

NBER WORKING PAPER SERIES

MIGRATION RESPONSES TO CONFLICT:
EVIDENCE FROM THE BORDER OF THE AMERICAN CIVIL WAR

Shari Eli
Laura Salisbury
Allison Shertzer

Working Paper 22591
<http://www.nber.org/papers/w22591>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2016

We thank Ran Abramitzky, Jeremy Atack, Martha Bailey, Hoyt Bleakley, Leah Boustan, Ann Carlos, William Collins, Paul David, Rowena Gray, Tim Guinnane, Eric Hilt, Naomi Lamoreaux, Frank Lewis, Josh Lewis, Paul Rhode, Alex Whalley, seminar participants at Guelph, Michigan, Stanford, U.C. Merced, Vanderbilt, and Yale, as well as participants at meetings of the Canadian Network for Economic History and the Social Science History Association for helpful comments. We also thank Zvezdomir Todorov for excellent research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2016 by Shari Eli, Laura Salisbury, and Allison Shertzer. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Migration Responses to Conflict: Evidence from the Border of the American Civil War
Shari Eli, Laura Salisbury, and Allison Shertzer
NBER Working Paper No. 22591
September 2016, Revised September 2016
JEL No. J61,N31,R23

ABSTRACT

The American Civil War fractured communities in border states where families who would eventually support the Union or the Confederacy lived together prior to the conflict. We study the subsequent migration choices of these Civil War veterans and their families using a unique longitudinal dataset covering enlistees from the border state of Kentucky. Nearly half of surviving Kentucky veterans moved to a new county between 1860 and 1880. There was no differential propensity to migrate according to side, but former Union soldiers were more likely to leave counties with greater Confederate sympathy for destinations that supported the North. Confederate veterans were more likely to move to counties that supported the Confederacy, or if they left the state, for the South or far West. We find no evidence of a positive economic return to these relocation decisions.

Shari Eli
Department of Economics
University of Toronto
150 St. George Street
Toronto, ON M5S 3G7
CANADA
and NBER
shari.eli@utoronto.ca

Allison Shertzer
Department of Economics
University of Pittsburgh
4901 WW Posvar Hall
230 South Bouquet Street
Pittsburgh, PA 15260
and NBER
shertzer@pitt.edu

Laura Salisbury
Department of Economics
York University
Vari Hall 1092
4700 Keele Street
Toronto, ON M3J 1P3
CANADA
and NBER
lsalisbu@yorku.ca

1 Introduction

More than half of the nations around the world have faced armed conflicts or civil wars during the last fifty years. Recent literature has shown that civil war is linked to low per capita incomes and slow economic growth.¹ While there has been a substantial increase in scholarship on the topic of civil war in the economic development literature, persistent challenges associated with following survivors over time make it difficult to measure the long-term economic effects of such conflicts on participants. Moreover, this literature emphasizes the direct negative impact of civil conflict on physical infrastructure, health, or human capital accumulation rather than the negative impact on a community's social fabric.² These effects are nevertheless important to study because ideological divisions between victors and the defeated may lead to a lack of economic integration, even after hostilities have ceased. The lingering social consequences of civil conflict may be particularly acute in settings where opposing groups live in close physical proximity to one another.

In this paper, we investigate the economic impact of social frictions generated by civil conflict. In particular, we study the impact of the American Civil War on individuals from the border area between the Union and the Confederacy, where families who would ultimately support different sides lived in the same communities. We focus on soldiers from the border state of Kentucky and ask how the conflict influenced where survivors chose to live later in life. This context provides a unique opportunity to investigate the consequences of social divisions between veterans from opposing sides of a civil conflict because Kentucky contributed a significant number of enlistees to both the Union and Confederate Armies.³ Moreover, soldiers from both sides of the Civil War can be identified using Kentucky enlistment records and followed through time using federal censuses.

Economists typically model migration as an investment: individuals migrate in order to maximize their expected lifetime earnings net of mobility costs. In the case of civil conflict, animosity between former combatants may generate migrations which would not have happened for economic reasons in the absence of the conflict. In other words, by imposing social costs on participants, civil conflict may lead to economically inefficient migration behavior. This type of migration is another potential cost associated with civil war. On the other hand, conditional on a civil conflict

¹See Blattman and Miguel (2010) for a review of recent literature.

²Miguel and Roland (2011) measure the effect of exposure to bombing during the Vietnam War on human capital attainment; Bundervoet et al (2009) look at the effect of exposure to conflict in Burundi on child health; Annan and Blattman (2010) measure the effects of being conscripted into military service in Uganda on human capital attainment and income in later life.

³While there was enlistment on both sides in all border states – Maryland, Delaware, Kentucky and Missouri – enlistees were most evenly split between the Union and Confederate army in Kentucky.

having occurred, migration may reduce prolonged exposure to enemy combatants, and may thus limit ongoing violence. In any case, post-war migration is an important vehicle through which civil conflict may affect the later-life outcomes of participants.

We construct a novel longitudinal database of Union and Confederate recruits from Kentucky by matching military records from the state's regiments to the Federal Census of 1860. We then link recruits forward to the Federal Census of 1880. This allows us to measure, as well as control for, selection into each army on the basis of observable socioeconomic status. In addition, we are able to observe recruits' county of residence prior to enlistment, which allows us to infer whether recruits were living in places where Union or Confederate status would have been socially rewarded. Thus, we can determine whether Union recruits were "pushed" out of counties that were socially aligned with the Confederacy and "pulled" toward counties that were socially aligned with the Union. The longitudinal nature of our database allows us to address concerns of differential migration patterns that were the result of differences in skill as opposed to social rewards or penalties from military service. For example, if Union recruits were systematically less skilled, and the return to skill is higher in counties sympathetic to the Confederacy, we should expect Union veterans to leave Confederate-leaning counties for economic reasons alone.⁴ Therefore, the ability to observe ex-ante characteristics of recruits, such as occupational attainment and wealth, is a major advantage of our research design.

We document a series of facts about how ideology, socioeconomic characteristics, and participation in the Civil War interacted to shape the later life of Kentucky veterans. First, although Confederate enlistees tended to come from wealthier families, there was no difference in the propensity to migrate by side. However, Union soldiers were more likely to migrate the greater the support for the Confederacy in their home counties. These veterans settled in counties that were more pro-Union on average. Confederate soldiers, for their part, were more likely to choose Confederate-leaning counties or states in the far West or South if they moved. More than half of Kentucky enlistees migrated out of their home county between 1860 and 1880, so the degree of resorting was significant. We find that the drive to migrate was stronger for veterans themselves compared with their families: relatives who did not fight were both less likely to migrate and less responsive to the ideology of their home county. If these family members did move, however, they exhibited a similar preference for like-minded destination counties.

Our findings about the skill selectivity of migrants also point to social rather than purely

⁴See Borjas (1987) for a discussion of selective migration and the return to skill.

economic motives for migration among veterans. Most studies of internal migration during the mid-19th century find that migrants were negatively selected on skill, measured by occupational status (Ferrie 1997; Stewart 2006; Salisbury 2014). While we find some evidence that Union and Confederate veterans who migrated out of their their pre-Civil War county were negatively selected on family wealth, we find no evidence of negative selection on occupational attainment. Our results suggest that the economic drivers of migration among veterans were different from other contemporaneous episodes of internal migration.⁵ We also investigate the gains associated with this ideological resorting, and we find little evidence that moving out of one’s home county led to an increase in occupational income for either Union or Confederate veterans. While our result is more consistent with the existing literature on short-distance moves during this period (Salisbury 2014), these findings suggest that the economic gains from socially-motivated migration after the Civil War may have been minimal.⁶ However, there may have been non-pecuniary benefits associated with moving to an area with more ideologically similar residents such increases in emotional wellbeing that are difficult to detect using census data.

Our findings relate to the literature in economic history on the post-Civil War outcomes of Union Army veterans. Costa (1995, 1997) used Union army veterans to study the impact of pension income on retirement and living arrangements, Eli (2015) studied income effects on the health of Union Army veterans, and Salisbury (2016) investigated the impact of Union Army widows’ pensions on remarriage. Bleakley, Cain and Ferrie (2014) study labor market discrimination among Union army veterans. Costa and Kahn (2008) explicitly measure the impact of the war on veterans by examining how unit cohesion affects later life outcomes and find that deserters were more likely to leave their home towns after the war. In more recent work, Costa, Kahn, Roudiez, and Wilson (2016) find that Union Army veterans co-located with men from their former companies. Given this literature, an advantage of our paper is that we are able to study the migration behavior of both Union and Confederate Army veterans. Union and Confederate men both preferred to live in like-minded communities.

⁵Long and Siu (2016) study migration prompted by a natural disaster – the 1930s Dust Bowl – and also find little evidence that migrants were selected on skill.

⁶Most studies of historical migration to or within the United States find evidence of large economic gains; however, these are rarely expressed in terms of occupational attainment alone. Rather, they embed gains migrants experience by moving to areas with higher average wages (or land availability in the case of the frontier). For instance, frontier migration during the mid to late 19th century was associated with significant wealth accumulation (Ferrie 1997; Stewart 2006). Similarly, pre-WWI Norwegian immigration to the United States (Abramitzky, Boustan, and Eriksson 2012) and pre-WWII black migration from the American South (Collins and Wanamaker 2014) were associated with wage gains of 55 to 70 log points, respectively. Our findings are not strictly comparable to these studies; however, they are suggestive that the economic gains from migrating out of Kentucky after the Civil War were not large.

2 Historical Background

The Civil War began on April 12, 1861 when Confederate ships attacked the Union Army at Fort Sumter, South Carolina and ended on April 9, 1865 when Robert E. Lee surrendered at Appomattox Courthouse in Virginia. Approximately 2.2 million men served on the side of the Union (North) and 1.1 million men served for the Confederacy (South). Kentucky was one of four “border states,” or slave-owning states that did not secede from the Union; Missouri, Maryland, and Delaware were the others. In this section, we provide historical background on Kentucky’s role in the Civil War, as well as what is known about the post-war experiences of Civil War veterans.

In general, the literature on the economic consequences of the Civil War emphasizes the experience of the southern U.S. Much less is written about the consequences of the war in border regions, where individual communities contributed troops to both sides. In the border states, the “civil” nature of the Civil War is most apparent. This paper contributes to the historical literature on the American Civil War by offering new insight into the experience of border areas.

2.1 Kentucky during the War

During the Civil War, Kentucky – a border and slave state – did not secede from the Union. As in other border states, pro-Confederate and pro-Union supporters lived alongside each other (both Union President Abraham Lincoln and Confederate President Jefferson Davis were born in Kentucky). Tobacco, whiskey, snuff and flour produced in Kentucky were exported to the South and Europe via the Ohio and Mississippi rivers and to the North by rail. Therefore, Kentucky’s economy relied on markets in the Union and the Confederacy. In addition, though most Kentuckians owned no slaves, a share of them were heavily involved in the profitable export of slaves to the Deep South.

In general, antebellum Kentucky was solidly proslavery but decidedly more moderate than most states in the Deep South. This was common among border states, where all governors elected during the late 1850s were proslavery Democrats (Phillips 2013, p 5). At the same time, as Astor (2012, p 9) argues, Kentuckians and other border residents “rarely viewed the national debates over slavery as irreconcilable,” due to social ties with the Midwest and the frequency with which free wage labor and hired slave labor interacted in factories and farms.⁷ This relative moderation is

⁷Astor (2012) notes that most Kentucky slaveholders owned fewer slaves than their counterparts in the Deep South. It was common practice to “hire out” slaves to factories, or to hemp or tobacco plantations, during harvest season. As such, the institution of slavery differed in many respects in the border region compared to the South. Astor (2012) argues that this encourage the Kentucky electorate to believe that “northern” and “southern” modes of production could easily exist side by side, leading to a political culture that split the difference between North and South.

clear from Kentucky's behavior in the 1860 presidential election, in which the westward expansion of slavery was a major campaign issue. While northern states voted overwhelmingly in favor of Abraham Lincoln's Republicans, who explicitly favored banning slavery in all U.S. territories, and southern states voted overwhelmingly in favor of John C. Breckenridge's Southern Democrats, who explicitly favored the protection of slavery in the territories, Kentucky voted in favor of John Bell, a moderate with respect to slavery. Bell headed the Constitutional Union party, which consisted largely of moderate ex-Whigs who found the Republican party too "radical." His party's platform avoided the question of slavery in the territories altogether. Bell carried 60 of Kentucky's 110 counties, and Breckenridge placed second, carrying 43 counties (Harrison 1975, p. 4).

