Wealth, Slave Ownership, and Fighting for the Confederacy: An Empirical Study of the American Civil War*

Andrew B. Hall
Stanford University

Connor Huff
Harvard University

Shiro Kuriwaki
Harvard University

February 10, 2018

Abstract
How did personal wealth affect the likelihood southerners fought for the Confederate Army in the American Civil War? We offer competing accounts for how we should expect individual wealth, in the form of land, and atrociously, in slaves, to affect white men’s decisions to join the Confederate Army. We assemble a dataset on roughly 3.9 million white citizens in Confederate states, and we show that slaveowners were more likely to fight in the Confederate Army than non-slaveowners. To see if these links are causal, we exploit a randomized land lottery in 19th-century Georgia. Households of lottery winners owned more slaves in 1850 and were more likely to have sons who fought in the Confederate Army than were households who did not win the lottery. Our results suggest that for wealthy southerners, the stakes associated with the conflict’s threat to end the institution of slavery overrode the incentives to free-ride and to avoid paying the costs of war.

*Authors contributed equally and are listed in alphabetical order. Andrew B. Hall (andrewbenjaminhall@gmail.com, http://www.andrewbenjaminhall.com) is an Assistant Professor in the Department of Political Science at Stanford University. Connor Huff (cdezzanihuff@fas.harvard.edu, http://connordhuff.com) is a Ph.D. Candidate in the Department of Government at Harvard University. Shiro Kuriwaki (kuriwaki@g.harvard.edu, http://www.shirokuriwaki.com) is a Ph.D. Candidate in the Department of Government at Harvard University. For research assistance, we thank Nishant Karandikar, Mikhail Koiganov, and Judy Pintor. We are grateful to Ran Abramitzky, Lisa Blaydes, David Broockman, Jennifer Eggert, James Feigenbaum, Jeffry Frieden, Vicky Fouka, Steve Haber, Federica Izzo, Tyler Jost, Sergiy Kudelia, Christopher Lucas, Shom Mazumder, Christoph Mikulaschek, Ian Morris, Jon Rogowski, Robert Schub, Jaume Sempere, Ken Shotts, Pavi Suri, Monica Duffy Toft, participants at the Harvard Experimental Political Science Graduate Student Conference, MPSA 2017, the Political Violence Workshop at Harvard University, the Harvard-MIT-Tufts-Yale Political Violence Conference, the Stanford, CIDE, and Colmex joint conference in Mexico City, the NYU Politics and History conference, and the LSE Historical Political Economy conference for helpful comments and suggestions.
1 Introduction

Civil wars are pervading features of human society, despite their profound costs. Between World War II and the new millennium alone, there were over 70 civil wars resulting in more than 16 million deaths worldwide (Fearon and Laitin 2003). These conflicts take lives, destroy property, and prevent the success of stable governments. Underlying the macro-level phenomena of civil wars are the individual decisions of millions of people to participate in these violent conflicts.¹ What leads someone to abandon the political process and take up arms against the state, risking personal life, property, and security for uncertain gains? In this paper we study this question in the context of the American Civil War, among the most destructive civil wars ever fought, and “the most horrific war in United States history” (Costa and Kahn 2003, 520). Motivated by historical research on this defining period in America’s development, paired with insights from conflict studies about why individuals participate in rebellions, we focus on how personal wealth and slave ownership affected the likelihood that individuals fought for the Confederate Army. Were wealthier people more or less likely to fight in the Confederate Army?

Research drawn from political science and history offers countervailing predictions. On the one hand, one of the most famous historical sayings about the American Civil War was that it was “a rich man’s war, but a poor man’s fight.”² The saying captures the idea that poorer white southern men, most of whom did not own slaves, were more likely to fight in the Confederate Army than their wealthier slaveowning peers, and it is consistent with research in political science that argues that individuals participate in conflict in part because they gain greater material benefits from fighting than from not fighting (Berman et al. 2011; Collier and Hoeffler 2004; Dasgupta, Gawande, and Kapur 2017; Dube and Vargas 2013; Fearon and Laitin 2003; Humphreys and Weinstein 2008; Miguel, Satyanath, and Sergenti 2004; Olson 1965). By this logic, we might expect wealthier

¹ For recent work on individual decision-making in violent conflict, see for example Getmansky and Zeitzoff (2014); Hazlett (2013); Rozenas, Schutte, and Zhukov (n.d.); Weinstein (2006).
² The origins of this saying are unknown, but it is generally thought to refer to wealthy southerners—particularly plantation owners—who agitated for rebellion yet avoided military service (Wallenstein 1984).
southern men to be less likely to participate in the conflict, both because their wealth raises the
opportunity costs to fighting and because of the potential diminishing marginal utility of money.\(^3\)

On the other hand, prior historical work shows that free, white men of even modest means
throughout the Antebellum South often invested their excess capital in land and slaves (e.g.,
McPherson 2003; Wright 1978). Because the American Civil War was fundamentally fought over
the institution of slavery, we might expect that, as some poor white farmers throughout the Antebel-
lum South became wealthier and became slaveowners, their incentives to preserve the Confederacy’s
slavery system likewise rose—perhaps making them more willing to fight for the Confederacy. This
would also be consistent with research throughout political science highlighting how individuals are
motivated to fight due to grievances against the state (Cederman, Gleditsch, and Buhaug 2013;
Gurr 1970; Humphreys and Weinstein 2008; Paige 1978)—in this case, grievances against a federal
government they saw as threatening an institution that they had been socialized into and upon
which their future livelihood depended.

Empirical historical work that attempts to resolve this debate comes to conflicting conclusions.
Studying Georgia, Harris (1998, 153) writes: “When men of the same age and family status (such as
household head, son, or boarder) are compared, those who did not serve in the army were wealthier,
and owned more slaves, than those who did serve.” Studying Mississippi, Logue (1993) shows that
Confederate combatants were significantly poorer than non-combatants, on average (though the
paper also shows additional evidence in the other direction). Studying Harrison County, Texas,
Campbell (2000) finds that wealthier individuals were more likely to fight for the Confederacy.
Reviewing some of this literature, Logue (1993) describes these varying conclusions and discusses
“obvious problems of scale in investigating enlistments” (Logue 1993, 612). Limited to manual
inspection, scholars have almost exclusively studied small samples of Confederate soldiers, often in

\(^3\) Throughout this paper we use a relatively low bar to define wealthy and are not simply referring
to the small strata of ultra-rich plantation owners. Doing so is in line with prior research
exploring how modest increases of wealth among even the poorest individuals affects their
propensity to fight (Berman et al. 2011; Dasgupta, Gawande, and Kapur 2017).
a single county (or, in the case of Logue (1993), a single state). The difficulties of working with samples probably explains why empirical work in this area has proved inconclusive.

To solve this problem, we take advantage of recently digitized datasets on the entire population of the Confederacy, rather than a particular sample. We assemble individual-level data on roughly 3.9 million free citizens in the Confederate states alive prior to the outbreak of the Civil War. Our data records information about each citizen’s wealth, the number of slaves owned, occupation, family relationships, and, for men, an estimate of whether or not each fought in the Confederate Army. Using this data, we show that households that did not own slaves fielded fewer Confederate Army soldiers, on average, than did households with slaves. Households reporting no real-estate wealth in the 1850 census likewise fielded fewer Confederate soldiers, on average. To understand these patterns—and, in particular, to gauge whether wealth and slave ownership drove people to fight for the Confederacy, or was simply a correlate of other attributes that made people more likely to fight—we present experimental evidence for the causal effect of economic wealth on the propensity to rebel. Using the results of Georgia’s 1832 land lottery, which formally randomized a meaningful amount of wealth across white male citizens (Bleakley and Ferrie 2016; Williams 1989), we show that lottery winners’ households owned more slaves, and subsequently fielded significantly more Confederate Army members, than lottery losers’ households. These results are robust to examining effects on the fraction of sons who fight in the Army, a measure which accounts for fertility effects.

In sum, modest increases in wealth in 19th-century Georgia caused individuals to be more likely to fight for the Confederate Army, probably in part due to the fact that these increases came mostly in the form of slaves, deepening these individuals’ attachment to the institution of slavery.

In the final part of the paper, we try to shed light on some of the mechanisms for these effects. Even though slaveowners as a whole had incentives to fight against the Union, any individual slave-owner would still face incentives to shirk and avoid risking death in war. Inevitably, these analyses are more speculative, but we document patterns that suggest local communities organized to encourage collective action. There is a strong, positive association between the county-level fighting

---

4 One notable exception is Costa and Kahn (2003). Studying the Union Army, and focusing on different questions, the paper finds that higher-income soldiers are less likely to desert (the paper does not compare combatants and non-combatants).
rates of slaveowners and non-slaveowners, suggesting locality-level effects, like social pressure, on fighting rates. Counties with a higher percentage of slaveowners also exhibit higher fighting rates, on average.

The paper contributes to the broader literature on the relationship between political institutions and violent conflict by assessing an important historical debate in a case that seems to have fallen in between political science’s subfields. The main finding of this paper—that wealthier slaveowners were on average more likely to fight for the Confederate Army than non-slaveowners—demonstrates how the individuals who had the greatest stake in the continuation of the institution of slavery were the most likely to fight in defense of it.

2 Slave Ownership, Wealth, and Fighting for the Confederacy

We now turn to summarizing a range of historical research which provides countervailing evidence for whether we should expect wealthier individuals to be more or less likely to fight for the Confederate Army. Before describing these divergent predictions in more detail, we first need to emphasize that the decision to fight for the Confederate Army was an actual choice. The Confederacy implemented the first compulsory draft in North American history (McPherson 2003, 430), so it would seem at first that southerners had little choice in fighting. In reality, there were many ways to avoid the draft. Wealthy citizens could pay to avoid service, and the draft “was not uniformly administered” (Ambrose 1962, 264). Later in the war, men who owned more than 20 slaves were exempt from the draft altogether (Martin 2012).

Perhaps more importantly, many people rich and poor simply avoided the draft. In Northwest Georgia—a region important to this paper, because it is home to the land allocated through the land lottery we study later—draft dodging was so rampant that the Confederate Army was dispatched to round up reluctant soldiers (Sarris 2006, 88). These efforts were largely futile. Some men joined only to desert at the first opportunity; some simply refused and chose to serve time in prison; others were hidden in their homes or by neighbors; and many others melted into the woods when the army came near (McPherson 2003, 432). As recounted in Sarris (2006, 89), W.A. Campbell, a former

5 For more information on the history of conscription in America and other Anglo-Saxon countries, see Levi (1996, 1997).
Confederate officer, returned home to Fannin County, Georgia, to find “a very large majority of the people now here, perhaps two-thirds, are disloyal...not 1/2 dozen men have gone into the service.” In many parts of the Confederacy, especially in areas more ambivalent towards secession, concerns that individuals could “sabotage the Confederate conscription” were widespread (Sarris 2006, 90). Indeed, the state government of Georgia, led by Governor Joe Brown and his lieutenant, Adjutant and Inspector General Henry C. Wayne, “obstructed conscription in every way they could” (Scaife and Bragg 2004, 3).

In a landmark study on fighting in the Civil War, McPherson (1997, 5) sums up how men joined the fight: “[M]ost Union and Confederate soldiers were neither long-term regulars or draftees, but wartime volunteers from civilian life whose values remained rooted in the homes and communities from which they sprang to arms and to which they longed to return.” Thus despite the Confederacy’s best efforts to conscript soldiers, for many, the decision to fight for the Confederacy was a choice. Given this choice, how then should we expect an individual’s wealth to shape his decision to fight?

