Regional Science and Urban Economics 70 (2018) 289–299

Contents lists available at ScienceDirect

Regional Science and Urban Economics

journal homepage: www.elsevier.com/locate/regec

Persistent social networks: Civil war veterans who fought together co-locate in later life

Dora L. Costa [a,b], Matthew E. Kahn [b,c,*], Christopher Roudiez [b], Sven Wilson [d]

[a] UCLA Department of Economics, 9272 Bunche Hall, Los Angeles, CA 90095-1477, United States
[b] NBER, United States
[c] USC, Department of Economics, Los Angeles, CA 90089, United States
[d] BYU Department of Political Science, United States

ARTICLE INFO

JEL classification:
R23
Z13

ABSTRACT

We demonstrate the long reach of early social ties in the location decision of individuals and in their older age mortality risk using data on Union Army veterans of the US Civil War (1861-5). We estimate discrete choice migration models to quantify the trade-offs across locations faced by veterans. Veterans were more likely to move to a neighborhood or county where men from their same war company lived and were more likely to move to such areas than to areas where other veterans were located. Veterans also were less likely to move far from their origin and avoided urban immigrant areas and high mortality risk areas. They also avoided areas that opposed the Civil War. This co-location evidence highlights the existence of persistent social networks. Such social networks had long-term consequences: veterans living close to war-time comrades had a 6% lower probability of dying.

1. . Introduction

Do social capital and ties built in early life persist? Do they lead to individuals living near each other in later life? Recent work (see Ioannides, 2013 for a review) has investigated the impact of interactions on the location decisions of both individuals and firms and how these decisions affect the spatial form of cities. We demonstrate the long reach of early social ties in the location decision of individuals using data on Union Army veterans of the US Civil War (1861-5). It is unusual to be able to study the long-run effects of social capital, but our data permit us to examine a 40 year time span.

More than 2 million white men served in the Union Army in the US Civil War out of a total of more than 5 million men of military age in 1861-1865 (Fogel, 1993). Throughout the war, they fought in companies with the same men and the social capital created in the war determined their probability of deserting and surviving prisoner of war camps (Costa and Kahn, 2003, 2007b). Each surviving veteran faced the decision of whether to move back home or to a new county or city. Men would have been pulled to areas where their fellow veterans were living and the resulting geographic clustering would increase interactions (Helsley and Zenou, 2014) thus reinforcing social capital. Which men would be more likely to live close to fellow veterans? Helsley and Zenou (2014) emphasize that geographic clustering should be higher among those who are the most central to the social network. It also should be higher among those who would benefit the most from geographic proximity; however, individuals will face a trade-off between economic benefits and tastes for proximity (Selod and Zenou, 2006).

We study the correlation between social networks and where veterans lived in later life. Because veterans faced a menu of locational choices, which differed by proximity to their origin location, climate, industrial structure, immigrant structure, and quality of life, we estimate discrete locational choice models so that we can control for location characteristics. We document that veterans were more likely to live in small geographic areas (city wards) together with their former company comrades. We cannot observe the timing of their moves nor their interactions with their fellow veterans but we argue that the geographic clustering of veterans is unlikely to be a coincidence. We recognize that there are alternative explanations for such "co-ordinated migration," such as correlated preferences for local public goods (Manski, 1993) or correlated employment destinations. Imagine a case in which all men from the same war company have similar tastes for avoiding immigrant communities. In this case, two veterans may choose to live in the same geographic area not because they are friends but because they independently decided that a community was best for them. To address this

* Corresponding author at: USC, Department of Economics, Los Angeles, CA 90089, United States.
E-mail addresses: costa@econ.ucla.edu (D.L. Costa), kahnme@usc.edu (M.E. Kahn), croudiez@gmail.com (C. Roudiez), sven_wilson@byu.edu (S. Wilson).

https://doi.org/10.1016/j.regsciurbeco.2017.09.005
Received 20 May 2017; Received in revised form 23 August 2017; Accepted 21 September 2017
Available online 11 October 2017

concern, we compare the locational choice of veterans from the same home towns (and also of all veterans) with that of veterans from the same companies. Veterans, particularly those from the same home towns, arguably have similar tastes for locations. We document that all veterans and veterans from the same towns are less likely to cluster in the same areas compared to men from the same company. We also find that migrants are more likely to co-locate with former company comrades. These findings are consistent with the importance of stronger network ties and of more beneficial network ties for clustering.

