterans. Unpublished manuscript, Northwestern University.

Lebergott, S., 1964. *Manpower in economic growth: The American record since 1800* (p. 125). New York: McGraw-Hill.

Leland, A., 2010. American war and military operations casualties: Lists and statistics. DIANE Publishing.

Linker, B., 2011. *War's Waste: Rehabilitation in World War I America*. University of Chicago Press.

Mazumder, S., 2017. Becoming White: How Mass Warfare Turned Immigrants into Americans.

Mincer, J., 1974. Schooling, Experience, and Earnings. Human Behavior & Social Institutions No. 2.

Olson, K.W., 1974. *The GI Bill, the veterans, and the colleges*. University Press of Kentucky.

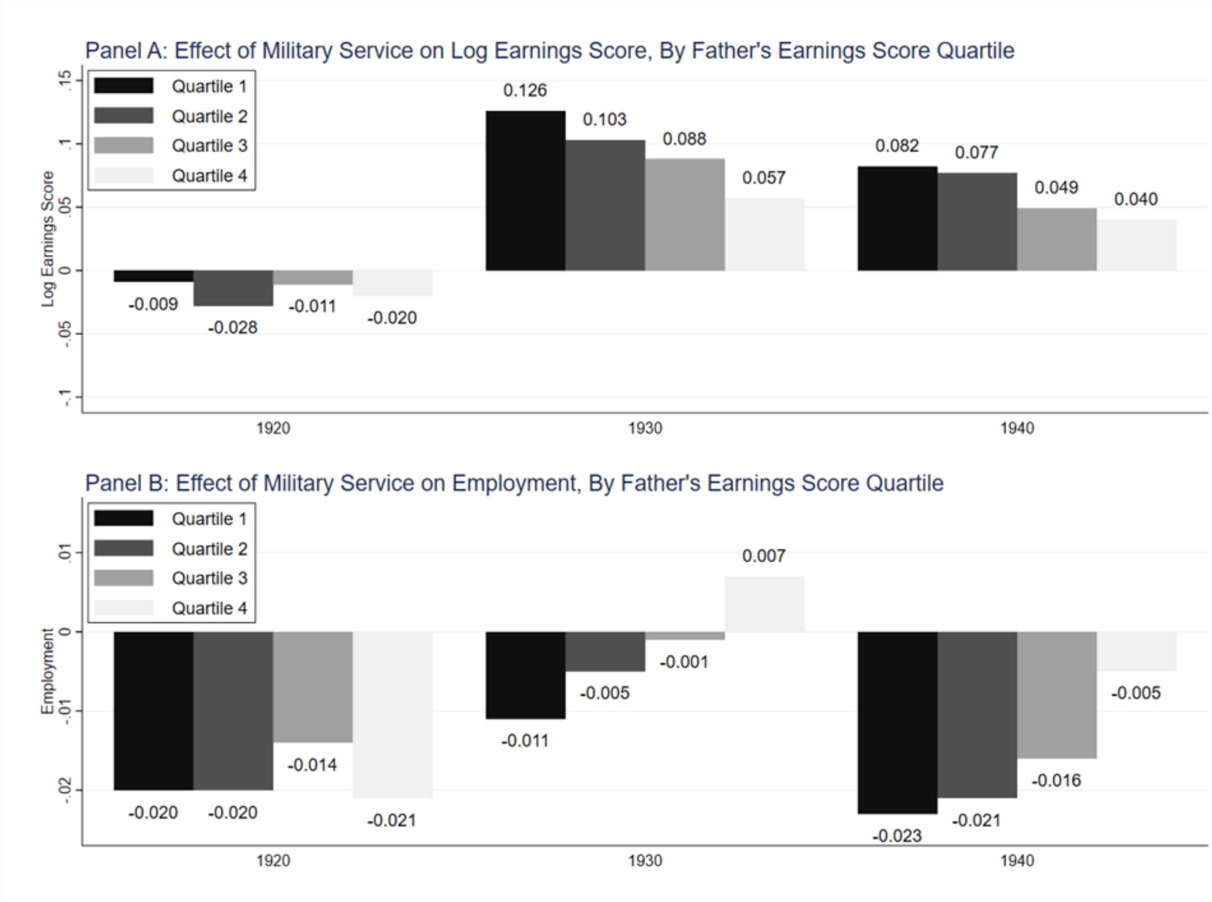
Ruggles, S., Flood, S., Goeken, R., Grover, J., Meyer, E., Pacas, J., and Sobek, M., IPUMS USA: Version 8.0 [dataset]. Minneapolis, MN: IPUMS, 2020. https://doi.org/10.18128/D010.V8.0

Schmick, E., 2018. Re-integrating Wounded Veterans into the Civilian Labor Market: Evidence from World War I. Unpublished Manuscript.