2.1 Why the Wealthy Might Fight Less

A range of historical research on the American Civil War suggests that wealthier individuals were on average less likely to fight than their poorer compatriots. This research highlights a number of historical institutions seemingly created to incentivize poorer individuals to fight, while allowing wealthier southerners to avoid fighting. For example, throughout the Civil War both the Confederate and Union armies used enlistment bonuses to recruit soldiers, generally known as bounties. In early 1861, enlistees received a 10-dollar enlistment bounty (Scheiber 1969, 229). Later in the year, the Confederate Congress approved a 50-dollar bounty for individuals who re-enlisted (McPherson 2003, 430). These bounties clearly sought to induce poor individuals to fight by increasing the monetary payoff to enlisting. Such monetary rewards targeted toward the poor is consistent with research that argues higher levels of personal wealth make individuals less likely to participate in rebellion (Berman et al. 2011; Collier and Hoeffler 2004; Dasgupta, Gawande, and Kapur)

Indeed, Noe (2010, 2) estimates that as much as nine percent of Confederate enlistees were induced to serve as substitutes for other southerners who sought to avoid fighting.
2017; Dube and Vargas 2013; Fearon and Laitin 2003; Miguel, Satyanath, and Sergenti 2004; Olson 1965). These papers argue that individuals participate in conflicts after calculating the material benefits they can obtain from fighting against the benefits from staying out of the conflict. In this view, wealthier people may have more valuable outside options, making them more reluctant to fight. With diminishing marginal returns of additional wealth, the bounties would encourage poorer individuals more than wealthier individuals.

Other conscription rules appear to have provided wealthier individuals the opportunity to avoid fighting. The rules, which angered many poorer individuals throughout the Southern Confederacy (Harris 1998, 150), took a number of forms, including the ability to pay someone to fight in one’s stead and the exemption of individuals owning 20 or more slaves from the draft (Ambrose 1962, 265). These rules were explicitly set up to encourage the poor, most of whom did not own slaves, to join the Confederate Army and to excuse the wealthy from doing so. In short, wealthier individuals had both the means and opportunity to avoid directly shouldering the burden of war. This body of evidence then suggests that we should expect wealthier individuals to be on average less likely to fight.

2.2 Why the Wealthy Might Fight More

Alternative arguments suggest instead that increasing wealth made southerners more likely to fight. Historical accounts make clear that excess wealth in the Antebellum south was overwhelmingly invested in two areas: land and slaves. Most southern white men were farmers, and farmers by and large sought to own slaves. Surveying the economy of the Cotton South, Wright (1978, 141) concludes that “even the smaller [slave]holder would find his financial portfolio dominated by the value of his slave property.” Thus for southern white men in the Antebellum period, individual

---

7 This policy led some wealthy individuals to advertise in newspapers for substitute enlistees (Moore 1924, 30).
8 This latter policy generally exempted one man, owner or overseer, on plantations of twenty or more slaves” (Harris 1998, 150).
9 In 1860, 80 percent of total wealth held by Southerners were in the form of land and slaves (Bleakley and Ferrie 2016; Ransom and Sutch 2001).
increases in wealth would likely tie them even more tightly to the institution of slavery. Wright (1978, 143) writes: “Across the South, slaveowners formed a class of great wealth with a distinctive unity of economic interest—not necessarily on policies concerning the economics of slavery, but in slavery itself.”

Because the Civil War centered on the future of slavery, southerners who invested in this institution might perceive the stakes of a Confederate defeat to be higher than non-slaveowners. Seccessionists throughout the South worried a great deal about whether non-slaveowning whites would support secession (McPherson 2003, 242), precisely because they seemingly lacked the obvious economic motivation that slave owners possessed. For this reason, the motivations of non-slaveowners to fight a war over slavery have long been debated. Ambrose (1962, 259) speaks of “yeomen,” small farmers without slaves, who were the majority of the southern white population, as “often unresponsive or downright hostile to their country’s cause.” A poor North Carolina woman quoted in the paper wrote a letter to her governor requesting her husband be discharged from the Confederate army, writing: “I would like to know what he is fighting for...he has nothing to fight for...I don’t think that he is fighting for anything only for his family to starve” (267).

Studies of southerners who opposed seceding from the Union further support the argument that non-slaveowners perceived the stakes of the conflict to be lower than did slaveowners. Studying the “up-country” areas of the South, where opposition was most concentrated, Tatum (2000, 4) writes that “[a]lthough many of the up-country dwellers hoped some day to own slaves, they had relatively few at that time, and therefore had little to lose by emancipation.” Consistent with this claim, Wooster (1977) presents evidence that delegates from counties with more slaves were more likely to support secession, across the 15 secession conferences held by southern states at the

10 It is possible that, for many, this perception evolved from the time of Lincoln’s election through to the heights of the war. Initially, Lincoln and the Republicans were careful to confine their opposition only to the expansion of slavery, and not to its continued existence in slave states. Some, like the so-called “Fire-eaters” who had been advocating secession well before Lincoln’s election, found these claims non-credible. As the war grew, the threat became clearer.
beginning of the Civil War, while areas with fewer slaves were less likely to support secession.\footnote{McPherson (2003, 242) states that “In the conventions, delegates supporting delay or cooperation owned, on average, less wealth and fewer slaves than immediate secessionists.”} The secession, Wright (1978, 41) concludes, “was essentially a slaveholder’s movement.”

The logic of the preceding paragraph suggests that wealthier southerners were more likely to own slaves and thus perceive the stakes of the conflict to be higher than their poorer southern compatriots. Slaveowners might perceive the stakes of the conflict to be higher either because: (a) their expectations of their future economic well-being hinged on the perpetuation of the institution of slavery, or (b) their direct experience of being socialized or indoctrinated into the institution of slavery could increase their perception of the importance of perpetuating the institution.\footnote{For recent research on how socialization affects the choice to participate in violent conflict, see, for example, Horgan et al. (2017); Green (2017).} Regardless of the mechanism underpinning this stakes-based argument, we should expect slaveowners to perceive the stakes of the Civil War to be higher, and thus be more likely than poorer non-slaveholding southerners to fight for the Confederate Army.

### 2.3 Alternative Arguments for Why the Wealthy Might Fight More

Parallel to the body of evidence suggesting that wealthier individuals had a greater stake in the conflict, we might also suspect that their increased wealth reduced the costs associated with abandoning their farms to fight. Indeed, many in the South felt that the rich could better afford to fight: “Poor men and women, for obvious reasons, saw the issue of conscription in a different light from that of the rich. While one argument for a draft was that it would spread the sacrifices of war more equally among all citizens, it was bound to affect the families of slaveless farmers much more than those who still had someone to plow and harvest” (Harris 1998, 149). Through this reasoning, we might suspect that individuals who owned slaves were able to rely on the extra labor at home to reduce the potential costs associated with their absence for their ability to continue farming. This means that the costs of fighting—in terms of being able to provide for their families—is lower, suggesting that these individuals would then be on average more likely to fight. However, cutting against this logic is the argument that slaveowners could hardly leave their slaves unattended.
but one historical example of this reluctance, Harris (1998) documents a case where a slaveowner named David C. Barrows “expended considerable energy attempting to keep his overseers out of the draft, arguing that they were needed to police the slaves and raise corn” (Harris 1998, 151). We suspect that this incentive against leaving slaves unattended probably undoes much of the resource effect, but it remains an interesting possible mechanism.

In addition to the potential for wealthier individuals to have lower costs associated with leaving their farms unattended, we might also expect that wealthier individuals held higher ranks within the Confederate Army. If individuals with higher ranks suspect that they are on average less likely to be killed in combat, then we might expect them to have lower expected costs associated with their participation in the conflict. These perceptions of the differential costs of fighting will then make them on average more likely to fight. Taken together, both of these mechanisms about differential costs present important alternatives for why wealthier individuals might fight in the Confederate Army at higher rates independent of their perceptions of the stakes in the conflict.

We examine these possibilities empirically in Section 8 of the manuscript by looking at the rates of fighting for individuals who we know were in the Confederate Army in 1861. Since in the early stages of the war most southerners thought that the war would be easy and over fairly quickly, we should expect these costs concerns to be less prominent in the minds of individuals who chose to fight. In general, we find that wealthier individuals still fight at higher rates in these early stages of the war. While this by no means rules out these alternative mechanisms, the results do provide suggestive evidence consistent with the stakes mechanism proposed throughout this paper.

3 Historical Data on Confederate Citizens

To test the countervailing theoretical predictions for how wealth affected individual’s propensity to fight for the Confederate Army in the American Civil War, we constructed a new dataset of nearly the entire military eligible population of the Confederacy. The dataset links three sets of publicly available sources of individual-level information: The 1850 U.S. Census of free citizens, the 1850 Slave Schedule, and state rosters of Confederate Army membership. In the remainder of this section we outline each of the respective data sources and then discuss how we combined them, with the Appendix section A providing precise details.
The left panel of Table 1 summarizes our sources of data. We start from the full U.S. Census of 1850, which contains the names for all individuals, along with their age, occupation, and place of birth.\footnote{We obtained the preliminary version of the Census from IPUMS (Ruggles et al. 2015), who have digitized all fields of the 1850 Census. The U.S. Census bureau has conducted a federal census every 10 years since 1790, but the 1850 Census is notable for being the first census that recorded personal information of all free white members of a household, instead of only the household head. Enumerators also started to record many social statistics for the first time in 1850. Although the 1860 Census is closer to the Civil War, the extent of digitization lags behind that of 1850.} We limit our analysis to the 11 Confederate states, Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Texas, and Virginia. Next, we collected the full 1850 Slave Schedules of the same 11 states.\footnote{We gathered the slave schedule by scraping a freely available version on one of many genealogy websites, focusing on the Confederate states.} The Slave Schedule records minimal demographic information for 2,574,602 slaves under the name of 263,743 owners, each of whom is identified by last name, first name (or initial), and county of residence. Finally, we gathered the state rosters for all Confederate Army soldiers in the same states.\footnote{Like the slave schedule, we scraped the online digitized index of the National Archive’s Compiled Records (National Archives n.d.).} These rosters contain a limited set of information about soldiers including their name, state, year of enlistment or conscription,\footnote{The same individual found in multiple years in different units is given multiple records.} and unit. This data collection effort resulted in 1,496,931 records that contain 704,650 unique state and soundex-encoded name combinations.

We linked individuals across the three sets of data described above, starting from the 1850 Census. We declare an individual in the Census a match with a slaveowner name in the slave schedule if the two household head’s soundex-encoded full name,\footnote{We first match on full first name, and then match the remaining by first initial.} state, and county are the same.\footnote{Recent work in record linkage offers improvements over this basic matching procedure (Feigenbaum 2016).} Throughout, we take the unique serial number assigned to Census households in our digitized...
Table 1 – Overview of Data Collection. Each number is the number of observations for the corresponding variable. The left table summarizes the variables captured in our raw data, with the respective source in parentheses. The right table summarizes the counts in the dataset we construct after multiple matches. The dataset starts from the 1850 Census and then matches data from the Slave Schedules and the Confederate Rosters by name and geography. Two specifications account for the issue of name duplicates during the merge – 1-over-N and M-over-N, and result in different estimates of the effective size of merged data. Appendix section A provides detailed documentation.

