Supplementary Appendix

"Let our ballots secure what our bullets have won": Union Veterans and the Making of Radical Reconstruction

Full table results for all regression analyses reported in tables and figures in the Supplementary Appendix may be found in the Supplementary Tables document in the replication file.

A Enlistment Rates Difference-in-Differences: Design and Robustness

A.1 Design

In this paper, I apply a difference-in-difference design to examine the effect of enlistment rates in the Union Army on county-level voting for Republicans and Black Suffrage. Here, the treatment, D_i is the enlistment rate in the Union Army among military-aged males in county i. The the time variable T_t is an indicator for whether county i is observed before the Civil War (0), and thus before the county experienced the treatment, or after the start of the Civil War (1), when the county was affected by enlistment.

I include county-specific intercepts μ_i and state-year intercepts γ_{st}. This deviates from the classical model in that it assumes parallel trends **within states**, rather than nationwide.

This application of the DID differs in two key ways from the classical formulation.

1. Continuous Treatment: In the research design for this paper, treatment is the enlistment rate; it can take on values between 0 and 1. This contrasts with classical difference-in-difference approaches in which treatment is binary.

2. All units treated: In this design, all counties have some enlistment. Thus, there are no units that go "untreated" in the post-treatment period. Nevertheless, the level of treatment that counties receive after the war is different.

These two differences raise questions: Do the assumptions required for identification differ? And what is the effect that is identified?[1]

[1]I am not here attempting to justify an ecological interpretation of these results. Though

A.1.1 Continuous Treatment

To understand what the parallel trends assumption means and what is identified, we need to expand the potential outcomes framework laid out above.

Rather than each unit in each time period having two potential outcomes (associated with treated or untreated), we can imagine that each unit i in time period t has a response schedule of potential outcomes: $Y(d)_{it}, d \in D$, where $Y(d)$ indicates the value Y would take if the continuous treatment D took on the value d.

We can represent this response schedule as $f_{it}(D)$: a function that tells us what the potential outcome of Y would be for unit i at time t for all possible values of D. This $f_{it}(D)$ could be non-linear and heterogeneous across i. The new effect we want to estimate is $\tau = E[f'_{i1}(D)]$: the average derivative of the response functions across all units i. As Angrist and Krueger (1999) point out, it is natural to view this average derivative of the response function as related to the conditional expectation function of Y as a function of D: $E[Y_i|D_i = d]$. It then makes it natural that we might estimate τ using least squares.

Just like the binary DID, we only observe $f_{i1}(D_i = d)$: the value of the response schedule for unit i for the actual value of treatment d that the unit takes. And because we do not observe the entire response function for unit i, $f'_{i1}(D)$ and $E[f'_{i1}(D)]$ are unknown.

Just as before, we can use a difference in differences to identify $E[f'_{i1}(D)]$, with some assumptions. The assumptions are easiest to see if we consider the average difference-in-differences in Republican voting between counties with an enlistment rate of d and counties with an enlistment rate of $d - \Delta$, were Δ is some arbitrary difference.

$$\{E[Y_{i1}(d)|D_i = d] - E[Y_{i0}(0)|D_i = d]\} - \{E[Y_{i1}(d-\Delta)|D_i = d-\Delta] - E[Y_{i0}(0)|D_i = d-\Delta]\}$$

This can be decomposed in terms of the average reponse function, as follows:

$$= E[f_{i1}(d) - f_{i1}(d-\Delta)|D_i = d] +$$

$$\{E[f_{i1}(d-\Delta) - f_{i0}(0)|D_i = d] - E[f_{i1}(d-\Delta) - f_{i0}(0)|D_i = d-\Delta]\}$$

This rearrangement gives us $E[f'_{i1}(d)]$ evaluated at d plus a bias term. This bias is the difference between the counterfactual trend units with enlistment rates of d would have had with enlistment rates of $d-\Delta$ and the actual trend units with enlistment rates of $d-\Delta$ had.

Angrist and Krueger (1999) note that assumption we have to make, in a regression context, is that conditional on covariates X, this bias term goes to zero. For the continuous difference-in-difference used in this paper, the assumption is that conditional on unit i fixed effects and state-year, dummies X_{it}:

I should note: conceiving of this model as an ecological regression sidesteps the issues of continuous treatments and all units being exposed to treatment, while introducing other problems.

$$E[f_{i1}(d-\Delta) - f_{i0}(0)|X_{it}, D_i = d] = E[f_{i1}(d-\Delta) - f_{i0}(0)|X_{it}, D_i = d-\Delta]\}$$

If this parallel trend assumption is correct: the unit fixed effects net out the baseline differences in the response functions: $f_{i0}(0)|D_i = d$ and $f_{i0}(0)|D_i = d-\Delta$, $f_{i1}(0)|D_i = d$ and $f_{i1}(0)|D_i = d-\Delta$. And the state-year fixed effects net out any changes in $f_{i1}(D) - f_{i0}(0)$ that are shared by units with different enlistment rates.

Note: this assumption still does not assume linearity in the response function $f_{i1}(D) - f_{i0}(0)$. And, this assumption does not require that the response functions are homogenous across units. It only requires that, within states, the response function of the change $f_{i1}(D) - f_{i0}(0)$ is, on average, the same across different rates of enlistment $D_i = d$. Thus, just as in the case where all units have some treatment, the key difference is that we must extend the parallel trends assumption: the average causal response functions of units with different levels of are parallel with each other.

When might this be violated?

1. These assumptions would be violented if units with different levels of D have, on average, different shifts in the average response function $f_{i1}(0) - f_{i0}(0)$ in the absence of treatment: pre-treatment trends are non-parallel. This is easy to test, as we can examine the pre-treatment period to assess this difference in average response function for several periods of time in which $D = 0$ for all units. This is true in the binary diff-in-diff

2. Additionally, these assumptions would be wrong if there were a confounder that is correlated with enlistment rates d that also determined Republican voteshare in counties in the post-war period (but not before). This is a kind of time-varying confounding, and it would not be detected by assessing parallel trends pre-treatment. This, too, is a threat in a binary diff-in-diff.

3. Even if units have parallel trends over time in the absence of treatment, and there is no time-varying confounder, these assumption would be violated if units with different levels of treatment d have different causal response functions for $f_{i1}(D) - f_{i0}(0)|D_i = d, d \neq 0$ *on average*. This implies that the EFFECTS of treatment D — relative to unit baseline and shared trends over time —are heterogeneous. And that this heterogeneity is not independent of the levels of treatment. In order for this to induce bias, there must be both heterogeneous effects and confounding (the heterogenous effects are related to treatment assignment).

How does this parallel trend assumption differ from the binary case? The only key difference is that it assumes that, within states, the average causal response function of the change in Y from $t = 0$ to $t = 1$ is either: (i) identical for all counties i, in which case selection based on potential outcomes into d is not a problem; or (ii) the average causal response functions are heterogenous, but selection into d is independent of these heterogeneous effects.

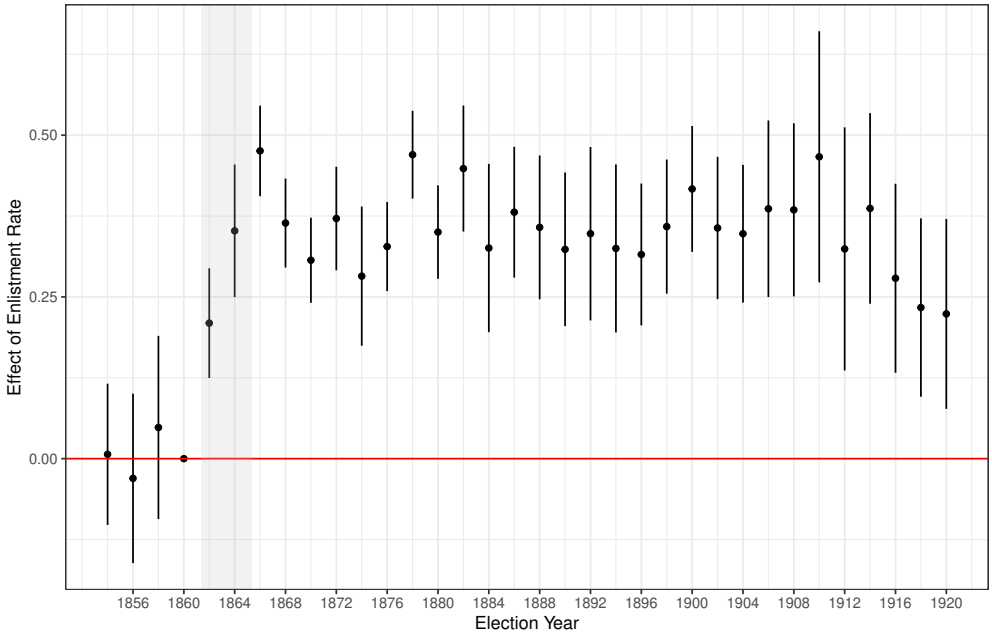
A.2 Checking Parallel Trends

I first assess pre-war parallel trends by estimating the following model, using the same data as Table 1.

$$\text{Republican Voteshare}_{ie} = \alpha_i + \alpha_e + \sum_{y=1854;\neq 1860}^{1920} \beta_y \text{Enlistment Rate}_i * \text{Year}_y + \epsilon_i + \epsilon_y \quad (4)$$

This estimates a separate slope on enlistment rate, for the difference between county Republican voteshare in each federal election between 1854 and 1920 and the 1860 federal elections. Figure A1, shows the results of these estimates. Prior to the Civil War, the over-time change in the average causal response functions of counties with different levels of enlistment is not different from 0.

Figure A1: Effect of veterans on Republican Vote-share in federal elections for each year between 1854 and 1920



This figure plots the year-specific effect of enlistment rates on Republican vote-share for federal elections with 1860 as the reference year. The model includes county and state-election fixed effects. Data from Congressional and Presidential elections across 384 counties between 1854 and 1880. Counties with election cycles in which the GOP does not contest an election cycle are dropped. Standard errors clustered by county and election year. Bars show 95 percent confidence intervals.

One might be concerned that this lacks credibility: do we really believe that counties in, say the top and bottom quartiles, of enlistment share parallel trends. In the body of the paper, I check this in Figure 1. Even when we split counties into the quartiles of enlistment in their state (reflecting the same with-in state parallel trends I specify in the model), the trends are visibly parallel across quartiles. Statistical tests of these differences confirm that

the change in Republican voteshare, relative to 1860, in counties with different quartiles of enlistment are not different from 0.

A.3 Addressing Time-Varying Confounding

Another concern might be that, even if there are parallel trends for counties with different levels of enlistment before the war, there may be some other attribute of counties with high enlistment that affected Republican voteshare once the war started. To address this possibility, I first explore the predictors of enlistment in Figure A2.

As discussed in the paper, enlistment rates are significantly correlated with several economic, demographic, and political attributes of counties in 1860. These could shape how counties respond to the war. To address this possibility, I estimate a new model, including an interaction between each of these pre-war attributes and the post-war period (Table A1) or election year (Figure A3).

$$\text{Republican voteshare}_{ie} = \alpha_i + \alpha_e + \beta \text{Enlistment Rate}_i \cdot \text{Civil War}_e + \mathbf{X_i} \cdot \text{Civil War}_e + \epsilon_i + \epsilon_y \tag{5}$$

The results are substantively unchanged, even though we allow for possible time-varying confounding.

A.4 Addressing Heterogeneous Effects

Another possible source of trouble would be if the effects of enlistment varied across counties with different levels of enlistment. One way to address this problem would be to directly estimate heterogeneous effects in the enlistment DD model for counties that were otherwise similar to each other on baseline characteristics. One way to do this is to use *generalized random forests* (Athey, Tibshirani and Wager 2018). Athey, Tibshirani and Wager (2018) 's causal forests estimate the heterogeneous effects treatment W for case i. If X_i is a set of covariates and W_i is a continuous or binary treatment, generalized random forests estimate $\theta(x)$, or the effect of W on Y for a set of values of X. This is done by generating weights that define the closeness of other observations to i within the space defined by X. These weights for the closeness of case j to case i are $\alpha_j(x_i)$. The sum of $\alpha_j(x_i)$ across all j is 1. Whereas many methods generate these weights using kernels which is subject to the curse of dimensionality, generalized random forests use random forests (repeated iterations of regression trees) to assign weights. In this case, θ is the locally weighted linear partial effect of W on Y.

$$\hat{\theta}(x_i) = \frac{\sum_{j=1}^{n} \alpha_j(x_i)(W_j - (\sum \alpha_j(x_i)W_j))(Y_j - (\sum \alpha_j(x_i)Y_j))}{\sum_{j=1}^{n} \alpha_j(x_i)(W_j - (\sum \alpha_j(x_i)W_j))^2}$$

These individual heterogeneous effects are then combined into average partial effect, with asymptotically unbiased standard errors. Using this method, I estimate the following difference-in-difference model for each election year y between 1854 and 1880. c indicates that I have centered the variable around the state mean, to impose the state-year fixed effects.

Figure A2: Predictors of County Enlistment Rates (Standardized Coefficients)

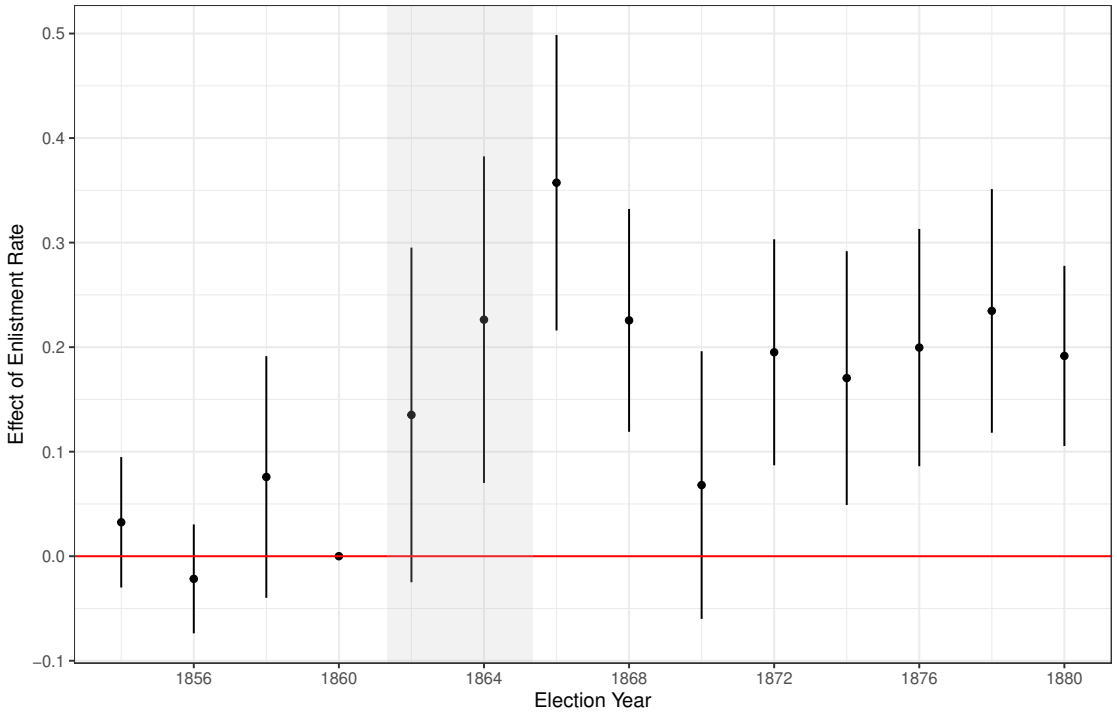
This figure plots the standardized coefficients from bivariates regressions of county enlistment rates on 32 different covariates, with state fixed effects. Data from 384 counties, dropping counties with values in top and bottom 0.5 percentiles of the respective covariate. HC1 Standard Errors. Bars show 95 percent confidence intervals.

Table A1: Difference-in-Differences Estimate (With Time-Interacted Controls) of Effect of County Enlistment Rate on Republican Voteshare

	Dependent variable:		
	Republican Voteshare		
	(1)	(2)	(3)
Enlist % · Postbellum	0.175***	0.192***	0.177***
	(0.031)	(0.043)	(0.036)
GOP no contest	included	dropped	dummy
County FE	X	X	X
State-Election FE	X	X	X
Covars · Postbellum	X	X	X
N Counties	330	257	330
Observations	6,930	5,397	6,930
R^2	0.785	0.770	0.840
Note:	*p<0.05; **p<0.01; ***p<0.001		

Data from Congressional and Presidential elections across 330 counties between 1854 and 1880. Standard errors clustered by county and election year. Counties with election cycles in which the GOP does not contest an election cycle either treated as a 0, the election is marked with a dummy, or the all observations for that county are dropped. The post-war indicator is interacted with 32 demographic, economic, and political covariates. Counties with missing data or extreme outliers in the covariates are dropped.

Figure A3: Effect of veterans on Republican Vote-share in federal elections for each year between 1854 and 1920, with Covariate-Time interactions

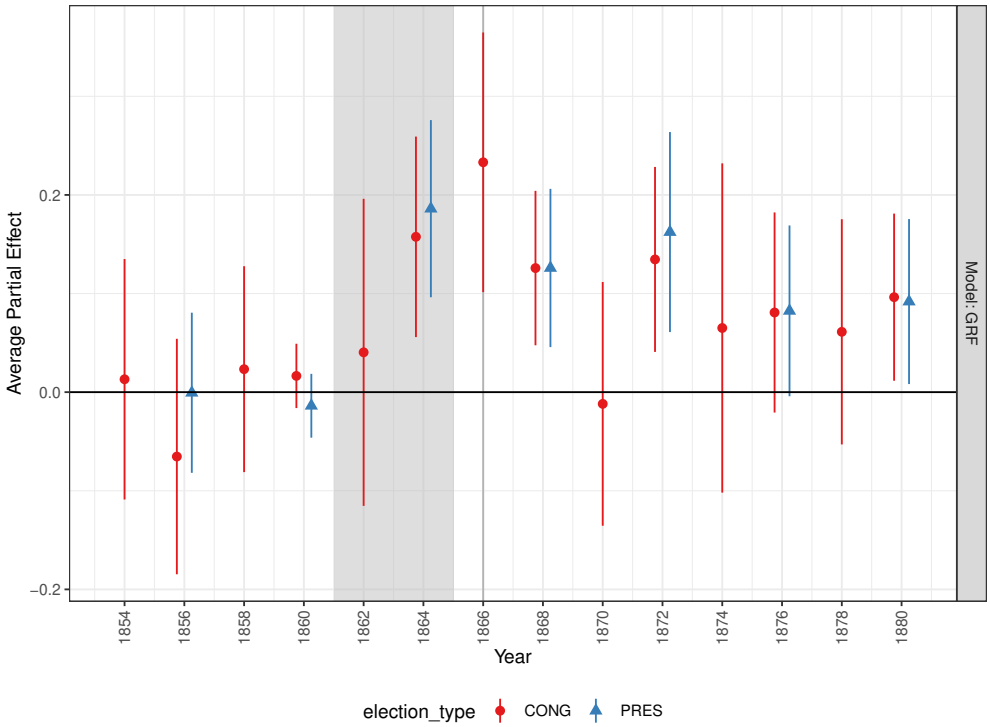


This figure plots the year-specific effect of enlistment rates on Republican vote-share for federal elections with 1860 as the reference year. The model includes county and state-election fixed effects. The year indicators, with 1860 as the reference, are interacted with 32 demographic, economic, and political covariates. Standard errors are clustered by county and year. Bars show 95 percent confidence intervals. Data from Congressional and Presidential elections across 257 counties between 1854 and 1880. Standard errors clustered by county and election year. Counties with election cycles in which the GOP does not contest an election cycle are dropped. Counties with missing data or extreme outliers in the covariates are dropped.

$$\text{GOP Voteshare}_{yc} - \text{GOP Voteshare}_{1860c} = \theta(x)\text{Enlistment Rate}_{ic} + \epsilon_i \quad (6)$$

The variables used in weighting case similarity for computing heterogeneous effects are: (a) The performance of the Republican party in every federal election (Congressional and Presidential) from 1854 to 1860. For any county in which a given party did not compete in an election, I set their vote-share as 0 and add a dummy indicating the party did not compete in this county-election or that the data is missing. (b) All non-political covariates shown above in Figure A2. (c) Indicators for the state in which the county is located. The key assumption here is that the effects of enlistment are effectively constant or independent of enlistment for cases that are similar in covariates X.

Figure A4: Average Partial Effect of Enlistment on GOP Voteshare: Diff-in-Diff



This reports the generalized random forests estimates of the average partial effect of enlistment for a difference-in-difference between Republican voteshare in each federal election between 1854 and 1880 and 1860, conditioning on political, demographic, and economic covariates. Standard errors are robust. Bars show 95 percent confidence intervals.

Figure A4 shows that the results obtained using generalized random forests are substantively the same, even though we explicitly and flexibly model heterogeneous effects of enlistment.

A.5 Robustness

The effects are unchanged when using surviving veterans as a fraction of 1860 military aged males, instead of all enlistments: Table A2.