Tan, H. R. (2020). Did military service during World War I affect the economic status of American veterans?. *Explorations in Economic History*, *75*, 101301.

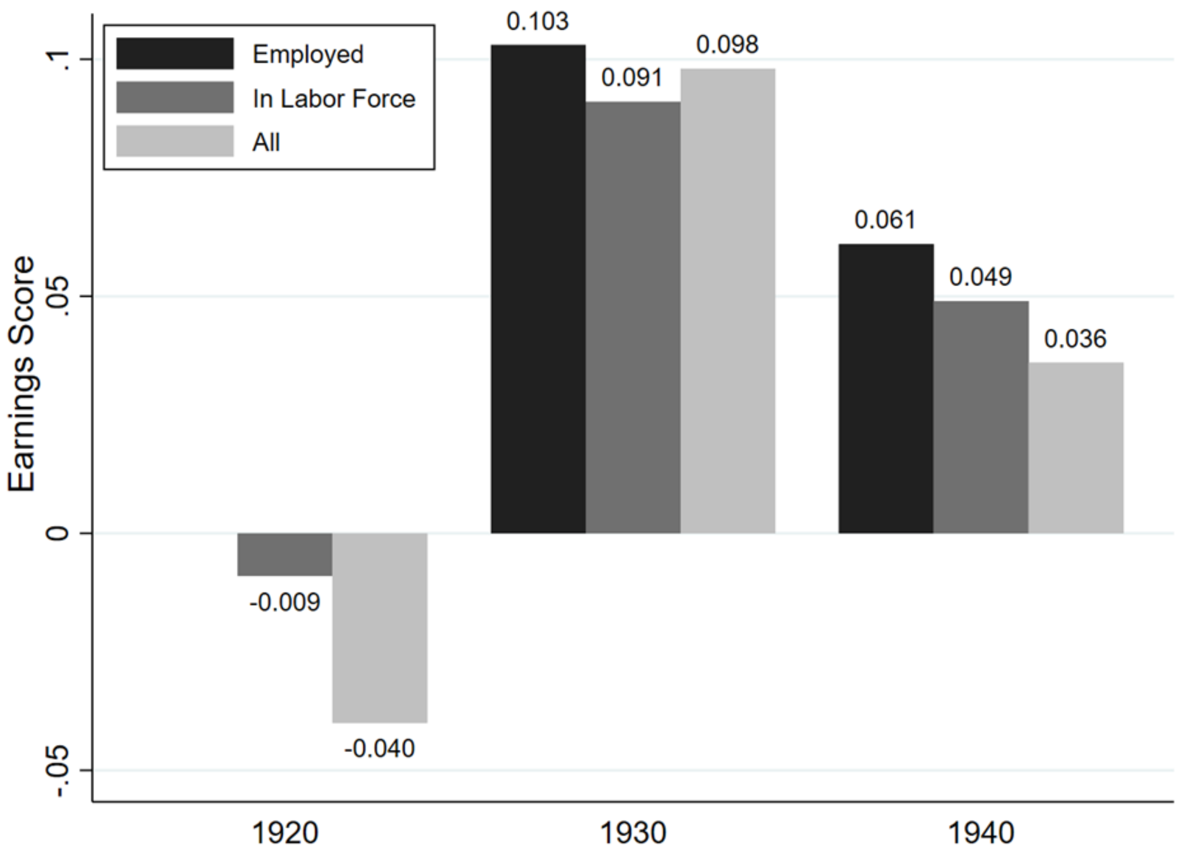
Figures

Figure 1: The Impact of Military Servers By Father's Earnings Score Quartile



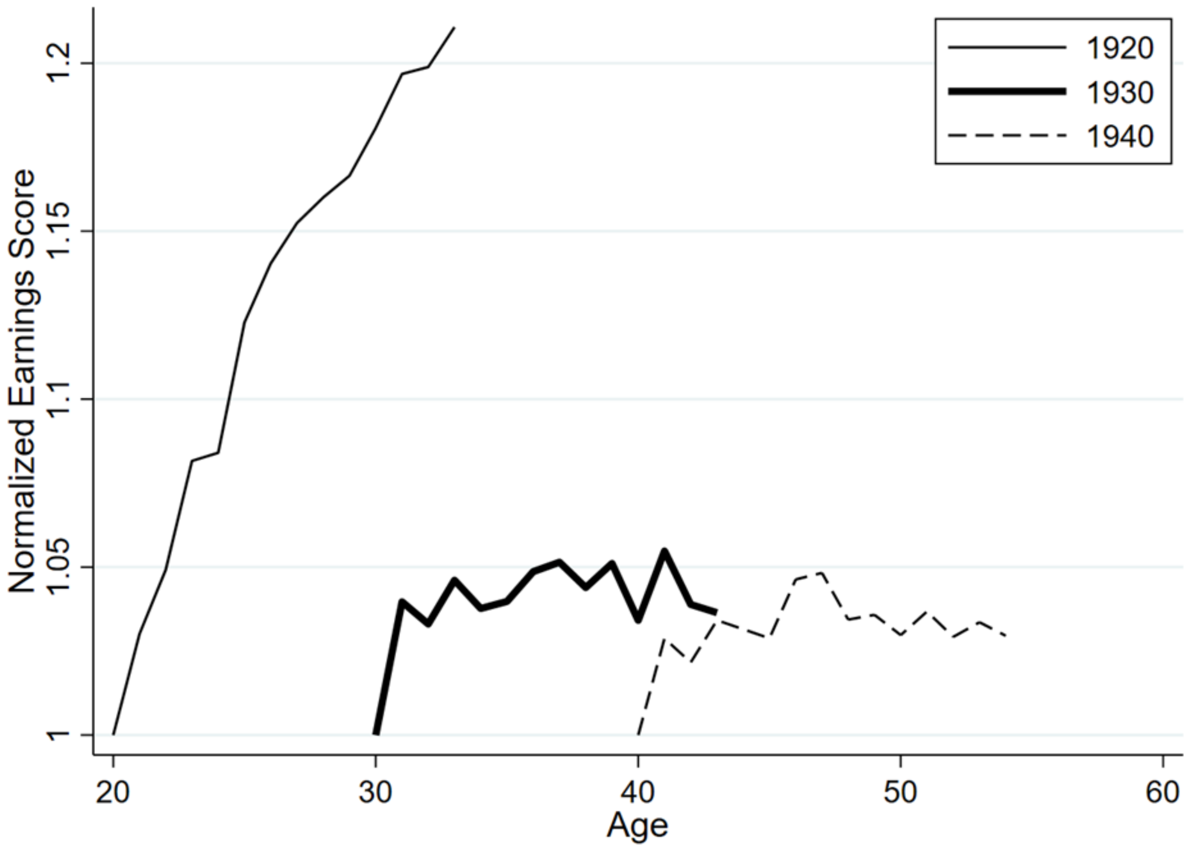
Note: This figure displays the heterogeneous estimates of WWI military service on log earnings score and employment. The results are obtained using the baseline family fixed effects model and are estimated using Equation 1. The sample is split into quartiles based on father's earnings score in 1900. For example, men whose father's 1900 earning score is equal to or below the 25th percentile are assigned to the first quartile.

Figure 2: Imputing Earnings Score for the Unemployed and Not in the Labor Force



Note: This figure displays the impact of WWI military service on multiple measures of the log earnings score. The results are obtained using the baseline family fixed effects model and are estimated using Equation 1. The dependent variable is the earnings score. To facilitate comparison to the baseline Table 2 estimates, with the logged earnings score, the marginal effect is calculated by dividing the coefficient on veteran status by the mean earnings score. The black bar replicates the baseline estimates with my preferred earnings score measure. The sample is limited to individuals who are currently employed. The dark grey bar expands the sample to those who are in the labor force, but are not currently employed by assigning them an imputed earnings score. Finally, the light grey bar covers the whole population, including those who are not in the labor force by imputing them a score of zero. Please refer to the text for more details.