While Kentucky considered the possibility of secession, most Kentuckians were "Conservative Unionists" (Astor 2012; Phillips 2013). It was not the case that all slaveowners wished to secede while all non-slaveowners wished to remain in the Union: many Kentucky slaveholders felt that their interests were better served within the Union than outside it. This desire was likely due to the fact that Kentucky shared its northern border with free states, and Kentucky slaveholders were concerned about relinquishing the protections they currently enjoyed under the Fugitive Slave Act. As prominent Kentucky attorney Joseph Holt argued, if Kentucky were to secede, it would "virtually have Canada brought to her door, denying the state's slaveholders legal protections to prevent enslaved people from fleeing northward to freedom" (quoted in Phillips 2013, p 12).

Kentucky initially tried to remain neutral; this proved impossible when the Confederate army invaded in the fall of 1861 and federal troops subsequently occupied the state. While the majority of Kentuckians initially favored remaining in the Union, public opinion changed over the course of the war. Many Conservative Unionists objected to Lincoln's troop call-up and the behavior of occupying federal troops in their state; even more objected to the Emancipation Proclamation and subsequent enlistment of former slaves in the Union Army. At the same time, many Kentuckians felt alienated by Confederate raids during 1862 and 1863 (Astor 2012; Harrison 1975). While, as Astor (2012) argues, few Kentuckians formally switched allegiance after early 1862, public opinion had become decidedly anti-Lincoln by the end of the war. Kentucky was one of the few states not to vote for Lincoln in the 1864 presidential election. In the end, approximately 100,000 Kentuckians served for the Union side, while 30-40,000 served on the Confederate side (Phillips 2013; Marshall 2010). Nonetheless, many whites of military age did not enlist – 187,000 by one estimate – which stands as further evidence of the state's ambivalence (Phillips 2013).

3 Data

Our dataset consists of linked military and census records. In this section, we describe the data from each source and the procedure by which records were linked.

3.1 Military Records

We begin with a collection of military records from the genealogical website fold3.com (U.S. War Department, 1890-1912). These data consist of indexes to compiled service records, which include muster rolls and other documents collected from the War Department and the Treasury Department. These records exist for both Union and Confederate soldiers; however, they are likely more complete for Union soldiers. The indexes to these record collections contain the recruit's regiment, full name, and (in some cases) age at enlistment. We extracted these indexes in their entirety for the state of Kentucky, with 107,589 entries on the Union side and 50,304 entries on the Confederate side.

Table 1 contains an illustration of the nature of the data extracted from these indexes. An obvious complication with using these indexes is that it is not clear when multiple entries refer to the same person. The first three entries in Table 1 are men from the 3rd Union Cavalry named John Ewbanks, John Ubanks, and John Ebanks, respectively. The 4th entry is a man from the 55th Union Infantry, who is also named John Ewbanks. These names are all phonetic variants of one another, and could easily refer to the same person. Soldiers frequently re-enlisted in multiple units, and if their names were spelled differently on different muster rolls or were duplicated for some other reason, they could easily appear in this index multiple times.

This poses a challenge for establishing the coverage of these records, or the fraction of all enlistees who appear in the indexes. In particular, estimating the coverage of these indexes will depend on assumptions that we make about which records are duplicates. In the top panel of Table 1, we illustrate the least conservative grouping, in which we assume that phonetically identical names from the same regiment are the same person. In the example in Table 1, this reduces the number of unique soldiers from 10 to 7. In the entire sample, this reduces the number of unique soldiers to 78,257 Union and 37,917 Confederate, for a total of 116,174 recruits from Kentucky (see panel A of Table 2 for relevant statistics).⁸ Another possibility is to assume that all Union or Confederate soldiers with phonetically identical names are the same person, as illustrated in panel

⁸These groupings are formed by creating NYIIS codes for both first and last names and grouping by these codes. When only first initials are given, they are grouped with full first names containing the same first initial.

B of Table 1; this reduces the number of soldiers in Table 1 to 5, and it reduces the number of records in the complete sample to 64,309 (44,976 Union and 19,333 Confederate). How do these sample sizes compare with the likely number of military recruits from Kentucky? An estimated 90,000 to 100,000 Kentuckians enlisted on the Union side, while only 30,000 to 40,000 enlisted on the Confederate side (Astor 2012; Marshall 2010). Thus, the most conservative estimate of the number of recruits included in these records implies a coverage rate of around 50%. This is, however, a conservative lower bound: it is likely that multiple men with similar names did enlist, implying a much higher coverage rate of up to 100%.⁹

The military indexes give us very little information other than the name of the recruit and the side on which he enlisted. Therefore, we need to match these indexes to other records in order to characterize these enlistees and their outcomes. A challenge is that the only information we can use to match military indexes to other records is first and last name. Although many enlistment records contain the recruit’s age at enlistment, this is substantially more common in Union records: more than 80% of Union records contain the recruit’s age at enlistment, while only about 15% of Confederate records contain this information. Accordingly, we cannot use age at enlistment to match records without introducing severe systematic differences in the accuracy of matches by Union or Confederate status. As we discuss in detail in Appendix A, it is imperative that we avoid introducing linkage error that is correlated with military side; this will severely impede our ability to draw inferences about the impact of military side on post-war outcomes.

An additional complication is that we cannot be sure how many individuals are covered by each unique name entry. Importantly, some names appear on both Union and Confederate rosters. To construct a list of names to match to census records, we group names by phonetic first and last name, defined using NYIIS codes (Atack and Bateman 1992), and military side, i.e. Union or Confederate. We restrict the sample to phonetic name groups that are uniquely identifiable as Union or Confederate, and we treat each phonetic group as a single individual. We also omit name groups that only include first initials, as we do not have sufficient information from these initials to accurately link our observations to other records.¹⁰ As an example, see panel C of Table 1, in

⁹The 50% figure is especially conservative because NYIIS codes tend to be over-inclusive, defining some names as “phonetically identical” when they are clearly not. For example, “John” and “James” have the same NYIIS code. On the other hand, NYIIS codes will usually fail to identify identical names with typing or transcription errors.

¹⁰This restriction reduces the number of Confederate recruits relative to Union recruits, since almost 20% of Confederate records list only a first initial, while very few Union recruits list only a first initial. This can be seen in panel A of Table 2. We find that, in our Confederate sample of names, the regiment that the soldier enlisted in explains about 12% of the variation in whether or not a full first name is reported, and we do not find evidence that the socioeconomic status of the soldier’s surname is related to the probability of reporting a full first name. We measure the socioeconomic status of a surname as the mean family wealth and occupational income among families

which only two of the five unique phonetic name groups listed would be included in our sample. As seen in panel A of Table 2, this leaves us with 49,180 unique phonetic name groups to match, 38,318 of which are Union and 10,862 of which are Confederate.

We select this method of constructing our database with specific empirical questions in mind. We are interested in comparing the ex ante characteristics and post-war outcomes of Union and Confederate soldiers. To do so, we first identify a sample of Union and Confederate recruits-to-be in 1860. When constructing this sample, our dual goals are the following: (i) maximize the accuracy of the Union or Confederate status assigned to individuals in our sample; (ii) minimize differences in the accuracy of Union and Confederate status. Goal (i) reduces attenuation bias in our estimates; goal (ii) reduces bias of unknown direction (see Appendix A for details). Selecting a sample of names whose phonetic variants do not appear on both sides will increase the accuracy of our assignment of Union or Confederate status. Because Confederate records are more likely to contain only first initials as compared to Union records, omitting records with first initials avoids introducing systematic differences in the accuracy of Union and Confederate status. While we believe these methodological choices best enable clean comparisons between Union and Confederate recruits, they introduce certain issues which are worth mentioning. Specifically, they cause people with uncommon names to be overrepresented, and they reduce the size of the Confederate sample.¹¹

3.2 Matches to 1860 census

We match our sample of uniquely Union and Confederate names to the Census of 1860 using records available from ancestry.com via the NBER. Again, our challenge is that the only linkable information we have in our military data is the soldier’s name. So, to facilitate matching to the census, we impose certain restrictions on our target sample of census records. First, because our sample of recruits comes from regiments of white males, we limit our search to white males in the census. A sample of Union Army veterans indicates that 99% of Union recruits were born between 1817 and 1847 (Fogel 2000). Assuming a similar age range in the Confederate army, and allowing for some error in the reporting of ages, we further restrict our search of the 1860 census to men born between 1815 and 1850. Finally, we restrict the geographic area in which we search for these

with this surname in Kentucky in 1850, using the 1850 full count census data from the North Atlantic Population Project (Ruggles et al 2010). As such, we believe that reporting only a first initial reflects record keeping practices of individual regiments or companies rather than systematic socioeconomic differences. As with age at enlistment, the fact that the availability of full first names differs by military side precludes using records with only first initials in our sample.

¹¹We do not find economically significant differences between Kentuckians with common and uncommon names.

soldiers.

We impose these restrictions on our target sample in order to maximize our linkage rate while minimizing linkage error. An unrestricted match to the 1860 census based on name alone would yield many potential matches, most of which would be incorrect. Using information about the prior probability that recruits have other characteristics can improve the accuracy of our matches. Take, for example, residential location. Given that our recruits enlisted in Kentucky regiments, it is overwhelmingly likely that they resided in Kentucky at the time of enlistment, which occurred between 1861 and 1865. Companies were typically organized locally, and regiments were named after the state that enlistees were from. So, we believe it is fair to assume that recruits were more likely to reside in Kentucky in 1860 than elsewhere; as such, matches residing in Kentucky are more likely to be correct than matches residing elsewhere. We perform two matching procedures: one in which we match military records to white men ages 10-45 residing in Kentucky (275,999 records in target sample), and one in which we match our military records to white men ages 10-45 residing in states surrounding Kentucky (3,610,482 records in target sample).¹² We match names by searching for exact phonetic first name and surname matches between the military records and the target census sample, then by comparing the similarity of the first and last names using the Jaro-Winkler algorithm (Ruggles et al 2010). We discard matches with a string similarity score of less than 0.9.¹³

Panel B of Table 2 contains information on matching rates using both approaches. Not surprisingly, matching to an expanded geographic area increases the fraction of military records matched to at least one census record, from around 43% to 66%. However, it decreases the fraction of records that are matched uniquely to the target sample, from around 25% to 18%. Moreover, it appears that the matches made exclusively to Kentucky are more accurate. Recall that we have information on age at enlistment for most of the Union recruits in our sample. While we do not perform matches to the census using this information, we can use it to check the accuracy of our results. Specifically, for individuals with an age of enlistment recorded on their military record, we can estimate

$$Age_{mil} = \beta_0 + \beta_1 Age_{1860} + u$$

¹²These states are: Kentucky, Tennessee, Missouri, Illinois, Indiana, Ohio, Virginia, Arkansas, Mississippi, Alabama, Georgia, North Carolina, and South Carolina.

¹³Approximately 75% of phonetic name groups contain a single entry. When a phonetic group contains multiple (differently spelled) entries, we select one entry to compare with potentially matched records in the 1860 census using the Jaro-Winkler algorithm. We use the following rule to select this entry: (i) we select the entry with information on age at enlistment, which will facilitate our test of linkage accuracy (described below); (ii) if there are zero or multiple entries with data on age at enlistment, we select the most frequently occurring spelling in the phonetic group; (iii) if multiple spelling occur with equal frequency, we select one at random.

If a match is correct, the age in the military record (Age_{mil}) should be more or less identical to the age in the census record (Age_{1860}). So, a sample of correct matches should yield an estimated intercept close to zero and a slope close to one. In the bottom panel of Table 3, we estimate this regression equation under two specifications: (i) using only records that are uniquely matched to the 1860 census; and (ii) using all matched records, weighting multiple matches by $1/N$, where N is the number of census records that match the military record in question. We estimate these specifications for three samples: (i) a sample matched to all states surrounding Kentucky; (ii) a sample matched to Kentucky only; (iii) and a sample matched to Missouri only as a placebo test.

The first four columns of panel C of Table 2 indicate that using unique matches between military records and the 1860 census introduces less error than using weighted multiple matches. And, these results indicate that matching to Kentucky is more accurate than matching to Kentucky *and* all states bordering Kentucky. Restricting the target sample to Kentucky will cause us to miss (or mis-match) recruits who migrated to Kentucky after 1860. However, it appears that expanding the target sample introduces enough false positives that we are better off with the restriction. The last two columns of the table indicate that matches to Missouri alone are extremely inaccurate, which gives us further confidence that our matches to Kentucky are of a high quality.¹⁴

The 1860 Census allows us to observe each man’s place of residence, the composition of his family, the occupation and literacy status of each family member, and the value of the family’s real and personal property. We assign 1950 occupational codes to each individual’s occupation (Ruggles et al 2010), and we assign a value of occupational income based on the 1900 occupational wage distribution with an imputed wage for farmers (Preston and Haines 1991; Abramitzky et al 2012; Olivetti and Paserman 2015; Salisbury 2014). Because some recruits are children in 1860, we assign each individual the socioeconomic indicator (occupational income or wealth) of the head of the individual’s household.