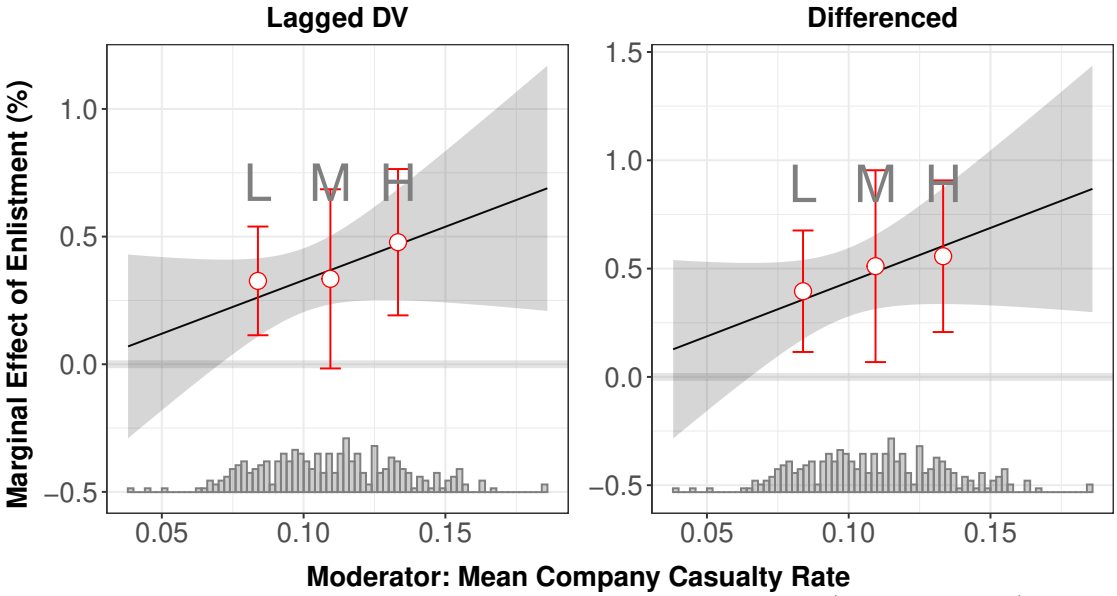
Table A2: Difference-in-Differences Estimate of Effect of County Enlistment Rate (Survivors Only) on Republican Voteshare

	Dependent variable:		
	Republican Voteshare		
	(1)	(2)	(3)
Survived % · Postbellum	0.414^{**}	0.323^{***}	0.322^{***}
	(0.104)	(0.076)	(0.062)
No. Rep. Candidate			-0.585^{***}
			(0.029)
GOP no contest	included	dropped	dummy
County FE	X	X	X
State-Election FE	X	X	X
Observations	6,153	4,473	6,153

Note: *p<0.05; **p<0.01; ***p<0.001

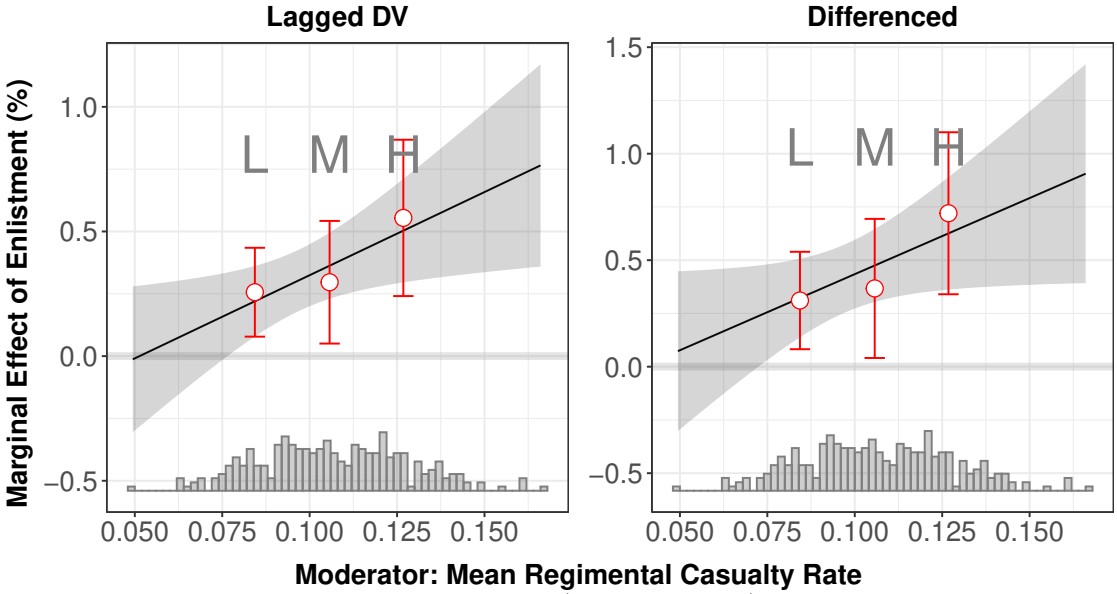
Data from Congressional and Presidential elections across 293 counties between 1854 and 1880. Standard errors clustered by county and election year. Counties with election cycles in which the GOP does not contest an election cycle either treated as a 0, the election is marked with a dummy, or the all observations for that county are dropped.

Figure A5: Marginal Effect of Enlistment Rates on Republican Voteshare Conditional on Company Casualty Rate)



This figure plots the marginal effect of County Enlistment Rate (survivors only) on Republican Voteshare in 1864-1868 elections vs 1854–1860, across Mean Company Casualty Rate, conditional on state fixed effects. Sample includes 293 counties, and excludes Indiana for which we cannot calculate mean casualty rates. Standard errors are robust (HC1).

Figure A6: Marginal Effect of Enlistment Rates on Republican Voteshare Conditional on Regimental Casualty Rate



Marginal effect of County Enlistment Rate (survivors only) on Republican Voteshare in 1864-1868 elections vs 1854–1860, across Mean Regimental Casualty Rate, conditional on state fixed effects. Sample includes 293 counties, and excludes Indiana for which we cannot calculate mean casualty rates. Standard errors are robust (HC1).

B Company Casualties in Infantry Regiments: A Natural Experiment

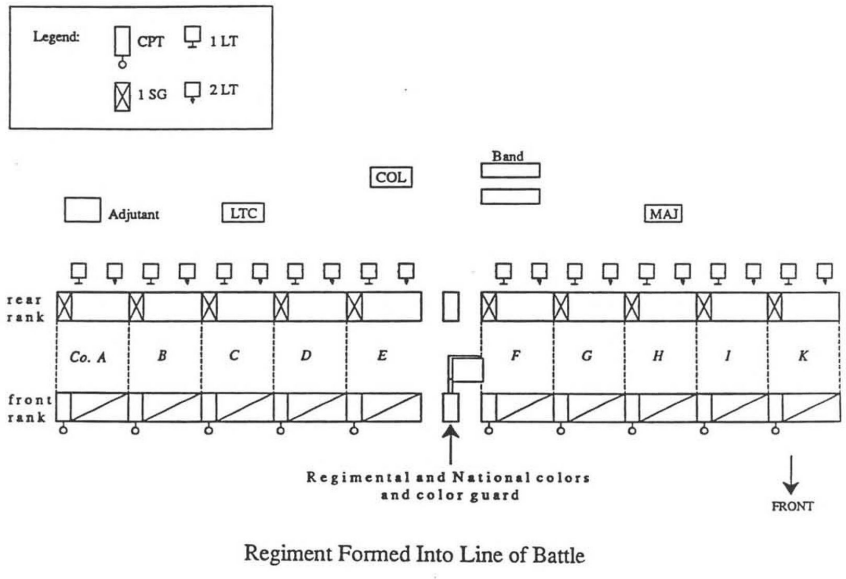
B.1 Company Exposure to Combat Fatalities

The vast majority of men serving in the Union Army served in infantry regiments. The basic unit in which infantrymen were mobilized, maneuvered, and went into combat was the regiment. While the Regular Army had independent companies of sharpshooters, volunteer regiments, which made up most of the army, operated as units. Regiments were further subdivided into ten companies of equal size, each commanded by a captain. Men in the same company were mustered in together and then trained, lived, worked, and fought directly alongside each other.

Men may have chosen when to enter military service, the branch of service, the term of their service, and even the specific regiment they joined. And Army commanders might decide based on regimental performance or perceived reliability to give certain regiments garrison duty or combat roles, or in battle, give more or less important or difficult objectives.

But, because regiments were the main unit around which military tactics of the time were designed, men in the same regiment went the same places, at the same times, and were located in the same place on the battlefield. In battle, infantry regiments typically formed up in a line of men, two deep, horizontally arranged by company, approximately 140 yards across (see Figure B1) (Hess 2015). In the chaos and smoke of battle, which companies in that line received more casualties was effectively arbitrary.

Figure B1: Regimental Battle Formation in the Civil War



Regiment Formed Into Line of Battle

This figure shows the standard battle formation of both Union and Confederate infantry regiments during the American Civil War. Ballard, Ted, and Billy Arthur, *Chancellorsville Staff Ride: Briefing Book*. Washington, DC: United States Army Center of Military History, 2002. Public Domain Image from Wikipedia

Because this logic applies only to infantry regiments, I exclude artillery and cavalry regiments from my analysis. This is because, artillery regiments deployed batteries to different places on the battlefield, and cavalry regiments, particularly early in the war, operated with detached companies.

Moreover, because the staff/headquarters company and regimental band were deployed behind the ten main companies, I exclude these from analysis. Similarly, regimental records often include "unassigned" troops, without a known company. I exclude these from analysis since we do not know which company they served with.

B.2 A Justification of As-If Random

Following Dunning (2012), I consider the information, incentives, and capabilities of soldiers, officers, and the enemy that might lead soldiers to experience more or fewer company casualties as a function of potential outcomes of partisanship. The evidence below comes almost exclusively from (Hess 2015), which is the only contemporary academic historical work devoted to Civil War infantry tactics.

B.2.1 Soldiers

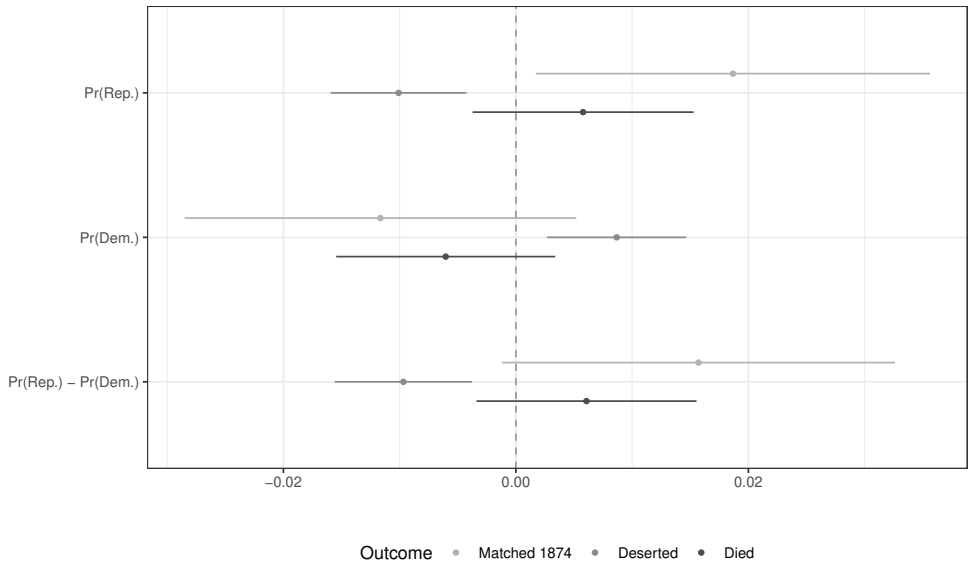
Information While soldiers could choose which company they joined when regiments were formed (Costa and Kahn 2008, 52–57), they could not have known which companies were more likely to receive casualties. (i) All companies stood together in a line across. (ii) While drill manuals specified which position companies should take in that line, soldiers entered the war unfamiliar with military drill instruction, and often the guidelines were ignored. This was true, because Army manuals called for companies to be arranged by the seniority of officers. This was irrelevant in almost all volunteer regiments, since all officers joined at the same time.

On the other hand, once in a regiment, soldiers could have known that tasks such as skirmishing (fighting in loose formation in advance of the main line) or serving in the color guard (protecting the flag bearers of the regiment) or color company (immediately to the right of the color guard) might be more dangerous. And soldiers may have known that their regiment was soon to enter battle, as well as whether some battles might be more dangerous.

Incentives In general, soldiers may have incentives to desert or use caution to avoid death. Conversely, soldiers may have wanted to seek glory or fight for the cause, leading them to take greater risks. And it is plausible that these incentives varied by partisanship. Republicans may have been more supportive of the cause and thus more motivated to take greater personal risks and fight harder, while Democrats may have been more likely to want to simply make it through the war alive. This is borne out in a raw comparison of predicted Democrats versus predicted Republicans: predicted Republicans were more likely to die, while predicted Democrats were more likely to desert (Figure B2.

Capacities Despite some information, soldiers had limited capacity to select themselves into treatment.

Figure B2: Differential Attrition by Predict Partisanship (Bivariate)



This figure shows the change in attrition, death, or desertion (related to attrition) associated with a 2 SD change in the predicted probability of Republican or Democratic partisanship in a bivariate regression model. Sample includes men serving in Indiana Regiments who were matched to the 1860 Census. Matches only include 'best' post-war links for each individual. Individuals are weighted by 1 over the number of matches. Latent partisanship is predicted probability of being Republican or Democrat based on the person's name, birth year, and birth place listed in the 1860 census. Standard errors are clustered by individual.

Before combat (i) Soldiers could work to be chosen for the color guard, which protected the regimental colors. Because the colors were essential for guiding the line of soldiers into battle and were source of morale, the enemy frequently targeted the colors. However, the members of the color guard were detached from their companies and were composed of corporals and sergeants. Excluding men of these ranks from analysis would address this form of selection.

(ii) Soldiers could not *choose* to join a company in advance based on its proximity to the colors. These companies might receive more casualties (though see below). These positions were often dictated by the company's ability to drill, which Hess notes was often a function of how skilled and motivated the captain was to drill his men. The companies best at drill would up on the flanks or in the center, near the color guard. This was to ensure that the regiment kept good formation in battle. While soldiers might attempt to shirk at drill to avoid being the color company, it seems unlikely that this would affect mortality. First, good performance at drill was also understood to be essential to for the regiment to survive under fire; even for reasons of survival, soldiers had reason to be competent at drill. Second, it is unlikely that individuals could conspire to get an entire company to be worse at drill. Third, even if they could get better at drill, it is implausible that this could be done with such precision that it would lead them to be placed on the flanks (potentially less exposure) vs. the center (more exposure). Moreover, the company directly to the left of the color guard was also likely in more danger. Yet this position was not dictated by drill performance.

(iii) Finally, soldiers could choose to desert rather than fight. This is a more serious problem of selection. If we simply were to use the actual "Exit" date from a company, soldiers who either deserted or died would determine their extent of exposure to combat casualties. And this could induce selection bias. But, I address this problem by constructing an intent-to-treat exposure to combat casualties.

During Combat Soldiers had limited capacities to alter their exposure to death during combat. Linear combat (fighting in lines) provided few opportunities for individuals or companies to take independent action. As discussed below, the whole point of having men fight in lines was to increase the effectiveness of massed rifle fire by largely unskilled marksmen under conditions of low visibility and limited means of command and control. Armies positioned officers and non-commissioned officers to keep men from lagging. And they practiced drill in order to keep their lines in good order. While men could elect not to fire, or fail in their aim due to fear, there were few opportunities to run or hide. In Hess's review of combat reports, he records very few mentions of companies operating independently at all. As discussed below, these instances were the result of battlefield idiosyncrasies.

Similarly, there were few opportunities for soldiers and companies to fight *harder*. While there are many accounts of units fighting to their last, which could induce selection bias, these accounts detail regiments, because companies simply did not operate independently. At the individual level, it could be that Republican soldiers took more personal risks. This appears to be the case, but even if this was true: (1) we don't include individuals in their own treatment variable and (2) if so, this could lead to attrition that reduces the number of Republicans observed after the war, because the most Republican men in a company might end up dying, increasing the "treatment'' for less Republican men who, because they survived, were located after the war. This would bias my estimates downward.

B.2.2 Commanding Officers

Information Officers likely knew, to some extent, the partisanship of their companies. They might infer this from where units were raised, the captains of those units, or the ethnic or occupational composition of the regiment. Or, more simply, given the vigorous political discourse within the Union Army, through conversation and debate (McPherson 1997). Based on this, officers may have been more suspicious of units that were less reliable.

Incentives There were three major considerations in how officers deployed companies in combat. **Keeping Unit Formation:** Army manuals suggested putting veteran companies on the ends and near the colors, but in an war where most regiments were of volunteers raised at the same time, this advice was rarely followed. Instead, regimental commanders debated whether to put the most well-drilled marchers on the ends of the line and in the center to help keep the regiment's line from falling apart.

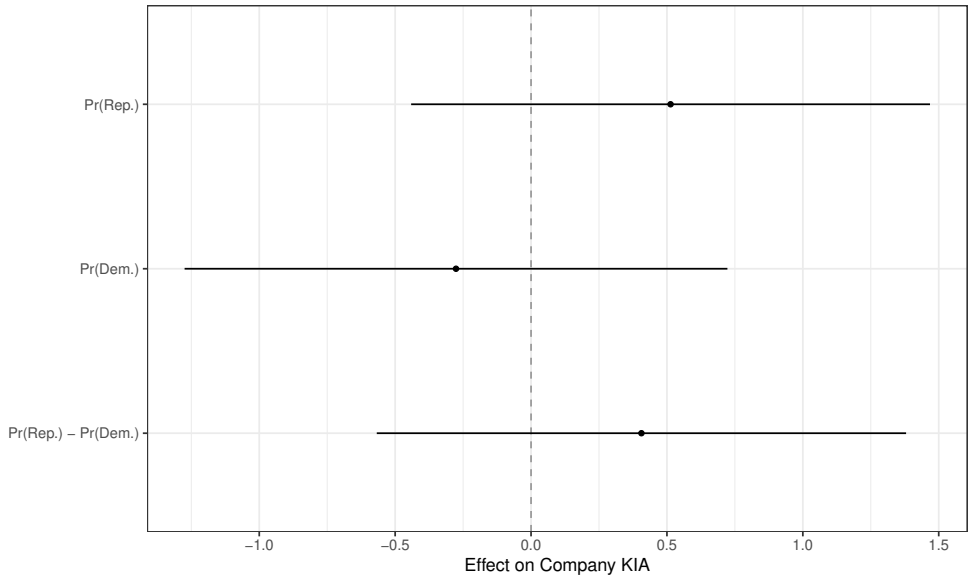
Another consideration was **terrain**. As a unit marched into combat, companies could become detached to the side or the rear of the regiment if they came upon obstacles. The affected companies moved to the rear of the rest of the regiment and retook their place in line when the obstacle was passed.

Finally, regimental tactics gave companies distinct orders based on **battlefield exigencies**. While most of the time, companies moved together in a line, deviations from this reflected emergencies. Sometimes companies were deployed as skirmishers, in loose formation, in front of a regiment. While this became more common over time, Hess notes that it was more common to use entire regiments as skirmishers for a brigade. Thus, it is unclear whether there was selection into skirmishing duty by partisanship. And it is also unclear whether skirmishers were in more or less danger, given that they were more spread out and encouraged to make use of cover. In other times, the left or right wings of regiments were used separately. This was done if the regiment were accidentally split in battle or when facing a flanking maneuver from the enemy. Again, it is hard to imagine how partisanship of companies would be used to inform how these decisions were made in the heat of battle.

If there were some selection, we'd expect companies with different baseline partisanship to have different combat casualties. In Figure B3, I report the within-regiment relationship between company casualty rates and average partisanship of men in those companies. There are no significant differences in casualty rates by company average partisanship, and the estimates are substantively small: within regiments, a 2 SD increase in mean Republican partisanship is associated with a 0.25 or smaller SD difference in casualties.

Capacity Other than choosing which companies would occupy the flanks and the center, officers could not position companies in line based on their partisanship. Fighting was done as a regiment. And when companies were detached, the reasons had to do with unexpected circumstances generated by chaos on the battlefield. One reason regiments fought together in cohesive lines was that it was difficult on a smoky battlefield, with rifle and artillery fire, to give specific directives to different companies. This is why regiments practiced drill together, in order to accomplish carefully choreographed maneuvers to be used in combat in order to fight more effectively and save their own lives in an attack.

Figure B3: Within-Regiment Relationship between mean Company Partisanship and mean Company KIAs



This figure the relationship between a 2 SD change in mean company predicted partisanship and the number of combat casualties in a company, conditioning on regimental fixed effects and weighting companies by the number of men use to calculate mean company partisanship. Results from least squares regression of mean company casualties, by company, on mean company partisanship, conditional regiment-enlistment year fixed effects. Sample includes serving in Indiana Regiments who were matched to the 1860 Census:10329 unique soldiers, in 896 companies, across 288 regiments. Companies are weighted by the number of soldiers. Standard errors HC robust.

When Hess discusses case studies of tactical errors that led to greater casualties, these occurred due to errors made by brigade or division commanders sending regiments into the wrong place, or when whole regiments found themselves in the wrong place on the battlefield.

B.2.3 Enemy

Finally, decisions made by the enemy about how to fire on regiments might create non-random company casualties.

Information Confederate units would be able to see color guards, since these were intended to be visible on the battlefield. It is also possible that they could distinguish officers.

Incentives Battle reports indicate that both armies sought to kill officers and color guards to both demoralize units and eliminate leadership to reduce the fighting effectiveness of the enemy.

Capacity While there was capacity, despite smoke, for Confederate units to see and target color bearers, and perhaps officers, this is mitigated by a few other considerations. First, officers were dispersed throughout the regiment (Captains and lieutenants were attached to each company). Second, while color bearers were targeted more, it is also the case that regiments were trained to fire against the enemy lines, not specifically at individuals. The primary reason that both armies kept using these linear tactics, despite the advent of the rifled musket, was that in order to maximize firepower, they wanted to mass rifle fire against enemy units. This was to increase the efficacy of rifle fire. Men in both armies lacked sufficient training to shoot accurately, the armies could not afford to expend munitions training them to shoot more accurately, and without smokeless powder, battlefields had limited visibility. This limited the effectiveness of individual firearms and prioritized large volumes of fire from large groups of men simultaneously.

B.2.4 Color Guards

Based on the preceding the discussion, the biggest potential source of selection effects in company-level casualties is related to color guards and color companies. Color guards, selected from corporals and sergeants of other companies, were at the center of the line. These men attracted more fire from the enemy. To their right, was the color company, often chosen for its ability to perform marching and drill well. Color guards may have been more likely to be Republican, given that it was a dangerous position. Color companies may have been more likely to be Republican, since good marching order might have been related to enthusiasm for the war.