The literature on social interactions is extensive (see Durlauf and Ioannides, 2010 for a review), In previous work two of us examined veterans' migration decisions across states and showed that deserters avoided pro-war areas (Costa and Kahn 2007a) and that black veterans were more likely to move to the home states of their fellow soldiers, particularly if they were illiterate, a likely proxy for the importance of network ties (Costa and Kahn 2006). Here we focus on white, honorably discharged veterans, whose ties to their former comrades would have been stronger than those of deserters and whose literacy would have permitted them to use written sources of information. Because we are able to examine more precise location choices we can allow for more heterogeneity in sorting across locations.

Were there any benefits to men of living near each other? We find that all else equal, veterans who live near men from the same company live longer. An emerging social networks literature has linked later life friendships to better health and quality of life of the elderly (e.g., Seeman, 1996). In the last section of this paper, we study the life expectancy of veterans who do and do not live with fellow veterans from their same war company in their geographic neighborhood. We find a positive association between a veteran's longevity and having at least one war time company tie in the local neighborhood but not between a veteran's longevity and the presence of a veteran from another company. Our findings thus suggest that it was the presence of company comrades not the neighborhood that affected longevity.

2. Veteran locational choice

Each veteran must choose where to live. Taking place of origin as given (in our case enlistment), the standard gravity model predicts that veterans in 1900 will be less likely to select destinations further from their origin. We will measure this effect and we will also follow standard revealed preference methods to study how other location specific attributes affect a veteran's propensity to choose to live in that location. Controlling for location specific variables, we are especially interested in the correlation between a veteran choosing a specific geographic location and his wartime comrades co-locating there.

We can imagine several different data generating processes that could yield such co-agglomeration patterns. Consider a case in which the surviving men in a war company, or a subset of them, choose collectively to live in the same location. Now imagine another case where there is asymmetry in the network such that there is one charismatic leader and he chooses where he wants to live and his follower follow him. In both of these cases, we would observe co-agglomeration taking place.

While we cannot disentangle these cases, we have the more modest goal of testing for excess co-agglomeration. Controlling for the physical attributes of a destination and its distance from the origin, do men who knew each other during war time cluster together? We document this excess clustering 25 years after the end of the war using conditional logit models. Our data allow us to test for spatial clusters of war company networks at late ages but we do not know the pathway which led to co-agglomeration.

Our conditional logit models follow the discrete locational choice literature by modeling locations as bundles of tied attributes. The locations we will examine are both counties of residence and neighborhoods within a city. A veteran, i, from company f and originally from location m chooses location j in 1900

$$\text{prob(soldier i chooses j)} = \frac{\exp(\beta_2 V_{fji} + \beta_1 Z_j + \beta_3 D_{jmi})}{\sum_{j=1}^{M} \exp(\beta_2 V_{fji} + \beta_1 Z_j + \beta_3 D_{jmi})} \quad (1)$$

where V_{fji} is the number of veterans from the same company, the same town, or any veterans, and is a veteran-specific attribute of the location, Z_j is a set of location specific attributes such as death risk and immigrant composition, and D_{jmi} is distance from origin, also veteran-specific to each location. Our analysis thus mirrors a standard discrete choice problem of locational choice, augmented to include soldier specific idiosyncratic factors such as the number of veterans from the same company. Throughout our study we assume that the veterans take locational attributes as given. All of our regressions examine where veterans were living in 1900.

We examine various network ties: the number of veterans in the same company, the number of veterans in the entire sample, the number of veterans from the same birth city, and the number of veterans from the same pre-war town. The impact of each of these networks will depend on the strength of these networks. If veterans had closer ties to men from the same company than to other veterans, they would be more likely to co-locate with veterans from the same company than with other veterans. We cannot determine if the company network was the strongest because of ties formed or strengthened during the war or because men from the same companies had the strongest pre-war ties.[1]

We do not know the exact timing of when veterans migrated and whether they followed their former comrades or independently chose to live in the same geographic area. Thus when we examine whether a veteran is more likely to live in an area in 1900 if his comrades are present, we confront Manski (1993) reflection problem. Former comrades may have correlated preferences for location specific attributes and thus any observed clustering may be caused by this sorting independent of the desire for social interactions. But if preferences for location specific attributes are determined by place of birth and drive locational correlations between former comrades then men from the same home town should be at least as likely to sort into the same areas as men from the same company.

Additional identification challenges for company-specific social networks include the presence of "veteran-intensive" jobs in some areas, veteran-specific benefits for local amenities, and any unobservables that determine both friendship formation and location choices. Although we cannot determine what jobs were "veteran-intensive" because the 1900 census did not identify veteran status and the only veteran-specific benefits were pensions and old-age homes (rarely used in 1900), we are comparing veterans from the same company with all other veterans.[2]

As we discuss below, our locational choice models will feature two different levels of geographic analysis. In one, we study the choice of wards where wards are small geographic units. Wards differ with respect to what major city they are located in, their distance from the origin, and other attributes such as the demographic mix of the people who live in the area and its location within the greater metropolitan area. In our other locational choice models, we study the choice of counties. Counties are large geographic units but our county data set provides additional power to study the impact of home town and allows us to control for different sets of attributes.