<table>
<thead>
<tr>
<th>N in Original Source</th>
<th>N in Merged Dataset 1-over-N</th>
<th>M-over-N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free Citizens (Census)</td>
<td>3,909,122</td>
<td>3,909,122</td>
</tr>
<tr>
<td>Households (Census)</td>
<td>746,506</td>
<td>746,506</td>
</tr>
<tr>
<td>Slaves (Slave Schedule)</td>
<td>2,574,602</td>
<td>1,811,224</td>
</tr>
<tr>
<td>Slave Owners (Slave Schedule)</td>
<td>263,743</td>
<td>193,785</td>
</tr>
<tr>
<td>Confederate Names (Roster)</td>
<td>704,650</td>
<td>1,393,589</td>
</tr>
</tbody>
</table>

The total number of slaves is then the number of slave records associated with a Census household. Next, we link individuals in the 1850 Census to the Confederate rosters by taking the male subset of the Census and declare that individual a match with a Confederate roster entry if his last name, first name, and state are exactly the same in both sources, again in soundex encoding. The total number of soldiers is the total number of members in any 1850 household who match to the Confederate rosters. Figures B1 and B2 illustrate the geographic distribution of our slave ownership and Confederate Army membership variables by county.

In the right panel of Table 1, we summarize the results of the matching process in terms of number of observations. The key challenge is to correctly locate slaveowners and Confederate soldiers in the Census. As the third row shows, we are able to locate 75 percent of slaveowners in the contemporaneous Census through our matching procedure. We discuss how the remaining matching error may affect our results in Appendix Section A. The final row shows the result of matching Confederate names to the Census. The roster has 704,650 distinct name-state combinations, yet we join these distinct names to almost 1.4 million records in the census. The discrepancy is largely

---

19 The data maintainer IPUMS (Ruggles et al. 2015) has constructed the serial numbers that aim to determine unique households.

20 Because we take all men alive in 1850, some men may have been too young or too old to fight in the Confederate army ten years later. However, because there exists no strict age limit on eligibility, we consider all ages. Results are similar using the eligibility requirements for conscription, i.e. restricting to men who would be 18 to 45 during the war.
due to common names within state. We address the issue of multiple roster entries matches by creating two weighting schemes that down-weight duplicate names. The 1-over-N column in Table 1 provides the effective number of soldiers we find in the census once we downweight each duplicated record by how many times it is duplicated in the Census. The M-over-N column does the same, except here we also up-weight the 1-over-N specification based on duplicates in the Confederate roster as well. In the Confederate rosters, the number of name duplicates within the geographic unit is so large that we prefer either the unweighted specification or the M-over-N specification.

4 Who Fought for the Confederacy? Descriptive Facts.

The data described in the previous section provides us important individual-level information about Southern residents in 1850. In the first column of Table 2 we present key descriptive statistics about our population based sample. The population is largely comprised of rural farmers. Households of free citizens on average include 5.5 family members, and about a fifth of all free households are estimated to own at least one slave. To examine the association between wealth and fighting, we first focus on two key pre-war measures of wealth: the number of slaves owned by an individual’s household and the value of real-estate property, both reported in the 1850 Census.

Figure 1 presents the average number of people who served in the Confederate Army, per household, across bins of the number of slaves owned by the household. The four bins roughly correspond to the four quartiles of the range of household slave ownership. Because many households owned no slaves, the first bar represents just this subset of the data. The next three bars divide the slave-owning households into thirds.

Three patterns emerge from this simple analysis. First, the propensity to fight in the Confederate Army is lowest for those households owning no slaves. Second, while the jump from being a non-slaveowner to being a slaveowner predicts a large increase in army membership rates, further increases in the number of slaves are associated with more modest rates. Indeed, using a multivariate regression that enters slave ownership both as a binary variable and linearly interacted with the total number of slaves a household owns (and controls for property values), we estimate that becoming a slaveowner is associated with an increase of 0.12 soldiers, whereas conditional on own-
Table 2 – Descriptive Statistics of the Two Populations Examined. Summary statistics on the demographic variables. The full-count data of free citizens in the 1850 Census explored in section 4 shown in the left column. The subset explored in section 5 shown in the next two columns. For these two columns the binary specification of treatment is used to separate the treatment and control; see text for other specifications. All values are from 1850.

<table>
<thead>
<tr>
<th></th>
<th>All Confederacy States</th>
<th>Lottery Sample Control Group</th>
<th>Lottery Sample Treatment Group</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Family</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion Men</td>
<td>0.499</td>
<td>0.526</td>
<td>0.53</td>
</tr>
<tr>
<td>Prop. Child (of Household Head)</td>
<td>0.556</td>
<td>0.709</td>
<td>0.713</td>
</tr>
<tr>
<td>Average Household Size</td>
<td>5.527</td>
<td>8.235</td>
<td>8.387</td>
</tr>
<tr>
<td><strong>Geography</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prop. Households in Georgia</td>
<td>0.119</td>
<td>0.731</td>
<td>0.726</td>
</tr>
<tr>
<td>Prop. Households in Alabama</td>
<td>0.098</td>
<td>0.155</td>
<td>0.155</td>
</tr>
<tr>
<td>Prop. Households in Mississippi</td>
<td>0.069</td>
<td>0.039</td>
<td>0.042</td>
</tr>
<tr>
<td><strong>Socioeconomic</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prop. White</td>
<td>0.971</td>
<td>0.999</td>
<td>0.999</td>
</tr>
<tr>
<td>Prop. Households in Urban Area</td>
<td>0.071</td>
<td>0.012</td>
<td>0.017</td>
</tr>
<tr>
<td>Prop. Headed by Farmer</td>
<td>0.548</td>
<td>0.87</td>
<td>0.874</td>
</tr>
<tr>
<td><strong>Wealth and Slavery</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Total Property Wealth</td>
<td>1,141.572</td>
<td>2,087.693</td>
<td>2,124.769</td>
</tr>
<tr>
<td>Median Total Property Wealth</td>
<td>0</td>
<td>700</td>
<td>660</td>
</tr>
<tr>
<td>Prop. Households Owning at least One Slave</td>
<td>0.218</td>
<td>0.374</td>
<td>0.406</td>
</tr>
<tr>
<td><strong>Sample Size</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Individuals</td>
<td>3,909,122</td>
<td>90,438</td>
<td>15,862</td>
</tr>
<tr>
<td>Number of Households</td>
<td>746,506</td>
<td>11,439</td>
<td>1,975</td>
</tr>
<tr>
<td>Number of Counties</td>
<td>692</td>
<td>388</td>
<td>234</td>
</tr>
</tbody>
</table>
Figure 1 – Slave Ownership and the Propensity to Fight for the Confederacy. Households without slaves provided Confederate Army soldiers at a lower rate than households with slaves. Error bars indicate 95 percent confidence intervals.

- **No Slaves** (N = 552,721)
- **1 − 3 Slaves** (N = 74,067)
- **4 − 9 Slaves** (N = 60,612)
- **10 − 961 Slaves** (N = 59,106)

The average number of soldiers in a household decreased as the number of slaves increased:

- **No Slaves**: 1.78
- **1 − 3 Slaves**: 2.04
- **4 − 9 Slaves**: 2.1
- **10 − 961 Slaves**: 2.19

Buying slaves at all an additional slave is associated with an increase of only 0.0017 soldiers (Appendix Table C1). Still, both coefficient estimates are distinguishable from a null relationship.

Third, the overwhelming majority of the Confederate Army were not slaveowners. Indeed, according to our data, there were more non-slaveowning households who nevertheless provided Confederate soldiers (410,646 households) than there were slaveowning households in total (193,785 households). This prevalence of non-slaveowners in the army may have led some observers to infer that non-slaveowners were just as likely, if not more likely, than slaveowners to fight. Instead, non-slaveowner’s rate of fighting in the Confederate Army was substantially lower.

Next, we consider wealth as proxied by real-estate property value—the only indicator of wealth besides slave ownership available for 1850. We again divide households into four bins, the first for non-property owners and the latter three for three terciles in terms of real-estate property wealth. Figure 2 shows that households reporting no real-estate value fought in the Confederate Army at a lower rate than those possessing some real-estate wealth. Like slave ownership, the starkest difference is between households owning no property and owning at least some property.
The patterns apparent from the population averages are informative, but many other variables may confound the association between a household’s wealth and slave ownership in 1850 and the number of soldiers the household fields a decade or more later. In Appendix Section C, we attempt to account for observable confounders by estimating regressions with other controls, fixed effects, and allowing for correlated errors within relevant groups. Our findings are robust to fixed effects by last name, a proxy of family heritage, as well as to fixed effects for county. The three patterns in Figures 1 and 2 are robust to including fixed effects clustering standard errors by last name, a proxy of the socioeconomic status of a family’s heritage, as well as to fixed effects and clustered standard errors for county. The positive association between wealth in the form of slave ownership and fighting also holds when we take as our outcome variable the fraction (instead of counts) of a household’s male residents that are soldiers. This accounts for the understandable concern that larger households simply have more sons available to fight. In the following section, we further rule out unmeasured confounders by focusing on a subset of our sample that received a random shock of wealth.
5 Wealth Increases Propensity to Fight: Evidence From the Georgia Land Lottery

Did southern white men fight in the Confederate Army because they were wealthy, or are these variables simply correlated with other attributes that made these men more likely to fight? We now turn to an experimental design that leverages a large-scale lottery in the state of Georgia in which plots of land, worth a considerable amount of money, were randomly allocated to white men 29 years prior to the outbreak of the Civil War.

Conducted six times during 1805–1833, the lotteries were unique, widely popular to its white citizens, and covered a vast area of contemporary Georgia (Weiman 1991). We focus on the sixth lottery of Georgia, held in 1832, because of its outsized scale and its proximity in time to the Civil War. The 1832 land lottery, also known as the Cherokee Land Lottery, distributed Cherokee Nation territory (the northwest of contemporary Georgia) to white citizens of Georgia. The lands were forcefully taken from the Cherokees in infamous fashion, and the rights to the land became even more controversial after a gold vein was discovered in the area around 1829 and white settlers inundated Cherokee territory to strike gold. The 1832 lottery distributed land which the Georgia state legislature declared as “Cherokee County” in December 26, 1832, shortly after President Jackson won re-election. Following carefully stipulated rules set by the state legislature, the surveyor office carved out land lots roughly 160 acres each to be made available to any lottery winner for a processing fee of $18. Unclaimed land would eventually return to the ownership of the state of Georgia.

After eligible citizens registered for the lottery by sending their names to the governor’s office in Milledgeville, Georgia, commissioners conducted the lotteries in a doubly-randomized process (Williams 1989). Lottery organizers set up two large drums called “wheels.” One wheel for names contained slips of paper for every registrant for the lottery. The second wheel for land contained a slip of paper for every parcel of Cherokee Nation land up for lottery. Commissioners simultaneously drew one slip of paper from each of the two wheels, such that both whether or not a person won

---

21 For the full text of the resolution, see Georgia Legislative Documents (1830).

22 For reference, the wholesale price of raw cotton in 1832 was $9.40 per 100 pounds (United States Census Bureau 1975, 209, E126).
the lottery, and the exact location of the land were randomized. Eligibility requirements remained roughly consistent across all six lotteries. Household heads 18 years or over, with a minimum 3-year residence in Georgia and U.S. citizenship, were all covered (Georgia Archives n.d.). Winners of previous lotteries were excluded. Certain groups, most notably those with a spouse and children, were advantaged by having two draws at the lottery.