It is impossible to reconstruct the membership of the color guard or color companies. The composition of the color guard changed and is not usually recorded on company rosters in the data available. Companies chosen to be color companies changed in regiments throughout the war: regiments often held drill competitions between companies to choose which were best.

While we cannot easily test whether color companies or selection into the color guard led to bias, a couple points suggest this is less of a problem.

(1) If members of the color guard were more likely to be Republican, then the remaining men in their companies would be exposed to more company deaths, if color guard members were more likely to die. But this would produce selection where the remainder of the company was, on average, less likely to be Republican, more likely to have more combat death exposure, and more likely to be observed after the war. This differential attrition would bias the effects I estimate downward.

(2) Companies with better marching skills might be chosen to be the color company. But better marching companies were also chosen to be on the left and right flank, far away from the color guard. And the company to the left of the color guard was not chosen based on its skills. Thus, even if companies that were more Republican on average were better motivated at drill, they could not easily determine whether they ended up in the center or on the flanks. And companies with poorer marching skills might still end up near the color guard. Moreover, the survival of a regiment in the face of the enemy depended on effective drill. Being able to complete complex choreographed maneuvers under fire often made the difference between holding off an enemy attack or being flanked and suffering great losses. Thus, efficacy in drill may have been motivated by enthusiasm for the war or enthusiasm to make it home alive.

For these reasons, I consider it implausible that soldiers could have selected into different levels of company casualties based on partisanship.

B.3 Data

In the following sections, I describe the data used in the analysis of this natural experiment and how it was collected.

B.3.1 Post-war Partisanship

This analysis is only possible because there are data on post-war partisanship for individuals who served in the war. I draw on the *People's Guides* of nine Indiana counties, published in 1874.[2] These were published by Cline and McHaffie. These guides report the history of the county and township, and include a directory of people residing there. An example of one page is located in Figure B4. While not fully exhaustive, they include a large fraction of adult men in these counties. For each person, their name, occupation, location of residence, birth place and year, date of settlement in the county, religious affiliation, and political partisanship is listed. A sample from these guides was previously collected and analyzed by (DeCanio 2007; DeCanio and Smidt 2013). Similar guides exist for 6 counties in Illinois. These are harder to digitize and contain less data on which to match individuals (no birth year), so I did not prioritize this data collection.

[2]Bartholomew, Boone, Hamilton, Henry, Hendricks, Johnson, Montgomery, Morgan, and Vermillion Counties

Figure B4: Example Page from 1874 People's Guide

Apel, Chas.; carpenter and builder; Columbus. Born in Prussia 1842; settled in B. C. 1865. Rep. Lutheran.

Aikens, David; farmer; 2 m s Columbus. Born in Va. 1835; settled in B. C. 1860. Dem. Protestant.

Aikens, James; at leisure; 2 m s Columbus. Born in Va. 1828; settled in B. C. 1867. Rep. Protestant.

ABBETT, W. A.; justice of the peace and farmer; 3 m s w Columbus. Born in B. C. 1832. Indpt. Methodist.

Abbett, O. P.; farmer; 2 m s w Columbus. Born in B. C. 1834. Indpt. Methodist.

Abbett, Henry; cooper and farmer; 2 m w Columbus. Born in B. C. 1832. Dem. Christian Union.

Abbett, Washington; farmer; 2½ m w Columbus. Born in Ky. 1827.

Arwin, John S.; physician and surgeon; Columbus. Born in Tenn. 1824; settled in B. C. 1868. Dem. Protestant.

Anderson, S. F.; 4 m n Columbus. Born in B. C. 1833. Rep. United Brethren.

ARNHOLT, WM.; farmer; 4 m s e Columbus. Born in Germany 1846; settled in B. C. 1868. Dem. Lutheran.

Arnold, Thomas; farmer; 3½ m n w Columbus. Dem. Chris.

Anthony, Joseph; farmer; 3 m n Columbus. Born in Tenn. 1830; settled in B. C. 1849. Rep. Protestant.

Armet, Charles; farmer; 2½ m n w Columbus.

Akins, C. E.; collector of Singer Sewing Manufacturing Co.; Columbus. Born in Ohio 1846; settled in B. C. 1873. Rep. Christian.

Adams, D. W.; druggist; Columbus.

This figure shows an example page listing biographical entries in one of the 1874 People's Guides.

To collect the names, birth years, birth places, and partisanship of individuals listed in these guides, I did the following. (1) I downloaded all pages of the guides from archive.org. (2) I converted each page into a gray-scale image. (3) I passed each gray-scale image to Google Cloud Vision to detect the text on each page. (4) Using custom R and Python scripts, I combined letters into words, words into rows, and rows into biographical entries using the x-y coordinates of text on the page.

Then, I used regular expressions to extract the names, birth places, birth years, years of settlement, and partisanship of each individual. I then had workers on Amazon Turk correct errors in machine transcription for each biographical entry. While some minor errors likely remain, I have key biographical details for 27169 people across these 9 counties.

B.3.2 Union Soldiers from Nine Indiana Counties

To find the effects of wartime experiences on post-war partisanship, I needed a baseline sample of soldiers to track over time. While some soldiers undoubtedly left their county of residence, a comparison of soldiers from the same county seems more plausible than to include soldiers who later moved into these counties. Moreover, by considering, say, all Union Soldiers or all Union Soldiers from Indiana, I was increasing the risk of making incorrect matches.

Based on records from the Indiana Adjutant General, 21056 men enlisted in the Union Army from these nine counties. However, a large fraction of Indiana soldiers in the ACWRD lack places of residence. To link individual soldiers to these nine counties I did the following: (1) I linked each listed place of residence given by Indiana soldiers in the ACWRD to a list of all place names (county, township, city/village, post-office) reported in the 1860 and 1870 Census in Indiana. In some cases, the same place name appears in multiple counties. And for some soldiers, no place of residence is listed. (2) For soldiers whose residence uniquely matches one of the 9 counties, they were linked to that county. (3) If soldiers' places of residence matched to multiple counties in that list of 9, they were linked (if possible) to the one county to which the majority of men in his company were uniquely matched. Otherwise, they were considered as possible residents of multiple counties (4) For soldiers with no residence listed, they were assigned to one of the 9 counties if the majority of the men in their company came from that company.

Table B.3.2 shows the number of enlistees recorded in each county by the Adjutant General of Indiana and the number of individual soldiers I match to each county (weighted by 1 over the number of counties to which they match).

For these set of Union soldiers, I have the following variables: rank at enlistment, term of enlistment, type of regiment (infantry, artillery, cavalry), date of enlistment, regiment of service (and muster in/muster out dates), company of service, how this person entered the regiment (volunteer, drafted, substitute), how they left the regiment, and exposure to casualties as described in the paper.

B.3.3 Linking Soldiers to 1874 People's Guides

I link Union Soldiers to the 1874 People's Guides using a deterministic matching procedures. While Enamorado, Fifield and Imai (2019) argues that there are gains to be made by using

Table B1: Number of Union Army Servicemen by County

County	**Soldiers (Adjutant General)**	**Soldiers (ACWRD)**
Bartholomew	2813	2490
Boone	2442	3196
Hamilton	2272	2756
Hendricks	2416	2747
Henry	2549	2633
Johnson	2033	1500
Montgomery	2850	2903
Morgan	2114	2046
Vermillion	1567	1029

an automatic and probabilistic method, applying the fastLink algorithm to this data yielded poor results. The primary reason for this are the few characteristics on which to match and a substantial number of cases in which only first initials are available.

My deterministic matching procedure uses the following attributes: county of residence, first name, last name, and birth year.

For each soldier who *survived the war*, I implement these steps to identify a set of matches:

1. I identify a set of people in the 1874 People's Guides who (a) reside in the same county the soldier resided in at enlistment and (b) are reported as moving to that county before 1866 (since soldiers may have returned home as late as that). 2. I clean first names by linking reported names, which may be misspelled or abbreviated to full names, using a crosswalk created by Abramitzky et al. (2019). 3. I then generate potential matches under the following conditions:

- If the first name and last name match exactly
- If first name matches exactly and lastnames match exactly using the metaphone sound code or have a distance of less than 0.1 on a Jaro-Winkler score.
- First names match exactly on metaphone sound code or have a Jaro-Winkler distance is less than 0.1 and last name matches exactly
- For soldiers with only a first initial listed, I consider as matches those with the same first initial and an exact match on the last name.

Within these, I consider as the best matches those that are closer on all name matching metrics and have the closest year of birth.

This procedure generates matches for 4895 Union soldiers. On average, each soldier is matched to 1.51 people post-war, with a max of 14. When restricting to best matches, the average number of matches is 1.18, with a max of 12. In cases where soldiers have more than one match, I weight them in subsequent analyses with $1/m$ where m is the number of matches.

B.3.4 Linking Soldiers to the 1860 Census

In addition to linking soldiers to the 1874 People's Guides, I also link them to the 1860 US Census. I do this for three reasons. (1) Soldiers may vary in how easily they can be located, in general, in historical records. This may have to do with how consistent they were in providing biographical details, and it may reflect whether soldiers are in fact from the county I match them to using the procedure used above. Linking soldiers to the 1860 US Census provides confirmation that soldiers were in fact from the county in question and provide a pre-treatment measure of "findability".

(2) In addition to military enlistment details, this provides additional demographic characteristics on which to test balance and use as conditioning variables. (3) Most importantly, linking to the US Census permits me to generate a measure of predicted partisanship for these soldiers, even if they are not located in the 1874 People's Guides. This permits conditioning on predicted partisanship and checking for differential attrition by baseline partisanship.

I link soldiers using the FastLink algorithm, blocking on county of residence (Enamorado, Fifield and Im 2019). I generate matches between soldiers and people in the Census using on cleaned first names, last name, birth year, and the metaphone sound codes of the first and last name. I kept matches in which the posterior probability of a correct match exceeded 0.8.

In using this procedure, I match 12103 soldiers to the 1860 Census. The mean number of matches per soldier is 1.14, and the maximum number is 14. When using Census data for soldiers, I take an average of the data for each Census match, weighted by the match probability given by fastLink.

B.4 Predicting Partisanship

In order to examine imbalance on partisanship by treatment and to investigate the possibility of differential attrition by partisanship, I need to have a pre-treatment measure of partisanship for soldiers. While it is not possible to find individual partisan affiliation for soldiers in this period of time, the next best option, if imperfect, is to generate predicted partisanship based on demographic attributes of soldiers.

To do this, I use demographic data available in the 1874 to predict partisanship. Because I need to generate these predictions for soldiers who don't appear in the 1874 People's Guides and these predictions shouldn't suffer from post-treatment bias, the demographic details need to exist in both the 1874 guides and the 1860 Census. I use names, birth years, and places of birth.

I trained a machine learning algorithm using these variables to predict partisanship in 1874. I then use this model to predict the partisanship of soldiers in 1860. In what follows, I describe this process in more detail and provide validation that the predict partisanship measure meaningfully captures variation in partisanship.

B.4.1 Training Data

To generate predictions of partisanship, I use all people listed in the 1874 People's Guides who were *not matched to any soldiers*. I then created the following features corresponding to these people:

- The first name of the person as listed in the guide/census
- The last name of the person as listed in the guide/census
- The cleaned first name of the person, using the name variant/abbreviation crosswalk created by (Abramitzky et al. 2019).
- The metaphone sound encoding of the first name as listed in the guide/census
- The metaphone sound encoding o the last name as listed in the guide/census
- The birth year listed in the guide/census
- The place of birth listed in the guide/census. Here, I used state of birth in the US or country of birth (collapsing many regions in Germany/Ireland to the country).
- Partisanship, collapsed to: Republican, Democrat, Other, or None. The vast majority of people are in Republican or Democrat. The most common "other'' is the Grange.

Note that the county of residence or township and occupation, although also in the census, are not included. These may be affected by war-time experience, and so I do not include them as features to predict partisanship.

B.4.2 Machine Learning Classifier

I then used the `fastText` algorithm to classify individuals' partisanship. This algorithm is used for text classification and, rather than using a bag of words to classify texts, represents words and subsets of characters in that word in a lower dimensional space, and then uses these word representations to classify documents. I used this approach, because name spellings are too numerous to include as binary-encoded features and I would be predicting with new datasets that might even contain new name variants. `fastText` permits making predictions even with names/words that have not been seen in the training data.

To train the algorithm, I split this data into 5 different training/testing groups and trained the model on each group. Each time, I used `fastText`'s auto-tune functions to select the optimal model parameters, as chosen by performance on the test group.

B.4.3 Validating Predictions

To validate the performance of these models for four sets of data. (1) I generated predictions for each model on the "test'' group (not used in training, only in choosing model parameters). (2) I also check the performance of taking the average of predictions from all five models (bagging) for the people in the 1874 guides that were matched to soldiers, and not used in the training at all ("validation'' group). (3) I averaged the predictions from the five models for all men of voting age in residing in Indiana counties in 1860, using data from the 1860 Census. This is the "aggregate census'' group. (4) For analyses in the paper, I generate latent partisanship for soldiers by (a) averaging predicted probabilities of partisanship from all five models for people in the Census and (b) averaging these scores across all matches a soldier has in the Census, weighting by the match probability derived using fastLink. This is the "individual census'' group.

Predicting Partisanship in 1874 I assess the predictive power of these classifiers in two ways.

ROC/PR Curves The ROC curve plots the trade-off between false positives rates and true positive rates, given different probability thresholds for making a binary classification. A poor classifier performs about as good as or worse than random guessing (which would have the performance of the diagonal line). The PR curve plots the trade-off between precision (the fraction of positive classifications that are true positives) and recall (the fraction of the true positives that are correctly identified by the classifier). Here, we assess the performance of the curve against the performance of the "no-skill'' classifier of guessing every case to be a match. The more the area under the PR or AUC curve, and the more these curves are higher than the "no-skill'' curves, the greater the predictive power of the classifier.

Predicted Probability vs. Actual Probability I also plot the (smoothed) probability of being a Republican/Democrat across predicted probabilities of being a Republican/Democrat. This should be nearly perfectly correlated, if the classifier performs well.

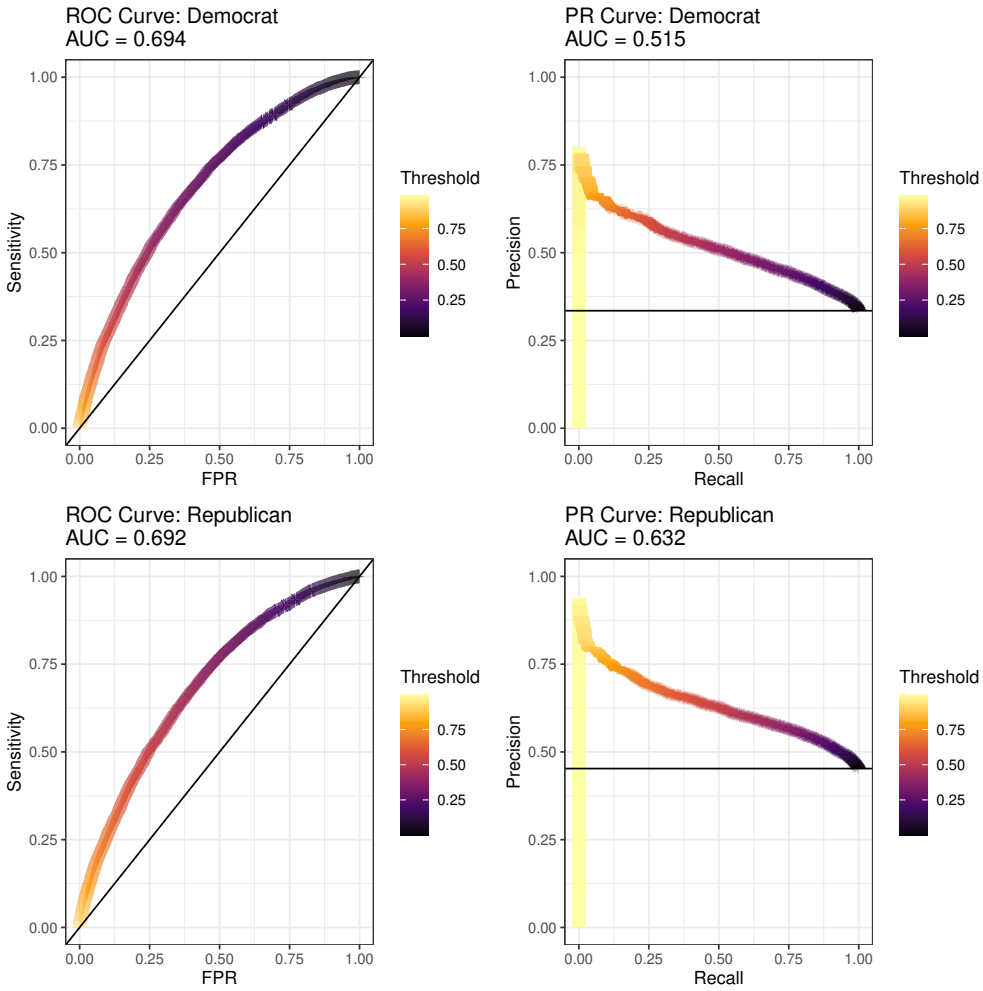
Performance in the "test'' set In figure B5, I show that the ROC and PR curves for the classifier in the "test'' set, while imperfect, substantially outperforms the naive classifier for both Republicans and Democrats. And in figure B6, I show the predicted probability of being a Republican and Democrat versus the actual probability. Again, while the classifier tends to over-predict at the high and low end, there is clearly a strong relationship between predicted and actual probability of Partisanship.

Performance in the "validation'' set In figure B7, I show the ROC and PR curves for this classifier in the "validation'' data: the 1874 biographical records linked to Union Soldiers. These were never used in training the classifier. Despite that, and despite there being substantial differences in the fraction who are Republican vs Democrat (soldiers are more Republican and non-soldiers), the classifier still outperforms the naive "no-skill'' classifier. And again, the predicted probability of partisanship is strongly related with actual probability of partisanship (Figure B8). Like the test set, though, the relationship isn't one to one. It should be noted that that the classifier does a worse job at predicting Democrats among veterans: veterans predicted with to be Democrats with nearly 100 percent probability are only actually Democrats less than 70 percent of the time. This is likely reflecting the effects I find that people were converted to being Republican.

Performance in the "aggregate census'' set This provides evidence that the classifier does a fairly good job predicting partisanship in 1874 for people whose demographic data was measured in 1874. But does it do a good job predicting partisanship during or before the Civil War, when selection into service and, potentially, combat experiences took place?

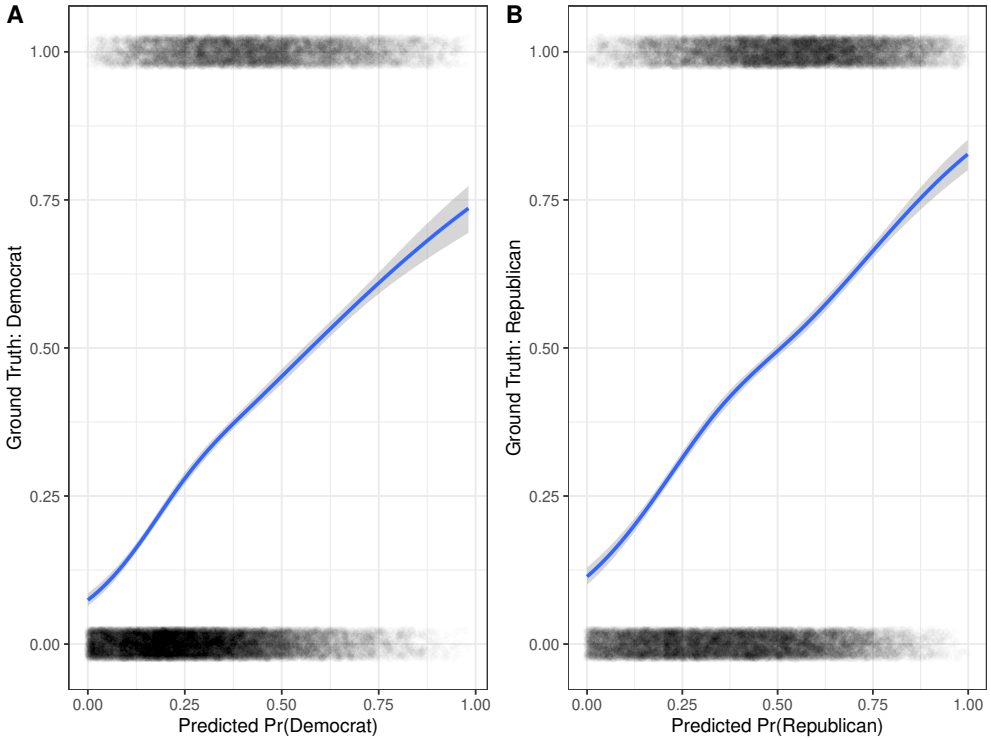
To assess this, I aggregate the predicted probabilities of Democratic and Republican partisanship for men eligible to vote in the 1860 elections and living the 9 Indiana counties,

Figure B5: AUC/PR Curves for "Test" Set



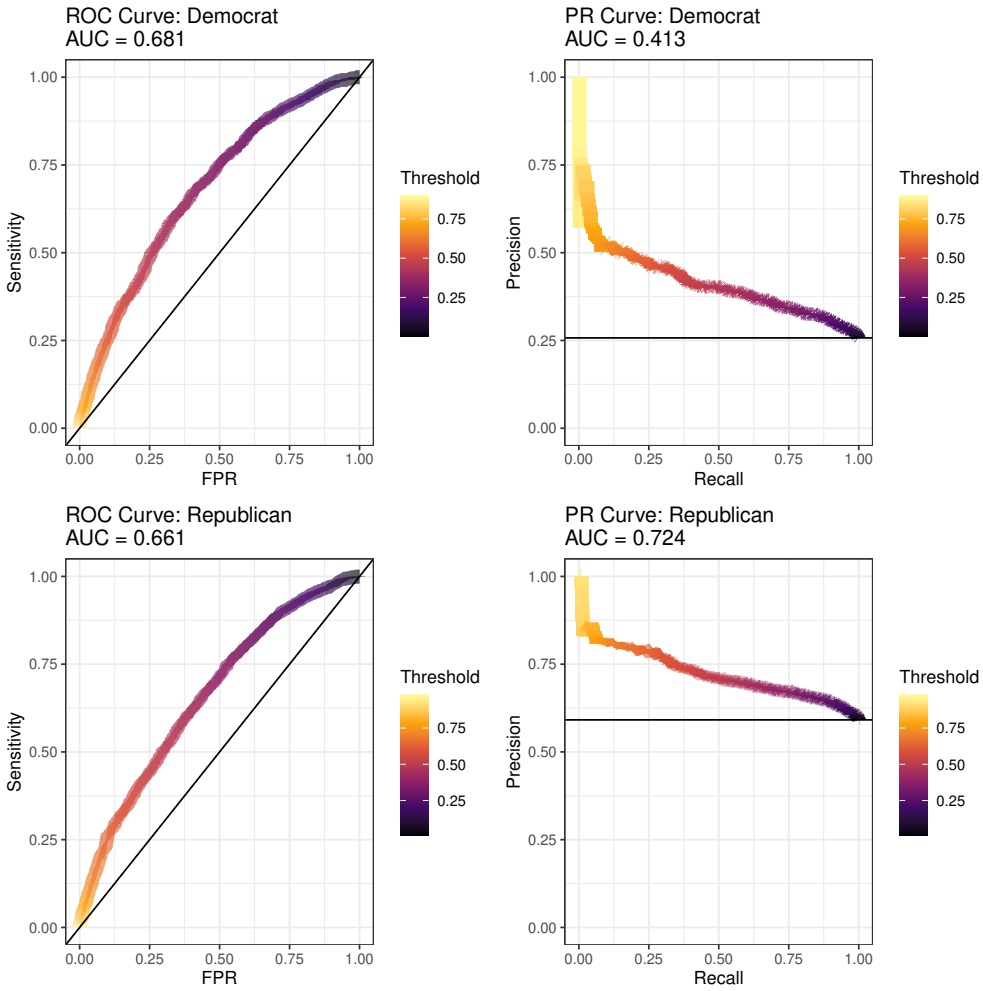
This figure shows the ROC and PR curves for the fastText classifier of Republican and Democratic partisanship in 1874, for the "test" set or cases held out in each fold when training the classifier.