3.3 Matches to 1880 census

We match our recruits from the 1860 census (12,440 in total) to the 1880 100% census sample (NAPP). Here, we make use of the demographic information we obtain from the 1860 census in

¹⁴This “check” on the accuracy of our matches is necessarily driven by Union recruits, as they comprise the overwhelming majority of records with age information. However, we have no reason to believe that Confederate recruits were less likely to come from Kentucky than Union recruits. When we match to all states surrounding Kentucky, we end up finding a greater fraction of Confederate matches in Kentucky than Union matches: 30% of our Confederate matches reside in Kentucky, whereas 26.5% of our Union matches live in Kentucky. So, we are confident that restricting our target sample to Kentucky improves match accuracy overall.

order to locate recruits in 1880. We search the entire 1880 census for records that exactly match our 1860 census records on the following dimensions: birth place, phonetic first and last name codes, sex, and race. We restrict birth year in the 1880 census to be no more than three years before or after birth year in the 1860 census. Finally, we discard matches in which the index measuring the similarity of names across census records (using the Jaro-Winkler algorithm) is less than 0.9. These procedures approximately follow Ruggles et al (2010). Using this procedure, we are able to uniquely match 30% of our Union soldiers and 29% of our Confederate soldiers. This match rate is comparable to other studies that perform automated record linkages (Ferrie 1996; Ruggles et al 2010; Abramitzky et al 2012).¹⁵

In addition to linking our sample of recruits to the Census of 1880, we link male relatives of recruits who are under the age of 45 in 1860. Male relatives are defined as males who are living in the same household as an individual linked to a military record, but who are not themselves linked to a military record. This does not necessarily mean that these relatives are civilians; only that they are more likely to be civilians than those linked to a military record. We link these men using an identical procedure to that used to link soldiers to 1880. We are able to identify 29,747 male relatives of recruits in the 1860 census, and we link 17% of these to the census of 1880 (the rate is similar for relatives of Union and Confederate soldiers). This linkage rate is substantially lower than the linkage rate among soldiers. This can be explained by the fact that soldiers in our database have uncommon names by construction, so very few records are discarded because they can be linked to multiple records in the 1880 census.

¹⁵There is a growing body of research into record linkage using machine learning (Feigenbaum 2016; Bailey et al 2016). The primary benefit of this approach is that allows for automated comparisons between alphabetic strings that more closely resemble comparisons made by the human eye; the “rule” for identifying matching strings can be more complex. Our rule for identifying matching strings – matching NYIIS codes and a Jaro-Winkler string similarity score of 0.9 or higher – is substantially less costly but also coarser. This may increase our error rate; however, because we are applying this linking algorithm to everyone in our sample, it should not introduce linkage error in a way that is correlated with military side. Thus, our relatively coarse string comparison rule may attenuate our estimates of the impact of military side on post-war outcomes; however, it should not affect the direction of our estimates (see Appendix A for proof). As our primary goal is to establish the direction of these effects, we maintain that our approach is appropriate. We also note that our likely error rate compares favorably with the IPUMS linked 1860-1880 sample (Ruggles et al 2010; this links the 1860 1% sample to the 1880 full count census). We make this inference by comparing the county out-migration rate in both samples: while using locational information to form links between censuses is inappropriate, linkage error should increase the measured migration rate, as an incorrectly linked individual is very unlikely to reside in the same county in both census years. The IPUMS linked sample of males shows a county out-migration rate of approximately 45% for Kentucky, while our data shows a migration rate of approximately 50%.

4 Empirical Approach

We are interested in understanding how service in the military affected locational choices after the Civil War. We typically model migration as an investment which maximizes a person’s expected lifetime earnings. Importantly, we usually think of migration as a welfare maximizing event: if people migrate in response to regional wage differentials, they are effectively sorting themselves into regions where labor is relatively productive. Moreover, if people migrate to regions that complement their individual skill profiles, as much of the existing research on migrant selection contends, they are sorting into the regions in which they are individually most productive. In the case of civil conflict, migration may occur for other reasons. In particular, animosity among combatants may generate migrations which would never have happened for economic reasons in the absence of the conflict.

Our aim with this paper is to measure the degree to which conflict among recruits from opposing sides generated a migration response after the Civil War. To fix ideas, suppose a person i ’s earnings in county c (Y_{ic}) depend on both individual ability (A_{ic}) and “social capital” (S_{ic}), both of which are person and location specific. A person will choose to locate in the county that maximizes $Y(A_{ic}, S_{ic})$ net of migration costs; if this is the person’s home county, he will choose not to migrate. Our hypothesis is that the Civil War affected S_{ic} . In particular, in counties more sympathetic to the Union, S_{ic} should have fallen for Confederate recruits and risen for Union recruits; in counties more sympathetic to the Confederacy, S_{ic} should have fallen for Union recruits and risen for Confederate recruits. This will affect observed migration behavior in two key ways: (1) Recruits who have experienced a reduction in S_{ic} in their home county should be more likely to migrate; (2) Migrant recruits should be more likely to select a destination in which S_{ic} has increased for recruits from their side.

We use two county-level measures to infer relative “social capital” to Union and Confederate recruits after the war. First, we use the share of recruits we identify in a given county that enlisted in the Confederate army. Second, we use the county’s share of the presidential vote going to John C. Breckenridge – the southern Democratic candidate – in the 1860 election. As Breckenridge carried all southern states in the 1860 presidential election – and explicitly supported the westward expansion of slavery – counties that favored Breckenridge can be presumed to be more aligned with the South than the North. As we will show in the next section, this metric is a strong predictor of military side. Conditional on other characteristics, a 10 percentage point increase in vote share to

Breckenridge generates a 6 percentage point increase in the probability of serving in the Confederate army. This result is significant at the 1 percent level. A county’s Confederate enlistment share is perhaps the most obvious measure of that county’s sympathy with the Confederacy. However, because we measure this using our sample of soldiers linked to the 1860 census, it is likely measured with error. In particular, we cannot be sure that we are sampling Union and Confederate soldiers from every county at the same rate. Moreover, this measure embeds a certain amount of linkage error. As an indicator of public opinion, Breckenridge vote share is likely measured with less error. However, the link between voting behavior and Confederate sympathy may be more attenuated than the link between Confederate enlistment and Confederate sympathy. As neither measure is perfect, we present results using both.¹⁶

We test whether the propensity to leave a county depends differently on these indicators for Union and Confederate recruits. We also test whether migrants from Union and Confederate sides sorted differentially into places more sympathetic to the South, measured as Confederate enlistment share and Breckenridge vote share in 1860 for intrastate migrants, and region of residence in 1880 for interstate migrants. Lastly, we explore differences in the selection of migrants, as well as the return to migration, by military side.

4.1 Migration Propensity

To determine whether Union recruits were more likely to leave more “Confederate” counties, we estimate the following equation using OLS:

$$M_{ij,1880} = \alpha + \beta_1 U_{ij} + \beta_2 S_{j,1860} + \beta_3 U_{ij} \times S_{j,1860} + \gamma X_{i,1860} + \delta_j + u_{ij} \quad (1)$$

Here, M_{ij} is an indicator equal to one if person i from county j had migrated by 1880; U_{ij} is equal to 1 if this person served in the Union army; $S_{j,1860}$ is a measure of sympathy for the Confederacy in county j in 1860; $X_{i,1860}$ is a matrix of individual characteristics observed in 1860, including age and birthplace fixed effects; δ_j is an 1860 county fixed effect. We expect to find $\beta_3 > 0$.

A complication with this approach is that Union and Confederate recruits are drawn from systematically different parts of the skill distribution: Confederate recruits are more skilled on

¹⁶The other strong predictor is the fraction of a county’s population that is enslaved: a 10 percentage point increase in the fraction enslaved generates a 6.5 percentage point increase in the probability of joining the Confederate army. However, we do not use this as one of our baseline measures of Confederate sympathy because it is very clearly linked to the direct economic impact of the Civil War, which may differ by military side if Confederate recruits are more likely to own slaves in slaveholding counties than Union recruits. This paper emphasizes social (rather than direct economic) determinants of migration.

average. Thus, we may find that Union recruits are more likely to leave “Confederate” counties if these counties are more complementary to skilled individuals (Borjas 1987). In other words, differences in migration propensities may work through differences in individual skill and not differences in social capital. Because we observe indicators of recruits’ socioeconomic status in 1860 – namely, occupation (or occupation of the household head in the case of children) and family wealth – we can include interactions between ex ante socioeconomic status and our indicator of social alignment with the Confederacy. If β_3 is robust to the inclusion of these controls, then selection on skill is unlikely to explain differential migration behavior by military side.

It is also possible that Union and Confederate recruits are differently selected on unobservable skill, so controlling for 1860 socioeconomic status is not sufficient to show that differential migration behavior is driven by social capital and not skill. One way to address this problem is to use recruits’ family members to control for systematic unobserved skill differences by military side. In addition to our sample of recruits, we link male family members of recruits, who are under the age of 45, to the Census of 1880. We then estimate the difference in the difference between a soldier’s and a related civilian’s migration propensity by military side. Specifically, we estimate the following:

$$M_{ijk,1880} = \alpha + \beta_1 V_{ijk} + \beta_2 V_{ijk} \times UF_{jk} + \beta_3 S_{j,1860} + \beta_4 UF_{jk} \times S_{j,1860} + \beta_5 V_{ijk} \times S_{j,1860} + \beta_6 V_{ijk} \times UF_{jk} \times S_{j,1860} + \gamma X_{i,1860} + \delta_j + \phi_k + u_{ijk} \quad (2)$$

Variables are generally defined as above, with i indexing individuals, j indexing county of origin, and k indexing families. The variable V_{ijk} is equal to one if person i from county j and family k is a veteran and zero if this person is a civilian family member. The indicator UF_{jk} is equal to one if family k from county j is a “union family” and zero otherwise. The parameter ϕ_k is a family fixed effect. For Confederate civilians, the marginal effect of $S_{j,1860}$ on the probability of migrating is β_3 ; for Confederate veterans, this marginal effect is $\beta_3 + \beta_5$; for Union civilians, this marginal effect is $\beta_3 + \beta_4$; and for Union veterans this marginal effect is $\beta_3 + \beta_4 + \beta_5 + \beta_6$. The parameter we are most interested in is β_6 : if $\beta_6 > 0$, this means that Union soldiers respond more to $S_{j,1860}$ than their family members, and by a greater margin than Confederate soldiers relative to their family members. The family fixed effect ensures that between-family variation in skill is not driving the result.

We note that, while $\beta_6 > 0$ is evidence that skill is not the sole driver of differential migration behavior among Union and Confederate veterans, $\beta_6 = 0$ is not sufficient to prove that it is. If soldiers and soldiers’ family members are treated similarly after the war, then civilian family

members should be equally encouraged to leave counties hostile to their side. Thus, even if social forces (and not skill) drive migration, we could observe that $\beta_5 = \beta_6 = 0$. As such, in addition to arguing that social forces guide migration decisions, $\beta_6 > 0$ informs us about the way in which these social forces work. In particular, these forces are more powerful for combatants than non-combatants. It is also important to note that the military status of recruits' family members is measured with a substantial amount of error. We define "civilian" family members as family members who are not linked to a military record. This does not necessarily mean that these family members did not fight, only that they are less likely to have fought than those linked to a military record. This matching error will tend to attenuate these estimates.