3. Data

Our data are from three samples collected by the NIA funded Early Indicators project (Costa, PI; Fogel, Original PI) and available at uadata.org.[3]. The first sample consists of roughly 39,338 Union Army soldiers

[1] We have too few men to examine differences between those from the same company and the same pre-war city and those from the same company and different pre-war cities.

[2] The 1910 census identifies veteran status but is better suited to studying retirement than to studying occupations.

[3] For a description of currently available and future data see Costa et al., 2017

in 330 companies, where the companies (of roughly 100 men each) were randomly selected. The sample is representative of the US male population of military age in 1860 (Fogel, 1993). The second sample consists of over 12,671 Union Army soldiers who enlisted in the largest Northern cities in the US in 1860 (Baltimore, Boston, Brooklyn, Chicago, New York City, and Philadelphia).[4] The companies were drawn in proportion to city size in 1860. All men were linked to their pension records, including detailed surgeons' exams. The first sample (original Union Army) was linked to the 1850, 1860, 1900, and 1910 manuscript census schedules and the second sample (urban) was linked to all 1850-1940 manuscript census schedules with the exception of 1890 which was destroyed in a fire.[5] The third sample consists of city maps, with wards, and ward-level characteristics for Baltimore, Boston, Brooklyn, Chicago, Cincinnati, New York City, and Philadelphia from 1850 to 1930.[6]

Civil War companies contained roughly 100 men and were generally not replenished with new men when disease, military casualties, and desertions whittled down its numbers. Although recruitment was local, companies were diverse.[7] At the beginning of the war, men would enlist with one or several friends but rarely with fifty and once companies were full, they would take no more men. Later in the war, men might enlist in a distant town to receive a large bounty. The need to travel to recruiting stations, particularly for farmers' sons, increased geographic diversity as well (Costa and Kahn, 2008: 59-60).

In conducting our analysis, we must confront the following limitation. Suppose that 50 men in a company survive the war. Let 40 of them return to civilian life without interacting further with their comrades while 10 of them continue to stay in touch. This smaller subset would be more likely to cluster near each other in later life. Our data sampling captures their later life geographic distribution. We cannot claim that one subset of the 10 caused the others to cluster near them. Instead, if we observe correlated clusters, this is consistent with the claim that this group sought to benefit from continued interactions. Because wards were walkable continued interactions were possible. Because men could choose among wards with similar characteristics, they could have picked a ward without men from their company.

We used our three samples to create two different data sets, both of which we restricted to the uninstitutionalized. Our first data set consists of all veterans for whom we know county of enlistment and county of residence in 1900 and for which we have information on county characteristics in 1900, i.e., 7600 men who could choose among 2752 US counties.[8] Fig. 1 shows the number of veterans in each county at enlistment and in 1900. The movement west mirrors national trends. Roughly 3% of native-born veterans were in the Pacific states in 1900 compared to 4% of all native-born men born in a Union state and of military age during the Civil War.[9]

When we examine county of choice for the original Union Army sample, our primary variables of interest are, in the county, the number of veterans from the same company, the number of veterans in the sample, the number of veterans from the same 1860 town, and the number of veterans from the same enlistment town. The highest correlation (0.3) among these network measures was between the number of men in the county from the same company and the number of men from the same pre-war town in the county. Additional county characteristics we constructed were county population, the fraction of workers in manufacturing, the fraction of "new" immigrants, distance from origin county to destination county (measured using county centroid), dummy variables indicating southern and coastal counties, mean February temperature, and voting in the 1864 and 1900 presidential elections. Unfortunately, we can observe neither local unemployment rates nor local area economic shocks. See the Data Appendix for details.

The second data set consists of everyone we could place in a ward in 1900 in Baltimore, Boston, Chicago, Cincinnati, New York City (all boroughs, including Brooklyn), and Philadelphia and for whom we knew city of enlistment. These six cities were among the top 10 cities by population size in the United States in 1900 and contained 44% of the population of the top 100 cities.[10] This sample contains 1,387 men who could choose from 222 wards. The smallest ward contained 1,488 people and the largest 476,602 (the median was 24,048). Of the men in the sample, 853 are from the urban sample and 534 are from the original Union Army sample.[11] Among the men, 13% are from Baltimore, 7% from Boston, 18% from Chicago, 4% from Cincinnati, 30% from NYC (including all boroughs), and 27% from Philadelphia. Fig. 2 shows the number of veterans by ward in each city. The median size of a ward was 1,870,661 square meters, implying that within a median size ward a veteran was no more than 24 minutes by foot from another person.[12]

Our primary variables of interest are the number of veterans from the same company in the ward, the number of veterans in the sample in the ward, the number of veterans in the ward from the same 1860 town, and the number of veterans from the same birth city in the ward. In our city sample, the highest correlation is between the last two measures (0.5) and the other correlations are 0.3 or less.