Figure B6: Actual Probability of Partisanship vs. Predicted Probability of Partisanship: "Test" Set



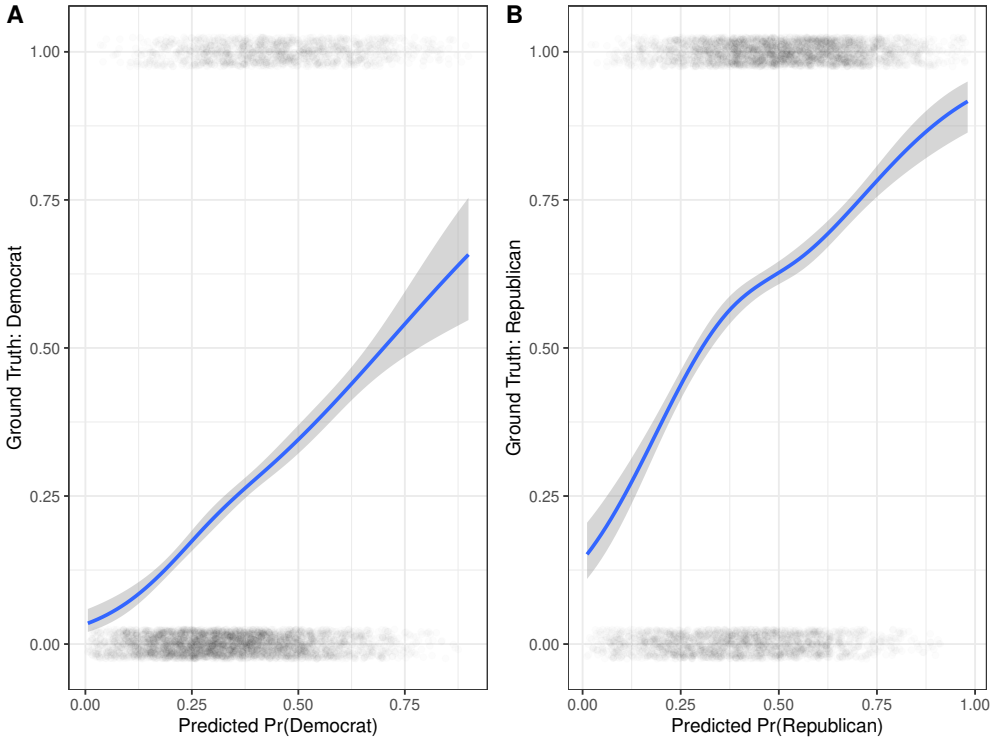
This figure shows the calibration curve of predicted partisanship vs. actual partisanship from the fastText classifier of Republican and Democratic partisanship in 1874, for the "test" set or cases held out in each fold when training the classifier.

Figure B7: AUC/PR Curves for "Validation" Set



This figure shows the ROC and PR curves for the fastText classifier of Republican and Democratic partisanship in 1874, for the "validation" set or cases linked to veterans and not used in training the classifier.

Figure B8: Actual Probability of Partisanship vs. Predicted Probability of Partisanship: "Validation" Set

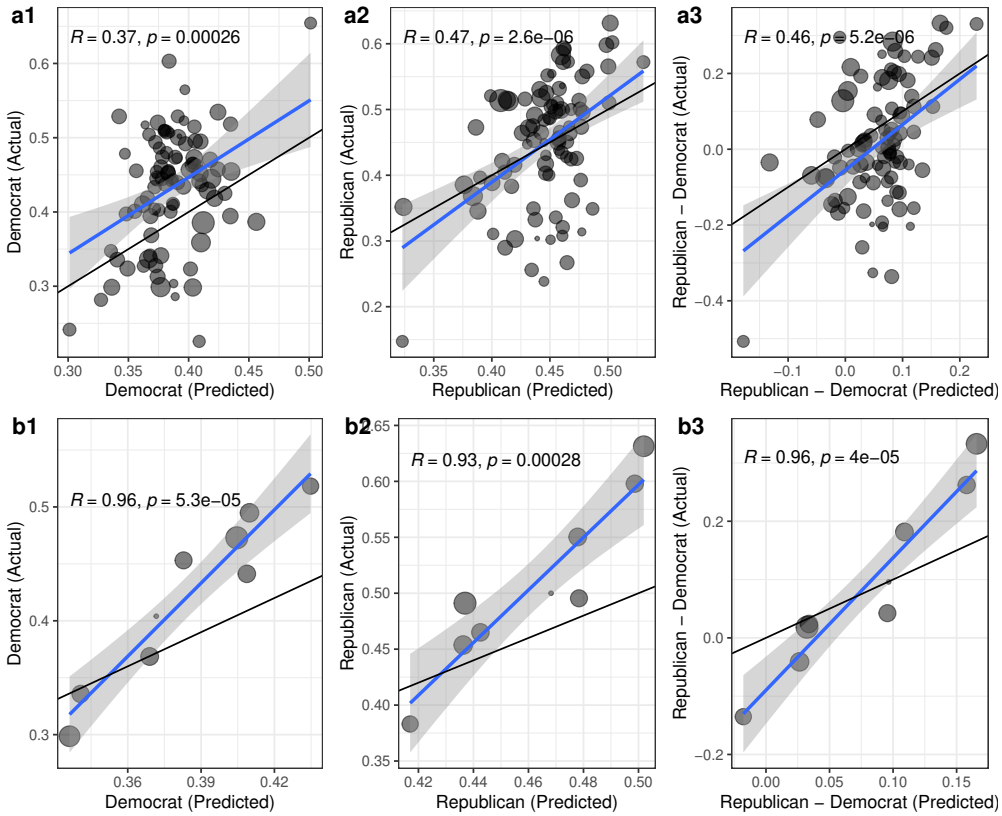


This figure shows the calibration curve of predicted partisanship vs. actual partisanship from the fastText classifier of Republican and Democratic partisanship in 1874, for the "validation" set or cases linked to veterans and not used in training the classifier.

according the 1860 Census. I then take the average Republican probability, average Democratic probability, and the difference of these two for all men in each county (and townships in 4 counties).

In figure B9, I show the correlation between the predicted partisanship of these 9 counties in 1860 and the actual 1860 vote share for Democrats and Republicans. While, again, the predicted probability does not predict voteshare at a one-to-one level, the correlation is nearly 1. This is remarkable given that the classifier was not provided any county labels.[3]

Figure B9: Predicted vs. Actual 1860 Voteshare for 9 Indiana Counties

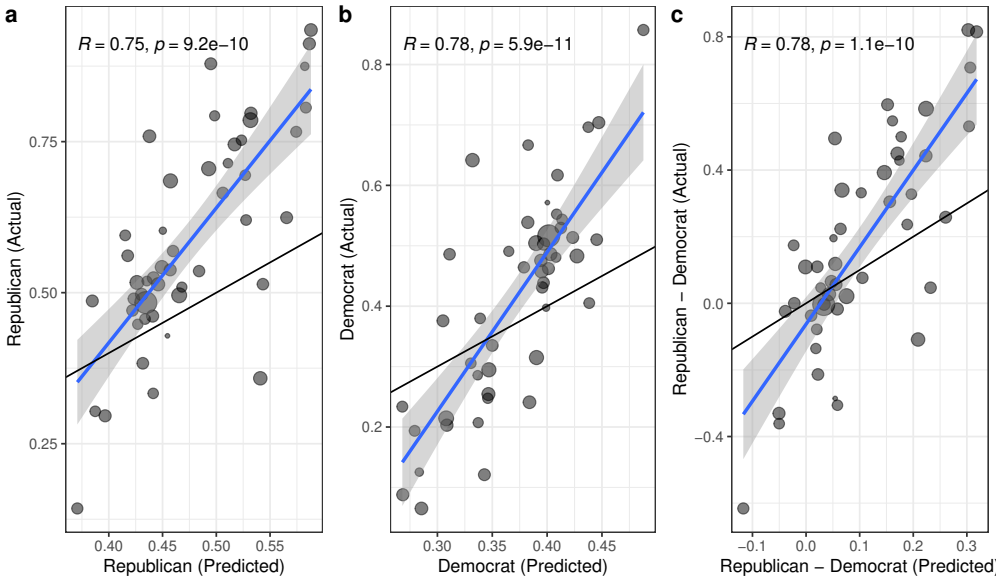


This figure shows the calibration curve of predicted partisanship vs. actual partisanship from the fastText classifier of Republican and Democratic partisanship in 1874, for the Indiana Counties in the 1860 Presidential election. Predicted partisanship was obtained by applying the fastText classifier to all voting-eligible males in each county and taking the mean probability of Democratic and Republican partisanship.

I repeat the same procedure using township level returns in Hendricks, Henry, Montgomery, and Morgan counties. These results are reported in Figure B10. Again, while the relationship is not one-to-one, predicted partisanship of townships in 1860 is correlated with the difference in Democratic and Republican vote share at nearly 0.8. Clearly, the model generates predictions of partisanship that correspond to real pre-war differences.

[3]The correlation between predicted partisanship and voteshare is weaker for all Indiana counties, but still highly significant.

Figure B10: Predicted vs. Actual 1860 Voteshare for Townships in Hendricks, Henry, Montgomery, and Morgan Counties



This figure shows the calibration curve of predicted partisanship vs. actual partisanship from the fastText classifier of Republican and Democratic partisanship in 1874, for the townships in Hendricks, Henry, Montgomery, and Morgan counties in the 1860 Presidential election. Predicted partisanship was obtained by applying the fastText classifier to all voting-eligible males in each township and taking the mean probability of Democratic and Republican partisanship.

Performance in the "individual census" set Finally, I evaluate whether the specific measure I use in the paper for the latent partisanship of individual soldiers (averaging predicted partisanship across all census matches for that soldier) is meaningfully predictive. In Table B2, I show that predicted partisanship significantly predicted actual 1874 partisanship for the set of soldiers who are linked to both the census and to 1874. Moreover, I also find that predicted partisanship significantly predicts features of military service that correlate with partisanship. For example, in the 1874-linked sample, substitutes are more likely to be Democrats. I find that this is also true using predicted partisanship. Similarly, I find that predicted Democrats are more likely to desert and predicted Republicans are more likely to die (Figure B2).

Table B2: Indiana Veterans: Predicting Partisanship with Latent Partisanship

	Dependent variable:					
	Rep.	Dem.	Party Diff.	Rep.	Dem.	Party Diff.
	All Matches			Best Matches		
	(1)	(2)	(3)	(4)	(5)	(6)
Rep. Prob.	0.442***			0.460***		
	(0.041)			(0.043)		
Dem. Prob.		0.408***			0.417***	
		(0.039)			(0.041)	
Diff Prob.			0.228***			0.236***
			(0.019)			(0.019)
Observations	5,965	5,965	5,965	4,541	4,541	4,541

Note: *p<0.05; **p<0.01; ***p<0.001

Sample includes men serving in Indiana Regiments who were matched to the 1860 Census and 1874 People's Guides. Restricted sample includes only the best 1874 matches. Individuals are weighted by 1 over the number of matches. Latent partisanship is predicted probability of being Republican or Democrat based on the person's name, birth year, and birth place listed in the 1860 census. Standard errors are clustered by individual.

B.5 Balance Tests

In this section, I report the results of balance tests on pre-war covariates. Figure B11 reports the relationship between company casualties and average partisanship, net regiment fixed effects. Figure B3 shows the relationship between a 2 SD change in each pre-war covariate and the change in company casualties, using the design with regiment fixed effects. There is balance across nearly all variables. The exception is age and status as household head. People

who are household heads or older had slightly lower exposure to company casualties. This imbalance is unconcerning for two reasons. (1) The within-regiment relationship between age and household head and post-war partisanship is precisely estimated and 0. (2) This likely reflects the fact that younger men (less likely to be heads of household) are less risk averse and thus either take actions that might endanger themselves or their company. Moreover, these changes are substantively small. The SD of within-regimental variation in company casualties is 2.3. A 2 SD change in age at enlistment is related to less than a tenth of a SD change in company casualties. The dashed vertical lines indicate 0.1 SD change in company casualties.

While there appear to be large differences by draft and substitute status, this is because there is little variance in draft/substitute status within regiments.

B.6 Robustness

Tables B3, B4, and B5 show that the results are robust to considering different samples: using the best or all matches in 1874, and using all soldiers or only those found in the 1860 census.

Figure B12 shows the distribution of t statistics on company casualties across different specifications: including all soldiers (where partisanship for those without a match is 0), only non-attriters; Census matched/All soldiers; Best post-war match or all post-war matches; No covariates, army covariates, or army and census covariates; measuring treatment as company casualties or casualty rate.

B.7 Attrition

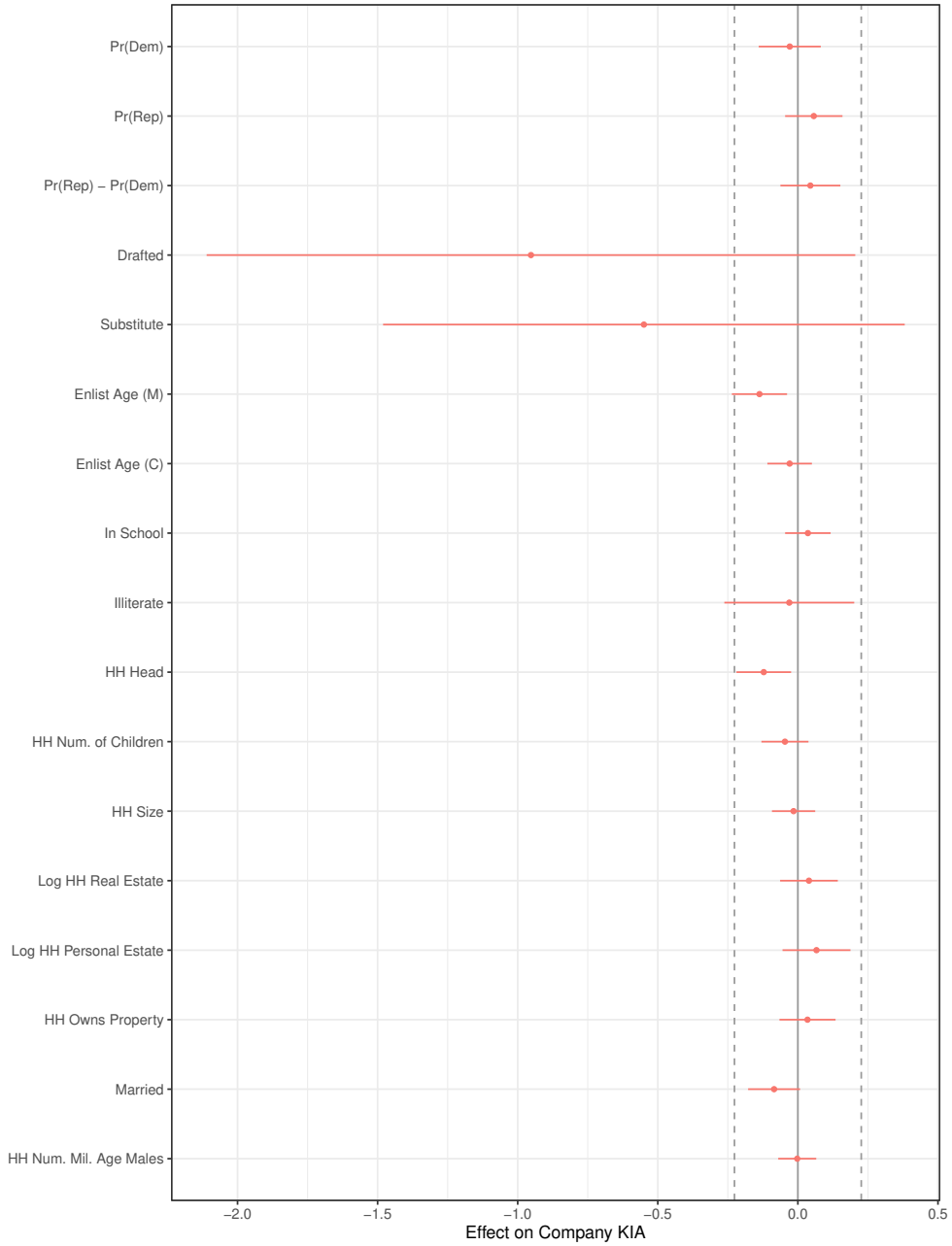
Only 33 percent of soldiers found in the 1860 Census were matched to the 1874 People's Guides and only 23 percent of soldiers over-all were matched. While this is not far off from other attempts to link historical records (see, e.g. Abramitzky et al. 2019), there are reasons to be concerned about attrition bias. It is necessary to assess whether there is any relationship between attrition and potential outcomes, any relationship between attrition and treatment, and any interaction between treatment, partisanship, and attrition. In this section, I investigate these possibilities and then directly assess the robustness of my results to the use of IPW adjustments, assuming Missingness Independent of Potential Outcomes conditional on covariates.

B.7.1 Sources of Attrition

I start by examining what explains variation in attrition. In the analysis sample (infantry soldiers matched to the 1860 census), I compare modes of entering and exiting service and find that, other than death, the only significant predictor of attrition was desertion. Deserters were about 5 ppt less likely to appear post-war.

In unadjusted comparisons (no regiment fixed effects), partisanship is related to attrition (Figure B2). Latent Republicanism increases the probability that a person is observed after the war, decreases the chance that a person deserts, and is positively but not significantly related to dying. Democrats, by contrast, were more likely to desert.

Figure B11: Balance on pre-war attributes across levels of treatment



This plots the standardized effect of a 2 SD change in pre-war attributes on company casualties, conditional on regiment fixed effects.Vertical dashed lines indicate 1/10th of within-regiment standard deviation of company casualties. Estimates derived from regression of Company KIA on pre-muster covariates, conditioning on regiment fixed effects. Sample includes men serving in Indiana Infantry Regiments who were matched to the 1860 Census: 10288 unique soldiers, in 889 companies, across 288 regiments. Individuals are weighted by 1 over the number of census matches. Standard errors are clustered by company.

Table B3: Effect of Company Casualties on Post-war Partisanship (Best Matches)

	Dependent variable:					
	Rep.	Dem.	Party Diff.	Rep.	Dem.	Party Diff.
	Baseline			Army Controls		
	(1)	(2)	(3)	(4)	(5)	(6)
Company KIA	0.006	-0.008^{**}	0.007^{*}	0.007	-0.006^{*}	0.007^{*}
	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)
Regiment FE	Y	Y	Y	Y	Y	Y
Army Controls	N	N	N	Y	Y	Y
Census Controls	N	N	N	N	N	N
Observations	4,659	4,659	4,659	4,659	4,659	4,659

Note: *p<0.05; **p<0.01; ***p<0.001

Sample includes men serving in Indiana Regiments who were matched to the 1860 Census and their best match, if any, in the 1874 People's Guides. Baseline and control models, respectively, include data on 3885 individual soldiers, serving in 617 companies, across 234 regiments. Regiment fixed effects includes a dummy for each group of soldiers who joined a regiment in the same year. Observations are weighted by 1 over the number of post-war matches for each unique soldier.Standard errors are clustered by company.