4.2 Migration Destination

To determine whether Union recruits were more likely to sort into less "Confederate" counties, we estimate the following, using a sample of internal migrants within Kentucky:

$$S_{il,1860} = \alpha + \beta U_{ijl} + \gamma X_{i,1860} + \delta_j + u_{ij} \quad (3)$$

Here, $S_{il,1860}$ is a measure of social alignment with the Confederacy in 1860 in county l , where person i is residing in 1880; and U_{ijl} is an indicator equal to one if person i who migrated from county j to county l between 1860 and 1880 served in the Union army. The remaining variables are defined as above. Here, we expect to find $\beta < 0$. We also estimate a multinomial logit model of residence in each region in 1880 on a sample of migrants. We predict that enlistees on the Union side should be more likely to move north than enlistees on the Confederate side.¹⁷

This analysis is again complicated by systematic differences in skill between Union and Confederate recruits. Because we are able to control for ex ante occupational attainment and family wealth, we can rule out the hypothesis that differences in locational choices are entirely driven by observable skill. We can also use non-veteran family members to control for family-specific unobservable skill. One additional concern is that Union veterans were eligible to acquire land under the Homestead Act of 1862, while Confederate veterans were excluded until 1867. So, Union veterans may have disproportionately migrated to areas with better land, since they had the first opportunity to do so. To argue that the Homestead Act does not explain our regional location

¹⁷Phillips (2013, pp 108-109) notes that the West was a popular destination for ex-Confederates from border states, citing several well-known and high ranking Confederates from the region who "chose expatriation – Cuba, Mexico, Brazil, or Canada – or headed for the Far West rather than face postwar retributive violence or trial for treason in their home states."

results, we control for 1880 farm value per acre and the rate of farm ownership, which are available at the county level in 1880. If we still estimate a significant β , then regional locational differences cannot be explained by regional differences in land quality.

4.3 Differences in Migrant Selection and Returns

In addition to establishing that Civil War veterans migrated due to social pressure, we investigate *who* chose (or was able) to migrate due to social pressure. In particular, we estimate the following regression separately for Union and Confederate veterans:

$$M_{ij,1880} = \alpha + \beta_1 Y_{i,1860} + \beta_2 Y_{i,1860} \times S_{j,1860} + \gamma X_{i,1860} + \delta_j + u_{ij} \quad (4)$$

Here, Y_i denotes a measure of socioeconomic status in 1860 (log occupational earnings or log family wealth), and other variables are defined as above. In equation (3), the parameter β_2 captures the selectivity of migrants from counties with higher S . In particular, if $\beta_2 > 0$, then migrants from counties with larger S are more positively selected than migrants from counties with lower S . If $\beta_{2,Union} > \beta_{2,Confed.}$, then Union migrants from counties more aligned with the Confederacy are more positively selected than Confederate migrants from these counties. This may indicate that S affects the earnings of skilled individuals more than the earnings of unskilled individuals. For instance, occupations higher in the skill distribution – such as managers and officials – may benefit more from social capital than occupations that are lower in the skill distribution. This finding is also consistent with the notion that higher SES individuals are better able to absorb the economic costs associated with migrating for social reasons.

Finally, we consider how military side affects the return to migration, in terms of occupational income.¹⁸ We are interested in the overall return to migration, the difference in the returns of Union and Confederate veterans, and heterogeneity in this difference by county of origin. We estimate the following separately for Union and Confederate veterans:

$$Y_{ij,1880} = \alpha + \beta_1 M_{ij,1880} + \beta_2 M_{ij,1880} \times S_{j,1860} + \gamma X_{i,1860} + \delta_j + u_{ij} \quad (5)$$

¹⁸Here, we use the 1900 occupational wage distribution with an imputed wage for farmers, assigned to 1950 occupational codes (Preston and Haines 1991; Abramitzky, Boustan and Eriksson 2012; Olivetti and Paserman 2015; Salisbury 2014). Occupational attainment is the only measure of socioeconomic status available in the 1880 census. By using a national occupational income measure, we are abstracting away from migration returns experienced by moving to a place with a higher wage in a given occupation.

If $\beta_{2,Union} > \beta_{2,Confed.}$, then Union migrants from Confederate-leaning counties experience a larger return to migration than Confederate migrants from Confederate-leaning counties.

Our predictions for this final test are ambiguous. If enlistees migrate for purely social reasons, then it is not clear that they should realize an economic return to migration. Predictions about heterogeneity in the return to migration by military side and county of origin are similarly ambiguous. On the one hand, if men living in unsympathetic counties are inclined to migrate for social rather than economic reasons, we may see more migrants with lower potential returns from these counties: this should tend to make $\beta_{2,Union} < \beta_{2,Confed.}$. On the other hand, if migrants from unsympathetic counties are positively selected on unobservable skill, this will tend to make $\beta_{2,Union} > \beta_{2,Confed.}$.

5 Results

We present preliminary results using our sample of soldiers' names that are matched uniquely to Kentucky in 1860 (12,440 individuals). We are able to uniquely match 3,693 of these men to the Census of 1880, in addition to 5,064 male family members (age 45 or younger) of our 12,440 soldiers in 1860. We use these samples in the results that follow.

5.1 Characteristics of Soldiers in 1860

Panel A of Figure 1 illustrates the fraction of recruits linked to each county that enlisted on the Confederate side. There are relatively more Union recruits in coal-producing areas of the state, specifically in the lower portion of the eastern mountains and coalfields, and in the western coalfields. There are relatively more Confederate enlistees from the northeastern agricultural ("Bluegrass") region and around the Mississippi Plateau in the southwest portion of the state, which is also an agricultural region. There is also a concentration of Confederates in the eastern part of the state along the border with Virginia. In panels B and C of Figure 1, we plot the share of the vote in the 1860 presidential election to John C. Breckenridge and the fraction of the population that was enslaved in 1860, respectively. It is clear from this figure that Confederate enlistment is positively correlated with both of these characteristics.

In Table 3, we compare average characteristics of Union and Confederate soldiers. The first column contains mean values of each variable for Union soldiers, the second column contains means for Confederates, and the third column contains means for all white men in Kentucky between the ages of 10 and 45. The fourth column contains results from an OLS regression of an indicator for

Union status on all characteristics together. As a group, soldiers were younger and less likely to be married than the general population, which is not surprising. They were also more likely to be native to Kentucky or native to the United States.

Comparing Union and Confederate soldiers, a number of differences are apparent. On average, Union soldiers were older, more likely to be married, and less likely to live with a parent. Table 3 also indicates large differences in nativity. Confederate enlistees were much more likely to be born in Kentucky or in the South generally. Union soldiers were much more likely to be born in the Northeast, Midwest, or abroad.

Evidence also points to differential selection of Confederate soldiers on socioeconomic characteristics. Confederate soldiers systematically came from counties with more slaves, greater value of property per family, and more people employed in agriculture. While we do not have data on the individual wealth of everyone in our sample, we find that Confederate soldiers typically had surnames that were associated with greater value of real estate and more white collar employment in 1850.¹⁹ These findings are consistent with men who had greater ties to slavery being more likely to join the Confederate army. We also find significant differences in voting patterns. Men who joined the Confederate army tended to live in counties with a greater vote share going to the Democratic Party in the 1860 presidential election. Conversely, Union soldiers came from counties more likely to vote for John Bell, an alternative candidate from the Constitutional Union party who did not explicitly favor the westward expansion of slavery.

5.2 Outcomes for Soldiers in 1880

Table 4 contains additional summary statistics for our sample of soldiers who are matched to the 1880 census. This table includes average outcomes in 1880, as well as average 1860 characteristics that we have only collected for this sample. These are largely consistent with Table 3 in that they point to Confederate recruits being of higher socioeconomic status *ex ante*. The table also indicates systematic differences in locational outcomes for Union and Confederate soldiers. We discuss these differences in detail below.

¹⁹Our 1860 full count data does not contain information on wealth or occupation. We enter this information by hand from census manuscripts for the 3,693 men who we link between 1860 and 1880; however, we do not have this information for all 12,440 recruits linked to 1860. To infer socioeconomic status for these men, we calculate the mean value of real estate wealth, as well as the fraction in each occupational class, among male household heads with a particular surname in Kentucky in 1850, and we link this information with our sample of recruits. We use the 1850 full count data from NAPP (Ruggles et al 2010), which contains information on real estate wealth and occupation.

5.2.1 Migration Propensity

From Table 4, we can see that roughly 70 percent of our sample still resided in Kentucky as of 1880; however, approximately 50 percent of our sample had moved between counties by 1880. This is true of both Union and Confederate veterans. As can be seen in column (1) of Table 5, Union soldiers were no more or less likely to migrate than Confederate soldiers. However, migrants appear to be negatively selected from the overall population of Kentucky soldiers in terms of family wealth in 1860.

In Table 5, we estimate the impact of the home county Confederate enlistment share (panel A) and Breckenridge vote share (panel B) on the propensity to migrate among Union and Confederate recruits. The key variable is the interaction between Union soldier and our indicator of alignment with the South: in column (2) of panel A, the coefficient can be interpreted to mean that the marginal effect of Confederate enlistment share on the probability of migrating is 0.419 (0.145) greater for Union soldiers than Confederate soldiers. Our results are very similar when we use Breckenridge vote share to instead measure social alignment with the Confederacy. To address concerns that this differential is driven by county-level differences in the return to skill, we include log family wealth and the household head's log occupational income in 1860 and interactions between these variables and Confederate measures in column (3). The inclusion of these variables has a moderate impact on our results: the coefficient on the interaction between Union soldier and social alignment with the Confederacy decreases somewhat, and it is not quite significant at the 10 percent level when we use Breckenridge vote share. However, we find little clear evidence that migrants from counties more sympathetic to the Confederacy are negatively selected on skill, which is necessary for skill differences among Union and Confederate recruits to drive our results.

In columns (4) and (5), we estimate equation 2, in which we use non-recruit family members to control for differences in unobservable skill by military side. In column (4), we omit family fixed effects and include UF_{jk} as an explicit control. This specification allows us to include all linked soldiers and family members, not just pairs of soldiers and relatives from the same family. The necessary assumption here is that the distribution of unobservable skill is the same among Union veterans and Union family members. Similarly, the distribution of skill must be the same among Confederate veterans and Confederate family members. In these columns, the variable "Union soldier" is the interaction $UF \times V$, which appears in equation 2. The key variable is the interaction between Union soldier and $S_{j,1860}$: in column (5) of panel A, the interpretation is that

the difference in the marginal effect of Confederate enlistment share on the probability of migrating between soldiers and civilian family members is 0.690 (0.199) higher for Union than Confederate families. This result tells us two things: first, the different response to Confederate enlistment share by military side cannot be explained entirely by differences in skill; second, soldiers are more responsive to social pressure to migrate than their relatives. The results in panel B are similar, albeit less conclusive.

Figure 2 illustrates these results graphically. In panel A, we plot the predicted probability of migrating (based on results in column (3) of Table 5 for Union and Confederate soldiers with mean characteristics, in a county with 5 percent Confederate enlistment share and in a county with 65 percent Confederate enlistment share, which is the maximum in our sample). The probability of a Confederate soldier leaving a county with 5 percent Confederate enlistment share is approximately 10 percentage points higher than the probability of an otherwise identical Union soldier leaving; however, the probability of a Confederate soldier leaving a county with 65 percent Confederate enlistment share is almost 15 percentage points lower than the probability of Union soldier leaving. In panel B, we illustrate the results from column (4). In a county with a 5 percent Confederate enlistment share, Union soldiers and family members are equally likely to leave, while Confederate soldiers are almost 10 percentage points more likely than their relatives to leave. Conversely, in a county with a 65 percent Confederate enlistment share, Union soldiers are around 5 percentage points more likely than their relatives to migrate, while Confederate soldiers are approximately 5 percentage points *less* likely to leave than their relatives.

5.2.2 Migration Destination

In Figure 3, we map the locations of Kentucky veterans who left Kentucky in 1880. There appear to be clear locational differences: Union recruits are more likely to move north, and Confederate recruits are more likely to move south and west. In panel A of Table 6, we estimate differences in locational choices of migrants by military side. We estimate a multinomial logit model of region of residence in 1880 (South, West, or Northeast, relative to the Midwest). Our explanatory variables include an indicator for having served in the Union army, as well as other 1860 characteristics including age, birthplace and county of residence fixed effects. We find that Union recruits were significantly more likely to migrate to the Midwest than either the South or the West. This result is robust to controlling for ex ante occupational attainment and family wealth (column 2), so differences in destination region cannot be explained by differences in observable skill. This finding

is consistent with recruits moving to areas of the country where the social returns to their military service are highest.²⁰

A concern is that these results are not driven by regional differences in the social returns to military service but by the Homestead Act of 1862. Because Confederate veterans were excluded until 1867, they may have been excluded from acquiring the best available land, if this land was claimed first. If the best land was in the Midwest, this mechanism could generate our results. To address this concern, we control for two salient 1880 county characteristics: average farm value per acre, and the ownership rate in agriculture. If differences in destination region are totally explained by the fact that Union recruits could access better quality land, then including these controls should wipe out any systematic regional differences in location by military side. We show this is not the case, as can be seen in column (3). In column (4), we include non-recruit family members to address the possibility that Union and Confederate soldiers are differently selected on unobservable skill. Because we have very few linked soldier-relative pairs who are *both* migrants, we do not implement a family fixed effects model similar to equation 2. Rather, we include UF_{jk} as a control and omit the family fixed effect. The results are very similar, although the impact of being a Union soldier on the probability of moving to the South (relative to the Midwest) is not quite significant.