We constructed several variables describing ward characterisics: an adjusted death rate, distance to the city center, population density, and the fraction of "new" immigrants (immigrants born in Eastern or Southern Europe), blacks, Irish, and Germans. We also constructed city-level variables: a dummy variable equal to one if the current city of residence was the same as the city of enlistment, the distance from the city of enlistment to the current city, and city population in 1900. Details are provided in the Data Appendix.

We also created individual level variables which we used for interaction terms: a dummy equal to one if the veteran was a professional, proprietor, or artisan at enlistment, dummies equal to one if the veteran was born in Ireland or Germany, pension amount collected in 1900 (an exogenous income transfer), a dummy variable indicating poor health in 1900, and a dummy equal to one if health status was unknown. Details are given in the Data Appendix.

The accompanying Data In Brief article, "Using the Union Army Samples to Study Locational Choice and Social Networks," provides additional details on the construction of the original Union Army sample of 39,341 men and of the urban over-sample. We present sample means showing the differences between the two samples and present means for our county and ward samples. Our means show that veterans were more likely to live in either counties or wards with former comrades from their companies and with other fellow veterans. In the next section, we examine these correlations, controlling for other ward and county characteristics.

4. Results

We examine both county choices and ward choices within a city using

[4] City of enlistment does not necessarily equal city of residence in 1860. Many men were from outlying areas.

[5] Linkage was done by hand. This form of linkage has the advantage of low false positive rates (see Bailey, 2016). Our linkage rates to the 1900 census of all men known to be alive in 1900 are 84% for the original Union Army sample and 80% for the urban sample. The high linkage rate for the urban sample is probably due to improvements in linkage technology.

[6] This collection complements the urban sample. Cincinnati is not in the urban sample because the majority of its enlistees were from outside the city.

[7] Roughly 95% of all Union Army soldiers were volunteers, with the remainder divided between draftees and substitutes.

[8] US territories are excluded.

[9] Estimated from the sample and from the 1900 census of Ruggles et al. (2015).

[10] Calculated from https://www.census.gov/population/www/documentation/twps0027/tab13.txt.

[11] There is no difference in the magnitude of our main results if we restrict the sample to either the urban sample or the original Union Army sample.

[12] We assume 80 m for 1 min of walking time and a square ward.