Figure 3: The Age-Experience Profile In 1920, 1930 and 1940



Note: This figure reports the relationship between age and a normalized earnings score measure for men in the sample in 1920, 1930 and 1940. To ease comparison across different census years and to take into account differences in educational attainment across cohorts, the earnings score is normalized. The normalization is the following; the average score for the youngest men in each census year for each category of educational attainment is set to one. Then the score for older men, with comparable educational attainment, is calculated in relation to the youngest.

Tables

Table 1: Sample Summary Statistics by Year

	Census Data	Matched Census Data		Matched Siblings	
	All	All	Siblings	Veterans	Non-Veterans
Sample Size					
1920	10,317,664	2,042,728	556,633	169,005	387,628
1930	10,077,236	3,894,903	1,471,350	440,083	1,031,267
1940	9,313,196	2,189,881	566,912	171,293	395,619
In Labor Force					
1920	0.93	0.93	0.94	0.92	0.94
1930	0.97	0.98	0.98	0.98	0.98
1940	0.93	0.94	0.94	0.94	0.94
Employed					
1930	0.90	0.91	0.91	0.91	0.92
1940	0.80	0.83	0.83	0.83	0.84
Farmer					
1920	0.15	0.17	0.19	0.10	0.23
1930	0.16	0.18	0.20	0.13	0.24
1940	0.15	0.17	0.20	0.14	0.22
Age					
1920	26.10	26.01	26.00	25.48	26.22
1930	36.17	36.23	36.27	35.73	36.50
1940	46.03	46.50	46.51	45.97	46.75
Black					
1920	0.10	0.05	0.03	0.03	0.04
1930	0.11	0.08	0.07	0.05	0.08
1940	0.10	0.07	0.05	0.04	0.06
Literate					
1920	0.96	0.98	0.98	0.99	0.98
1930	0.97	0.97	0.97	0.99	0.97
1940	-	0.98	0.98	0.99	0.98
Years of School					
1940	8.39	8.70	8.59	9.25	8.30
Urban Residence					
1920	0.52	0.49	0.46	0.55	0.42
1930	0.57	0.55	0.53	0.62	0.49
1940	0.55	0.52	0.49	0.56	0.47
Veteran					
1920	-	0.29	0.30	1.00	0.00
1930	0.28	0.29	0.30	1.00	0.00
1940	-	0.30	0.30	1.00	0.00
Wage Income					
1940	1402.30	1437.93	1501.98	1628.83	1441.17
Earnings Score					
1920	769	799	760	798	747
1930	997	1,014	922	1,084	866
1940	1,149	1,230	1,149	1,297	1,093

Note: This table displays sample summary statistics for the 1920, 1930 and 1940 census years. It does this for the full count census data, the matched census data, and the sibling sample in the matched data. In the matched sibling sample summary statistics are also provided separately for veterans and non-veterans.

Table 2: Baseline Estimates

	In the Labor Force / Employed			Log Earnings Score		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: 1920						
Veteran	-0.022***	-0.027***	-0.020***	0.066***	0.003	-0.017***
	(0.001)	(0.001)	(0.001)	(0.003)	(0.002)	(0.003)
Observations	556,633	556,633	556,633	255,746	255,746	255,746
R-squared	0.002	0.045	0.598	0.004	0.409	0.788
Panel B: 1930						
Veteran	-0.004***	-0.005***	-0.003***	0.229***	0.138***	0.096***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)
Observations	1,471,350	1,471,350	1,471,350	589,196	589,196	589,196
R-squared	0.000	0.011	0.527	0.023	0.364	0.746
Panel C: 1940						
Veteran	-0.009***	-0.015***	-0.016***	0.177***	0.101***	0.060***
	(0.001)	(0.001)	(0.002)	(0.004)	(0.003)	(0.005)
Observations	566,912	566,912	566,912	164,706	164,706	164,706
R-squared	0.000	0.012	0.536	0.017	0.361	0.743
Controls		X	X		X	X
Family FE			X			X