In panels B and C of Table 6, we estimate differences in destination county characteristics within Kentucky. We take a sample of men living in Kentucky in 1880 but in a different county from their county of residence in 1860. We regress the Confederate enlistment share (panel B) or Breckenridge vote share (panel C) in the person's 1880 county of residence on a Union indicator, adding the same controls as in panel A. In columns (1)-(3), we find that Union soldiers who migrate within Kentucky are significantly less likely to end up in a county more sympathetic to the Confederacy. In column (4), we add non-combatant relatives, as in panel A. We find that internal Kentucky migrants from Union families are significantly less likely to end up in Confederate-sympathizing counties than those from Confederate families; however, there is no differential effect for soldiers relative to civilians. This result may mean that these results are driven by systematic differences in unobservable skill by military side; or, it may be that both soldiers and family members are equally attracted to counties sympathetic with their families' side. This test does not differentiate between these two possibilities.

²⁰As the West was sparsely populated, migration to the West should have been associated with a smaller "penalty" for serving on the Confederate side. We include the Northeast in the model, but the results are omitted for brevity; there is no systematic difference in the propensity to move to the Northeast relative to the Midwest by military side.

5.3 Differences in Migrant Selection and Returns

In Table 7, we explore differences in the selection of migrants from the Union and Confederate side. In panel A, we consider differences in the selection of migrants overall. We regress an indicator for having migrated between 1860 and 1880 on measures of observable skill in 1860 separately for Union and Confederate recruits, and we test whether or not our coefficients of interest are significantly different for Union and Confederate recruits. We find little evidence that migrants from the Union side were systematically differently selected than movers from the Confederate side while migrants from both sides were negatively selected on family wealth in 1860.

In panels B and C, we estimate equation 4. Here, we consider whether Union recruits from more “Confederate” counties were differently selected than Confederate recruits from “Confederate” counties in terms of ex ante observable skill (measured as the household head’s log occupational income in 1860 and log family wealth in 1860). We find evidence consistent with positive selection: Union migrants from “Confederate” counties are more positively selected than Confederate migrants from “Confederate” counties. While this difference is not significant when we measure alignment with the Confederacy using the Confederate enlistment share, it is significant at the 10 percent level when we use Breckenridge vote share. Thus, we have some evidence that more skilled recruits migrated in response to social pressure.²¹ It may be that skilled workers are more sensitive to social pressure than unskilled workers. This notion is consistent with social capital being a more important determinant of skilled worker’s earnings than an unskilled worker’s earnings. A final potential explanation is that skilled workers were better able to absorb the costs of migrating in response to social pressure than unskilled workers.

Finally, in Table 8, we explore the return to post-Civil War migration for Union and Confederate soldiers. In panel A, we report overall differences in migration returns. In particular, we regress occupational income in 1880 on an indicator for the person having moved counties between 1860 and 1880, an indicator for serving in the Union army, and an interaction between these two variables. We find that Union soldiers have poorer occupational outcomes than Confederate soldiers, even conditional on the occupational income and wealth of the soldier’s household head in 1860. This finding may reflect positive selection into the Confederate army on unobservables or a positive

²¹We should also note that Union recruits may have been more positively selected in predominantly Confederate counties, perhaps because skilled men were more likely to “go against the grain.” We test whether or not Union recruits were more skilled ex ante (measured by occupational attainment and family wealth) when they came from counties more aligned with the Confederacy. We do not find any evidence that this is the case. However, they may have been selected on unobservables.

causal effect of service in the Confederate army on occupational attainment. In any case, Union soldiers were unable to overcome their initial economic disadvantage over the next two decades despite having emerged victorious in the conflict.

In column 3, we include an indicator for having migrated out of a soldier's home county in 1860. The coefficient on the main effect of migrating and the interaction between migrating and Union status are both close to zero and insignificant, indicating no return to migration in terms of occupational income for soldiers on either side. We note that our measure of occupational income obscures any gains to migrants who earned higher wages in their destination county despite remaining in the same occupation. It is thus possible that there was some economic return we are unable to capture. Nonetheless, since the majority of these migrants were moving within or very near Kentucky (see Figure 3 and Table 4), we believe wage differentials across destinations were generally small.²² These findings underscore that the migration of former Kentucky soldiers was likely driven by social or ideological factors.

In panels B and C, we explore the heterogeneity of migration returns by characteristics of the county of origin by estimating equation 5. In particular, we are interested in whether soldiers who moved out of ideologically incompatible home counties realized an economic return to migration. We find that the return to migration was larger for Union soldiers from more Confederate counties, but this effect is not significant at traditional levels and is only present for the Breckenridge vote share measure. This finding could reflect weakly positive selection of such migrants on unobservable skill rather than a positive return; we cannot distinguish between these mechanisms. Overall we find no robust or significant evidence that even soldiers who migrated out of counties that were a poor ideological match realized an economic return on their relocation decision.

6 Conclusion

Our results suggest that social divisions between Union and Confederate supporters following the Civil War induced a significant share of veterans to relocate for reasons of ideology rather than economic gain. Belonging to the winning side was not associated with a differential likelihood of migrating, and we do not find that Union soldiers were more likely to move than former Confederates. Rather, our results paint a picture of soldiers from both sides departing counties where they

²²Abramitsky, Boustan, and Eriksson (2012) and Collins and Wanamaker (2014) find a significant and positive return to migration using measures that account for both occupation and regional differences in the average wages associated with that occupation. Due to the short range of most moves in our sample, we do not adjust for occupational wage differences by geographic region.

were surrounded by their former enemies. Union soldiers were more likely to migrate the greater the support for the Confederacy in their home counties, and they settled in Kentucky counties and regions of the country that were more pro-Union on average. Confederate soldiers were more likely to choose Confederate-leaning counties or states in the far West or South if they moved. Although we are restricted to using occupational income to measure gains associated with these relocation choices, we estimate that such returns are close to zero for both Union and Confederate veterans. Even Union soldiers who departed Confederate-leaning counties saw no gains in occupational income by 1880, underscoring that these migration decisions were likely driven by social rather than economic considerations.

The implications of these findings are twofold. First, more research is needed to understand the potential individual return to migrating in the aftermath of a civil conflict. Increases in emotional wellbeing and feelings of personal safety may not be readily detectable using standard economic measures. Nonetheless, these benefits may greatly improve the quality of life for survivors of civil conflict. Future scholarship should search for ways to measure these benefits. Second, ideological resorting may be an important mechanism for maintaining peace after a war. The relative mobility of the nineteenth century American population could have facilitated the end of hostilities by allowing individuals and families to form new communities instead of trying to repair social divisions in towns fractured by conflict. Removing barriers to internal migration after a civil war could promote both individual welfare and reduce opportunities for continuing violence.

References

- [1] Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. “Europe’s Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration.” *American Economic Review* 102, no. 5 (2012): 1832-1856.
- [2] Annan, Jeannie and Christopher Blattman. “The Consequences of Child Soldiering.” *Review of Economics and Statistics* 92, no. (2010): 882-898.
- [3] Astor, Aaron. *Rebels on the Border: Civil War, Emancipation, and the Reconstruction of Kentucky and Missouri*. Baton Rouge, LA: Louisiana University Press, 2012.
- [4] Attack, Jeremay and Fred Bateman. “Matchmaker, Matchmaker, Make Me a Match: A General Personal Computer-Based Matching Program for Historical Research” *Historical Methods*. 25, no. 2 (1992):53-65.
- [5] Bailey, Martha and Morgan Henderson. “How Do Alternative Linking Methods Perform? Evidence from the LIFE?M Project.” Working Paper, 2016.

- [6] Blattman, Christopher, and Edward Miguel. "Civil War." *Journal of Economic Literature* 48, no. 1 (2010): 3- 57.
- [7] Bleakley, Hoyt, Louis Cain, and Joseph Ferrie. "Amidst Poverty and Prejudice: Black and Irish Civil War Veterans," in *Institutions, Innovation, and Industrialization: Essays in Economic History and Development*, Avner Greif, Lynne Keisling, John V.C. Nye (eds.), 2014. (Festschrift volume for Joel Mokyr.)
- [8] Borjas, George J. "Self-Selection and the Earnings of Immigrants." *American Economic Review* 77, no. 4 (1987): 531-553.
- [9] Bundervoet, Tom, Philip Verwimp, and Richard Akresh. "Health and Civil War in Rural Burundi." *Journal of Human Resources* 44, no. 2 (2006): 536-563.
- [10] Clubb, Jerome M., William H. Flanigan, and Nancy H. Zingale. *Electoral Data for Counties in the United States: Presidential and Congressional Races, 1840-1972*. ICPSR08611-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2006-11-13. <http://doi.org/10.3886/ICPSR08611.v1>
- [11] Collins, William J. and Marianne H. Wanamaker. "Selection and Economic Gains in the Great Migration of African Americans: New Evidence from Linked Census Data." *American Economic Journal: Applied Economics* 6, no. 1 (2014): 220-252.
- [12] Costa, Dora L. "Pensions and Retirement: Evidence from Union Army Veterans." *Quarterly Journal of Economics* 110, no. 2 (1995): 297-320.
- [13] Costa, Dora L. "Displacing the Family: Union Army Pensions and Elderly Living Arrangements." *Journal of Political Economy* 105, no. 6 (1997): 1269-1292.
- [14] Costa, Dora L, and Matthew Kahn. *Heroes and Cowards: The Social Face of War*. Princeton University Press. 2008.
- [15] Costa, Dora L., Matthew Kahn, Christopher Roudiez, and Sven Wilson. "Persistent Social Networks: Civil War Veterans who Fought Together Co-Locate in Later Life." NBER Working Paper no. 22397, 2016.
- [16] Eli, Shari J. "Income Effects on Health: Evidence from Union Army Pensions." *Journal of Economic History* 75, no. 2 (2015): 448-478.
- [17] Feigenbaum, James J. "Automated Census Record Linking: A Machine Learning Approach." Working Paper, 2016.
- [18] Ferrie, Joseph P. "A New Sample of Americans Linked from the 1850 Public Use Micro Sample of the Federal Census of Population to the 1860 Federal Census Manuscript Schedules." *Historical Methods*. 29 (1996): 141- 156.
- [19] Ferrie, Joseph P. "Migration to the Frontier in mid-Nineteenth Century America: A Re-examination of Turner's Safety Valve." Working Paper, 1997.
- [20] Fogel, Robert W. *Public Use Tape on the Aging of Veterans of the Union Army: Military, Pension, and Medical Records, 1860-1940, Version M-5*. Center for Population Economics, University of Chicago Graduate School of Business, and Department of Economics, Brigham Young University, 2000.

- [21] Haines, Michael R., and Inter-university Consortium for Political and Social Research. Historical, Demographic, Economic, and Social Data: The United States, 1790-2002 [Computer file]. ICPSR02896-v3. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2010-05-21. doi:10.3886/ICPSR02896
- [22] Harrison, Lowell H. *The Civil War in Kentucky*. Lexington: The University Press of Kentucky, 1975.
- [23] Long, Jason and Henry Siu. "Refugees from Dust and Shrinking Land: Tracking the Dust Bowl Migrants." NBER Working Paper no. 22108, 2016.
- [24] Marshall, Ann. *Creating a Confederate Kentucky: The Lost Cause and Civil War Memory in a Border State*. Chapel Hill: University of North Carolina Press, 2010.
- [25] Miguel, Edward and Gerard Roland. "The Long-Run Impact of Bombing in Vietnam." *Journal of Development Economics* 96, no. 1 (2011): 1-15.
- [26] Olivetti, Claudia and Daniele Paserman. "In the Name of the Son (and the Daughter): Intergenerational Mobility in the United States, 1850-1940." *American Economic Review* 105, no. 8 (2015): 1-31.
- [27] Phillips, Christopher. *The Civil War in the Border South*. Santa Barbara, CA: Praeger, 2013.
- [28] *Portraits and Sketches of the Lives of All the Candidates for the Presidency and Vice-Presidency, for 1860: Comprising Eight Portraits Engraved on Steel, Facts in the Life of Each, the Four Platforms, the Cincinnati Platform, and the Constitution of the United States*. New-York: J. C. Buttre, 1860. From Library of Congress, <http://hdl.loc.gov/loc.gdc/scd0001.00118958603> .
- [29] Ruggles, Steven J., Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. *Integrated Public Use Microdata Series: Version 5.0* [Machine-readable database]. Minneapolis: University of Minnesota, 2010.
- [30] Salisbury, Laura. "Selective Migration, Wages, and Occupational Mobility in Nineteenth Century America." *Explorations in Economic History* 53 (2014): 40-63.
- [31] Salisbury, Laura. "Women's Income and Marriage Markets in the United States: Evidence from Civil War Pensions," *Journal of Economic History* Forthcoming (2016).
- [32] Stewart, James I. "Migration to the Agricultural Frontier and Economic Mobility, 1860-1880." *Explorations in Economic History* 43, no. 4 (2006): 547-577.
- [33] United States War Department (1890-1912). *Carded Records Showing Military Service of Soldiers Who Fought in Volunteer Organizations During the American Civil War*. Record Group 94, National Archives Building, Washington, DC. Accessed at fold3.com.