Table B4: Effect of Company Casualties on Post-war Partisanship (Census-Linked, All Matches)

	Dependent variable:								
	Rep.	Dem.	Party Diff.	Rep.	Dem.	Party Diff.	Rep.	Dem.	Party Diff.
	Baseline			Army Controls			All Controls		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Company KIA	0.007*	−0.008**	0.008*	0.008*	−0.007*	0.008*	0.008*	−0.006*	0.007*
	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)
Regiment FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Army Controls	N	N	N	Y	Y	Y	Y	Y	Y
Census Controls	N	N	N	N	N	N	Y	Y	Y
Observations	4,669	4,669	4,669	4,669	4,669	4,669	4,669	4,669	4,669

Note: *p<0.05; **p<0.01; ***p<0.001

Sample includes men serving in Indiana Regiments who were matched to the 1860 Census and their best match, if any, in the 1874 People's Guides. Baseline and control models, respectively, include data on 3134 individual soldiers, serving in 540 companies, across 211 regiments. Regiment fixed effects includes a dummy for each group of soldiers who joined a regiment in the same year. Observations are weighted by 1 over the number of post-war matches for each unique soldier.Standard errors are clustered by company.

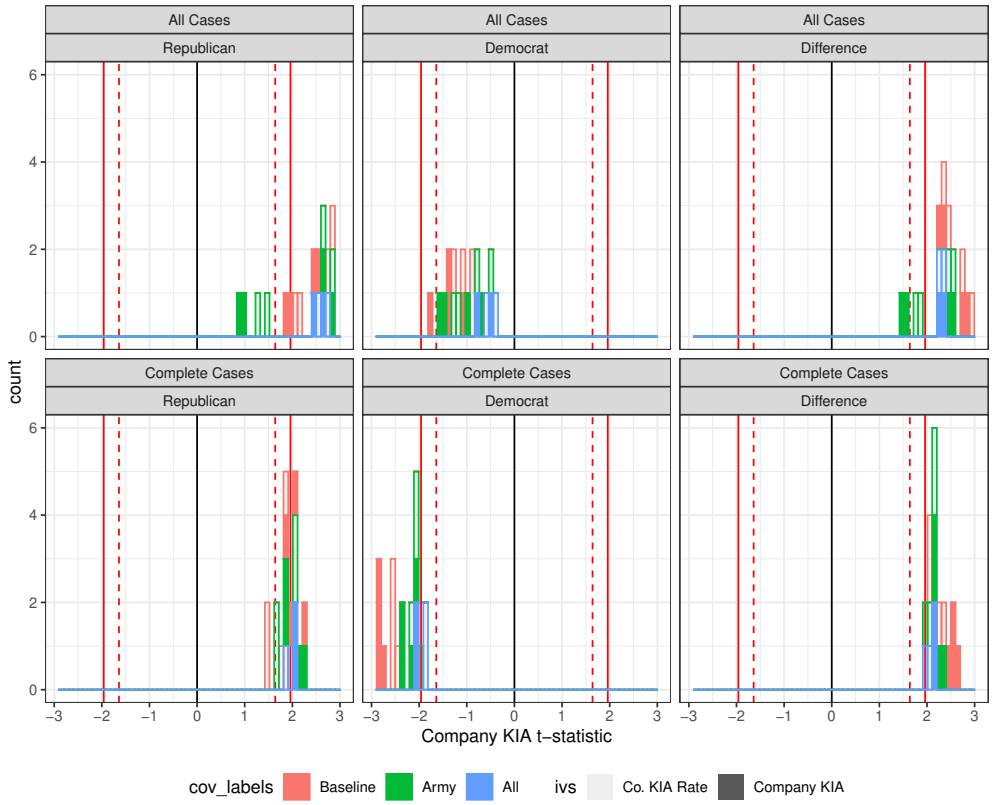
Table B5: Effect of Company Casualties on Post-war Partisanship (All Matches)

	Dependent variable:					
	Rep.	Dem.	Party Diff.	Rep.	Dem.	Party Diff.
	Baseline			Army Controls		
	(1)	(2)	(3)	(4)	(5)	(6)
Company KIA	0.005	-0.007^{**}	0.006^{*}	0.006	-0.006^{*}	0.006^{*}
	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
Regiment FE	Y	Y	Y	Y	Y	Y
Army Controls	N	N	N	Y	Y	Y
Census Controls	N	N	N	N	N	N
Observations	5,976	5,976	5,976	5,976	5,976	5,976

Note: *p<0.05; **p<0.01; ***p<0.001

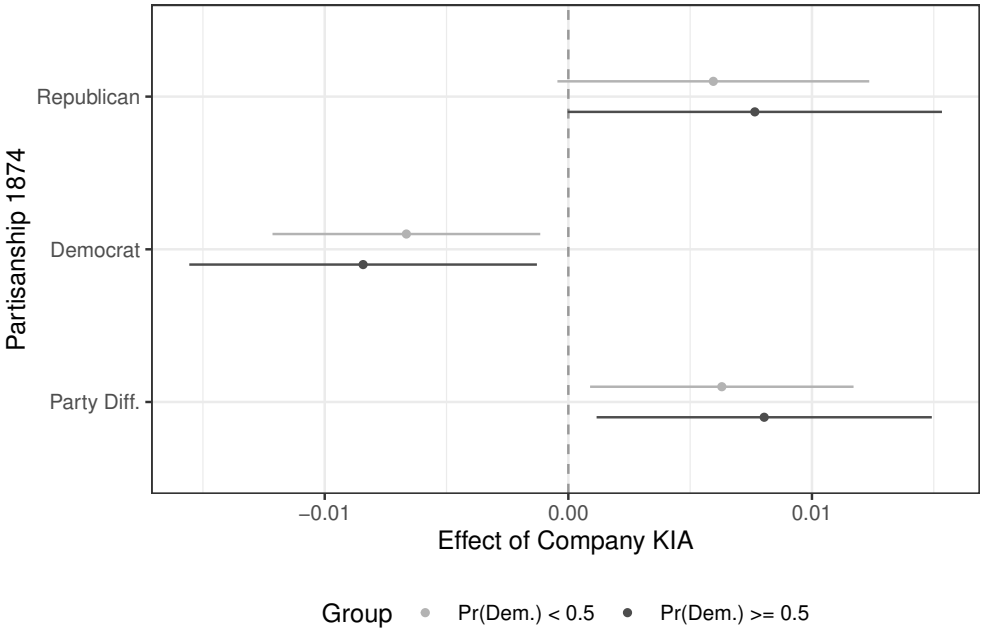
Sample includes men serving in Indiana Regiments who were matched to the 1860 Census and their best match, if any, in the 1874 People's Guides. Baseline and control models, respectively, include data on 3910 individual soldiers, serving in 618 companies, across 234 regiments. Regiment fixed effects includes a dummy for each group of soldiers who joined a regiment in the same year. Observations are weighted by 1 over the number of post-war matches for each unique soldier.Standard errors are clustered by company.

Figure B12: Robustness of Company Casualties Results to alternate specifications



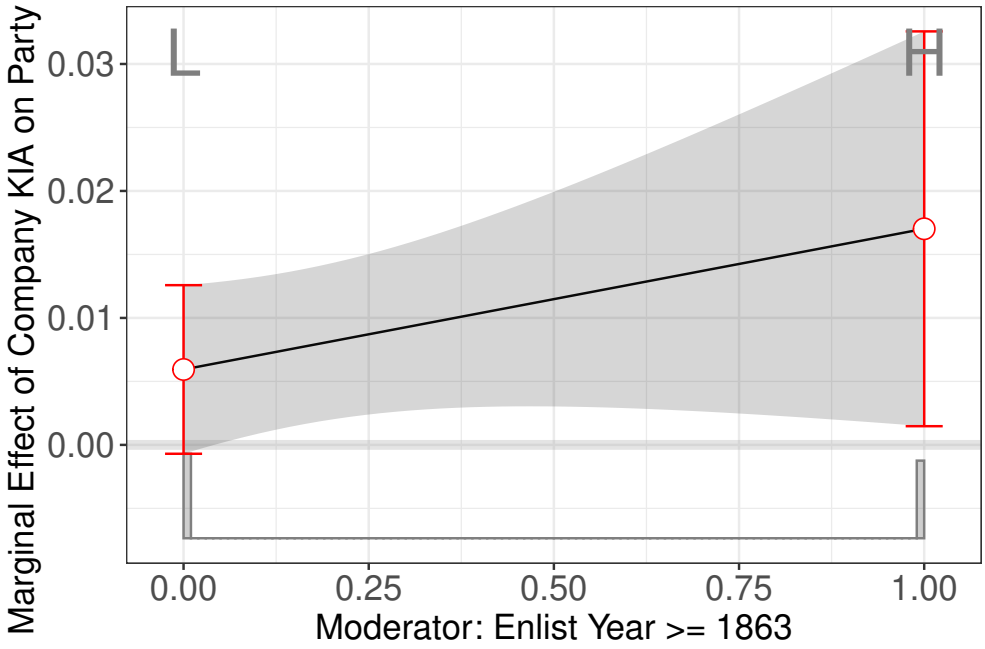
This reports t statistics for analyses of the company casualty natural experiment across different samples (all soldiers vs matched only; best matches vs all matches), covariates (none, army, army + census), and measures of the treatment (Company KIA and Company KIA rate). All models use standard errors clustering by company.

Figure B13: Effects of Company Casualties By Predicted Partisanship



This figure extends the results from Table 2, columns 1–3, by showing the marginal effect of Company KIAs on partisan swing, across predicted pre-war partisanship (probability of Democratic partisanship greater or less than 0.5), conditional on regiment fixed effects. Sample includes men serving in Indiana Infantry Regiments who were matched to the 1860 Census: 3259 unique soldiers, in 564 companies, across 216 regiments. Individuals are weighted by 1 over the number of 1974 Peoples Guide matches. Standard errors are clustered by company.

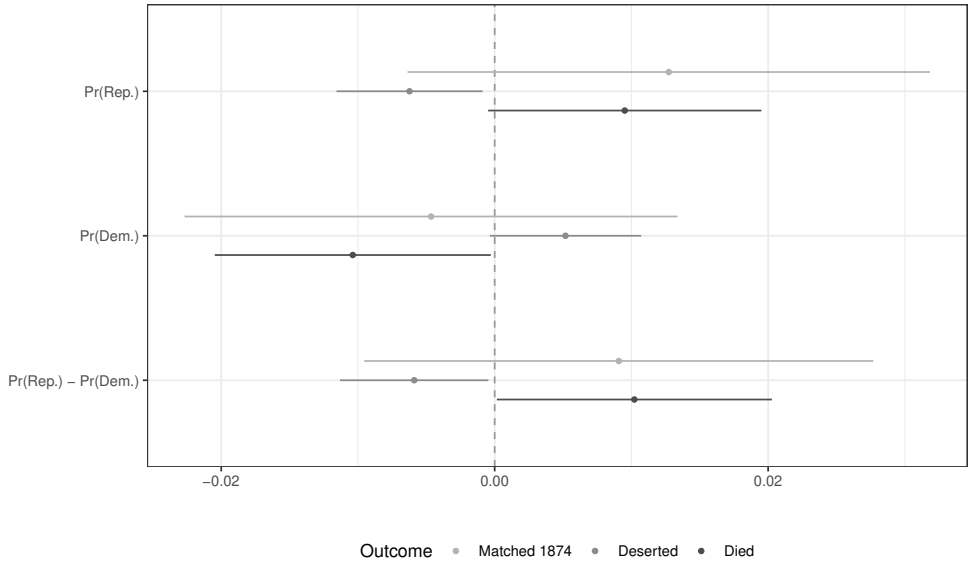
Figure B14: Effects of Company Casualties across early and late enlistments



This figure extends the results from Table 2, column 3, by showing the marginal effect of Company KIAs on partisan swing, for soldiers enlisting pre- and post- Emancipation Proclamation, conditional on regiment fixed effects. Sample includes men serving in Indiana Infantry Regiments who were matched to the 1860 Census: 3237 unique soldiers, in 563 companies, across 216 regiments. Individuals are weighted by 1 over the number of 1974 Peoples Guide matches. Standard errors are clustered by company.

Within regiments, partisanship is not significantly related to overall attrition (Figure B15). Latent Republicanism increases the probability that a person died and decreases their chance of dying. People who were more likely to be Democrats were more likely to deserts and less likely to die.

Figure B15: Differential Attrition by Predicted Partisanship (Within Regiment))



This figure shows the change in attrition, death, or desertion (related to attrition) associated with a 2 SD change in the predicted probability of Republican or Democratic partisanship, using the within regiment design.

B.7.2 Does treatment predict attrition?

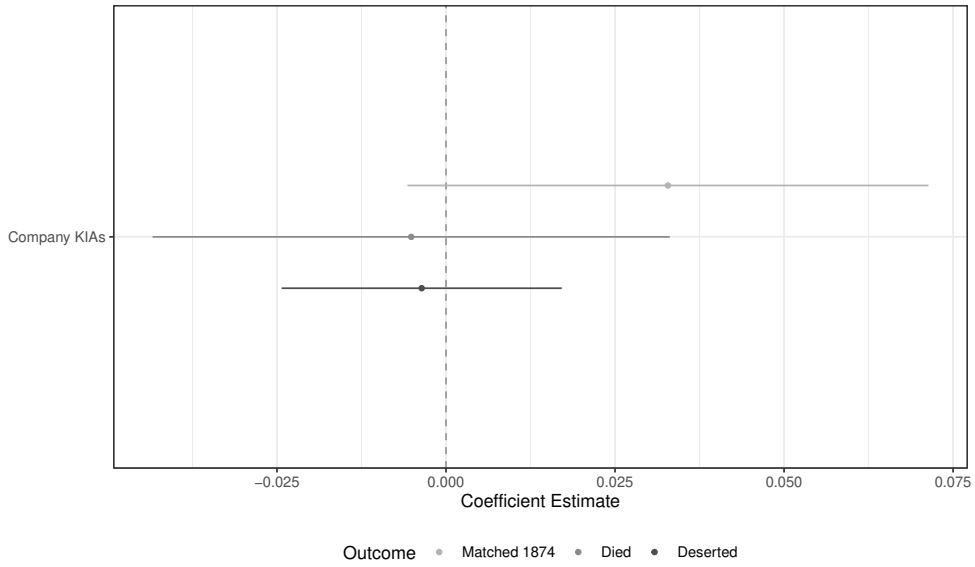
Attrition bias could arise if treatment is related to attrition and relationship between treatment and attrition is related to potential outcomes.

Using the design employed for analyzing the natural experiment (regiment fixed effects), do company-level casualties predict attrition? Figure B16 shows that neither death nor desertion (the two main sources of attrition) are related to treatment. However, a 2 SD increase in company casualties is positively (though not significantly) related to being found after the war (an increase of about 2.5 ppt). And Figure B11, above, shows that treatment is balanced across baseline partisanship.

Though not significant, I further test whether there is any relationship between the effect of treatment on attrition and latent partisanship. Figure B17 shows binned interaction effects between the combat casualties and predicted partisanship, obtained using the `interflex` package in R. There are no significant differences between any of the marginal effects of company casualties at high, medium, or low levels of Democratic, Republican, or difference in predicted partisanship.

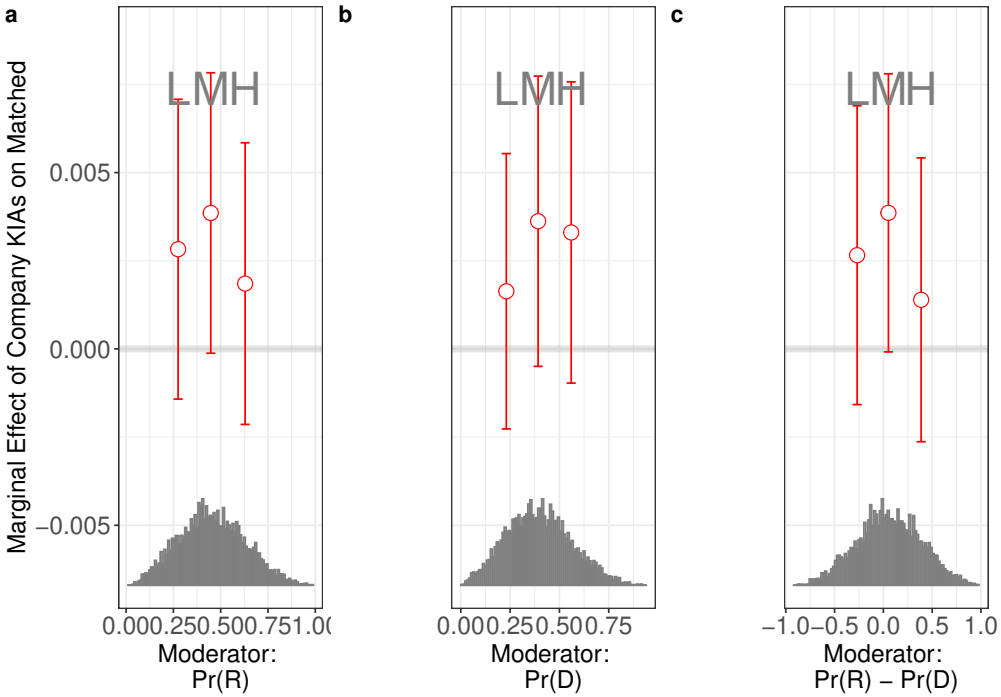
This is reassuring, as it does not provide any clear evidence of attrition by partisanship. If anything, the positive relationship between company combat deaths and attrition may reflect long-term trauma. On average, soldiers who were wounded or disabled were about

Figure B16: Differential Attrition by Treatment (Within Regiment)



This figure shows the change in attrition, death, or desertion (related to attrition) associated with a 2 SD change in company casualties, using the within regiment design.

Figure B17: Differential Attrition by Treatment (Within Regiment) and Partisanship



This figure reports the marginal effect of Company KIAs on post-war matching, across bins of predicted partisanship probability, conditional on regiment fixed effects. Sample includes men serving in Indiana Infantry Regiments who were matched to the 1860 Census: 10329 unique soldiers, in 896 companies, across 288 regiments. Individuals are weighted by 1 over the number of census matches. Standard errors are clustered by company.

3.8 ppt more likely to be found after the war. This is likely because they needed help from friends and family to cope with these injuries. Something similar might be happening with those who had more intense combat experiences.

Nevertheless, out of an abundance of caution, I also replicate the main analyses using IPW estimators to address any possible attrition bias.

B.7.3 MIPO conditional on covariates

If there is differential attrition, one way to address it is to estimate the effect of combat casualties, where Missingness is Independent of Potential Outcomes, conditional on covariates. The recommended way of addressing this is to weight complete cases as by:

$$\frac{1}{\pi(R_i|Z_i, X_i)}$$

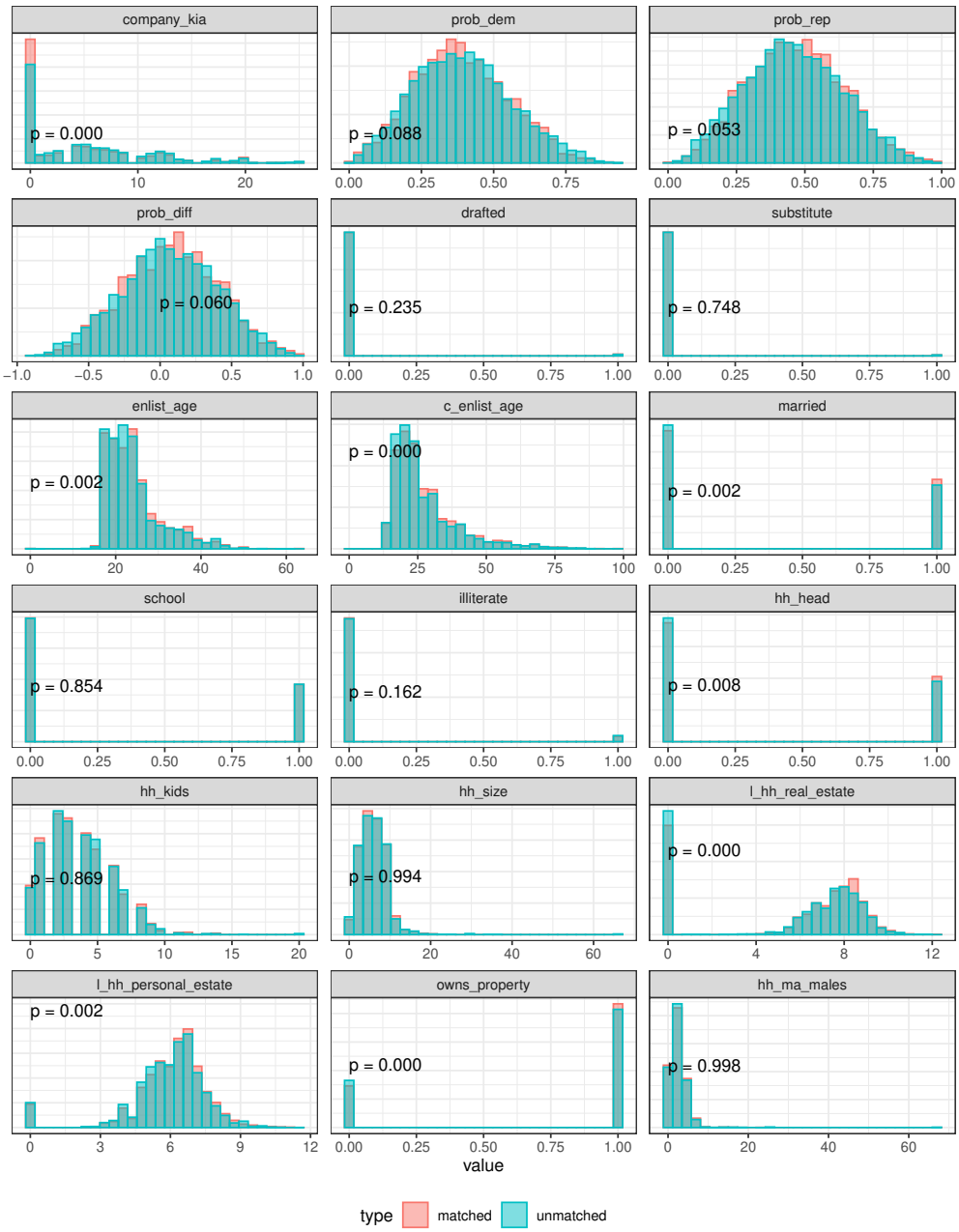
Where R_i is an indicator that case i is observed post-war; Z_i is level of treatment, and X_i is a vector of covariates (Gerber and Green 2012). This is an inverse probability weight estimator. A few issues arise when using inverse probability weights. First, these weights may perform poorly if there is a lack of positivity: there are sets of covariates for which missingness is near 0 or 100 percent. One way to investigate this is to look at overlap in covariates between the matched sample and the unmatched sample of soldiers. While there is imbalance in the distribution of covariates (both raw (Figure B18), and centered within regiments (Figure B19)), the distributions clearly overlap.