Note: This table uses matched US Census data. The sample is limited to men born in the United States who were between the ages 0-13 in the 1900 Census. Individuals without at least one male sibling in the 1920, 1930 or 1940 matched Census data are dropped. Controls include indicator variables for literacy, race, age, month of birth, place of birth and father's occupation in 1900. Due to data limitation, information on employment is not available in 1920. Therefore, the 1920 results are for labor force participation. Family fixed effect estimates cluster standard errors at the family level, other specification report robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3: Twin Fixed Effects and Pre-Treatment Controls

	Log Earnings Score				
	(1)	(2)	(3)	(4)	(5)
Panel A: 1920					
Veteran	0.005	-	-	-0.022***	-0.025***
	(0.025)	-	-	(0.005)	(0.005)
Observations	3,804	-	-	107,945	107,945
R-squared	0.824	-	-	0.785	0.791
Panel B: 1930					
Veteran	0.067**	0.081***	0.081***	0.098***	0.090***
	(0.028)	(0.011)	(0.011)	(0.005)	(0.005)
Observations	8,621	58,357	58,357	221,792	221,792
R-squared	0.777	0.746	0.750	0.748	0.759
Panel C: 1940					
Veteran	0.093*	-	-	0.059***	0.040***
	(0.053)	-	-	(0.005)	(0.005)
Observations	2,234	-	-	162,346	162,346
R-squared	0.762	-	-	0.743	0.776
Twin Sample	X				
1910 Earnings Score			X		
Education Controls					X

Note: This table uses matched US Census data. The sample is limited to men born in the United States who were between the ages 0-13 in 1930. Individuals without at least one male sibling in the 1920, 1930 or 1940 matched Census data are dropped. The first column limits the sample to pairs of twins. In columns (2) and (3), the linked 1900-1930 is additionally matched to 1920. Observations in the linked 1900-1910-1930 sample have to be at least 16 years of age in 1910 and be employed. In the last two columns, all observations are matched to both 1930 (veteran status information) and 1940 (years of schooling information). In column (1), the twin fixed effect estimates, only literacy status is controlled for. In column (3) the controls consists of indicator variables for literacy, race, age, month of birth, place of birth and a quartic of the 1910 earnings score. In column (5), the quartic of the earnings score is removed. It is replaced with educational attainment variables which consists of five indicator variables (less that middle school, middle school, some high school, high school and some college or more) and a linear term for years of schooling. Family fixed effect estimates cluster standard errors at the family level and are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Mechanisms

	Industry % Wage Δ, WWI	Farmer	Urban Residence	Log Earnings Score	Log Earnings Score
	(1)	(2)	(3)	(4)	(5)
Panel A: 1920					
Veteran	-1.429***	-0.094***	0.027***	-0.013***	-0.017***
	(0.114)	(0.003)	(0.003)	(0.003)	(0.003)
Observations	258,208	269,785	269,785	255,746	255,746
R-squared	0.677	0.688	0.788	0.829	0.788
Returns to Dep. Var.	-	4.9%	35.9%	-	-
Panel B: 1930					
Veteran	-2.283***	-0.073***	0.064***	0.041***	0.096***
	(0.081)	(0.002)	(0.002)	(0.002)	(0.003)
Observations	608,070	646,858	646,858	589,196	589,196
R-squared	0.610	0.636	0.661	0.835	0.746
Returns to Dep. Var.	-	-54.5%	62.0%	-	-
Panel C: 1940					
Veteran	-1.622***	-0.045***	0.036***	0.029***	0.060***
	(0.160)	(0.004)	(0.004)	(0.004)	(0.005)
Observations	169,980	178,329	178,329	164,706	164,706
R-squared	0.590	0.647	0.637	0.840	0.743
Returns to Dep. Var.	-	-61.9%	51.8%	-	-
Farmer FE				X	
State-by-Urban FE				X	