7 Tables and Figures

Table 1: Military Data: Example

Side	Regiment	Name
<i>Panel A: phonetic name + regiment groups</i>		
Union	3rd Cavalry	John Ewbanks
Union	3rd Cavalry	John Ubanks
Union	3rd Cavalry	John Ebanks
Union	55th Infantry	John Ewbanks
Union	1st Cavalry	Jefferson Eubanks
Confederate	Kirkpatrick's Battalion	John J Ewbank
Confederate	Kirkpatrick's Battalion	J J Eubank
Confederate	10th Infantry	Napolean Ewbanks
Confederate	12th Cavalry	Napolean Eubanks
Confederate	19th Infantry	F Eubanks
<i>Panel B: phonetic name + union/confederate groups</i>		
Union	3rd Cavalry	John Ewbanks
Union	3rd Cavalry	John Ubanks
Union	3rd Cavalry	John Ebanks
Union	55th Infantry	John Ewbanks
Union	1st Cavalry	Jefferson Eubanks
Confederate	Kirkpatrick's Battalion	John J Ewbank
Confederate	Kirkpatrick's Battalion	J J Eubank
Confederate	10th Infantry	Napolean Ewbanks
Confederate	12th Cavalry	Napolean Eubanks
Confederate	19th Infantry	F Eubanks
<i>Panel C: names included in final sample</i>		
Union	3rd Cavalry	John Ewbanks
Union	3rd Cavalry	John Ubanks
Union	3rd Cavalry	John Ebanks
Union	55th Infantry	John Ewbanks
Union	1st Cavalry	Jefferson Eubanks
Confederate	Kirkpatrick's Battalion	John J Ewbank
Confederate	Kirkpatrick's Battalion	J J Eubank
Confederate	10th Infantry	Napolean Ewbanks
Confederate	12th Cavalry	Napolean Eubanks
Confederate	19th Infantry	F Eubanks

Notes. Illustration of possible groupings of military records, under different assumptions about which records constitute a single person. In panel A, we assume that all records with phonetically matching first and last names from the same regiment refer to the same individual. In panel B, we assume that all records with phonetically matching first and last names from the same side (Union or Confederate) refer to the same person. In panel C, we illustrate the “individuals” who constitute our final sample of military records, obtained from the genealogical website Fold3.com, which we link to the census of 1860. These consist of phonetic first and last names that appear uniquely on the Union or Confederate side, excluding records with first initials only.

Table 2: Military Data

Summary Statistics and Test of Matching Accuracy

<i>Panel A. Characteristics of Military Data</i>			
	Union	Confederate	Total
Total # Records	107,589	50,304	157,893
% of total	68.1%	31.9%	100%
# unique phonetic last name + phonetic first name + regiment groups	78,257	37,917	116,174
# unique phonetic last + phonetic first + union/confederate groups	44,976	19,333	64,309
# groups with only first initials available	129	3,578	3,707
% of groups with only first initials available	0.3%	18.5%	5.8%
# unique phonetic last + phonetic first + union/confederate groups, excl groups with only first initials, identifiable as union/confederate only	38,318	10,862	49,180
% of total	77.9%	22.1%	100%

<i>Panel B. Linkage Rates to Census by Target Sample</i>				
	Matched to all states near KY		Matched to KY	
	Union	Confederate	Union	Confederate
# military records matched at least once to census	25,461	7,103	-	16,472
Mean # matches per military record (of matched)	11.86	9.00	***	2.12
# unique matches to census	6,503	2,008	-	9,529
% of military records matched to at least one census	66.4%	65.4%	**	43.0%
Conventional match rate	17.0%	18.5%	***	24.9%
% of matches living in Kentucky	26.5%	29.5%	***	26.8%
			-	4,588
			1.81	***
			2,911	-
			42.2%	***
			26.8%	***
			-	-

<i>Panel C. Test of Accuracy of Matched Records</i>						
	Matched to all states near KY		Matched to KY		Matched to MO (placebo)	
	Unique	Weighted	Unique	Weighted	Unique	Weighted
Sample:						
Coefficient on Military Age	0.501	0.214	0.632	0.474	0.051	0.048
Constant	10.3	18.0	7.0	11.4	22.3	22.5
R-squared	0.185	0.027	0.294	0.135	0.001	0.001
Observations	5,944	302,022	9,391	35,393	4,323	28,299

Notes. Statistics pertaining to military records and links to the census of 1860. Phonetic groups defined using NYIIS codes. Approximately 75% of phonetic groups contain a single entry. See text for details of matching procedures.

Table 3: Characteristics of Union and Confederate Soldiers, 1860

	Mean Comparison			OLS Regression	
	Union	Confederate		All men 10-45 in Kentucky, 1860	Dependent variable = 1 if Union
Age	22.398	22.031	**	23.895	-0.002** (0.001)
Married	0.319	0.270	***	0.370	0.019 (0.015)
Household head	0.302	0.251	***	0.371	0.020 (0.019)
Lives with parent	0.442	0.481	***	0.385	-0.013 (0.011)
Born Kentucky	0.749	0.820	***	0.722	
Born south (incl. Kentucky)	0.863	0.933	***	0.834	
Born northeast	0.018	0.010	***	0.021	0.100*** (0.026)
Born midwest	0.043	0.023	***	0.038	0.081*** (0.026)
Immigrant	0.077	0.034	***	0.106	0.153*** (0.021)
County % agricultural	0.754	0.788	***	0.766	0.041 (0.088)
County % urban	0.090	0.081		0.117	-0.254** (0.112)
County % slave	0.148	0.189	***	0.176	-0.650*** (0.226)
County % free black	0.008	0.009		0.009	1.229 (1.765)
Property per family (\$1000)	3.873	4.909	***	4.576	0.002 (0.012)
County ag value per acre	18.524	22.072		22.403	-0.001 (0.002)
County mean farm size	2.371	2.315		2.210	0.010 (0.019)
County churches per 100 people	0.193	0.187		0.187	0.249* (0.143)
County value per church	2.233	2.234		2.675	0.004 (0.008)
Vote share: Bell	0.454	0.419	***	0.451	
Vote share: Breckenridge	0.357	0.441	***	0.353	-0.597*** (0.101)
Vote share: Douglas	0.178	0.134	***	0.186	-0.189 (0.133)
Presidential voter turnout	0.667	0.699	***	0.671	-0.161 (0.182)
Surname: mean prop., 1850 (\$1000)	1.240	1.739	***	1.503	-0.002*** (0.000)
Surname: % white collar, 1850	0.057	0.066	***	0.065	-0.058 (0.053)
Surname: % farmer, 1850	0.627	0.625		0.623	-0.002 (0.025)
Surname: % laborer, 1850	0.064	0.061		0.062	
Constant					1.204*** (0.161)
Observations	9,529	2,911		275,999	10,346
R-squared					0.094

Notes. Stars next to mean comparison refer to significance of the coefficient on the 1860 characteristic in a univariate regression of union status on that characteristic. For county-level characteristics, standard errors are clustered at county level. The regression in the final column also clusters standard errors by county. Individual characteristics are obtained from the 1860 full count census; county characteristics are taken from Haines and ICPSR (2010); election returns data are taken from Clubb et al (2006); 1850 surname variables refer to mean characteristics of individuals with a particular surname living in Kentucky in 1850, obtained from the full count 1850 census (Ruggles et al 2015).

Table 4: Summary Statistics of Linked 1860-1880 Census Data

	Mean			Sample size	
	Union	Confederate		Union	Confederate
<i>Panel A. 1880 Characteristics</i>					
Age	42.170	41.960		2,848	845
Married	0.871	0.853		2,848	845
Lives in 1860 county	0.467	0.489		2,848	845
Lives in Kentucky	0.708	0.730		2,848	845
Lives elsewhere in south	0.059	0.089	***	2,848	845
Lives in northeast	0.021	0.009	**	2,848	845
Lives in midwest	0.203	0.150	***	2,848	845
Lives in west	0.009	0.020	**	2,848	845
White collar	0.080	0.118	***	2,848	845
Semi-skilled	0.110	0.098		2,848	845
Farmer	0.649	0.662		2,848	845
Laborer	0.135	0.095	***	2,848	845
No occupation	0.026	0.027		2,848	845
<i>Panel B. 1860 Characteristics</i>					
Family wealth (\$1000)	2.109	6.074	***	2,716	798
Parent white collar	0.048	0.085	***	1,308	411
Parent semil-skilled	0.088	0.075		1,308	411
Parent farmer	0.688	0.708		1,308	411
Parent laborer	0.051	0.039		1,308	411
Parent no occupation	0.088	0.058	*	1,308	411
White collar	0.044	0.065	*	1,540	434
Semi-skilled	0.112	0.088		1,540	434
Farmer	0.394	0.426		1,540	434
Laborer	0.242	0.214		1,540	434
No occupation	0.172	0.159		1,540	434

Notes. Characteristics of veterans linked between the census of 1860 and the census of 1880. See text for details about the linking process.

Table 5: Impact of County Characteristics on Migration Propensity

Dependent variable:	(1)	(2)	(3)	(4)	(5)
	<i>Moved counties, 1860-80</i>				
<i>Panel A. Confederate enlistment share</i>					
Union soldier	0.014 (0.025)	-0.108** (0.050)	-0.100* (0.052)	-0.098 (0.061)	-0.182** (0.077)
Union soldier X Confederate enlistment share, 1860		0.419*** (0.145)	0.315** (0.143)	0.330** (0.167)	0.690*** (0.199)
HH head log occ. income, 1860			0.045 (0.055)		
Log family wealth, 1860			-0.092*** (0.023)		
HH Head log occ. income X Confed. enlist. share, 1860			-0.026 (0.185)		
Log family wealth X Confed. enlist. share, 1860			0.137** (0.066)		
Soldier				0.083 (0.051)	0.162** (0.070)
Union family				-0.018 (0.043)	
Soldier X Confed. enlist. share, 1860				-0.205 (0.131)	-0.420** (0.161)
Union family X Confed. enlist. share, 1860				0.133 (0.125)	0.290 (0.654)
Sample	Soldiers	Soldiers	Soldiers	Soldiers & relatives	Soldiers & relatives, linked pairs
Observations	3,693	3,693	3,338	8,643	3,174
R-squared	0.100	0.103	0.136	0.085	0.053
<i>Panel B. Breckenridge vote share</i>					
Union soldier	0.014 (0.025)	-0.089* (0.050)	-0.095* (0.053)	-0.097 (0.060)	-0.038 (0.081)
Union soldier X Breckenridge vote share, 1860		0.247** (0.121)	0.199 (0.130)	0.212* (0.116)	0.103 (0.150)
HH head log occ. income, 1860			0.099* (0.053)		
Log family wealth, 1860			-0.056*** (0.018)		
HH Head log occ. income X Breckenridge vote share, 1860			-0.171 (0.130)		
Log family wealth X Breckenridge vote share, 1860			0.002 (0.046)		
Soldier				0.093* (0.054)	0.093 (0.078)
Union family				0.017 (0.038)	
Soldier X Breckenridge vote share, 1860				-0.173* (0.100)	-0.155 (0.140)
Union family X Breckenridge vote share, 1860				0.017 (0.084)	0.033 (0.425)
Sample	Soldiers	Soldiers	Soldiers	Soldiers & relatives	Soldiers & relatives, linked pairs
Observations	3,693	3,693	3,338	8,643	3,174
R-squared	0.100	0.102	0.135	0.084	0.048

Notes. “Union soldier” is an indicator for union status in columns 1-3 and an interaction between an indicator for union family status and veteran status in columns 4-5. Confederate enlistment share is the fraction of all soldiers linked to a particular county (12,440 soldiers in total) that are Confederate. Log family wealth is $\log(1 + W)$, where W is the sum of real and personal property. Breckenridge vote share from Clubb et al (2006). Occupational income is based on the 1900 occupational wage distribution with an imputed wage for farmers (Preston and Haines 1991; Salisbury 2014; Olivetti and Paserman 2015). All regressions contain controls for age (fixed effects in columns 1-3, quadratic in columns 4-5), birthplace fixed effects, 1860 county of residence fixed effects. Column (5) contains family fixed effects and includes 1,372 families with more than one individual linked between 1860 and 1880. Standard errors are clustered at the 1860 county level.