While the lottery distributed wealth in terms of land, its practical effect was to increase the winner’s general monetary wealth. Historical evidence suggests that richer citizens eagerly bought off lotteried land from winners of the lottery, effectively creating a private market for land. Indeed, “Even mediocre land could probably be sold for $25 or $50, probably enough to exempt [a yeoman household head who won the lottery] from ever paying taxes again ... If he hit the jackpot and won a piece of prime cotton land worth several hundred dollars, then he could quickly rise up the ranks and perhaps even buy a slave” (Weiman 1991, 857).

5.1 Identifying the Winners and Losers of the Land Lottery

To analyze the lottery, we first use a digitized dataset of lottery winners’ names (Smith 1838), which includes their self-reported place of residence and a symbol for whether the winner paid the fee and claimed the land he won at the time of publication. This results in information on 18,219 unique individuals. The records of lottery winners are an official census (not self-reported), “all carefully copied from the originals in the Executive Department and the office of the Surveyor General, designating also the lots which have been granted” (Smith 1838, iii). We will then define our experimental population as a specific subset of the 1850 Census which reasonably entered the lottery and had the same number of draws. Finally, we seek the names of the lottery winners among this subset of predicted entrants.

Identifying the counterfactual lottery losers is key to examining the causal effect of winning the lottery. In an ideal world, we would obtain the entire population of lottery entrants and then directly compare those that lost with those that won. Unfortunately there is no available list of state-wide lottery losers; instead, we adopt the same method used in recent research conducted in economics on the monetary and educational effects of the same land lottery (Bleakley and Ferrie 2016). We identify a subset of individuals who were eligible for exactly two draws from the land lottery in the 1850 Census, using a set of pre-treatment characteristics that all but ensures
the individual was eligible. Because the lottery was popular and estimates suggest more than 97 percent of those eligibles entered (Bleakley and Ferrie 2016, Appendix A), we then assume that such eligibles entered the lottery, or at least are equivalent in their pre-lottery characteristics and potential outcomes to those who entered and subsequently won. Specifically, we take all male household heads born before 1814 who had a child (alive in 1850) in Georgia between 1829 and 1832. Individuals who we find in the list of winners are labeled as treated, while those that do not match we label as control. In Appendix D, we show that treated and control households are similar in terms of the pre-treatment characteristics we are able to measure.

Astute readers will notice at this point that we are defining our counterfactual pool (1832 lottery losers) using post-treatment observations (1850 Census observations). In line with the empirical strategy presented in Bleakley and Ferrie (2016), we assuage these concerns in two ways. First, because we use a full-count dataset of the entire Southern United States, outcomes are observed even for individuals who moved away from Georgia. Second, treatment could affect individual wellness or longevity, so that lottery winners are more likely to still be alive and in the Census in 1850 than lottery losers. The resulting comparison among surviving lottery winners and losers could then be biased if, for example, the subset of losers who are still alive in 1850 are those who for other reasons became wealthier or higher status. We suspect this issue is relatively minor because the treatment was not so large as to affect longevity dramatically. Nevertheless, in order to address this concern, we also applied our assignment mechanism to the 1830 Census, finding similar rates of treatment assignment. Although the 1830 Census is not digitized to an extent that we can test our specific hypotheses, retrieving similar aggregate match rates from a pre-treatment dataset is encouraging evidence that treatment did not affect longevity in a way that would subsequently affect our estimates.

We identify lottery winners by computing the phonetic distance of both first and last names between the two data sources. This metric allows us to sidestep potential issues involving variation in spelling, transcription errors, and enumerator spelling errors, which might lead the same individual to have his name spelled slightly differently in the 1832 winners’ list and the 1850 cen-
Traditional exact matching will falsely declare such cases as negative matches. We use a phonetic
distance metric informed by the historical evolution of languages (Downey et al. 2008), implemented in the R package \texttt{alineR}. We declare two names a match if the \texttt{aline} distance of the first and last names are both less than 0.05 (where distance ranges from 0 to 1).

To account for multiple matches, we consider three specifications of treatment status based on this phonetic distance metric, roughly equivalent to our procedure in the construction of our observational dataset. In the \textit{binary} specification, we assign the treatment indicator by examining whether a name in our eligible population corresponds 1-to-1 with a name in the lottery winners’ list. The 1-over-N (following Bleakley and Ferrie (2016)) and M-over-N specification attempt to account for duplicate names by down-weighting multiple matches. These three specifications identify treated individuals among the eligible population with a range of coverage. The 1-over-N specification generates a 19.4 percent treated sample (i.e., the average of the weighted indicator is 19.4 percent), whereas the binary specification gives 14.7 percent and the M-over-N specification gives 25.9 percent. Among the winners we examine, 48.4 percent were recorded as having claimed the land.

All of this leaves us with 13,414 individuals (or about 2.6 percent of the population of Georgia in 1830) from the Census. These individuals combined with their 1850 household members amount to 106,300 individuals, or roughly 2.7 percent of the southern free citizens in our dataset. Figure 3 presents the geographic distribution of these eligible individuals across the Antebellum South in 1850. The center and right columns of Table 2 introduced earlier compares the demographic characteristics of this sample, separated by the binary treatment indicator. The lottery sample is concentrated in Georgia by design, although nearly a quarter of the lottery entrants are estimated to have moved to another state by 1850.

5.2 Estimation

Our experimental analyses compare 1850 household members whose household head won the 1832 land lottery to 1850 household members whose household head reasonably entered the 1832 lottery

---

\footnotesize{\textsuperscript{23} The computational costs of computing phonetic distance for the entire Census is prohibitive, and thus was not available for the data construction in section 3.}
Figure 3 – Distribution of Inferred Lottery Entrants in the 1850 Census analyzed. The map counts the number of 1850 Census household heads per county who we identified as lottery entrants (a strict subset of the total lottery entrant population). Counties outlined in red indicate former Cherokee county, the ownership rights of which the Georgia state government distributed by lottery in 1832. Most lottery winners’ households are still concentrated in Georgia 18 years after the land lottery, though some households have moved west.

![Map showing the distribution of inferred lottery entrants](image)

Number, by county, of 1832 Georgia Land Lottery entrants examined [0] [100] [200]

Source: 1850 U.S. Census and Smith (1838)

but lost. The randomized nature of the lottery allows us to estimate causal effects at the household level using a simple OLS equation of the form

\[ Y_i = \beta_0 + \beta_1 \text{Won Lottery}_i + X_i^\top \gamma + \varepsilon_i, \] (1)

where \( Y_i \) is an outcome variable for household \( i \). The treatment of the 1832 father winning the lottery, denoted by \( \text{Won Lottery}_i \), is also realized at the household level. This variable stands in for any of the three treatment variable specifications discussed above. The vector \( X_i \) stands in for an optional vector of control variables. The coefficient \( \beta_1 \) estimates the average effect of winning the lottery. Because 52 percent of those who won did not claim, the coefficient represents a dilution of
Figure 4 – 1832 Lottery’s Effects on Slave Wealth and Real Estate Property Wealth in 1850. Winning the land lottery increased both the average number of slaves a household owned, and the proportion of households that owned at least one slave. In contrast, the land lottery’s effects on overall wealth seem to be driven by slavery wealth, as there is no discernible increase in real-estate wealth. Each panel shows the estimated means for a given dependent variable, among each treatment group. Error bars show 95 percent confidence interval from robust standard errors.

<table>
<thead>
<tr>
<th></th>
<th>Lottery losers</th>
<th>Lottery winners</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of slaves</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>4.4</td>
<td>5.4</td>
</tr>
<tr>
<td>Holds at least one slave</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>37%</td>
<td>44%</td>
</tr>
<tr>
<td>Estimated real estate property ($)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>2038.2</td>
<td>2158.7</td>
</tr>
</tbody>
</table>

the average effect of actually experiencing the wealth shock. In the following analyses we present estimates of winning the lottery ($\beta_1$), but note that the estimated effect of wealth among those who actually consumed it is roughly twice that of $\beta_1$, computed by instrumenting the choice to claim with the randomized result of the lottery.

6 Effects of Randomized Land Lottery on Slave Ownership and Fighting

We now present experimental estimates.24 First, we find a large positive effect of winning the land lottery on slave ownership at the household level. The first two panels of Figure 4 shows that households whose fathers won the 1832 lottery had on average 0.9 more slaves in 1850, and were 6.2 percentage points more likely to own slaves at all, compared to households whose fathers did not win. This represents a substantial difference in the number of slaves owned by individuals who won the lottery compared with those who lost. Substantively, this provides direct evidence supporting

24 When not otherwise stated, we use the 1-over-N indicator for treatment though as we show in Table 3, results are substantively similar regardless of the indicator used.
accounts of individuals throughout the Antebellum South investing their newly acquired wealth in purchasing slaves.

The next two panels of Figure 4 show modest effects on the amount of real estate property wealth owned by individuals who won the Georgia land lottery as recorded in the 1850 Census. We find that average household real-estate value in the Census was roughly $150 higher for lottery winners than for lottery losers. In contrast, the proportion of households with any property wealth is almost approximately the same for lottery winners and losers. These modest effects on real-estate wealth again bolster the suggestion that most excess capital was invested in slavery. Indeed, Bleakley and Ferrie (2016) estimate a meaningful effect of the lottery on total wealth. These effects are primarily concentrated in slavery rather than real estate. Like southern white men more generally, winning families in our sample appear to have invested much of their newfound capital in slavery.

Next, we trace these effects through to fighting in the Confederate Army. Figure 5 shows our main result, demonstrating that households whose fathers won land in 1832 on average have 0.33 more men fight in the Confederate Army, and were 4.2 percentage points more likely to have a male household member fight in the army at all. The point estimates are substantively large and statistically distinguishable from zero.

Table 3 presents our main results across specifications.25 The first column in the top panel shows the estimated effect on the average number of Confederate soldiers in the household, for the binary and 1-over-N treatments, respectively. This estimate shows that winning the lottery led to an average increase of 0.31 in the number of soldiers in the household who fought for the Confederate Army. To address concerns that results are driven by the conduciveness of certain names to match across data sources, we also estimate the same quantity with fixed effects for the last name, or with fixed effects for the first name, following Bleakley and Ferrie (2016). The goal

---

25 We also note that our main results are also robust to our two approaches to the issue of record linkage, as presented in Appendix E. Similar to the previous section, we re-estimate the same quantities only using the set of households in the 1850 census whose head had a name unique in the state (Table E7). We also use the M-over-N weighted estimator (Table E8). The estimated effects of winning the lottery on fighting in the Confederate army are similar using this with these two sets of data.
Table 3 – Effect of Winning 1832 Lottery on Household Confederate Army Membership.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Number of Confederate Soldiers in Household</th>
<th>Probability at Least One Son Fights</th>
<th>Fraction of Sons Who Fight</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treatment No FEs Last Name FEs First Name FEs</td>
<td>Treatment No FEs Last Name FEs First Name FEs</td>
<td>Treatment No FEs Last Name FEs First Name FEs</td>
</tr>
<tr>
<td>Lottery Winner</td>
<td>0.31 0.16 0.29</td>
<td>Lottery Winner ((\frac{1}{n}))</td>
<td>0.04 0.02 0.04</td>
</tr>
<tr>
<td></td>
<td>(0.05) (0.05) (0.06)</td>
<td></td>
<td>(0.01) (0.01) (0.01)</td>
</tr>
</tbody>
</table>

Each cell is a regression coefficient. Sample size in all regressions is 13,414 households. Outcome variable in top panel is the number of registered Confederate soldiers in the household. Outcome variable in middle panel is fraction of sons in household who fight in Confederate Army. Outcome variable in bottom panel is indicator for whether at least one son in household fights in Confederate Army. Estimates in first row of each panel use binary treatment indicator based on unique name matches. Estimates in second row of each panel include non-unique name matches, where treatment variable takes the value \(\frac{1}{n}\) for a lottery winner name matched to \(n\) households in 1850 Census. Robust standard errors in parentheses.
Figure 5 – 1832 Lottery’s Effects on Confederate Army Membership in 1860s. Winning the land lottery increased both household’s average number of men in the Confederate Army (both as an average count and as a fraction of men in the household) as well as the proportion of households containing at least one soldier. Each panel shows the estimated means for a given dependent variable, among each treatment group. Error bars show 95 percent confidence interval from robust standard errors.