Note: Columns 1 to 3 report estimates from the baseline family fixed effects model. The first dependent variable is the mean wage growth the individuals current industry of employment experienced during WWI. In Column (2) it is an indicator for working in a farming occupation and in Column (3) it is an indicator for residing in an urban area. The last column augments the baseline model with controls for state-by-urban fixed effects and a farmer indicator variable. The returns to working in a farming occupation, residing in an urban area and working in an industry that experience a one percent increase in wages are reported in the Returns to Dep. Var. row. The returns are calculated using a family fixed effects model. All results are estimates on matched US Census data. The sample is limited to men born in the United States who were between the ages 0-13 in the 1900 Census. Individuals without at least one male sibling in the 1920, 1930 or 1940 matched Census data are dropped. Controls include an indicator variables for literacy, race, age, month of birth, place of birth and father's occupation in 1900. Family fixed effect estimates cluster standard errors at the family level. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: The Impact of Strucutral Economic Changes

	Log Earnigns Score in 1940	
	(1)	(2)
Veteran	0.009***	0.003***
	(0.001)	(0.001)
Observations	1,049,414	1,049,414
R-squared	0.592	0.647
1930 Cell FE	X	X
Controls		X

Note: This table uses matched US Census data. The sample is limited to men born in the United States who were between the ages 0-13 in the 1900 Census and are linked to both the 1930 and the 1940 censuses. The 1930 Cell Fixed Effects are indicator variables based on the cross interaction of the following 1930 characteristics: occupation, industry, state of residence, urban residence, and educational attainment in 1940. Controls include indicator variables for literacy, race, age, the month of birth, place of birth, and father's occupation in 1900. Family fixed effect estimates cluster standard errors at the family level. *** p<0.01, ** p<0.05, * p<0.1.

Appendix Tables

Appendix Table 1: Alternative Earnings Score Measures

	1940 Occupation & Industry Score	1940 Occupation Score	1950 Occupation Score	Industry Score	Wage/Salary Income
	(1)	(2)	(3)	(4)	(5)
Panel A: 1920					
Veteran	-0.017***	0.001	0.007*	0.048***	-
	(0.003)	(0.004)	(0.004)	(0.004)	-
Observations	255,746	255,746	255,660	255,746	-
R-squared	0.788	0.786	0.757	0.778	-
Panel B: 1930					
Veteran	0.096***	0.087***	0.071***	0.112***	-
	(0.003)	(0.003)	(0.002)	(0.004)	-
Observations	589,196	589,196	588,259	589,196	-
R-squared	0.746	0.751	0.699	0.696	-
Panel C: 1940					
Veteran	0.060***	0.058***	0.046***	0.064***	0.041***
	(0.005)	(0.005)	(0.004)	(0.006)	(0.011)
Observations	164,706	164,706	164,367	164,706	79,117
R-squared	0.743	0.746	0.695	0.695	0.614

Note: This table uses matched US Census data. The sample is limited to men born in the United States who were between the ages 0-13 in the 1900 Census. Individuals without at least one male sibling in the 1920, 1930 or 1940 matched Census data are dropped. Controls include an indicator variables for literacy, race, age, month of birth and place of birth. The description of the dependent variables is available in the text. Family fixed effect estimates cluster standard errors at the family level and are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Appendix Table 2: The Impact of Lost Civilian Labor Market Experience

	Log Earnings Score		
	1920	1930	1940
	(1)	(2)	(3)
Potential Experience	0.0254***	-0.0053	-0.0101**
	(0.0021)	(0.0033)	(0.0050)
Potential Experience^2	-0.0004***	0.0001	0.0001
	(0.0001)	(0.0001)	(0.0001)
Avg. Log Score Loss	-0.017	0.001	0.002
Observations	107,945	221,792	162,346
R-squared	0.790	0.757	0.775

Note: This table uses matched US Census data. The sample is limited to men born in the United States who were between the ages 0-13 in 1930. Individuals without at least one male sibling in the 1920, 1930 or 1940 matched Census data are dropped. All observations are matched to both 1930 (veteran status information) and 1940 (years of schooling information). Potential experience is constructed as age – (6 + years of schooling). The experience variable is included in the regression as a quadratic. Additional controls include an indicator variables for literacy, race, age, month of birth, place of birth, educational attainment (less that middle school, middle school, some high school, high school and some college or more) and a linear term for years of schooling. Family fixed effect estimates cluster standard errors at the family level and are reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.