Second, inverse probability weights depend on estimating a propensity score model. If the propensity score model is incorrectly specified, bias may persist. To flexibly estimate the propensity that soldiers were matched after the war, I use generalized regression forests proposed by Athey, Tibshirani and Wager (2018). This permits both non-linearities and interactions between covariates in estimating the propensity score. In this model, I estimate propensity to be matched as a function of the treatment and a vector of covariates.[4] I include all of these variables in as both differences from the regiment mean and as the regimental mean.

Third, propensity scores do not ensure balance on covariates (in either the first or higher moment). Thus, I also estimate propensity scores for being matched using covariate balancing propensity scores that both predict propensity but also are constrained to maximize balance on covariates (Imai and Ratkovic 2014). I employ the CBPS method for the ATE,

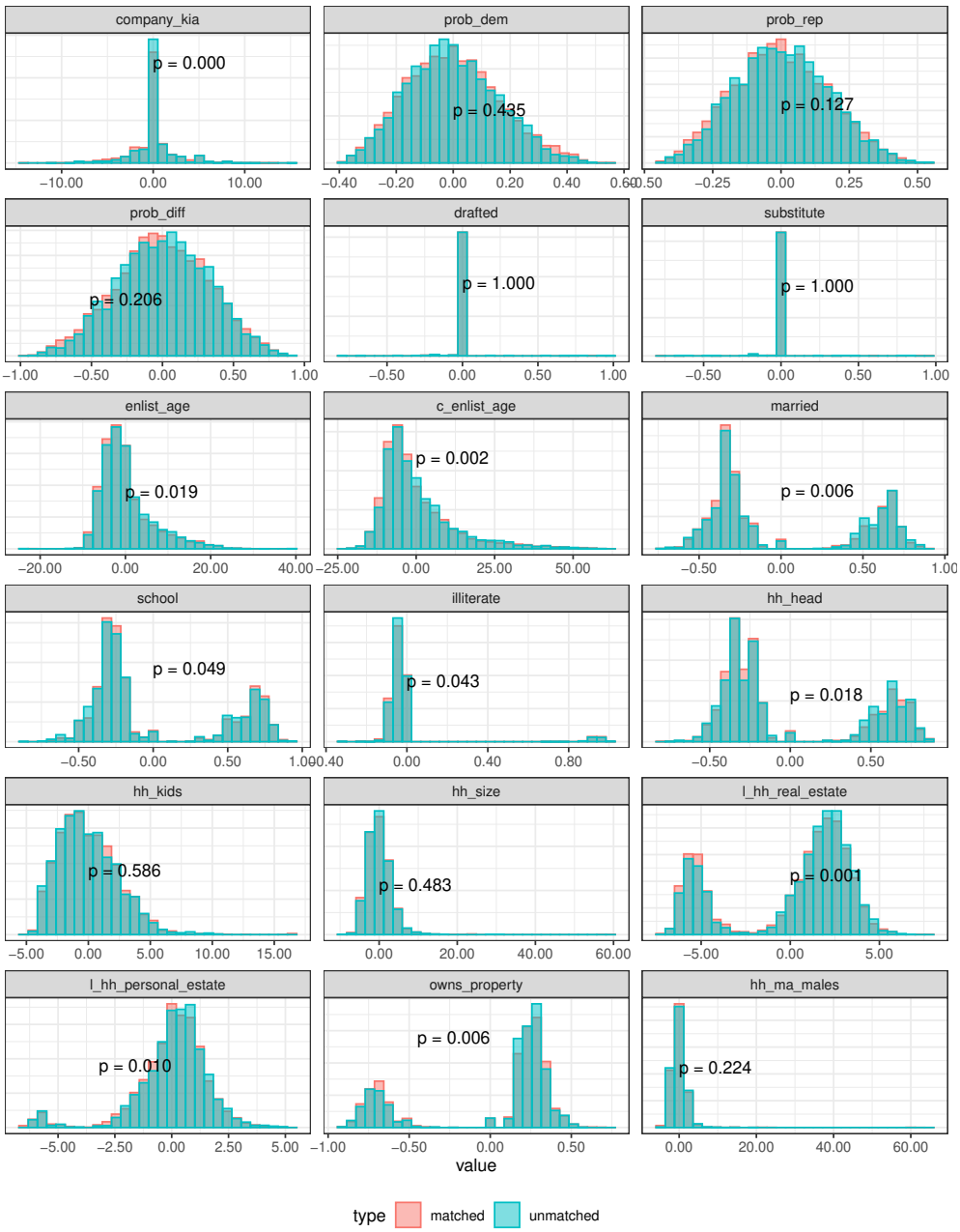
[4]Combat casualties, other deaths, company casualty rate, other death rate, number of men in company, date of muster into company, date regiment was organized, rank at enlistment, pr(Democrat), pr(Republican), pr(No party), drafted, substitute, Age (census), attended school in 1860, illiterate in 1860, HH head in 1860, children in HH in 1860, logged HH real estate value, logged HH personal estate value, HH owns property, married in 1860, HH size in 1860, military aged males in HH, birth place, county of residence in 1860.

Figure B18: Overlap on (raw) covariates for the soldiers matched and unmatched to 1874 People's guides.



This figure shows overlap in the distribution of covariate values for the unmatched sample of soldiers and those matched to the 1874 Guides. p values are for differences in means for binary variables, and for KS tests for continuous variables. Sample includes men serving in Indiana Infantry Regiments who were matched to the 1860 Census: 10288 unique soldiers, in 889 companies, across 288 regiments.

Figure B19: Overlap on (regiment-mean-centered) covariates for the soldiers matched and unmatched to 1874 People's guides.



This figure shows overlap in the distribution of covariate values, centered on the regiment mean, for the unmatched sample of soldiers and those matched to the 1874 Guides. p values are for differences in means for binary variables, and for KS tests for continuous variables. Sample includes men serving in Indiana Infantry Regiments who were matched to the 1860 Census: 10288 unique soldiers, in 889 companies, across 288 regiments.

which returns identical weights to those defined above. Here, I constrain balance on all of the same covariates listed above, interactions between the treatment variable and every other covariate, and the square of every variable, including treatment. I consider two different models: one in which I include the covariates, interactions, and squared terms both centered on regiment means and at the regiment mean (CBPS 1); and for only the covariates centered on regimental means (CBPS 2). Figures B20 and B21 show the improvement in balance on post-war missingness achieved.

Figure B22 shows the results of replicating the analysis in the body of the paper (Table 2, columns 1–3) using no weights, regression forest weights, and the CBPS weights. The results are virtually identical.

C Veterans and Suffrage

In this section, I discuss the plausibility of drawing ecological inferences about the effect of military service on veterans' support for African American suffrage, using a difference-in-differences ecological regression.

C.1 Ecological Diff-in-Diff

In the conventional ecological "accounting identity'', pre-war suffrage support in a township i, S_{i0}, can be expressed as a function of (a) the fraction of men who will enlist, X_i, (b) the fraction of enlistees who supported suffrage at $t = 0$, β^e_{i0}, and (c) the fraction of non-enlistees who supported suffrage at $t = 0$, β^n_{i0}.

$$S_{i0} = X_i\beta^e_{i0} + (1 - X_i)\beta^n_{i0}$$

Assuming that β^e_{i0} and β^n_{i0} are independent of X_i, Goodman's ecological regression $S_{i0} = \alpha + X_i\beta + \epsilon_i$ returns an unbiased estimate of $\beta = E(\beta^e_{i0}) - E(\beta^n_{i0})$.

If we extend this to a second time period, time = 1, then:

$$S_{i1} = X_i\beta^e_{i1} + (1 - X_i)\beta^n_{i1}$$

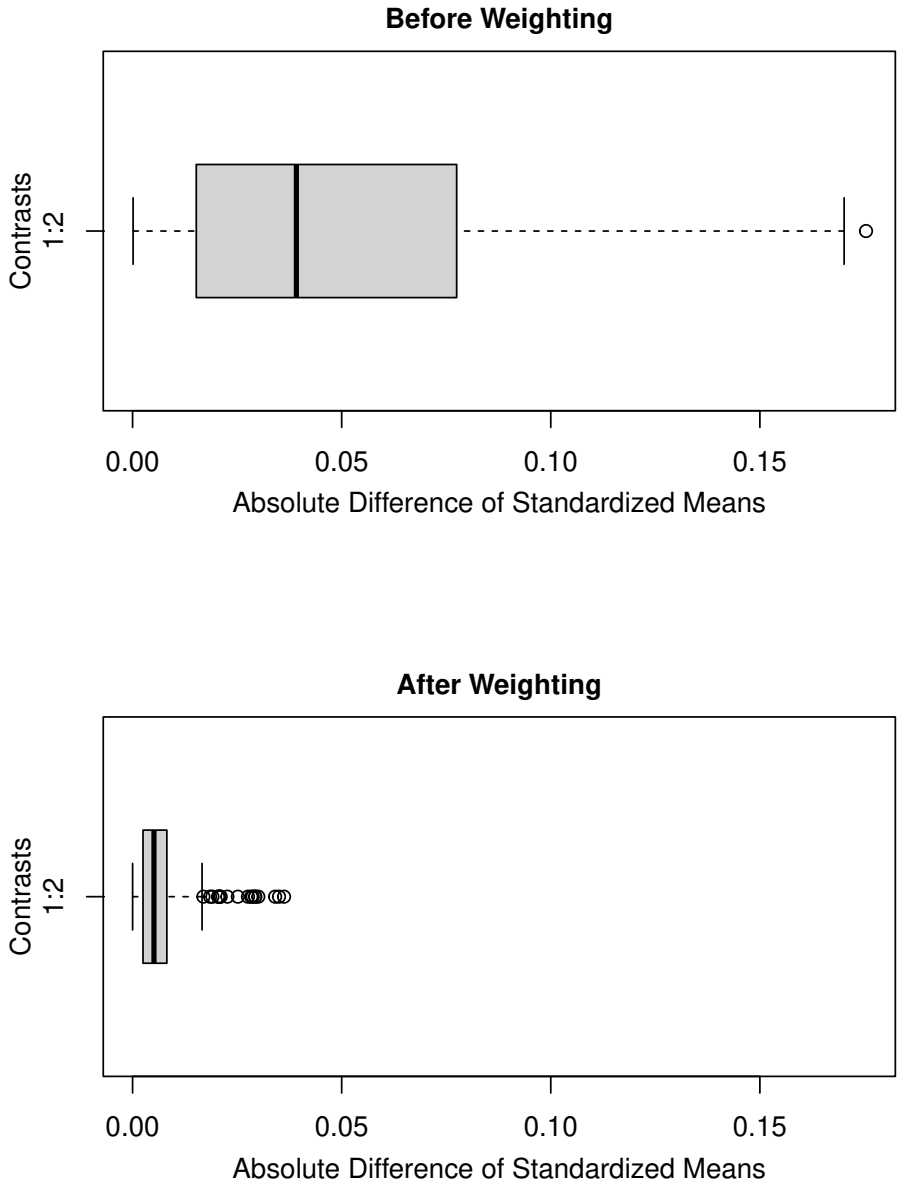
And the difference Δ in support for suffrage can be expressed as a function of the change in support among enlistees, and change in support among non-enlistees.

$$S_{i\Delta} = X_i(\beta^e_{i1} - \beta^e_{i0}) + (1 - X_i)(\beta^n_{i1} - \beta^n_{i0})$$

And, here, $S_{i\Delta} = \alpha + X_i\beta + \epsilon_i$ then returns an unbiased estimate of $\beta = (\beta^e_{i1} - \beta^e_{i0}) - (\beta^n_{i1} - \beta^n_{i0})$ if $E(\beta^e_{i1} - \beta^e_{i0})$ and $E(\beta^n_{i1} - \beta^n_{i0})$ are independent of X. However, there is no requirement that β^e_{i1}, β^e_{i0}, β^n_{i0}, or β^n_{i1} are independent of X.

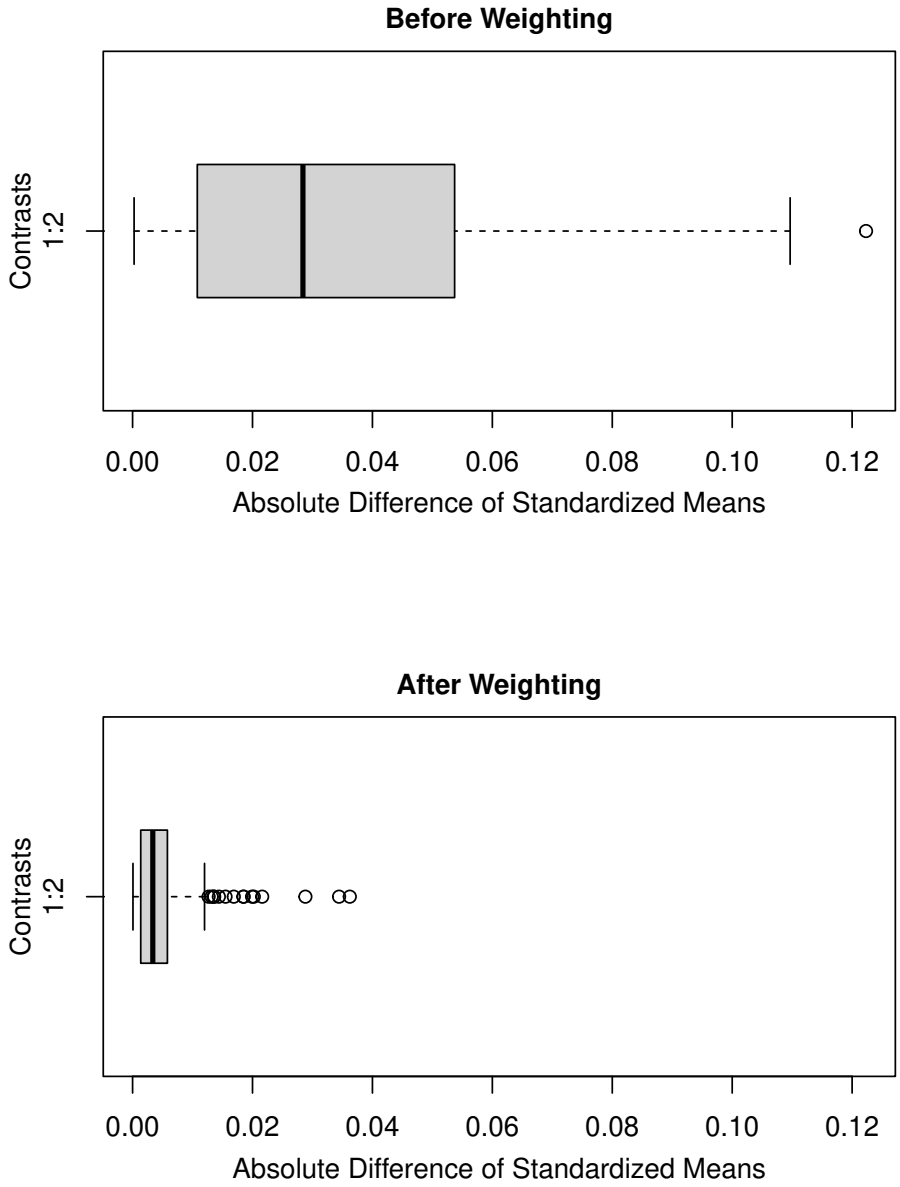
For further proof of this, the replication file includes code to simulate this process, allowing there to be contextual differences in the cross-sectional support for suffrage among enlistees and non-enlistees. The simulations show that the ecological difference-in-differences regression is an unbiased estimator if the *shifts* in support for suffrage among enlistees/non-enlistees are uniform across X or independent of X.

Figure B20: Balance between matched and un-matched soldiers before and after weighting using CBPS (option 1)



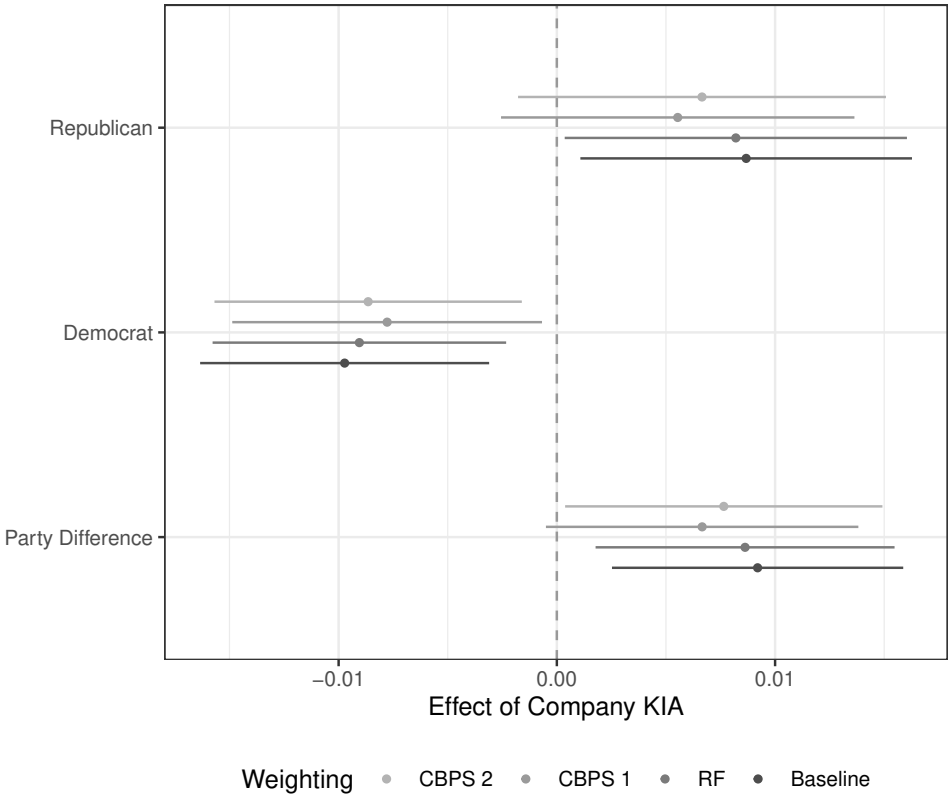
This figure shows balance between matched and unmatched soldiers before and after applying inverse probability weights generated using CBPS balancing on within- and between- regiment differences in treatment, covariates, interactions between treatment and covariates, and squared terms

Figure B21: Balance between matched and un-matched soldiers before and after weighting using CBPS (option 2)



This figure shows balance between matched and unmatched soldiers before and after applying inverse probability weights generated using CBPS balancing on differences from within-regiment means of treatment, covariates, interactions between treatment and covariates, and squared terms

Figure B22: Robustness of Effects of Company Casualties to Weighting to address Attrition



This shows a replication of the analyses in Table 2, columns 1–3, using no weights, regression forest weights, and the CBPS weights. Estimates come from the least squares regression of post-war partisanship on Company KIA, conditioning on regiment fixed effects, using different weights for MIPO. Sample includes men serving in Indiana Infantry Regiments who were matched to the 1860 Census and the 1874 People's Guides: 3087 unique soldiers, in 521 companies, across 195 regiments. Individuals are weighted by 1 over the number of matches multiplied by their inverse probability weight of being matched post-war. Standard errors are clustered by company.

C.2 Contextual Effects: Conditioning

I address the possibility that there are contextual effects, where the shifts in support for suffrage are related to X, in two ways. First, I condition on a set of covariates W that are plausibly related to X and to the shift in support for suffrage over time. If, after conditioning on W, $E(\beta^e_{i1} - \beta^e_{i0})$ and $E(\beta^n_{i1} - \beta^n_{i0})$ are independent of X, then the ecological difference-in-difference estimator is unbiased.

For the suffrage referenda in Wisconsin townships, it is possible for me to condition on county fixed effects, the difference between Republican support and suffrage support in 1857, and the change in the fraction of people eligible to vote between 1857 and 1865. If, within counties, or conditional on the gap between Republican and suffrage support, in 1857, enlistment rates are independent of the shift in support for enlistees and non-enlistees ($E(\beta^e_{i1} - \beta^e_{i0})$ and $E(\beta^n_{i1} - \beta^n_{i0})$), then the estimated slope on enlistment rates are an unbiased estimator of the difference-in-difference ATT for enlistment.[5] The results of these analyses are reported in Table C1. Even in the most restrictive specification, I find that veterans increased support for suffrage by 8.4 ppt.

C.3 Contextual Effects: Bounding

The second approach to dealing with contextual effects is to place mathematical bounds on their size, and then use that to derive the implied bounds for the ATT of service on enlistees. This approach essentially admits that we may not be able to condition away contextual effects, and instead seeks to find a region of identification within which the contextual effects must be. This is more believable, because it is plausible that there are contextual effects: the changes in support for suffrage among enlistees and non-enlistees may vary with the fraction of men who enlisted. Jiang et al. (2020) show that, if we assume that these contextual effects are linear,

$$E(\beta^e_{it}|X_i) = e_{0t} + e_{1t}X_i$$

$$E(\beta^n_{it}|X_i) = n_{0t} + n_{1t}X_i$$

then, we can partially identify the following model:

$$E(S_{it}|X_i) = n_{0t} + (e_{0t} - n_{0t} + n_{1t})X_i + (e_{1t} - n_{1t})X_i^2)$$

I apply the bounds that Jiang et al. (2020) derive to bound the ATT on enlistees' support for suffrage. Importantly, these bounds, because they use more information than the Duncan-Davis bounds, are usually narrower. Where L_{nt} and U_{nt} are the lower and upper bound for support among non-enlistees in time t, and L_{et} and U_{et} are the lower and upper bound for support among enlistees in time t, derived using the partial identification of the linear-contextual effects ecological regression proposed by (Jiang et al. 2020).

$$ATT_L = (L_{e1} - U_{e0}) - (U_{n1} - Ln0)$$

[5] Also assuming parallel trends for enlistees and non-enlistees.