The middle and bottom panels repeat these same specifications for our alternative outcome variables—the probability that at least one male member fought (middle panel), and the fraction of free men who fought in the Confederate Army (bottom panel). Again, we find consistent and large results.

The outcome measure in the bottom panel is particularly important because they help refute an alternative, fertility-based explanation of our findings. Wealthier households might simply have had more sons, and therefore more chances to field soldiers. Winning the lottery indeed increased fertility, which we present in Table 4. While fertility effects are smaller in magnitude than the

---

26 Results are also highly similar including both sets of fixed effects. We do not present these for brevity.
Table 4 – Effect of Winning 1832 Lottery on Fertility.

<table>
<thead>
<tr>
<th>Lottery Winner</th>
<th>Number of Sons</th>
<th>No FEs</th>
<th>Last Name FE</th>
<th>First Name FE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0.11</td>
<td>0.09</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.04)</td>
<td>(0.05)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>Lottery Winner</td>
<td></td>
<td>0.10</td>
<td>0.08</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.04)</td>
<td>(0.05)</td>
<td>(0.05)</td>
</tr>
</tbody>
</table>

Each cell is a regression coefficient. Sample size in all regressions is 13,414 households. Outcome variable is total number of the household head’s sons. Estimates in first row of each panel use binary treatment indicator based on unique name matches. Estimates in second row of each panel include non-unique name matches, where treatment variable takes the value \( \frac{1}{n} \) for a lottery winner name matched to \( n \) households in 1850 Census. Robust standard errors in parentheses.

fighting effects in the top panel of Table 3, directly addressing the extent to which this alternative mechanism is driving our results requires some care. One approach is to control for fertility as an additional explanatory variable, but because fertility is an intermediate outcome (1850) between the treatment (1832) and the main outcome of interest (1861-1865), doing so leads to post-treatment bias (e.g., Acharya, Blackwell, and Sen 2016). Instead of a regression-based adjustment, we simply take the fraction of the number of soldiers as a proportion of the household’s total number of men, a better identified quantity. The bottom panel of Table 3 shows the result: positive effects on fighting hold even accounting for the increase in the denominator of the fraction. As a share of all men in each family, lottery winners’ male members joined the Army at a rate roughly 6 percentage points higher than that of lottery losers’.

Whether this particular shock of wealth is generalizable to the population of Confederate citizens is a methodologically complex but important question. Table 3 suggests that the uptake of the wealth shock induced a household to obtain 1.8 more slaves, and also induced about 0.60 more men to join the army, or about a 12 percentage point increase in the fraction of men who join. The ratio of the effect on fighting over the effect on slavery mirror our observational estimates of the association between obtaining one slave: 0.32 more men (Table C1) and 5.6 percentage points in fraction of men (Table C2). The similarity suggests that our lottery-based estimates in this section is generalizable as well as internally valid. That said, we acknowledge that such comparisons
account neither for unmeasured confounding nor the effects of the lottery that did not flow through slavery, and are by no means definitive. All together, our observational and experimental findings are consistent with the argument that wealthier individuals were on average more likely to own slaves, increasing their perceptions of the stakes of the American Civil War, thus making wealthier slaveowners on average more likely to fight.

7 Addressing an Alternative Mechanism: Differential Costs of Fighting

As discussed in Section 2.3, an alternative potential explanation for why wealthier slaveowners are on average more likely to fight is that the costs of being away from home might be lower than for non-slaveowners. While there is a counter-argument cutting against this, since we suspect that there is also an incentive to avoid leaving slaves unattended, it is an interesting possible mechanism which merits more scrutiny. Moreover, we might also expect wealthier slaveowners to estimate the costs of participating in conflict to be lower, perhaps because they would enter the military at higher ranks than their poorer southern compatriots (Logue and Barton 2007, 256). This would then make them fight at higher rates. Taken together, both of these arguments about differential estimated costs of fighting suggest an interesting an important alternative for why we might expect wealthier individuals to fight at higher rates.

In order to better understand whether this possible mechanism is solely responsible for the wealth effect we observe, we can look at the rates of fighting for individuals who we know were in the Confederate Army in 1861. At the onset of the conflict, many people on both sides of the conflict believed that the war would be both short and easy, likely lasting just a few battles. This sentiment is captured by the writings of an Alabama soldier fighting for the Confederate Army, who asserted that the war would be over by the next year since “we are going to kill the last Yankey before that time if there is any fight in them still. I believe that J.D. Walker’s Brigade can whip 25,000 Yankees. I think I can whip 25 myself” McPherson (2003, 333). Given this sentiment, we should expect individuals who joined the Confederate Army at conflict onset to be less concerned about the costs associated with leaving their farms unattended for extended periods of time, since they expected to be home again by the following year. Moreover, this optimism at conflict onset
Table 5 – Effect of Winning 1832 Lottery on Fighting at Different Stages of the War. Lottery effects on fighting throughout the War are similar compared to the fighting only recorded in 1861.

<table>
<thead>
<tr>
<th>Number of Soldiers</th>
<th>Fraction of Men</th>
<th>At Least One</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall 1861</td>
<td>Overall 1861</td>
<td>Overall 1861</td>
</tr>
<tr>
<td>0.33</td>
<td>0.06</td>
<td>0.04</td>
</tr>
<tr>
<td>(0.05)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
</tbody>
</table>

Each cell is a regression coefficient. Each pair of columns compare the effect of winning the 1832 Lottery on one of the three metrics of army membership, corresponding to the panels in Table 3. The comparison is between fighting for all years (as already presented in Table 3) vs. fighting only in 1861, as recorded by the roster record. All specifications do not use fixed effects and use the $\frac{1}{n}$ treatment specification.

would lead individuals selecting into the military to perceive the costs of fighting to be low. As one Mississippian put it, he joined the Confederate Army “to fight the Yankies—all fun and frolic” (McPherson 2003, 333).

Given that in 1861 many southerners rich and poor suspected the war to be over quickly and easily, if this differential cost argument is correct, then we would expect to observe a null effect for the relationship between wealth and fighting, since the estimated costs of leaving home to fight would be roughly similar for slaveowners and non-slaveowners. Contrary to this argument, Table 5 shows that the main finding that wealthier individuals were on average more likely to fight holds when we look at the effect of the land lottery only among individuals fighting in 1861. Substantively, this suggests that for the period of the war when we would not expect differential estimates of costs to be present, we observe the same difference in outcomes. It is important to emphasize that this evidence is imperfect and by no means rules out the idea that differences in perceptions of the costs of fighting may have affected the choice of whether to fight for the Confederate Army. However, the fact that the main results of the paper remain the same when looking at a period in the conflict when we would think that expectations of the costs of leaving home are roughly similar, does lend further support to the argument that slaveowners perceived the stakes of the conflict to be higher, and thus they were on average more likely to fight.
8 Did Local Communities Encourage Fighting?

Having laid out both descriptive and experimental evidence for the causal links between wealth, slaveowning, and fighting in the Confederate Army, we now turn to some more speculative evidence for how and why the increasing stakes of the conflict associated with the potential end to the institution of slavery overrode the incentives to free-ride and avoid paying the costs of war. Although we have explained why owning slaves might increase their perception of the stakes of the conflict, the results do not directly reveal why any individual actually chose to fight. Indeed, for any one individual the risk of death in war would seem to make joining the Army fundamentally at odds with self interest, no matter the increased stakes.

Historical accounts suggest that local communities organized together to encourage white southern men to fight, while also isolating and punishing shirkers. Sarris (2006, 52) describes for example how in Georgia, “communities rallied to support the soldiers,” offering “parades and public displays supporting their departing soldiers, ceremonies that were designed to reinforce the commonality of interests among the men, women and children...” The account continues, “Indeed, one of the first acts of the secession convention was to establish a formal definition for treason. Confederate loyalty was a statewide obsession...” (Sarris 2006, 57). As the main supporters of secession, and as leaders within their communities, slaveowners were at the center of these efforts. We might then suspect that slaveowners may have found themselves unable to induce non-slaveowners to fight without personally contributing to the war effort.

To investigate the possibility of community-level mobilization, we examine some simple patterns concerning who fought and where. Figure 6 compares fighting rates for slaveowning and non-slaveowning households at the county-level, for all 692 counties in our population dataset. If, at the local level, the decision to fight is driven by contextual factors beyond slaveowning, in places where slaveowners were on average more likely to fight, non-slaveowners would also be more likely to fight. This means we would see a positive correlation between the fighting rates of slaveowners and non-slaveowners. In contrast, if slaveowners and non-slaveowners coordinate their behavior

27 For work on how rebels overcome this collective-action problem, see, for example, Lichbach (1998) and Moore (1995).
Figure 6 – Comparing the Fighting Rates of Non-Slaveowners and Slaveowners. Each point is a county’s rate of fighting in the Confederate Army among nonslaveowning households, compared to the county’s fighting rate among slave-owning households. A 45-degree line, which would denote that slaveowning and non-slaveowning households fought at equal rates, is added for clarity. Slaveowning households consistently fielded more Confederate soldiers than the non-slaveowning households in the same county, but the two are tightly correlated.

Separately, their rates of fighting may be uncorrelated, and if wealthy slaveowners compelled non-slaveowners to fight while avoiding fighting themselves, the two rates may be negatively correlated.

Figure 6 shows a tight correlation of 0.82. In counties where more non-slaveowners fight, more slaveowners fight, too. Consistent with our main findings, slaveowners also fought at higher rates than the non-slaveowners in their own counties, as the clustering of points above the 45-degree line makes clear. These two patterns suggest that while wealth in the form of slavery encouraged households to fight, coordination between slaveowners and non-slaveowners also may have played a widespread role.

We also consider the link between the share of a county that owns slaves and the rate of fighting that occurs. If it is local community pressure that leads slaveowners and non-slaveowners to fight at similar rates, then counties with more slaveowners ought to be more effective at inducing more individuals to fight. Figure 7 presents the observed relationship, separately by state. The horizontal axis of each plot shows the percent of households in a given county that own slaves; the vertical
**Figure 7 – Fighting Rates Across Slaveowning Contexts.** In most states, counties with high rates of slave ownership saw higher rates of fighting than counties with lower rates. The northern Confederate states of Arkansas, North Carolina, and Virginia are exceptions. Each point in the figure is a county in 1850, placed by the proportion of slave owners (the proportion of slaveowners estimated to own at least one slave) and the proportion of all free men in the county who are estimated to have fought. The latter variable uses the specification of Confederate Army membership which does not down-weight for duplicates, consistent with the rest of the analyses in the main text. For each state, numbers are regression slope coefficients with standard errors in parentheses and states ordered by the magnitude of the slope coefficient. The entire distribution of counties reappears in the background for comparison.