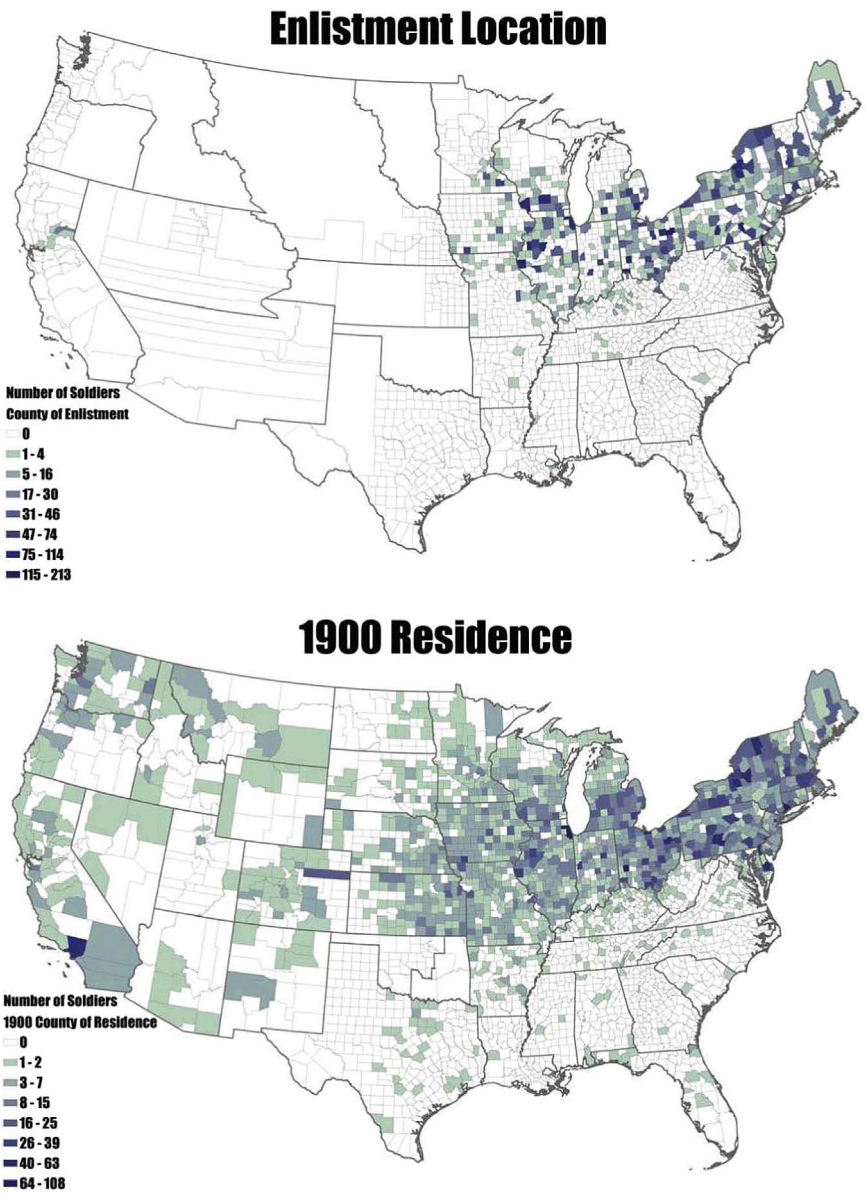

Fig. 1. County Location of Veterans at Enlistment and in 1900.

a smaller, urban sample. A ward was a political unit used to elect city councilmen and also a unit for city neighborhood statistics. An advantage of examining ward rather than county locations is that there are fewer potential confounders that could lead to correlated locations. Our ward regressions control for city population.[13] Our identification thus comes from ward differences within cities. Another advantage of examining wards is that we observe ward mortality rates whereas we do not know county mortality rates. An advantage of studying counties is that we can use a larger dataset and thus have greater power to compare hometown and company social networks. County-level data also provide a different set of observable controls, including ideology, which we found to be an important determinant of migration in our past work (Costa and Kahn, 2007a).

We expect veterans to live in areas where they share commonalities with the residents. One commonality is ideology. Veterans should thus avoid the South and counties where McClellan, the "peace without victory" candidate in the 1864 election won a greater share of the vote.[14] Another commonality is ethnicity. Veterans would share little in common with the new immigrants arriving from southern and eastern Europe. In addition, if they were not Irish, we would expect that they would not want to live with the Irish, a group dominated by poor laborers. Whites (and the Irish and new immigrants were not yet regarded as white) have a distaste for living with minorities (Cutler et al., 2008; Shertzer and Walsh, 2016).

We also expect veterans to seek out economic opportunities, such as those provided by more populous areas, particularly if the veteran was in a skilled occupation and to seek out amenities such as warmer winter temperatures and lower disease risk, but also, because of the costs of migration, to remain close to home, as predicted by the classic gravity

[13] When we ran specifications with city fixed effects rather than city population, we obtained very similar results.

[14] Lincoln won 76% of the soldier's vote compared to 55% of the popular vote (Burnham, 1955: 260-83). Costa and Kahn (2007a) find that Union Army deserters avoided pro-Lincoln (in the 1864 election) areas. Eli et al. (2016) find that Union Army veterans from Kentucky were more likely to leave pro-Confederate counties for pro-Union areas whereas Confederate veterans did the opposite.