Table C1: Effect of Enlistment on Support for Black Suffrage (Wisconsin Township Returns)

	Dependent variable:									
	(Yes/Elig.)		(Yes/Elig.)			(Yes/Elig.)		(Yes/Elig.)		
	Full Sample					Restricted Sample				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Enlistment Rate	0.210***	0.184***	0.191***	0.149***	0.113***	0.179***	0.147***	0.160***	0.109***	0.084**
	(0.031)	(0.030)	(0.036)	(0.032)	(0.030)	(0.032)	(0.027)	(0.037)	(0.029)	(0.030)
Lagged DV	Y	Y	N	N	N	Y	Y	N	N	N
Differenced	N	N	Y	Y	Y	N	N	Y	Y	Y
Controls	N	Y	N	Y	Y	N	Y	N	Y	Y
County FE	N	N	N	N	Y	N	N	N	N	Y
Observations	362	362	362	362	362	321	321	321	321	321

Note: *p<0.05; **p<0.01; ***p<0.001

Enlistment rate is number of men serving over those eligible to vote in 1865. Suffrage vote totals come from state constitutional referenda returns in stable clusters of townships/counties in Wisconsin. Control variables include: (i) the fraction eligible to vote in 1857 over those eligible to vote in 1865; and (ii) The fraction voting for the Republican gubernatorial candidate in 1857 minus the fraction voting for suffrage in 1857. Townships are weighted by number of white men. The restricted sample includes only townships where the population eligible to vote in 1865 changed by less than 50 percent between 1860 and 1870. Standard errors are robust.

$$ATT_U = (U_{e1} - L_{e0}) - (L_{n1} - Un0)$$

Using this approach (and excluding outliers, as the authors suggest), the bounds for the ATT on veterans in Wisconsin are $-0.396, 0.686$, which includes the estimated effect, 0.265. Figure C1 shows the size of the contextual effects needed for the true effect on veterans to be less than 0. For veterans and civilians, it plots the $t_1 - t_0$ changes in support for suffrage at the high and low end of enlistment rates that are possible given the bounds on linear contextual effects. For each point on the graph, it also uses a diverging color scale to show what the true ATT would be, for contextual effects of that magnitude.

For there to be no effect on veterans, we'd have to believe that veterans' support for suffrage in low-enlistment townships dropped by 6 ppt or more, while veterans' support in high-enlistment townships *increased* by 20 ppt. These are strongly heterogeneous effects: substantially more than the 7.5 ppt shift toward suffrage state-wide and larger in absolute magnitude than the shifts toward suffrage in more than 96 percent of townships, and 1.9 SD larger than the average overtime change within townships. If the absolute difference in over-time change in support for suffrage among veterans in high and low enlistment townships were any smaller (than 26 ppt), then the ATT must be above 0.

Another approach to addressing the issue of bounds is to look at the bounds for subsets of the data. I look at bounds for townships where pro-suffrage votes in the pre-war period were near 0. These are natural to consider, because they are likely to have the tightest bounds. This is because, when pre-war suffrage support is near 0, the upper and lower bounds of pre-war support for suffrage among enlistees and civilians must be narrowly bounded near 0. When calculating the bounds on the ATT, then, we mostly rely on the bounds from $t = 1$. Figure C2 reports the lower and upper bounds of the ATT for all towns below different thresholds of pre-war suffrage support. There is a region where pre-war suffrage support was lower than approximately 4 percent[6] for which the the *ATT* of military service on support for suffrage is bounded above 0. Since these are deterministic bounds, this approach does not amount to "*p*-hacking."

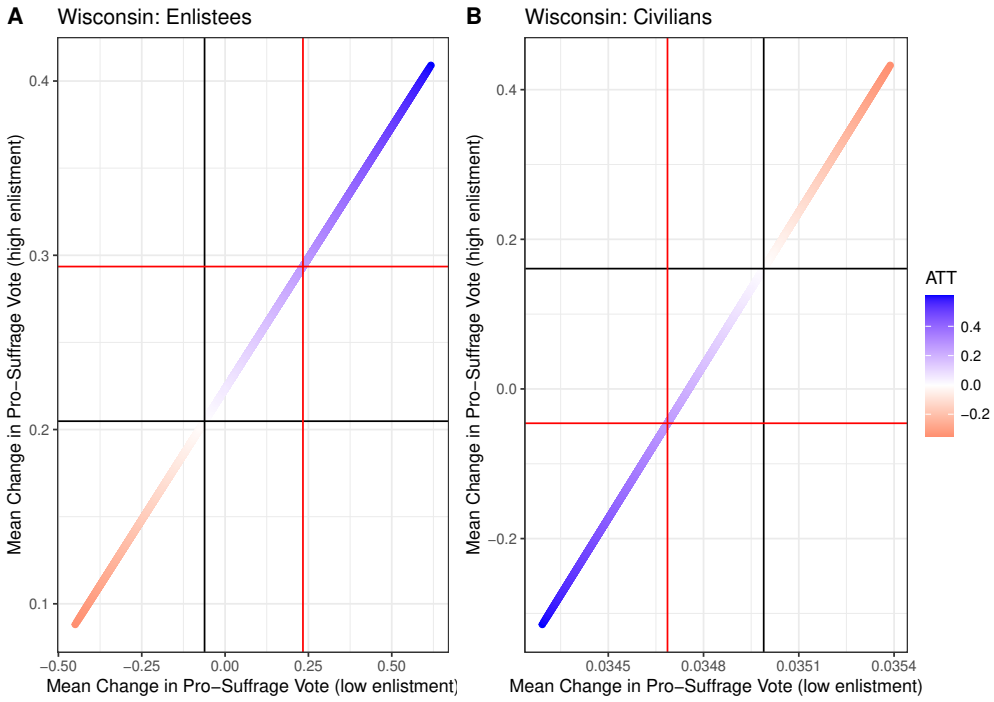
C.4 Bounds: Unobserved Confounding

Despite showing that partisan selection into service cannot explain why veterans were more likely to trend toward suffrage, it could be that there was some other form of unobserved confounding that led people who were going to trend toward suffrage anyway over the course of the war to enlist at higher rates. This would bias the ATT of enlistment on support for suffrage. The central concern is that people who were prone to become supporters of suffrage joined the army at higher rates.

To engage with this potential bias, I partially identify the effects of military service on support for suffrage in the presence of unobserved confounding, using data from the low-suffrage townships examined above. I first obtain the ecological bounds on how people who served and did not serve voted for suffrage, before and after the war. Then, I apply the bounds for difference-in-difference estimates of the ATT under unmeasured confound-

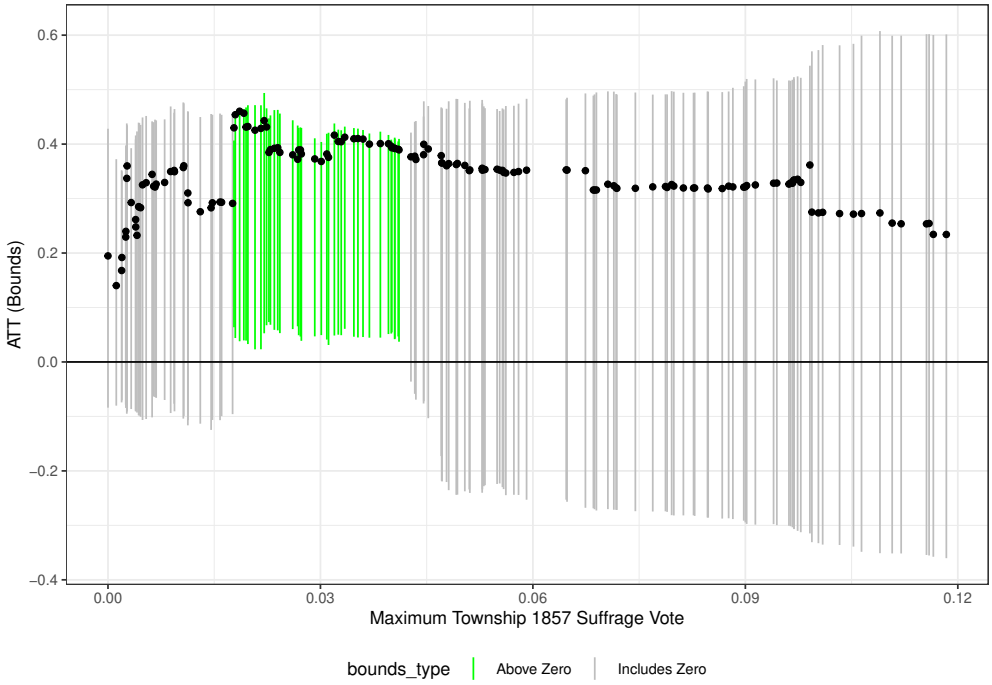
[6]This includes between 54 and 93 townships.

Figure C1: ATT of enlistment on suffrage support across possible values sizes of linear contextual effects (Wisconsin Referenda)



This figure shows the size of the contextual effects needed for the ATT on veterans to be less than 0. The x axis shows the $t_1 - t_0$ changes in support for suffrage in low enlistment townships and the y axis shows the $t_1 - t_0$ changes in support for suffrage in low enlistment townships that are possible given the bounds on linear contextual effects. The color of the dots shows the ATT associated with the contextual effect associated with the values of x and y. The black lines show the contextual effects implied for an effect size 0; the red lines show the contextual effects implied by the estimate obtained using the difference ecological regression.

Figure C2: Deterministic bounds on the ATT of military service on suffrage for townships with low pre-war support for suffrage



This figure show the Jiang et al. (2020) bounds for the ATT of military service on support for suffrage for towns below different thresholds of pre-war support for suffrage. Black dots indicate the estimated ATT from an ecological difference-in-difference.

ing developed by the (Athey and Imbens 2006, 151–3), with one minor change. I add the assumption of monotone treatment response (that is, enlistment can not hurt support for suffrage). This seems plausible given that virtually no one supported suffrage (so cannot reverse positions) and it reflects the concern that soldier's support for suffrage reflected a predisposition to become supportive of suffrage.

$$ATT = [Y(1)_1|D = 1] - [Y(0)_1|D = 1]$$

To put bounds on the ATT given unmeasured confounding, we put maximum and minimum values on what soldiers would have done after the war, even if they had not enlisted. Modifying the bounds from (Athey and Imbens 2006, 151–3) using the monotone treatment response assumption, we find that:

$$max([Y(0)_1|D = 1]) <= [Y(1)_1|D = 1]$$

and

$$min([Y(0)_1|D = 1]) >= max\{[Y(1)_0|D = 1], [Y(0)_1|D = 0]\}$$

That is to say, in the absence of enlistment, soldiers would have supported suffrage after the war at least as much as the people who stayed at home and at least as much as they did before they went to war. And they would have supported suffrage after the war no more than they actually did after serving.

The most important thing to note about these bounds is that they set the lower extreme bound of the ATT to be 0. Conversely, the upper bounds (determined by the ecological bounds discussed above) imply that the ATT was an increase in support for suffrage of between 13.4 and 36.1 percentage points. Only in the most extreme case where all changes over time among soldiers were due to selection effects does the effect of enlistment disappear.

These bounds also help us interrogate how plausible this extreme scenario would be. Overall, it seems very unlikely.

- First, we can establish the proportion of the population that would have to be "trending" toward suffrage in order for the extreme case of $ATT = 0$ to be true. In this set of cases, it would imply that whereas only 1.7 percent of people were pre-war supporters of suffrage, 10 percent were going to support suffrage anyway. This seems implausible. Support for suffrage across all of Wisconsin increased by only 7.5 points. And in this sample, this is nearly equal to the proportion of the population most like to trend toward suffrage (Republican voters who did not vote for suffrage and Democratic voters who did not vote against suffrage made up 10.8 percent of people). Moreover, ecological bounds show that, at most, 65 percent of that group could have switched to vote for suffrage.

- Second, we can also assess the plausibility of the implied magnitude of the selection bias. At the *lowest* ecological bound on the veteran voting for suffrage, for the ATT to be 0, we'd have to believe that people trending toward suffrage enlisted with a probability of 0.42, while those who were not trending toward suffrage enlisted with probability of 0.23 (1.8 times more likely). At the *highest* ecological bound on veteran

voting, we'd have to believe that those trending toward suffrage enlisted with probability of 0.85 while those who were not enlisted with probability of 0.18 (4.7 times more likely). For comparison, the *maximum possible* ratio of Republican enlistment rates to Democratic enlistment rates implied by ecological bounds was 3.1 (the lowest possible ratio was 0.38).

Thus, in order to claim that confounding entirely explains the effect of service on soldier's support for suffrage, the most plausible scenario still requires that we assume that:

- veteran support for suffrage after the war was actually at the lowest possible value implied by ecological bounds
- that this was driven by some unobserved and unknown confounding that was larger than the most plausible source of confounding (Republicans not supporting suffrage trending toward support over time)
- that differences in selection into service by party were as large as possible
- that this unknown confounder induced differential selection into service almost as large as partisanship.

Overall, without a plausible account of what this confounder might be, these worries seem unreasonable.

C.5 Interpretation

In the paper, I note that it could be that something other than military service is driving the different changes in support for veterans and non-veterans. Perhaps selection based on Republican Partisanship into military service makes veterans more susceptible to changing views on suffrage as the party changes its position. To address this, I estimate the difference-in-differences effect of enlistment on suffrage, conditioning on 1857 Republican support for townships with very low support for Republicans and Suffrage in 1857. Figure C3 shows the results of these tests across a range of thresholds. Even where there is no possibility of selection based on partisanship or support for suffrage, we find a strong positive effect of enlistment on suffrage support after the war.

C.6 Heterogeneous Effects

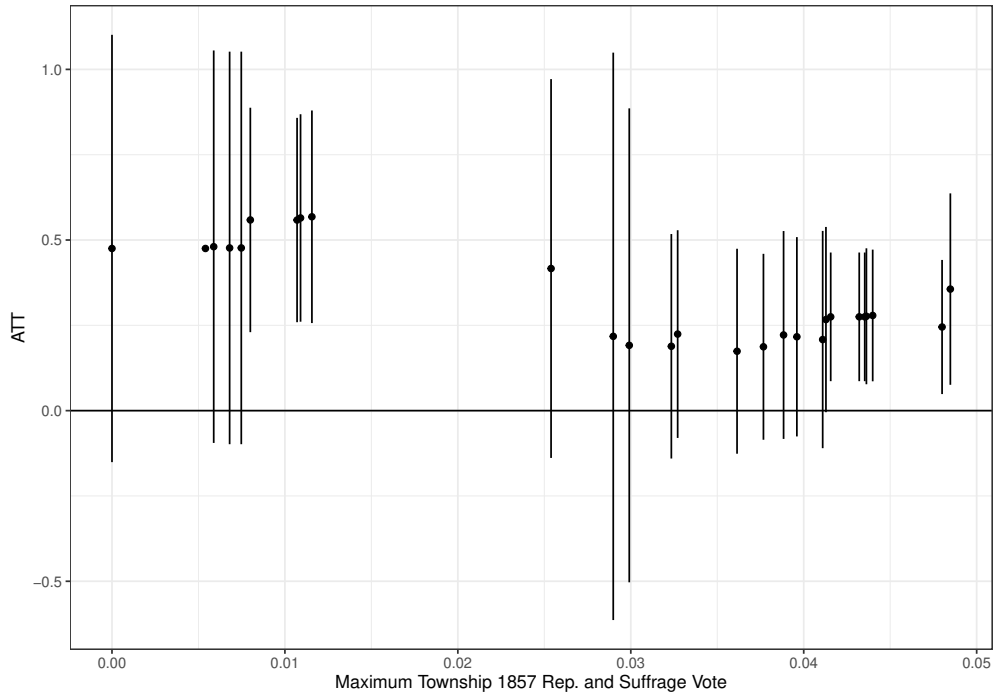
C.7 Effects of Legislators

D Data

D.1 Demographic and Economic Data

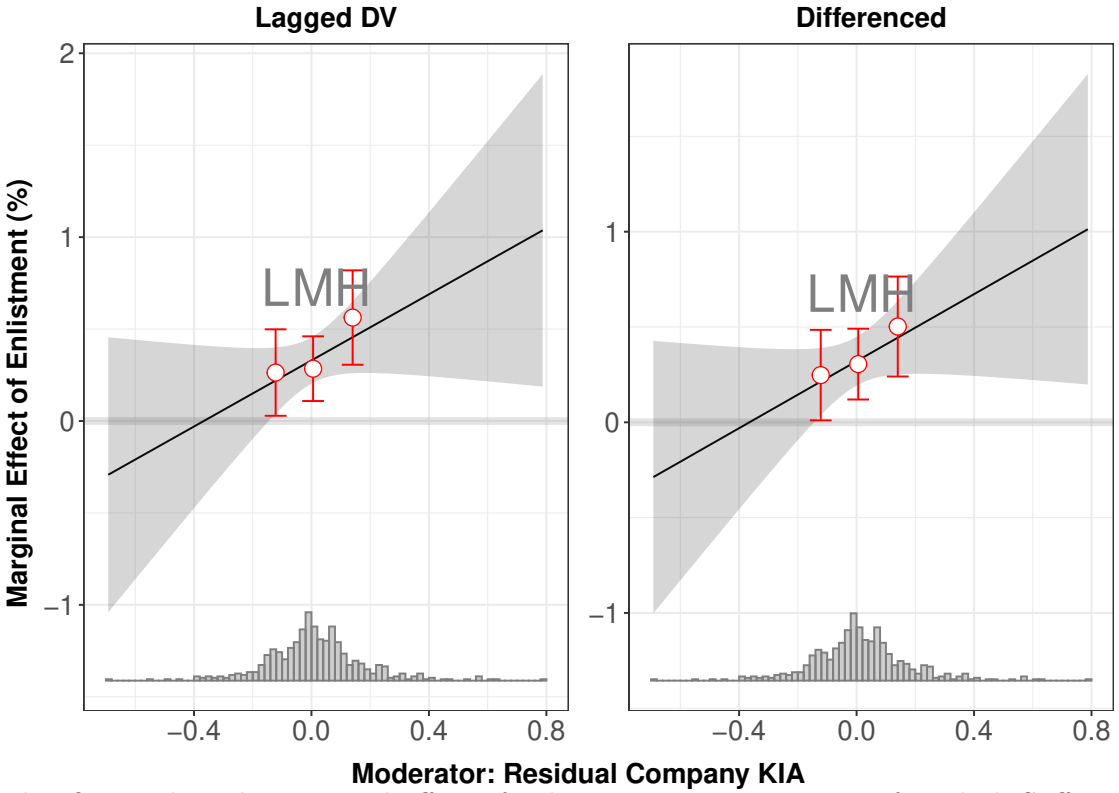
Data on economic and demographic attributes come from the United States census county-level reports available on the ICPSR. Using data from 1860, I calculate the following characteristics of counties:

Figure C3: Difference-in-Difference effects of Enlistment on Support for Suffrage in Townships with Low Republic/Low Suffrage Support (Wisconsin)



This figure shows the difference-in-differences effect of enlistment on suffrage for townships with pre-war support for Republicans and Suffrage less a given threshold. These models also include a control for pre-war support for Republicans. Standard errors are the maximum of homoskedastic or HC3 variance estimates.

Figure C4: Marginal Effect of Enlistment Rates on Votes for Black Suffrage Conditional on Company Casualty Rates



This figure plots the marginal effect of enlistment rates on support for Black Suffrage across mean experience of combat casualties in the Iowa and Wisconsin state constitutional referenda in 1857, 1865, and 1868. All models include state dummies. Enlistment rate is number of surviving men who enlisted over those eligible to vote in 1865. Residual Company KIA is the standardized difference between combat deaths in a soldier's company and the rest of his regiment, averaged by town. Suffrage vote totals come from state constitutional referenda returns in 545 stable clusters of townships/counties in Wisconsin and Iowa. Townships are weighted by number of white men. Standard errors are robust.