![Figure 7](image)

axis shows the estimated fraction of white men in the county who fight in the Confederate Army. Different states exhibit varying relationships between county-level slaveholding and fighting. Slave ownership correlates negatively with fighting in the northern Confederacy states of Arkansas, North Carolina, and Virginia. But in the majority of states fighting rates correlate positively with the prevalence of slaveowners. The association between county-level fighting rates and slave ownership, after netting out within state factors by fixed effects, is clearly positive. Moving from a county with no slaveowners to a county of only slaveowners is associated with an increase the fraction of white men in the county who fought in the Confederate Army by 9 percentage points ($t = 5.47$). Louisiana, Texas, and Florida exhibit especially positive slopes.
A full account of how and why individuals chose to fight for the Confederacy is beyond the scope of this paper. However, in this section we have presented suggestive evidence that local community activity encourage individuals to fight.

9 Conclusion

In this paper, we explored the incentives for individuals to participate in violent conflict and rebellion in the context of the American Civil War. We presented two main empirical analyses. In the first, we used data on almost every citizen of the Confederacy to show that slaveowners fought in the Confederate Army at higher rates than non-slaveowners. To understand whether these descriptive patterns were causal, we then focused on a randomized lottery run by the state of Georgia. We showed that the households of lottery winners owned more slaves, and, perhaps as a consequence, were more likely to have sons fight in the Confederate Army. Taken together, the analyses show that in the case of the American Civil War, wealthier citizens were more likely to fight against the federal government.

Our results contribute to a longstanding historical debate about who fought for the Confederacy. The old saying that the Civil War was a “rich man’s war but a poor man’s fight” may hold water, perhaps when applied to the very wealthy, but our evidence suggests that modest increases in wealth and in slave ownership actually made southern white men more likely to fight, not less. A majority of the Confederacy did not own slaves, but those who owned slaves fought at higher rates. The familiar observation that many soldiers were non-slaveowners largely reflects that most southerners were not slaveowners, but this does not imply that non-slaveowners supported the Confederacy more.

In addition to helping shed light on historical understandings of the Antebellum South and Confederacy, the paper highlights at least two main potential areas for future research. First, future research could more explicitly theorize and empirically test when and why we should expect individual perceptions of the increasing stakes of a conflict to overcome the incentives to free-ride, or when instead we should expect the incentives to free-ride to win out. In the case of the American Civil War, our results demonstrate that for wealthier slaveowners, the increasing stakes of the conflict effectively overrode their incentives to free-ride and avoid paying the costs of war.
However, in other circumstances we might expect the opposite relationship to hold true. Second, our paper highlights the potential for further research into the relationship between individual investments in political institutions, perceptions of the stakes of maintaining those institutions, and the implications of this for an individual’s willingness to participate in armed conflict. In cases where individuals’ wealth is not tied to the current institutional structure under threat, we might expect that their wealth likely makes them less willing to fight. However, in cases where their future wealth depends on preserving a particular institutional structure, and where the present civil conflict threatens this structure, these individuals might have a stronger incentive to fight.

We also hope that this paper demonstrates the benefits of linking the study of conflict to the study of American politics and history. We argue that there is much to be gained from using this and other similar historical cases. Perhaps because America’s internal violent conflicts fall into an unclear zone between the fields of American Politics, Comparative Politics, and International Relations, the subject often seems neglected. But the violent conflicts of America’s past have much to teach political scientists, both about conflict and about American political development. Moreover, as demonstrated by our ability to gather individual-level information on roughly 3.9 million Confederate citizens for this study, new efforts by archivists and genealogists offer unprecedented ways to study American political history on a comprehensive scale. In our view, this represents an opportunity to study fundamental questions about individual-level decision-making in key historical moments at both a scope and level of detail not previously possible.

The American Civil War was an incredibly destructive conflict, fought over a singularly horrifying institution, slavery. As research throughout political science shows, the ideas and motivations shaping why southerners fought for the Southern Confederacy did not die when the Civil War ended; they played a central role in shaping the long-run development of modern America (Key 1948; Acharya, Blackwell, and Sen 2018). Understanding why individuals would willingly risk their own lives to fight on behalf of this ugly system remains a question of the utmost importance in

---

28 Similar in spirit, though on a different topic, Harvey and West (2016) uses Civil War data on Union soldiers to examine the effects of discrimination in public accommodations on health, and Acharya, Blackwell, and Sen (2018) study the long-term effects of slavery on modern-day political behavior.
the social sciences, so that we can prevent such atrocities in the future, and also so that we can understand the conditions under which political extremism, in all its forms, does or does not lead to violent conflict. No doubt, there are many explanations, and no single answer. In this paper, we have used large-scale data on the Confederacy and the Confederate Army to shed light on one key component of the answer: the harsh reality of individual stakes in the institution of slavery. Contrary to some conventional wisdom, wealthier people joined the Confederate Army at higher rates than poorer people, probably because they had the most to gain from preserving a system of slavery that prioritized their own well-being over the freedom and well-being of others, and the most to lose from the system’s destruction.
References


Georgia Archives, University System of Georgia. n.d. “1832 Land Lottery.”


Key, V. O. 1948. *Southern Politics in State and Nation*.


Moore, Albert Burton. 1924. *Conscription and Conflict in the Confederacy*. University of South Carolina Press.


Sarris, Jonathan Dean. 2006. A Separate Civil War: Communities in Conflict in the Mountain South. University of Virginia Press.


Online Appendix

*Intended for online publication only.*

**Contents**

A  Data Construction Procedure 40
B  Sample Descriptive Statistics 45
C  Additional Observational Analyses 48
D  Randomization Validation 53
E  Additional Lottery Analysis 53
F  Data on Civil War Battle Locations 57
A Data Construction Procedure

This section describes the steps with which we combined the different datasets. Execution was done in the R programming language.

Dataset Overview

The paper deals with four individual-level, full name-containing datasets:

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Units</th>
<th>Smallest Geographic Information</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1850 Census</strong></td>
<td>All Free Citizens</td>
<td>County</td>
</tr>
<tr>
<td><strong>The 1850 Slave Schedule</strong></td>
<td>Slaveholders and slaves</td>
<td>County</td>
</tr>
<tr>
<td><strong>Confederate State Rosters</strong></td>
<td>Confederate soldier - entry</td>
<td>State</td>
</tr>
<tr>
<td><strong>1832 Winner’s List</strong></td>
<td>Lottery winner</td>
<td>County of origin</td>
</tr>
</tbody>
</table>

Several features of the data are worth further highlighting. Only the census comes with a set of unique numeric identifiers. We refer to the *serial ID* as IPUMS’ variable *serialp*, which is constructed to identify unique households. The availability of a identifier makes the Census the preferred dataset upon which to base others.

On the other hand, the confederate state rosters pose a challenge to measurement because each record in this dataset is an entry in the roster. Thus, the same soldier may appear twice on the list if at different times he served in different units or disconnected windows of time. There is no perfect way of distinguish name duplicates due to different people holding the same name with the same person entering under different units. As we detail in the following section, this challenge both makes the ‘M-over-N’ weighting meaningful but also introduces other concerns.

The terms involving merging can vary by statistical language. In this section we mostly rely on terminology from SQL or the SQL-based R package *dplyr*, and define them as follows:

- “Table”: Any rectangular dataset with records (rows) and variables (columns).
- “Key”: A set of variables that together uniquely defines each record in the table.
- “Duplicates/Is Unique”: Duplicates and unique rows with regard to the table’s key, unless otherwise noted.
- “Group A by K”: To collapse rows of A with the same key (K) values and summarize them by a statistic, such that the resulting table is unique in terms of its key.
- “Summarize Y by f”: To apply a function f such as *mean* or *max* to a variable during grouping (collapsing), in order to define a new variable Y.
- “Subset A to B”: To define a new table B by a subset of the rows of A based on some criteria.
- “Bind A and B”: To stack the rows of A and B together. Row-bind.
- “Left join from $A$ to $B$ on $K$”: For each row of $A$, find matching rows in $B$, where “matching” means that keys $K$ of both records are all the same. Subsequently append those records as new variables to the corresponding $A$ records. Note that this will never drop any rows from $A$. Duplicate rows of $B$ will lead to duplicating the corresponding rows of $A$, so often we want $B$ to be unique. Thus in practice this is similar to merge m:1 in Stata.

Pre-processing

We pre-process the raw census as follows, and use the resulting table throughout.

1. Remove any spaces and apostrophes from the first name. Do the same for last name.
2. Drop entries where names contain the characters * or ?. These characters were due to illegible census forms.
3. Drop entries where either the first name or last name is only two characters or less.
4. Drop entries with a birth year earlier than 1750, to remove outliers.

Combine Census and Slave Schedule

(1) Group the raw slave schedule, originally keyed on slaves, by their slave owners' full name and county. Summarize total slaves as the number of slave entries.

(2) Further group the slave count table by using soundex names rather than recorded names. Summarize total slaves again as the mean of previous slave totals.

(3) Subset the census to household heads, where these are defined by the relative codes \{ 101, 201, 501, 502, 601, 602 \}.

(4) Left join from (3) to (2) on soundex full name, county, and state.

(5) Group the unjoined slave schedule rows of (4) by first initial (instead of first soundex name), last soundex name, county, and state. Summarize total slaves again as the mean of previous slave totals.

(6) Left join from the unpopulated census rows of (4) to (5) on the keys of (5). Assign total slaves = 0 to all un-populated rows.

(7) Bind the unjoined rows of (4) and (6) itself.

(8) Group (7) by serial ID, and summarize the following household-level variables:

- total household slaves by the maximum of the number of the slave count
• *total 1-over-N slaves* by the maximum of the member’s slave count weighted by 1 over the number of duplicates in (3), which we shall refer to as \( N^{29} \).

• *M-over-N slaves* by the same idea as 1-over-N but where weights are further multiplied by the number of duplicates \( M \) in either (2) or (5).

• *at least one slave* by an indicator of whether *total household slaves* > 0.

• *1-over-N at least one slave* as the sum of the member’s 1-over-N weighted slave count, capped at 1.

• *M-over-N at least one slave* as equivalent to *1-over-N at least one slave* but with using the *M-over-N* weights.

We end up with a slave schedule table keyed on serial ID and the above six types of variables.

**Combine Census and Confederate Rosters**

1. Bind all confederate rosters, and generate soundex first names and last names.

2. Group (1) by soundex full name and state. All *in roster* entries for each row is 1.

3. Subset the census to men.

4. Left join from (3) to (2) on soundex full name and state. All *in roster* entries for each row is 1.

5. Group the unjoined roster rows of (4) to by first initial (instead of soundex first name), soundex last name, and state. Summarize *in roster* again as 1.

6. Left join from the unpopulated census rows of (4) to (5) on the keys of (5). Assign *in roster* = 0 to all unpopulated rows.

7. Bind the joined rows of (4) and (6) itself.

8. Group (7) by serial ID, and summarize the following household-level variables:

• *total confederate count* by the sum of individuals who have *in roster* = 1.

• *1-over-N confederate count* by the sum of individuals who have *in roster* = 1, but each weighted by 1 over the number of duplicates in (3) corresponding to that name.

• *M-over-N confederate count* by the same idea as 1-over-N but where the weights are further multiplied by the number of duplicates \( M \) in either (2) or (5).

• *at least one confederate soldier* by an indicator of whether *total confederate count* > 0.

---

\(^{29}\) We generally refer to the duplicates of the census as \( N \) and the duplicates of the other datasets as \( M \).
• 1-over-N at least one confederate soldier by 1 minus the product of 1 minus the 1-over-N weights for each row of (7)^30.