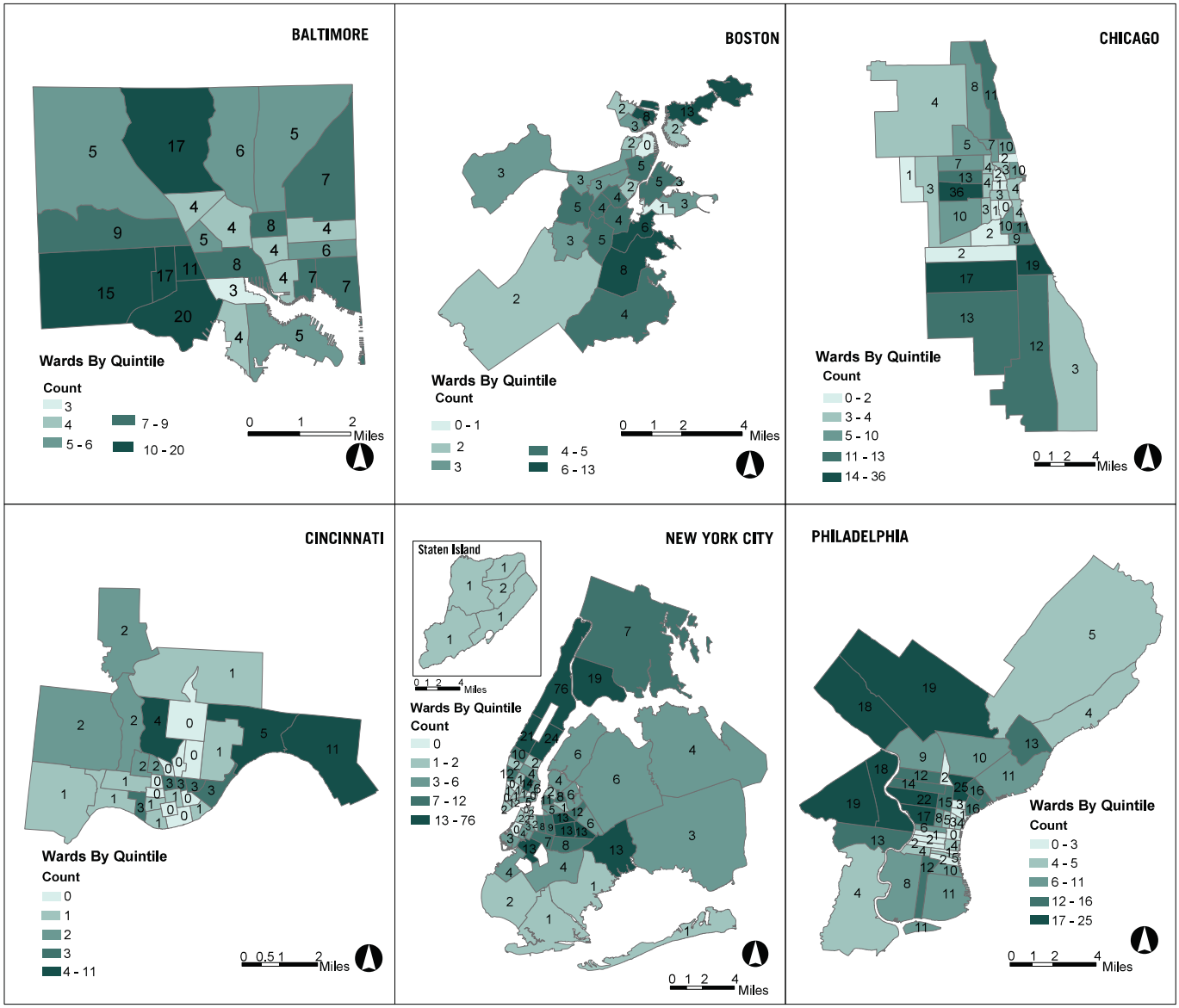

Fig. 2. Location Within City Wards of Veterans in 1900.

model of migration. We do not have any predictions as to whether a veteran would prefer to live close or further from the central business district. Although rents were presumably lower and commuting costs higher further from the central business district, the rise of the streetcar enabled workers (and most veterans were still in the labor force) to live far (in distance) from the central business district but still obtain quick access to it (Gin and Sonstelie, 1992; LeRoy and Sonstelie, 1983). After World War II, rising household income explains much of post-war suburbanization (Margo, 1992).

We investigate how pension income affected veterans' choices. Controlling for health, a Union Army pension was an arguably exogenous income transfer (Costa, 1995). We expect that richer veterans would be less likely to live with minorities. Unfortunately, we cannot examine how much more they were willing to pay in rent to avoid minority areas. One of the drawbacks of examining this time period is that information on rents is unavailable.

4.1. Locational choice: counties

Twenty-six percent of all veterans in 1900 were living in their county of enlistment. Controlling for county characteristics, a veteran was 17 times as likely to live in a county if he had enlisted in it (see the first specification in Table 1 which presents the odds ratios from a conditional logit model of county choice). His probability of living in a county that was in the former Confederacy was 95% lower compared to his probability of living in a former Union county or territory controlling for distance from enlistment.

4.1.1. Veteran networks

Veterans were more likely to pick a county where there were other men from their company (see Table 1). An extra man from their war-time company increased their probability of living in the county by 21%. An extra veteran who was not from the same company increased the probability of living in the county only by 1%. An extra veteran from the same pre-war town increased this probability by 8% and an extra veteran from the same birth city increased this probability by 3% (see the first specification). The odds ratio on the number of men from the same company was statistically significantly different from each of the odds ratios on the other network measures at the 0.1% level. We find evidence that fellow company members were mainly valuable to men who had not enlisted in the county. The impact of a fellow company member increased the probability of staying in the county by 51% for men who were not from

Table 1
County Locational Choice Regression.