Table 6: Locational Choices of Migrants

	(1)	(2)	(3)	(4)
<i>Panel A. Multinomial Logit, dep. var. 1880 region of residence (relative to Midwest)</i>				
South				
Union soldier	-0.286** (0.137)	-0.342** (0.147)	-0.284* (0.157)	-0.318* (0.163)
Soldier				0.312** (0.145)
Union family				0.093 (0.106)
West				
Union soldier	-1.123*** (0.371)	-0.984** (0.415)	-0.988** (0.443)	-1.131** (0.468)
Soldier				0.750* (0.395)
Union family				-0.013 (0.330)
Observations	1,951	1,780	1,777	4,906
<i>Panel B. OLS, dep. var. 1860 Confederate enlistment share in 1880 (Kentucky) county</i>				
Union soldier	-0.054*** (0.017)	-0.057*** (0.019)	-0.055*** (0.018)	-0.016 (0.015)
Soldier				-0.001 (0.013)
Union family				-0.036*** (0.012)
Observations	787	694	694	1,832
R-squared	0.398	0.422	0.495	0.347
<i>Panel C. OLS, dep. var. 1860 Breckenridge vote share in 1880 (Kentucky) county</i>				
Union soldier	-0.050*** (0.017)	-0.053*** (0.019)	-0.058*** (0.019)	-0.013 (0.020)
Soldier				-0.007 (0.019)
Union family				-0.032** (0.015)
Observations	805	711	711	1,792
R-squared	0.451	0.477	0.540	0.261
Controls:				
Age, birthplace & 1860 county	Y	Y	Y	Y
1860 SES	N	Y	Y	N
1880 county agricultural land value & farm ownership rate	N	N	Y	N
Sample:	Migrant soldiers	Migrant soldiers	Migrant soldiers	Migrant soldiers & relatives

Notes. “Union soldier” is an indicator for union status in columns 1-3 and an interaction between an indicator for union family status and veteran status in column 4. Null results about the impact of military side on the probability of migrating to the Northeast (relative to the Midwest) are omitted from Panel A for brevity. In panels B and C, the sample consists of individuals who left their county of origin for another county in Kentucky between 1860 and 1880; standard errors clustered by 1880 county. Controls for 1860 SES include log family wealth and household head’s log occupational income (see notes to Table 5 for details). Controls for the value per acre of agricultural land and the farm ownership rate in 1880 are from Haines and ICPSR (2010).

Table 7: Differential Selection of Union & Confederate Migrants

Dependent variable	(1)	(2)	
Sample	Union	Migrated counties Confederate	p(u=c)
<i>Panel A. Overall</i>			
HH head log occ. income, 1860	0.014 (0.039)	0.103 (0.086)	0.337
Log family wealth, 1860	-0.052*** (0.014)	-0.048** (0.019)	0.858
Observations	2,569	769	
R-squared	0.143	0.305	
<i>Panel B. Heterogeneity by Confederate enlistment share</i>			
HH head log occ. income, 1860	0.015 (0.040)	0.159 (0.097)	0.152
Log family wealth, 1860	-0.052*** (0.014)	-0.068*** (0.021)	0.478
HH head log occ. income X Confed. enlist. share, 1860	-0.059 (0.350)	-0.542 (0.427)	0.450
Log family wealth, 1860 X Confed. enlist. share, 1860	0.196* (0.116)	0.169 (0.113)	0.869
Observations	2,569	769	
R-squared	0.145	0.309	
<i>Panel C. Heterogeneity by Breckenridge vote share</i>			
HH head log occ. income, 1860	0.010 (0.042)	0.141* (0.084)	0.158
Log family wealth, 1860	-0.051*** (0.015)	-0.045** (0.019)	0.816
HH head log occ. income X Breckenridge vote share, 1860	-0.070 (0.180)	-0.761** (0.358)	0.100
Log family wealth, 1860 X Breckenridge vote share, 1860	0.051 (0.060)	-0.019 (0.084)	0.473
Observations	2,569	769	
R-squared	0.143	0.311	

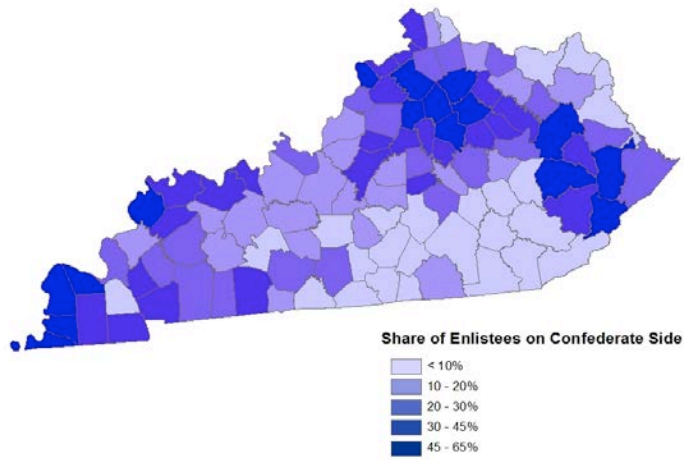
Notes. P-values are from a test of equality of coefficients across regressions. Sample consists of soldiers only. See notes to Table 5 for variable definitions. All regressions include age, birthplace, and county of origin fixed effects. Standard errors are clustered by 1860 county.

Table 8: Differential Migration Returns for Union & Confederate Soldiers

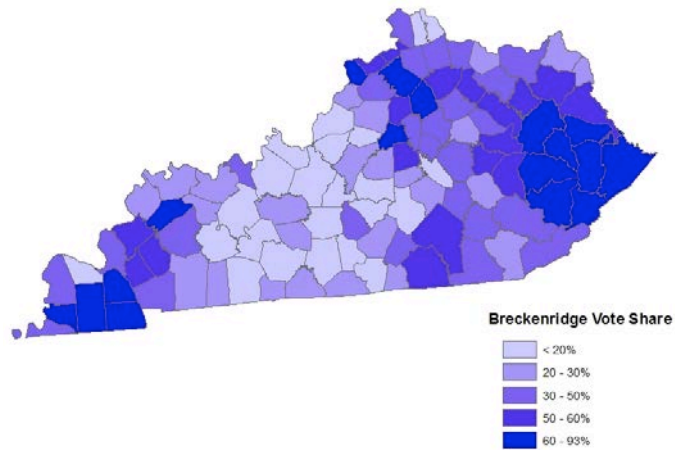
Dependent variable:	(1)	(2)	(3)
	<i>Log occupational income, 1880</i>		
	<i>Panel A. Differences in Return to Migration</i>		
Union	-0.059*** (0.012)	-0.041*** (0.012)	-0.039** (0.017)
Moved counties			0.003 (0.020)
Union X moved counties			-0.003 (0.023)
HH Head log occ. Income, 1860		0.162*** (0.021)	0.162*** (0.021)
Log family wealth, 1860		0.026*** (0.006)	0.026*** (0.006)
Observations	3,606	3,261	3,261
R-squared	0.068	0.100	0.100
	<i>Panel B. Heterogeneity by Confederate Enlistment Share</i>		
	<i>Union</i>	<i>Confederate</i>	<i>p(u=c)</i>
Moved counties	-0.004 (0.014)	0.010 (0.032)	0.653
Moved counties X confed. enlist. share	-0.048 (0.095)	-0.060 (0.149)	0.948
HH Head log occ. Income, 1860	0.137*** (0.025)	0.204*** (0.064)	0.331
Log family wealth, 1860	0.031*** (0.009)	0.025** (0.011)	0.68
Observations	2,509	752	
R-squared	0.105	0.302	
	<i>Panel C. Heterogeneity by Breckenridge Vote Share</i>		
	<i>Union</i>	<i>Confederate</i>	<i>p(u=c)</i>
Moved counties	-0.001 (0.012)	0.009 (0.031)	0.757
Moved counties X Breckenridge vote share	0.073 (0.059)	-0.088 (0.122)	0.207
HH Head log occ. Income, 1860	0.137*** (0.025)	0.201*** (0.064)	0.35
Log family wealth, 1860	0.031*** (0.009)	0.024** (0.011)	0.671
Observations	2,509	752	
R-squared	0.106	0.302	

Notes. P-values are from a test of equality of coefficients across regressions. Sample consists of soldiers only. See notes to Table 5 for variable definitions. All regressions include age, birthplace, and county of origin fixed effects. Standard errors are clustered by 1860 county.

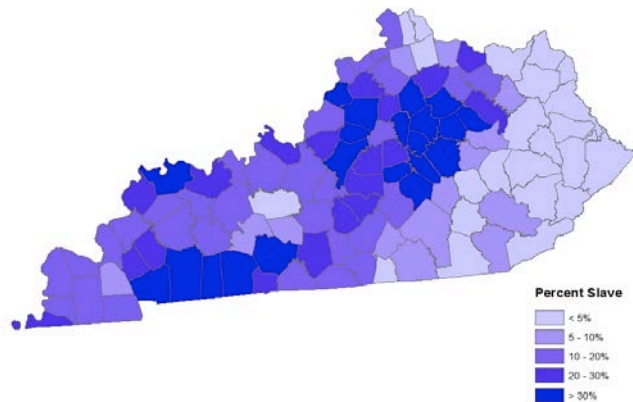
Figure 1: Relevant Kentucky County Characteristics, 1860