• M-over-N at least one confederate soldier as equivalent to 1-over-N at least one slave, but with using the M-over-N weights.

We end up with a confederate table keyed on serial ID and the above six types of variables.

Combine Census and the Lottery Winner’s List

(1) To create the subset of lottery entrants, subset the 1850 census to individuals with the criteria defined in Bleakley and Ferrie (2016). Essentially this finds fathers with at least one co-residing, 1829 - 1832 Georgia-born child.

(2) Subset the census to all household members with a household head in (1).

(3) Group (1) by serial ID, randomly choosing an entrant in the rare case that there are multiple entrants in a single household. Keep track of the last name and first name duplicates among these household representatives, refer to these as N.

(4) For every record in (3) (which is keyed on household and not necessarily on name), compute the aline phonetic string distance between the every record in the winner’s list. Specifically, first compute the string distance between the first names (not soundexed). Then compute the string distance between the last names (also not soundexed). Declare the pair a match if both distances are less than 0.05 (the distance metric ranges from 0 to 1). Keep track of the number of winner’s list matches each census record collects, refer to these as M.

(5) Summarize (4) the following household-level variables:

- treat 1-over-N by whether or not the entrant name has at least one match, divided by N the number of name duplicates in (3).
- treat M-over-N by whether or not the entrant name has at least one match, divided by N and multiplied by M the number of matches the pair has accrued.
- treat by an indicator of whether or not 1-over-N is 1. This is essentially assigning a 1 to names which accrued at least one match but erring on the side of limiting false positives for those winning names in which multiple potential entrants exist.

^30 For example, consider a serial ID containing three males with various assigned weights (they would be different because their names are different). If the three weights are \((\frac{1}{2}, \frac{1}{2}, \frac{1}{2})\), then this summary will give us \(1 - (0 \times 0 \times 0) = 1\). If the weights were \((\frac{2}{3}, \frac{1}{2}, \frac{1}{4})\), then this summary will give us \(1 - (\frac{1}{3} \times \frac{1}{2} \times \frac{1}{4}) = \frac{7}{8}\). If the three weights were \((\frac{1}{2}, \frac{1}{2}, 0)\), then this summary will give us \(1 - (0 \times \frac{1}{2} \times \frac{1}{2}) = 1\).
(6) Left join from the census subset in (2) to (5) on serial ID.

This leaves us with an individual-level table where treatment is assigned at the household level.

Putting the observational dataset together

(1) Group the pre-processed census by serial ID, summarizing socio-demographic variables such as the total real estate property value.

(2) Left join from (1) to the final output of section A on serial ID.

(3) Left join from (2) to the final output of section A on serial ID. The order between (2) and this is irrelevant.

Putting the experimental dataset together

(1) Group the final output of section A by serial ID, summarizing socio-demographic variables such as the total real estate property value.

(2) Prepare a new slave schedule and confederate roster dataset starting from the final output of section A as the census data, but otherwise following the exact same steps as in sections A and A. This will generate a table keyed on serial ID among the lottery entrant’s descendants.

(3) Combine the three datasets in the same way as section A

Sources of Merge Error and How They Affect Results

The task of merging together the 1850 census, the roster of Confederate soldiers, and the roster of slaveowners is not a perfect process. In addition to cleaning the raw data in a number of ways, we try to reconcile slight differences in spelling between names by using the soundex encoding of first and last names. Where possible, we also use geographic information to improve our merges; nonetheless, we know that merge errors remain. Before moving on, we need to consider how errors in this process could affect the conclusions we drew above. As a first problem, suppose that some names are simply easier to merge than others. Hard-to-merge names in the 1850 census will therefore be less likely to be found in the Confederate roster and in the slaveowners’ roster. This type of error could induce a false correlation between slave ownership and fighting, because it produces faulty observations where we mark people as both non-slaveowners and non-soldiers. To try to address this problem, we re-estimate our results with the addition of last-name fixed effects. When we do this, comparisons are only made among individuals with the same last name, which should help control for the ease of merging different names. Tables C3 and C4 present the results. The results on number of soldiers are quite similar; the results on the fraction of eligible men who fight also remain positive and meaningful, though they are smaller in magnitude than before.

A second problem is that some names, like John Smith, are very common. This is not a serious problem for matching the slave roster, because we also know the county each person resides in, but
for the Confederate roster we know only names and states, for the most part. Having the state helps to some degree, but there are still a significant number of duplicate matches. Imagine, as a hypothetical, that we had 10 John Smiths in the census, 5 of whom are slaveowners, and 5 John Smiths in the Confederate roster. Without more information that we do not have, we have no way of knowing if 5 of the John Smiths are slave-owning soldiers, or if 5 John Smiths are soldiers and non-slaveowners while the other 5 John Smiths are non-soldiers and slaveowners, or anything in between. To address this problem, we re-estimate our results only for the 547,747 households in the 1850 census for which the household head has a unique name within each state. Among this set of households, we can be virtually certain that, if we find them in the slave roster or in the Confederate army roster, it is the same household. Tables C5 and C6 present the results. We continue to find similar results; across all specifications, slave ownership continues to be strongly predictive of fighting in the Confederate army. Interestingly, when we account for fertility effects by looking at the fraction of sons, and we look only at unique names, the relationship with property ownership decreases further and becomes negative. However, these negative coefficients are smaller in magnitude than those on slaveowning.

B Sample Descriptive Statistics

Figure B1 shows the the geographic distribution of slaveowning households as of 1850, while Figure B2 shows the geographic distribution of Confederate army membership during the Civil War. Data for what became the state of West Virginia is not shown.
Figure B1 – Proportion of Slaveowning Households in 1850, by County

Source: 1850 U.S. Census and Slave Schedule
Figure B2 – Proportion of Households with at least one Member in the Confederate Army, by County

Households with at least one Confederate soldier: 40%, 60%, 80%

Source: 1850 U.S. Census and Confederate State Rosters
C Additional Observational Analyses

Multivariate Regressions

For a more formal treatment of the overall results when considering who fought for the Confederate Army as a whole, in Table C1 we run regressions to investigate the associations between slave ownership, property ownership, and fighting in the Confederate Army. These regressions reflect the same relationships as Figures 1 and 2 have already shown.

We present four specifications. The first three columns include, respectively, a dummy variable for slave ownership, a dummy for property ownership, and both dummies in one regression. Table C1 shows how households that own slaves send, on average, 0.12 – 0.32 more household members to the Confederate Army, conditional on other measures of wealth.

The fourth column if Table C1 includes linear specifications for number of slaves, scaled as 10s of slaves, and the value of property, measured as 1,000s of 1850 dollars. These specifications are included interacted with the indicators, so that the function does not reflect the jump from non-ownership to ownership. As we see, both slopes are small relative to the coefficients on their dummy variable counterparts. In particular, the associated increase in fighting from a marginal increase in 1000s of dollars of real estate property value is not distinguishable from zero. The increase in army membership for slave and property owners seems to be mainly related to the jump from non-ownership to ownership.

As we have stressed, these are observational relationships. There are many possible reasons slaveowners and property owners fought in the Confederate Army at higher rates. One possible explanation is that these patterns are merely the result of wealthier households having more sons. Could the positive association between wealth and army membership be explained by the fact that households who own slaves also contain more military-eligible sons, so that their army membership totals are higher even if the actual propensity for each individual son to fight is not? In Table C2, we re-estimate the same specifications from Table C1, but with a new outcome variable that measures the fraction of eligible men in a given household who fought in the Confederate Army. We still find that a higher fraction, roughly 6 percentage points more, of eligible men fought in the Confederate Army from households that owned slaves. However, the coefficient on property wealth drops significantly, suggesting that the large coefficients in Table C1 may have been biased upward by a household size confounder.

Robustness to Alternative Specifications

We next ask if the positive correlations of slave ownership and wealth ownership are robust to alternative specifications and potential consequences of merge error.

First, Tables C3 and C4 re-estimate the regressions in Tables C1 and Table C4, respectively, with fixed effects and clustered errors. We apply two types of fixed effects: one for each 1850 county of residence and another each soundex-encoded last name in the population. County fixed effects control away cross-geographic confounding in the relationship between wealth and fighting
by examining within-county differences in wealth. Last name fixed effects aim to control away confounding variables that varies across ethnic and socioeconomic status, proxied by lineage (last name), by examining within last name differences in wealth. We then allow for errors within each group to be correlated by clustering standard errors.\footnote{We do this in light of the possibility that our “treatment” of interest, slave-wealth, may exert an effect at the respective group-level. We do not cluster our standard errors in our subsequent lottery estimates because in that case our treatment variable is assigned at the individual level.}

Second, Tables C5 and C6 re-estimate the regressions in Tables C1 and Table C4, respectively, on a subset of the population that is unaffected by certain types of merge error (see Section A for the concern). We only keep people in the Census whose (soundex-encoded) full name is unique within their state, and compute all measures of household-level slaveownership and fighting on this subset. In this way, we account for the consequences of certain names being differentially matched to our dependent variable. We then apply the same regression specifications Tables C1 and Table C4.

In all four tables, we see our most important finding hold. That is, a household’s slave-ownership in 1850 is positively correlated with the level of fighting in the 1860s. These estimates are distinguishable from zero even after we allow for errors to be correlated within group.
Table C1 – Predictor of Fighting in the Confederate Army. Examines the full population of white, male Confederate citizens linked to the 1850 Census. Slaveowners and property owners fought in the Confederate army at higher rates than those without slaves or without property.

<table>
<thead>
<tr>
<th>Owns Slaves</th>
<th>Owns Property</th>
<th># Slaves (in 10s) · Owns Slaves</th>
<th>Property Value (in 1000s) · Owns Property</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.320</td>
<td>0.586</td>
<td>0.017</td>
<td>0.0000</td>
</tr>
<tr>
<td>(0.005)</td>
<td>(0.004)</td>
<td>(0.003)</td>
<td>(0.0004)</td>
</tr>
</tbody>
</table>

Each regression includes 746,506 households. Robust standard errors in parentheses. In column 4, number of slaves measured in 10s. Property value in thousands of 1850 dollars.

Table C2 – Predictors of the Fraction of Household Men who Fought in the Confederate Army. Estimates the same regressions as Table C1 with the fraction of men who fought as the dependent variable.

<table>
<thead>
<tr>
<th>Owns Slaves</th>
<th>Owns Property</th>
<th># Slaves (in 10s) · Owns Slaves</th>
<th>Property Value (in 1000s) · Owns Property</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.056</td>
<td>0.021</td>
<td>0.006</td>
<td>-0.002</td>
</tr>
<tr>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
</tbody>
</table>

Each regreession includes 746,506 households. Robust standard errors in parentheses. In column 4, number of slaves measured in 10s. Property value in thousands of 1850 dollars.
Table C3 – Number of Men in Households who Fought in the Confederate Army, with County and Last Name Fixed Effects. Examines the full population of households in Confederate states linked to the 1850 Census. Slave-owners and property owners fought in the Confederate army at higher rates than those without slaves or without property.