	(1)		(2)		(3)	
	e^{β}	Std. Err	e^{β}	Std. Err	e^{β}	Std. Err
Dummy = 1 if same as county of enlistment	16.834‡	2.705	102.337‡	9.502	102.730‡	9.590
Number of veterans from the same company	1.209‡	0.024	1.509‡	0.047	1.509‡	0.047
Number of other veterans	1.006 ‡	0.002	1.008‡	0.002	1.008‡	0.002
Number of veterans from the same 1860 town	1.077†	0.036	1.470‡	0.184	1.469‡	0.189
Number of veterans from the same birth city	1.026	0.019	1.065‡	0.017	1.066‡	0.017
Dummy = 1 if same as county of enlistment ×						
Number of veterans from the same company			0.710‡	0.023	0.710‡	0.023
Number of other veterans			0.977‡	0.023	0.976‡	0.003
Number of veterans from the same 1860 town			0.717‡	0.090	0.717‡	0.092
Number of veterans from the same birth city			0.965†	0.016	0.965‡	0.017
Percentage voting for McClellan in 1864	0.992‡	0.001	0.992‡	0.001	0.992‡	0.001
Percentage voting for McKinley in 1900	0.997	0.244	0.997	0.002	0.997	0.002
Logarithm of county population	1.717‡	0.067	1.641‡	0.053	1.650‡	0.053
Professional, proprietor, or artisan at enlistment ×						
Logarithm of county population	1.159‡	0.031	1.186‡	0.030	1.181‡	0.030
Mean February temperature (Fahrenheit)	1.010*	0.006	1.013‡	0.005	1.013†	0.005
Dummy = 1 if coastal county	1.071	0.118	0.971	0.084	0.968	0.084
Distance from enlistment county in miles	0.997‡	0.000	0.997‡	0.000	0.998‡	0.000
Dummy = 1 if former Confederacy	0.051‡	0.006	0.051‡	0.006	0.051‡	0.006
Fraction of wage earners in manufacturing	0.083‡	0.061	0.121‡	0.072	0.125‡	0.075
Fraction of “new” immigrants	0.005‡	0.007	0.020‡	0.021	0.693	0.943
Fraction of “new” immigrants ×						
Monthly pension					0.603‡	0.063
Poor health					12.249†	14.481

20,100,608 observations where each observation is each person’s choice of a county. The coefficients are exponents from a conditional logit model. The symbols ∗, †, and ‡ indicate statistical significance at the 10, 5, and 1 percent level. Robust standard errors, clustered on the company level. Additional controls include dummies for missing voting information in the 1864 and 1900 elections (the fractions voting for McClellan and McKinley were set equal to 0 if this information was missing) and a dummy indicating missing health information (poor health was set equal to 0 if this information was missing) interacted with the fraction of “new” immigrants.

the county but decreased the probability for men who were from the county (see the second specification). We found no evidence that wartime company cohesion, arguably a measure of network strength, mattered (results not shown). When we interacted the number of men from the same company with pension amount, we found a statistically significant increase in the odds of living with men from the same company but the magnitude of the effect was negligible.

4.1.2. County characteristics

Veterans were less likely to live in a county where a greater percentage of the electorate had voted for McClellan, the Democratic “peace without victory” candidate in the 1864 election. An increase of 1 percent in the share of the county voting for McClellan lowered their probability of living in that county by 1%. This is not an indicator of current party of affiliation – they were less likely (but not statistically significantly so) to live in a county where McKinley, the Republican presidential candidate and supporter of the high tariffs which financed Union Army pensions, had a larger percentage of the vote in 1900. Veterans also avoided the former Confederary (their probability of living there was 95% lower than that of living in the former Union or its territories) and counties with a high fraction of the population working in manufacturing.

Veterans were more likely to live in a county with a larger population, particularly if they were professionals, proprietors, or artisans at enlistment. They also were more likely to live in a county with a warmer February temperature and one that was close to county of enlistment, controlling for their propensity to stay in their county of enlistment.

Veterans also were more likely to avoid counties with a high fraction of ‘“new” immigrants. Our second specification shows that each increase of.1 in the fraction of new immigrants decreased their probability of living there by 10%. Those who received a higher pension, an exogenous income transfer, were more likely to avoid counties with “new” immigrants (see the third specification).