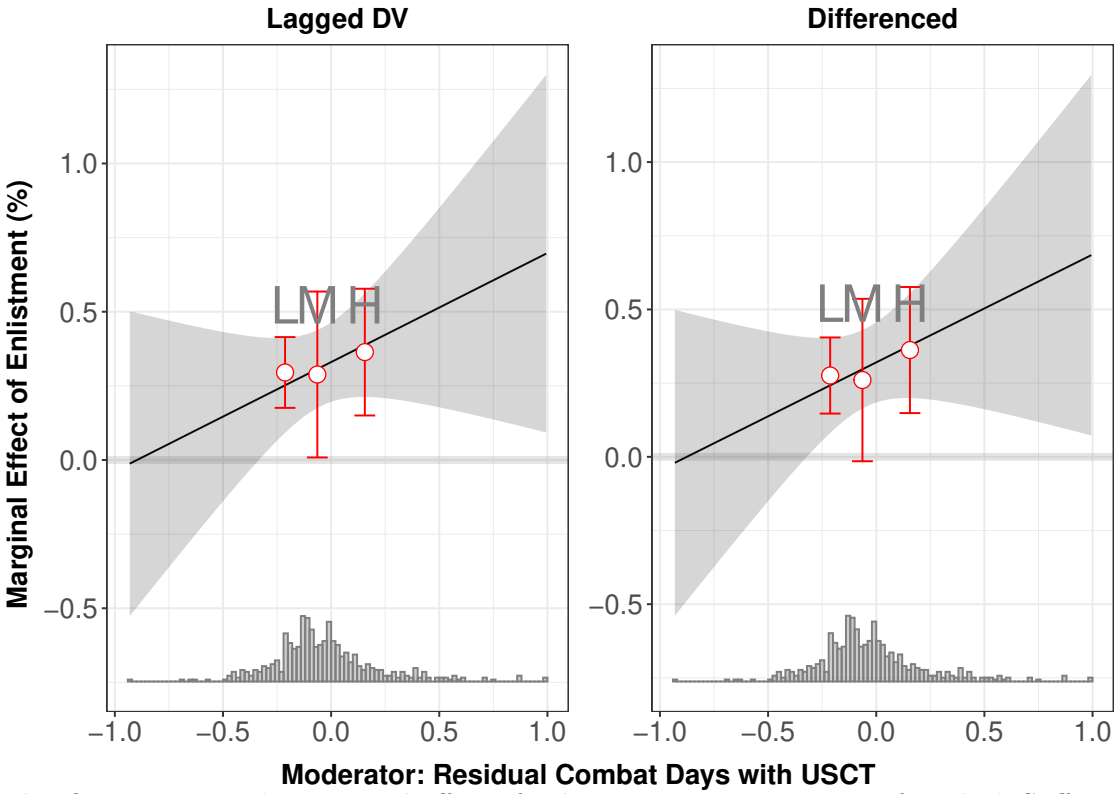
Table C2: Predictors of Iowa Republican Support for Removing Racial Qualifications

	Dependent variable:			
	Universal Suffrage		Full Equality	
	(1)	(2)	(3)	(4)
Enlistment Rate	6.292*	6.686*	3.183	5.132
	(3.446)	(3.760)	(3.297)	(3.772)
Pro-Suffrage VS (1857)		4.121		7.491**
		(3.655)		(3.733)
Republican VS (1857)		−0.985		0.034
		(3.576)		(4.167)
Observations	66	66	67	67

Note: *p<0.1; **p<0.05; ***p<0.01

Estimates from logit regression. Observations are Iowa General Assembly constituencies held by Republicans in 1866. "Universal Suffrage'' indicates a vote against amending the proposed constitutional amendment to grant limited suffrage to African Americans and then voting for suffrage extension. "Full Equality'' further indicates voting to eliminate all racial qualifications in the Constitution. Enlistment rate is number of men serving over those of military age in 1860. Suffrage voteshares come from state constitutional referenda returns in 1857. Republican voteshares come from the gubernatorial election of 1857.

Figure C5: Marginal Effect of Enlistment Rates on Votes for Black Suffrage Conditional on Days of Combat alongside USCT Regiments

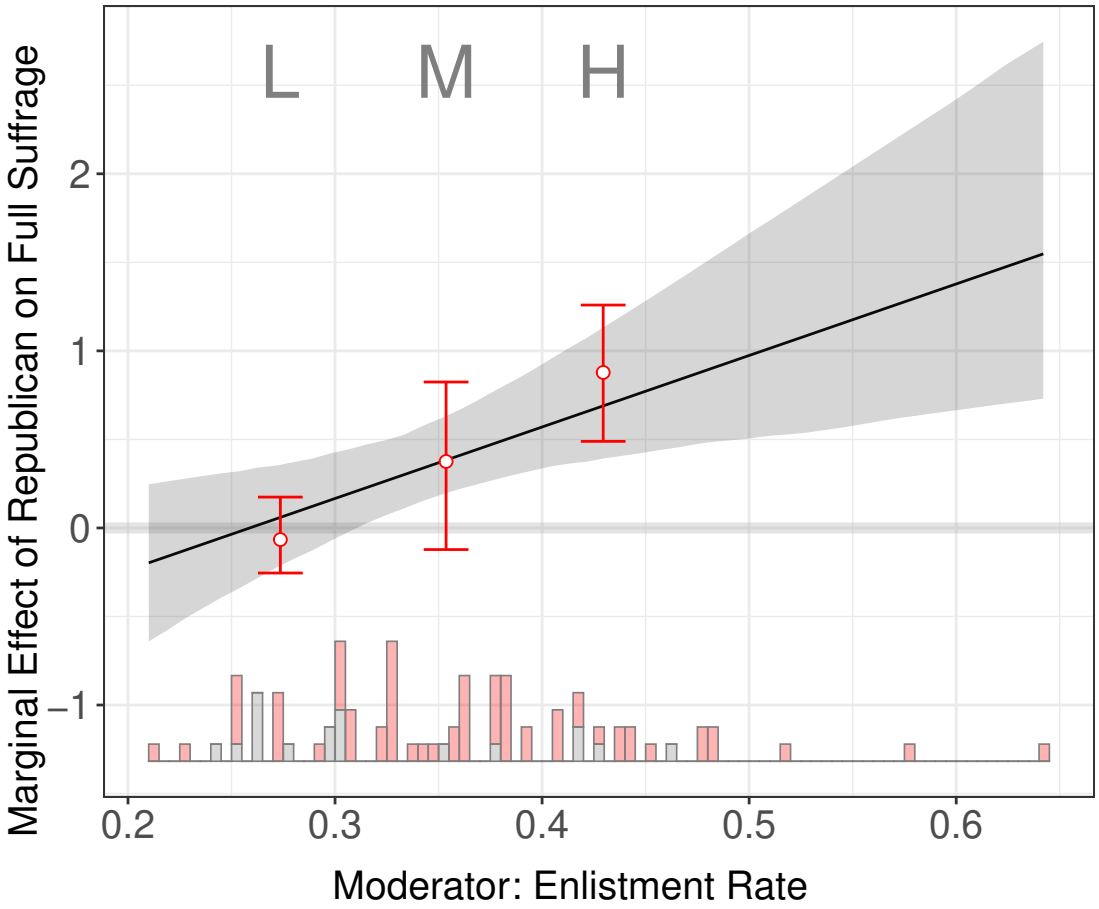


This figure reports the marginal effect of enlistment rates on support for Black Suffrage across mean experience of combat alongside USCT regiments in the Iowa and Wisconsin state constitutional referenda in 1857, 1865, and 1868.. All models include state dummies. Enlistment rate is number of surviving men who enlisted over those eligible to vote in 1865. Residual USCT Combat is the standardized difference between days a soldier spent fighting alongside USCT units compared to other soldiers who enlisted into the same branch of service and term of enlistment in the same year, averaged by town. Suffrage vote totals come from state constitutional referenda returns in 545 stable clusters of townships/counties in Wisconsin and Iowa. Townships are weighted by number of white men. Standard errors are robust.

Demographic: Logged total population, fraction of the population that is military-aged males, the ratio of military age males to females, fraction of the population belonging to religious denominations associated with abolitionism, fraction of females in the manufacturing jobs, fraction of men in manufacturing jobs, fraction of the population that was foreign born, fraction of the population that is black, fraction of the population that lives in an urban area, fraction of the population that is white, fraction of the population that was born in the South.

Economic: logged manufacturing output, logged manufacturing output per capita, logged agricultural output, log agricultural output per capita, logged agricultural output per acre of a improved land, mean farm acreage, farm value per acre, agricultural property gini coefficient, military-aged males per acre of improved land, per capita agricultural output.

Figure C6: Marginal Effect of Republican Partisanship on Legislator Support for Full Suffrage Conditional on Enlistment Rates



This figure plots the marginal effect of Republican partisanship of Iowa state legislators on their support for full African American suffrage, conditional on the enlistment rate in their constituency. Republicans in low-enlistment constituencies were more similar to Democrats, while Republicans in high-enlistment constituencies were significantly more likely to support equal suffrage for African Americans.

D.2 Election Data

Township election returns for Iowa came from Dykstra (1993) and are publicly available at (Dykstra 2005). Township election returns for Wisconsin come from (McManus 1998), but the data are not public and were provided by the author.

Because townships changed boundaries over time and election returns were not always available for every township in both elections, I created clusters townships that were stable in both geographic composition and reporting of election returns over time.

To construct these, I used the following procedure:

1. Identify a list of all townships in existence between 1857 and 1868 in Iowa and Wisconsin, using the 1870 Census reports.

2. Identify township boundary and name changes from the same.

3. Create clusters of townships which shared names over time.

4. Link election returns from Dykstra and McManus to townships.

5. For township clusters that have complete information on election returns in 1857 and 1865, calculate vote totals by cluster.

6. Subtract these the sum of vote totals across all complete township clusters from county returns.

7. Create new clusters out of the "remainder'' of each county.

D.3 Enlistment Data

I obtain enlistment data by county using the ACWRD except for the case of Indiana. The data from the ACWRD was matched to counties using the following procedures.

- First, I examined the ACWRD to find which states contained residences (not muster locations) for more than 90% of soldiers serving in their regiments. I elected to ignore muster locations as a source of data on residence they are not guaranteed to take place in every county and may have been in more urban areas as well as in areas where there was greater Republican loyalty.

- I then geocoded each residence using online geolocation tools. I first passed each residence to Geolocate, which includes a large number of US place names, and then to Google Maps.

- I then assign men to counties of residence by the 1860 county boundaries in which their georeferenced point is located.

- Finally, I take the enlistment rate to be the number of men counted as serving as a fraction of white men aged 10 to 39 in 1860. These are men that could plausibly have served during the war. Men below 10 in 1860 or over 39 were very unlikely to serve due to age restrictions.

The biggest potential source of bias comes from geocoding. There are likely errors induced by Geolocate or Google Maps excluding or misplacing historical names. The scope of this bias is likely limited. First, for towns and villages that no longer exist or changed names, there are often other features, like roads, in their vicinity that still bear their name. Google Maps frequently picks up these road and other features names as the best match. Second, I am able to test this geocoding technique against "ground truth'' data for Illinois. The Illinois Adjutant General (responsible for oversight of Illinois regiments during the war) produced a report indicating the number of enlisted men "credited'' to each county. While credits sometimes were mis-allocated and some credits reflected local units enlisting men from other states while in the field, this administrative data provides a good check for the quality of the geocoding. The number of soldiers from each Illinois counties, based on the administrative

and geocoded date have a correlation of 0.989. The fraction of eligible 1860 males who enlisted, based on these two sources of data, correlate at 0.83. This lends plausibility to using the geocoding procedure for other states.

Finally, I obtain data on county-level enlistment in Indiana through a different source. The Indiana Adjutant General reported enlistments and draftees by county for all periods of enlistment in the war exclude for a short interval between October 1862 and the Spring of 1863. This brief period excludes less than 10% of all enlistments in Indiana during the war, so it has about as good coverage as the geocoding approach.

D.3.1 Township Enlistment Rates

For Indiana and Wisconsin, I obtain enlistment rates for township clusters using the following procedure.

1. I created a list of all townships that existed in the states between their creation and 1870, using the 1870 Census.
2. I linked these townships into clusters connected by boundary changes between 1857 and 1868.
3. I combined these clusters of townships into larger units, if election results were not available for the township cluster in either 1857 or 1865.
4. I then created a list of unique enlistment places for soldiers in ACWRD.
5. For each enlistment place, research assistants matched that place to a county using lists of census townships/villages, records of old post office names in the GNIS database.
6. For each place, research assistants then matched it to township using 1875 maps of each county.
7. Because township names were often repeated in multiple counties, or people listed an entire county as their place of residence, I linked soldiers to each township election cluster they might possibly be in.

After linking soldiers to possible township election clusters, I then used the fastLink algorithm to link them to individuals.

1. Blocking on township-election clusters, I matched soldiers to people residing in those places in 1860 on first name, last name, cleaned first name, sound codes, and age.
2. Using matches with a probability threshold above 0.85, I assigned soldiers to township clusters using the following rules:
    - If a soldier's only matches were within a single township cluster, assign him to that township cluster.
    - If a soldier matched to multiple township clusters, assign his weight to each cluster in proportion to the probability score for his matches in those clusters.

- If a soldiers was not matched to the 1860 census, but his place of residence uniquely identifies a single township cluster, assign him to that cluster.
- If a soldier was not matched to the 1860 census, and his place of residence links him to more than one possible township cluster, assign his weight to each cluster in proportion to the military aged male population of each township cluster.

D.4 Wartime Experiences

I measure wartime experiences using a few different sources. First, I use the CWDB to identify experiences of combat and casualty rates. The database makes it possible to identify the number of combat injuries or fatalities within a unit on a given day, which makes it possible to count days of combat. It is also possible, because the CWDB includes data on who died (though this suffers from extensive undercounting), to calculate casualty rates for each unit during the war. Second, I also use CWDB to identify which regiments shared brigade duty at various times during the war. Third, the CWDB identifies officers within units and unit membership of individuals, making it possible to measure wartime organizational networks.

Using this data, I created the following measures

- *Exposure to African American Soldiers* I measure exposure to African Americans in the military based on a soldier's proximity African American units — either *(a)* Number of days in which a soldier's regiment saw combat in the same place and time as USCT (African American) regiments. *(b)* Number of day a regiment belonged to the same brigade as African American regiments. Brigades were larger military formations that coordinated the logistics and combat of groups of 2 to 5 regiments and the smallest mixed race units in the Union Army.
- *Wartime Sacrifice*: This is measured as *(a)* regiment-level combat experience (in days of combat) and *(b)* regiment-level casualty rates (combat and non-combat deaths).

D.5 Newspaper Coverage

To characterize the nature of Republican messaging on Reconstruction-related issues, I coded the coverage of the November 1865 Wisconsin suffrage referendum in 28 newspapers. I tracked down all Wisconsin newspapers that have digitized issues for the month between October 3 and November 7 1865. This amounts to 168 newspaper issues. I then had an RA read through each issue and identify articles that discussed African American suffrage generally and/or the upcoming referendum in particular. For each article, the RA classified whether the content was pro- or anti-suffrage in nature. Then, they recorded which arguments were used to support or oppose suffrage, as well as some other attributes. While this is not a random sample, it nevertheless covers 20 different Republican newspapers across 20 different counties (there were 60 counties in Wisconsin at this time). The list of Republican newspapers can be found below.

The presence of various pro-suffrage arguments in newspapers were coded using binary indicators, as follows:

Table D1:

publication	county
The Wisconsin State Register	Columbia
Janesville Weekly Gazette	Rock
The Appleton Motor	Outagamie
The Racine Advocate	Racine
Vernon County Censor	Vernon
Grant County Witness	Grant
Waukesha Daily Freeman	Waukesha
The Wisconsin Lumberman	Portage
Racine Journal	Racine
Mineral Point Weekly Tribune	Iowa
The Saturday Evening Press	Winnebago
Dodgeville Chronicle	Iowa
Waupun Times	Dodge;Fond Du Lac
Wisconsin State Journal	Dane
Lafayette County Union	Lafayette
Whitewater Register	Walworth;Jefferson
Door County Advocate	Door
Oconto Pioneer	Oconto
Shawano County Journal	Shawano
Eau Claire Daily Free Press	Eau Claire

- African American people loyal in the war and earned citizenship/voting rights
- African Americans soldiers loyal in the war and earned citizenship/voting rights
- African American voters needed to ensure end to slavery
- African American voters needed to ensure victories of the war remain won
- Former soldiers/officers endorse suffrage
- Need black voters to prevent return of rebels to power
- Need loyal voters in the South
- Opposition to suffrage by "Copperheads" or "traitors"
- White southerners/confederates have unfair advantage in representation
- Punish south for Rebellion

The presence of various anti-suffrage arguments in newspapers were coded using binary indicators, as follows:

- Suffrage leads to "social equality" or mixing of races
- Suffrage will establish "black supremacy"
- Suffrage tramples on the rights of white men
- Suffrage is insincere partisan ploy
- Suffrage/civil rights not the purpose of the war (only Union)
- Suffrage the work of "radicals" or "abolitionists" in pejorative sense
- Suffrage will alienate white voters from Republican party
- Suffrage will antagonize white Southerners
- Former soldiers/officers oppose suffrage

Other features of suffrage content was coded using binary indicators, as follows:

- Claim that loyal Republicans/Union party members must support suffrage
- Republicans/Union party members attacking members of their own party for NOT supporting suffrage
- Link voting for suffrage to fighting the rebels
- Republicans/Union party members opposing suffrage
- Republicans/Union party members personally support suffrage
- Demand for Republican/Union party members to state a position on suffrage

E Calculations

E.1 Veteran Share of Electorate

I start with the number of men who served in the Union Army. Army records suggest there were approximately 2.1 million men serving (this accounts for 2.6 million enlistments, subtracting approximately 500,000 reenlistments). I assume that all of surviving men were born citizens or had become citizens by 1870. Of these, 180,000 were African Americans and another 55,000 were from Confederate states. Subtracting these gives a total of 1.865 million men. Deaths due to combat and disease in the Union Army were approximately 360,000.[7] Assuming that the fatality rate is the same for Northern white troops as for Southern whites and the USCT, this suggests that 17.1 percent of Northern white soldiers died.[8] This gives approximately 1.545 million veterans after the war. Census records in 1870 show that there were 6.465 million naturalized white men in Northern states over the age of 21. This gives an estimate that 23.9 percent of the eligible voting population in states outside the Confederacy were Union veterans. In reducing soldier numbers due to re-enlistment, accounting for USCT and Southern Union soldiers, and calculating the distribution of deaths, I chose numbers that would deflate the overall count of white veterans in the North.

E.2 Indiana Effect Size

To give a more intuitive sense of the magnitude of the individual level effects of company casualties, I make a back-of-the-envelope calculation of how changes in exposure to company casualties would have affected the outcome of the 1874 Congressional Elections in Indiana. I do not claim that these calculations are valid counterfactuals. In actuality, Democrats won 8 out of 13 seats, while Republicans held 5.

I first approximate the fraction of the Indiana electorate that was a Union veteran. Using the 1870 Census, I find there are 380636 men in Indiana who were male citizens over the age of 21. Using the ACWRD, I find there were 149 thousand Indiana soldiers who are definitely recorded as having survived the war. This makes soldiers approximately 39 percent of the electorate.

For all Indiana soldiers, I then estimate the average partisan swing that would occur (using the estimates from Table 2, column (3) if I either (i) increased the combat casualties they experienced by one with-in regiment SD or to the maximum observed casualties in the regiment, choosing the smaller of the two; or (ii) decreased combat casualties they experienced by one with-in regiment SD or to the minimum observed casualties in the regiment, choosing the larger of the two. Increasing company casualties in this manner yields a 3.9 point swing

[7]Hacker (2011) estimates higher overall Civil War mortality, but his estimates for people born in Northern and Border states (and thus might be counted as in the "Union") is virtually identical.

[8]If anything, death-rates were much higher for African American regiments.

in the margin toward Republicans. Decreasing company casualties in this manner yields a 3.75 swing in the margin toward Democrats.

I then assume that this average affect can be extrapolated to all Indiana veterans in this manner, that all veterans remained living in the state, that they voted at the same rate as non-veterans, and comprised the same share of the electorate in each Congressional District. These are unreasonable assumptions about an actual counterfactual calculation, but the aim here is to give a sense of the effect size.

Using these assumptions, the increase in company casualties would have generated a 1.5 ppt increase in the overall margin for Republicans. Republicans lost Indiana's first Congressional district in 1874 by a margin of 1.3 ppt. This would have yielded 7 seats for Democrats and 6 seats for Republicans.

Using these assumptions, the decrease in company casualties would have generated a 1.47 ppt increase in the overall margin for Democrats. Republicans won Indiana's eight and thirteenth Congressional district in 1874 by smaller margins. This would have yielded 10 seats for Democrats and 3 seats for Republicans.

And in fact, several other elections had fairly close margins. Slightly larger changes in the company casualty rates would perhaps have given Republicans a majority or only one seat.

Appendix Sources

Abramitzky, Ram, Leah Boustan, Katherine Eriksson, James Feigenbaum and Santiago Perez. 2019. AUTOMATED LINKING OF HISTORICAL DATA. NBER Working Paper 25825 Cambridge, MA: .

Angrist, Joshua D and Alan Krueger. 1999. Empirical Strategies in Labor Economics. In *Handbook of Labor Economics*. Vol. 3.

Athey, Susan and Guido Imbens. 2006. "Identification and Inference in Nonlinear Difference-in-Differences Models." *Econometrica* 74(2):431–497.

Athey, Susan, Julie Tibshirani and Stefan Wager. 2018. "Generalized Random Forests." *eprint arXiv:1610.01271* .

Costa, Dora and Matthew Kahn. 2008. *Heroes and Cowards: The Social Face of War*. Princeton, NJ: Princeton University Press.

DeCanio, Samuel. 2007. "Religion and Nineteenth-Century Voting Behavior: A New Look at Some Old Data." *Journal of Politics* 69(2):339–350.

DeCanio, Samuel and Corwin Smidt. 2013. "Prelude to Populism: Mass Electoral Support for the Grange and Greenback Parties." *Party Politics* 19(5):798–820.

Dunning, Thad. 2012. *Natural Experiments in the Social Sciences: A Design-Based Approach*. New York, NY: Cambridge University Press.

Dykstra, Robert. 1993. *Bright Radical Star: Black Freedom and White Supremacy on the Hawkeye Frontier*. Cambridge, MA: Harvard University Press.

Dykstra, Robert. 2005. *White Attitudes Toward Black Civil Equality in the Nineteenth Century: Iowa's Equal Rights Referenda of 1857, 1868, and 1880 (ICPSR 4284)*. Inter-University Consortium for Political and Social Research.

Enamorado, Ted, Benjamin Fifield and Kosuke Imai. 2019. "Using a Probabilistic Model to

Assist Merging of Large-ScaleAdministrative Records.'' *American Political Science Review* 113(2):353–371.

Gerber, Alan and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*. New York, NY: W. W. Norton & Company.

Hacker, J David. 2011. "A Census-Based Count of the Civil War Dead.'' *Civil War History* 57(4):307–348.

Hess, Earl J. 2015. *Civil War Infantry Tactics: Training, Combat, and Small-Unit Effectiveness*. Baton Rouge, LA: Louisiana State University Press.

Imai, Kosuke and Marc Ratkovic. 2014. "Covariate Balancing Propensity Score.'' *Journal of the Royal Statistical Society* 76(1):243–263.

Jiang, Wenxin, Gary King, Allen Schmaltz and Martin Tanner. 2020. "Ecological Regression with Partial Identification.'' *Political Analysis* 28(1):65–86.

McPherson, James. 1997. *For Cause and Comrade: Why Men Fought in the Civil War*. New York, NY: Oxford University Press.