<table>
<thead>
<tr>
<th>Number of Soldiers in Household</th>
<th>Owns Slaves</th>
<th>Owns Property</th>
<th># Slaves (in 10s) · Owns Slaves</th>
<th>Prop. Value · Owns Property</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.368</td>
<td>0.537</td>
<td>0.015</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.009)</td>
<td>(0.004)</td>
<td>(0.001)</td>
</tr>
<tr>
<td></td>
<td>0.253</td>
<td>0.578</td>
<td>0.017</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.008)</td>
<td>(0.001)</td>
<td>(0.000)</td>
</tr>
<tr>
<td></td>
<td>0.208</td>
<td>0.484</td>
<td>-0.007</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.010)</td>
<td>(0.002)</td>
<td>(0.000)</td>
</tr>
<tr>
<td></td>
<td>0.062</td>
<td>0.562</td>
<td>0.012</td>
<td>-0.006</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.008)</td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td></td>
<td>0.190</td>
<td>0.475</td>
<td>-0.006</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.010)</td>
<td>(0.002)</td>
<td>(0.001)</td>
</tr>
<tr>
<td></td>
<td>0.048</td>
<td>0.558</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.008)</td>
<td>(0.003)</td>
<td></td>
</tr>
</tbody>
</table>

County Fixed Effects ✓ ✓ ✓ ✓ ✓ ✓
Last Name Fixed Effects ✓ ✓ ✓ ✓ ✓ ✓

Each regression includes 746,506 households. Clustered robust standard errors, where clusters are groupings used in fixed effects, in parentheses. In column 4, number of slaves measured in 10s. Property value in thousands of 1850 dollars.

Table C4 – Fraction of Household Men who Fought in the Confederate Army, with County and Last Name Fixed Effects. Repeats regressions in Table C3 but with fraction of men as a dependent variable.

<table>
<thead>
<tr>
<th>Fraction of Eligible Men in Household Who Fought</th>
<th>Owns Slaves</th>
<th>Owns Property</th>
<th># Slaves (in 10s) · Owns Slaves</th>
<th>Prop. Value · Owns Property</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.066</td>
<td>0.010</td>
<td>0.003</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.002)</td>
<td>(0.001)</td>
<td>(0.000)</td>
</tr>
<tr>
<td></td>
<td>0.027</td>
<td>0.017</td>
<td>-0.007</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.001)</td>
<td>(0.000)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.068</td>
<td>0.012</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.002)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.023</td>
<td>-0.006</td>
<td>0.021</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.002)</td>
<td>(0.000)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.067</td>
<td>0.014</td>
<td>0.044</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.002)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.021</td>
<td>-0.001</td>
<td>0.004</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
</tr>
</tbody>
</table>

County Fixed Effects ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓
Last Name Fixed Effects ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓

Each regression includes 701,900 households. Clustered robust standard errors, where clusters are groupings used in fixed effects, in parentheses. In column 4, number of slaves measured in 10s. Property value in thousands of 1850 dollars.
Table C5 – Number of Men in Households who Fought in the Confederate Army; Unique Names Only. Examines the subset of free citizens in the 1850 Census whose full names were unique within their state of residence.

<table>
<thead>
<tr>
<th></th>
<th>Number of Soldiers in Household</th>
</tr>
</thead>
<tbody>
<tr>
<td>Owns Slaves</td>
<td>0.292</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
</tr>
<tr>
<td>Owns Property</td>
<td>0.372</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
</tr>
<tr>
<td># Slaves (in 10s) · Owns Slaves</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
</tr>
<tr>
<td>Property Value · Owns Property</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
</tr>
</tbody>
</table>

Each regression includes 547,747 households. Robust standard errors in parentheses. In column 4, number of slaves measured in 10s. Property value in thousands of 1850 dollars.

Table C6 – Fraction of Household Men who Fought in the Confederate Army; Unique Names Only. Repeats regressions in Table C5 but with fraction of men as a dependent variable.

<table>
<thead>
<tr>
<th></th>
<th>Fraction of Men in Household Who Fought</th>
</tr>
</thead>
<tbody>
<tr>
<td>Owns Slaves</td>
<td>0.048</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
</tr>
<tr>
<td>Owns Property</td>
<td>-0.031</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
</tr>
<tr>
<td># Slaves (in 10s) · Owns Slaves</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
</tr>
<tr>
<td>Property Value · Owns Property</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
</tr>
</tbody>
</table>

Each regression includes 412,541 households. Robust standard errors in parentheses. In column 4, number of slaves measured in 10s. Property value in thousands of 1850 dollars.
D Randomization Validation

To validate the experimental analysis, we verify that our treated and control groups look similar on pre-treatment covariates. We conduct these balance tests for two main reasons—first, because we might be concerned that the lottery’s randomization was compromised, somehow, and second, to address the possibility that our phonetic string-matching procedure, by which we located individuals in the 1850 census, could differentially match people in a way correlated with their potential outcomes (including their wealth and their propensity to fight in the Confederate Army).

Figure D3 checks for differences between lottery winners and losers on the three main pre-treatment covariates we can observe. Each panel plots the standardized difference-in-means for each of the three treatment variable specifications (the simple binary treatment indicator, the 1-over-N treatment variable, and the M-over-N treatment variable).\(^{32}\)

All differences are negligible relative to their variance, on the order of 0.01 to 0.1 standardized differences. For instance, in the left panel, we examine whether lottery winners are systematically older or younger than lottery losers. Across all three treatment variables, we see that lottery winners are very slightly younger than lottery losers. The unstandardized difference is 0.4 years in terms of birth year, not a major difference. It is almost identical to the difference reported in Bleakley and Ferrie (2016, Table 1, 1467). In the remaining panels we also examine standardized differences of the number of characters in entrant’s names, the rate at which their last name starts with either ‘M’ or ‘O’, a common indicator for Celtic origin (Bleakley and Ferrie 2016), and the number of children and step-children in the household born in 1832 or earlier. Most differences are negligible, except perhaps that from the M-over-N indicator. As such, we will favor the other two specifications in our analyses.

E Additional Lottery Analysis

This section offers additional robustness checks for the lottery findings presented in the paper.

---

\(^{32}\) We use the standardized difference-in-means rather than the \(t\)-statistic as our test statistic of assessing covariate balance, following Imbens and Rubin (2015, 310-311). Given the large sample sizes, test statistics are likely to be statistically significant even if imbalances are not meaningfully large.
Figure D3 – Balance of Pre-Treatment Covariates in Lottery Sample.
Tests for imbalances in four observed pre-treatment covariates from the 1850 Census: year of birth (leftmost panel); number of characters in household head’s full name (center-left panel); whether or not the household head’s last name starts with ‘M’ or ‘O’, a common indicator of Celtic origin (center-right panel); and the number of children in the household born 1832 – the year of the lottery – or earlier (rightmost panel). Points represent standardized differences between lottery winners and losers, using one of three treatment specifications. No substantive imbalances are found.
Table E7 – Effect of Winning 1832 Lottery on Household Confederate Army Membership; Unique Households Only.

<table>
<thead>
<tr>
<th></th>
<th>Number of Confederate Soldiers in Household</th>
<th>Probability at Least One Son Fights</th>
<th>Fraction of Sons Who Fight</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No FEs</td>
<td>Last Name FEs</td>
<td>First Name FEs</td>
</tr>
<tr>
<td>Lottery Winner</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.29</td>
<td>0.14</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.05)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>Lottery Winner ($\frac{1}{n}$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.37</td>
<td>0.23</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.05)</td>
<td>(0.07)</td>
</tr>
</tbody>
</table>

Each cell is a regression coefficient. Sample size in all regressions is 11,681 households. Outcome variable in top panel is the number of registered Confederate soldiers in the household. Outcome variable in middle panel is fraction of sons in household who fight in Confederate Army. Outcome variable in bottom panel is indicator for whether at least one son in household fights in Confederate Army. Estimates in first row of each panel use binary treatment indicator based on unique name matches. Estimates in second row of each panel include non-unique name matches, where treatment variable takes the value $\frac{1}{n}$ for a lottery winner name matched to $n$ households in 1850 Census. Robust standard errors in parentheses.
Table E8 – Effect of Winning 1832 Lottery on Household Confederate Army Membership, where the dependent variable is weighted in the M-over-N specification.

<table>
<thead>
<tr>
<th></th>
<th>Number of Confederate Soldiers in Household</th>
<th>Probability at Least One Son Fights</th>
<th>Fraction of Sons Who Fight</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No FEs Last Name FEs First Name FEs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lottery Winner</td>
<td>0.35 (0.05)</td>
<td>0.05 (0.01)</td>
<td>0.07 (0.01)</td>
</tr>
<tr>
<td>Lottery Winner ((\frac{1}{n}))</td>
<td>0.28 (0.05)</td>
<td>0.03 (0.01)</td>
<td>0.05 (0.01)</td>
</tr>
</tbody>
</table>

Each cell is a regression coefficient. Sample size in all regressions is 13,414 households. Outcome variable in top panel is the number of registered Confederate soldiers in the household. Outcome variable in middle panel is fraction of sons in household who fight in Confederate Army. Outcome variable in bottom panel is indicator for whether at least one son in household fights in Confederate Army. Estimates in first row of each panel use binary treatment indicator based on unique name matches. Estimates in second row of each panel include non-unique name matches, where treatment variable takes the value \(\frac{1}{n}\) for a lottery winner name matched to \(n\) households in 1850 Census. Robust standard errors in parentheses.
Data on Civil War Battle Locations

Figure 3 indicates that the Atlanta Campaign invasion route appears to have run through the lands allocated in the Cherokee Land Lottery of 1832. This fact raises a concern for an alternative mechanism: households who won the lottery may have moved to areas in which they perceived more danger of property destruction, which in turn prompted them to fight. Under this logic, wealthier individuals are not fighting due to their perception that the stakes of the conflict are higher due to the potential abolition of slavery. Rather, they happen to live in areas that are more likely to be threatened by the Union armies moving south.

To address this possibility we identified which counties experienced at least one Union battle during the Civil War. While some county boundaries changed between 1850 and the Civil War, we use this variable to approximate the danger households may have perceived during the War.

We find that the land lottery made individuals approximately 1.7 percentage points more likely to live in a county that had a Civil War battle (Table F9), with about 8 percentage of the control group living in such counties. The small size of this effect paired with our historical knowledge about how Union generals chose their marching routes helps assuage concerns that our results are exclusively driven by this alternative mechanism. For example, as General Sherman was planning the route for his infamous march, he did so explicitly targeting wealthier areas (Feigenbaum, Lee, and Mezzanotti 2017). This means that at least part of the effect we observe for the relationship between wealth and likelihood of observing a battle will be a function of the strategic choices of the Union generals, rather than individuals throughout the Southern Confederacy. The small size of this effect paired with our historical knowledge about how Union generals chose their marching routes helps assuage concerns that our results are exclusively driven by this alternative mechanism.

### Table F9 – Effect of Winning the Lottery on Living in a Battle-Relevant Area.

<table>
<thead>
<tr>
<th>Lives in a County that Experienced Union Battle</th>
<th>No FEs</th>
<th>Last Name FEs</th>
<th>First Name FEs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lottery Winner</td>
<td>0.017</td>
<td>0.015</td>
<td>0.013</td>
</tr>
<tr>
<td>(0.007)</td>
<td></td>
<td>(0.008)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Lottery Winner ((\frac{1}{n}))</td>
<td>0.017</td>
<td>0.017</td>
<td>0.013</td>
</tr>
<tr>
<td>(0.007)</td>
<td></td>
<td>(0.007)</td>
<td>(0.008)</td>
</tr>
</tbody>
</table>

Each cell is a regression coefficient. The outcome variable is whether or not a 1850 household lives in a county which, during the Civil War, saw at least one Union battle. Robust standard errors in parentheses.

---

33 We thank a reviewer for making this point.

34 These data on the location of battles was recorded mainly by the National Park Service. We thank James Feigenbaum for sharing the tabulated data.