(a) Confederate Enlistment Share

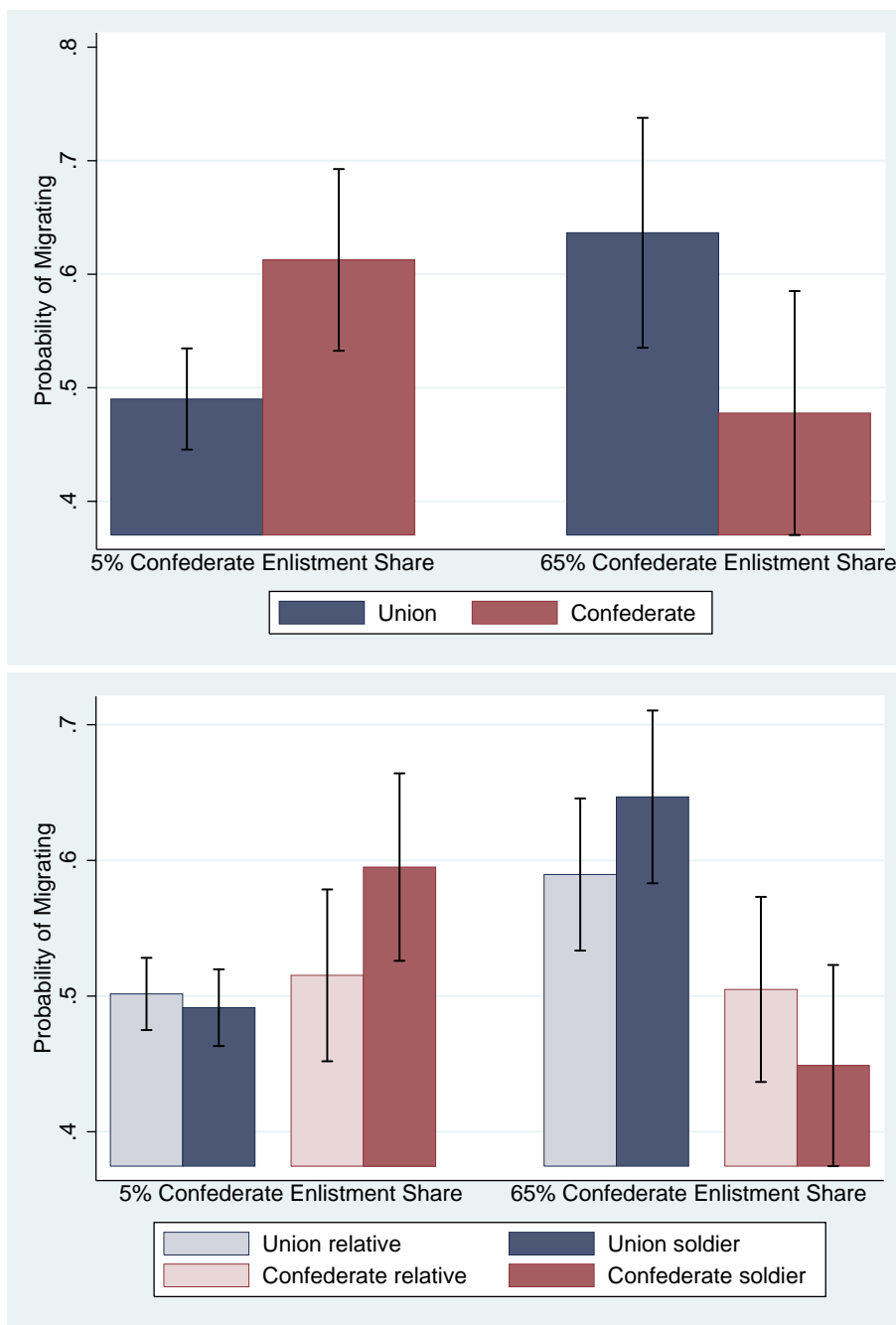


(b) Breckenridge Vote Share



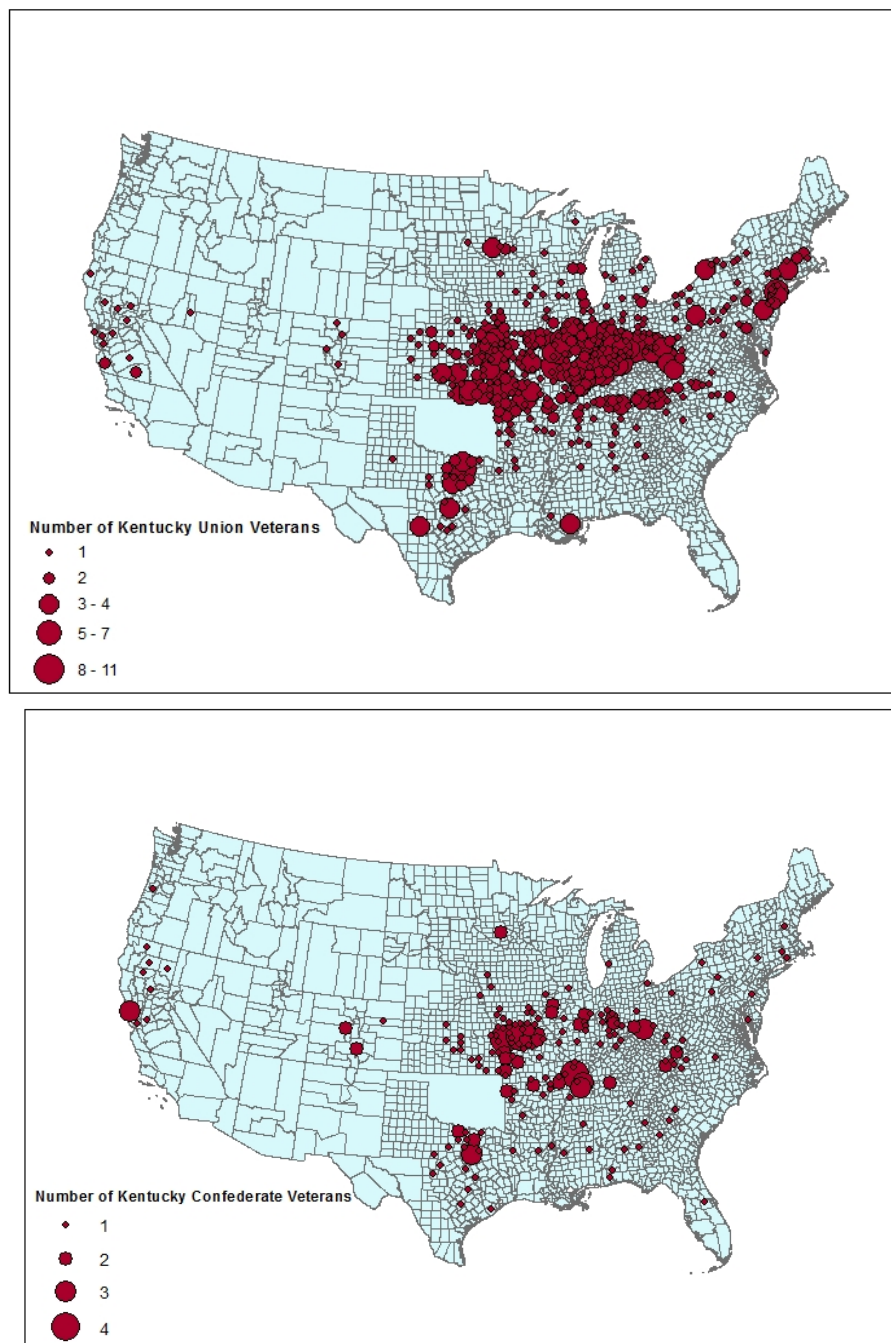
(c) Percent Slave

Figure 2: Impact of County Confederate Enlistment Share on Migration Propensity: Illustration



Notes. Top panel illustrates predicted probability of migrating for Union and Confederate recruits with mean characteristics, in counties with 5 and 65% Confederate enlistment share. Based on results from Table 5, panel A, column (3). Bottom panel illustrates predicted probability of migrating for Union and Confederate recruits and relatives with mean characteristics, in counties with 5 and 65% Confederate enlistment share. Based on results from Table 5, panel A, column (4).

Figure 3: Distribution of Interstate Migrants from Union and Confederate Armies, 1880



Notes. Illustrates county of 1880 residence of all interstate migrants from sample of Kentucky recruits.

A Note on Bias Introduced by Matching Error

Here, we justify certain features of our approach to matching Union and Confederate military records to the Census, which may reduce our sample size but are designed to make the rate of matching error similar for Union and Confederate veterans (specifically, discarding Confederate records with only first initials, and failing to use age at enlistment to generate matches). We show that, if the entire population could be classified as “Union” or “Confederate,” the optimal strategy is to minimize matching error, even if this generates differential matching error by military side. However, because the population can actually be classified as “Union,” “Confederate,” or “neither,” the optimal strategy is to equalize matching error among Union and Confederate recruits, even if this results in a higher error rate overall. The proof follows.

Population divisible into two categories

Suppose the population belongs to one of two group: A or B . Information on group membership comes from database 1, which we match to database 2. Database 2 contains information on some outcome variable y .

The causal effect of group membership on y is:

$$\beta = E(y|A) - E(y|B) \tag{6}$$

Matches are done with error. Suppose we observe variables a and b , which are equal to one if a person observed in database 2 is matched to a person in database 1 from group A or B , respectively. Matching is done with error, so $\phi_{A|a} \equiv Pr(A = 1|a = 1) < 1$ and $\phi_{B|b} \equiv Pr(B = 1|b = 1) < 1$.

If we assume matches are correct, we will estimate:

$$\begin{aligned} \hat{\beta} &= \frac{\sum_a y_a}{N_a} - \frac{\sum_b y_b}{N_b} \\ &= \frac{\sum_{a,A} y_{a,A} + \sum_{a,B} y_{a,B}}{N_{a,A} + N_{a,B}} - \frac{\sum_{b,B} y_{b,B} + \sum_{b,A} y_{b,A}}{N_{b,B} + N_{b,A}} \end{aligned}$$

Here, the subscript (i, J) refers to an observation whose true group is J but is assigned a group i , where $i \in \{a, b\}$ and $J \in \{A, B\}$.

Rearranging and taking expectations, it can be shown that

$$E(\hat{\beta}) = E(y|A)(\phi_{A|a} - \phi_{A|b}) - E(y|B)(\phi_{B|b} - \phi_{B|a}) = (\phi_{A|a} + \phi_{B|b} - 1)(E(y|A) - E(y|B))$$

This follows because there are only two possible groups, so $\phi_{B|a} = 1 - \phi_{A|a}$ and $\phi_{A|b} = 1 - \phi_{B|b}$. Then, bias will be:

$$\begin{aligned} E(\hat{\beta}) - \beta &= (\phi_{A|a} + \phi_{B|b} - 1)(E(y|A) - E(y|B)) - (E(y|A) - E(y|B)) \\ &= (\phi_{A|a} + \phi_{B|b} - 2)(E(y|A) - E(y|B)) \\ &= (\phi_{A|a} + \phi_{B|b} - 2)\beta \end{aligned}$$

If the probability that matches assigned to groups a and b are accurate is 1, then the estimator will be unbiased. If not, the estimate will be attenuated. Notice that the coefficient on β is always negative, so the estimator will be downward biased if $\beta > 0$ and upward biased if $\beta < 0$. As long as the probability of a match being accurate is more than 50%, this will not cause the sign of the estimated coefficient to be incorrect.

As such, if everyone in database 2 can be classified as either group A or group B , then the optimal approach is to maximize $\phi_{A|a}$ and $\phi_{B|b}$ individually – attenuation bias will always be reduced by raising $\phi_{A|a}$, regardless of the value of $\phi_{B|b}$, and vice versa.

Population divisible into three categories

Now, suppose there is a third category, C , which exists in the population; however, no individuals from group C appear in database 1. Matches between databases 1 and 2 are still done with error, so a person who is assigned group a or b could in fact be from group A , B , or C .

By a similar calculation to the one done above, we can write the expected value of our estimator for β as:

$$E(\hat{\beta}) = E(y|A)(\phi_{A|a} - \phi_{A|b}) - E(y|B)(\phi_{B|b} - \phi_{B|a}) + E(y|C)(\phi_{A|b} + \phi_{B|b} - \phi_{A|a} - \phi_{B|a})$$

And, bias of this estimator is:

$$E(\hat{\beta}) - \beta = E(y|A)(\phi_{A|a} - \phi_{A|b} - 1) - E(y|B)(\phi_{B|b} - \phi_{B|a} - 1) + E(y|C)(\phi_{A|b} + \phi_{B|b} - \phi_{A|a} - \phi_{B|a})$$

The direction of this bias is unclear. In particular, it will depend on $E(y|C)$, and the relative probability of assigning a person from group C to groups a and b . Consider, for example, the case in which matches to group a are perfectly accurate. So, $\phi_{A|a} = 1$, which means that $\phi_{B|a} = 0$. Then, the bias will be:

$$E(\hat{\beta}) - \beta = \phi_{A|b} \left(E(y|C) - E(y|A) \right) - (1 - \phi_{B|b}) \left(E(y|C) - E(y|B) \right)$$

The direction of bias in this case will depend on how the $E(y|C)$ relates to $E(y|A)$ and $E(y|B)$, which is not something we know.

Now, consider another example in which the probability of error is perfectly symmetric across groups. Specifically, $0 < \phi_{A|a} = \phi_{B|b} < 1$ and $0 < \phi_{B|a} = \phi_{A|b} < 1$. In this case,

$$\begin{aligned} E(\hat{\beta}) - \beta &= \left(\phi_{A|a} - \phi_{B|a} - 1 \right) \left(E(y|A) - E(y|B) \right) \\ &= \left(\phi_{A|a} - \phi_{B|a} - 1 \right) \beta \end{aligned}$$

This is just attenuation bias, and should not cause the sign of $\hat{\beta}$ to be incorrect as long as matches are more likely to be correct than incorrect.

If we are interested in correctly identifying the sign of β , the objective should not be to minimize match error individually for both groups, but to minimize match error conditional on match error remaining symmetric across groups. This provides us with a clear idea of what the bias in our estimator is.

Appendix Tables and Figures

Table A1: Impact of Military Side on Migration Propensity and Destination
 Different Measures of Occupational Attainment

	(1)	(2)	(3)	(4)
<i>Panel A. OLS, dependent variable = moved counties, 1860-80</i>				
Union soldier	-0.088* (0.052)	-0.104** (0.049)	-0.097* (0.052)	-0.083 (0.051)
Union soldier X Confederate enlistment share, 1860	0.291** (0.142)	0.333** (0.141)		
Union soldier X Breckenridge vote share, 1860			0.214* (0.123)	0.171 (0.121)
Observations	3312	3,514	3,312	3,514
<i>Panel B. Multinomial Logit, 1880 region of residence (relative to Midwest):</i>				
	<i>South</i>		<i>West</i>	
Union soldier	-0.337** (0.148)	-0.333** (0.143)	-0.975** (0.422)	-1.116*** (0.416)
Observations	1,762	1,885	1,762	1,885
<i>Panel C. OLS, Characteristic of 1880 (Kentucky) county:</i>				
	<i>Confederate enlistment share</i>		<i>Breckenridge vote share</i>	
Union soldier	-0.060*** (0.019)	-0.053*** (0.018)	-0.057*** (0.019)	-0.043** (0.018)
Observations	689	727	706	745
Occupational attainment measure	1950 occupational income	Occupational class indicators	1950 occupational income	Occupational class indicators