4.2. Locational choice: city wards

4.2.1. Vetren networks

Veterans were more likely to choose a ward popular with other veterans from the same company (see Table 2 which presents the odds ratios from a conditional logit model of ward choice). In our first specification, an extra man from a veteran’s war-time company increased the probability of choosing that ward by 32%. Veterans also were more likely to choose a ward popular with all veterans and with veterans from the same city. An additional veteran increased the probability of choosing a ward by 6% and an additional veteran from the same city increased the probability of choosing a ward by 12%. (The odds ratios on other ward characteristics provide some comparisons of magnitudes: the odds of picking a ward were 8.9 times greater if the veteran had enlisted in that city and each additional 1000 m from the central business district decreased the probability of choosing that ward by 2%.) The odds ratio on the number of veterans from the same company in the same ward was statistically significantly different from the odds ratios on the number of veterans in the ward and the number of veterans from the same birth city in the ward at the 1 and 10% level, respectively. We interpret the larger effect of an extra man from the same company as indicative of the strength of the network, not as indicative of shared tastes, which would likely be shared by veterans from the same city. We also investigated the impact of an extra man from the same company and birth city (arguably an even stronger network tie than the same company) but the effect was not statistically significant (see the first specification in Table 3), perhaps because of sample size issues.

We find that having a veteran in the same ward was less important to men who remained in their city of enlistment, suggesting that fellow veterans were a source of information or direct assistance for the nonnative. Our second specification in Table 2 shows that among those who were from a different city having a fellow veteran in the same ward increased men’s probability of choosing that ward by 78% but a fellow veteran in the same ward increased the probability of a native choosing

Table 2
City Ward Locational Choice Regression.

	(1)		(2)		(3)		(4)		(5)	
	e^{β}	Std. Err.	e^{β}	Std. Err.	e^{β}	Std. Err.	e^{β}	Std. Err.	e^{β}	Std. Err.
Number of veterans in ward from same company	1.315‡	0.044	1.781‡	0.162	1.760‡	0.163	1.759‡	0.007	1.748‡	0.162
Number of veterans in ward	1.059‡	0.005	1.058‡	0.006	1.061‡	0.007	1.061‡	0.007	1.061‡	0.007
Number of veterans in ward from same birth city	1.115‡	0.022	1.058‡	0.007	1.230‡	0.028	1.229‡	0.028	1.227‡	0.028
Same city as city of enlistment	8.924‡	1.673	10.508‡	2.024	10.282‡	1.984	10.332‡	1.994	10.399‡	2.012
Same city as city of enlistment ×										
Number of veterans in ward from same company			0.727‡	0.067	0.728‡	0.067	0.728‡	0.067	0.733‡	0.068
Number of veterans in ward			1.002	0.008	1.001	0.008	1.001	0.008	1.062‡	0.007
Number of veterans in ward from same birth city			0.847‡	0.025	0.848‡	0.028	0.849‡	0.025	0.849‡	0.025
Distance of city from city of enlistment	0.997‡	0.000	0.997‡	0.000	0.997‡	0.000	0.997‡	0.000	0.997‡	0.000
City population	0.917	0.066	0.925	0.067	0.940	0.068	0.941	0.068	0.941	0.000
City population ×										
Professional, proprietor, or artisan at enlistment	1.116	0.084	1.116	0.083	1.125	0.085	1.124	0.085	1.126	0.084
Adjusted ward death rate	0.996‡	0.002	0.996‡	0.002	0.996‡	0.002	0.996‡	0.002	0.996‡	0.002
Ward population density	5230.775‡	10732.20	5725.864‡	11,808.80	5293.559‡	10,945.74	5069.638‡	10,404.61	5391.181‡	11,082.09
Logarithm of distance from ward to city center	1.242‡	0.071	1.234‡	0.070	1.222‡	0.069	1.222‡	0.069	1.222‡	0.069
Fraction of blacks in ward	0.558	0.352	0.532	0.330	0.461	0.277	0.461	0.277	0.448	0.271
Fraction of "new" immigrants in ward	0.028‡	0.019	0.025‡	0.017	0.024‡	0.016	0.113‡	0.109	0.115‡	0.109
Fraction of Irish-born in ward	0.923	0.978	0.757	0.794	0.090†	0.101	0.091†	0.103	0.223	0.381
Fraction of German-born in ward	1.305	1.169	1.179	1.062	0.057‡	0.051	0.058‡	0.052	0.007‡	0.105
Fraction of Irish-born in ward ×										
Irish-born					112052.80‡	214643.10	112457.60‡	215658.00	127330.10‡	232586.20
Monthly pension amount									0.753†	0.094
Fraction of German-born in ward ×										
German-born					24,207.76‡	25,855.30	24,021.15‡	25,691.06	22,509.55‡	24,118.33
Monthly pension amount									1.058	0.094
Fraction of "new" immigrants in ward ×										
Monthly pension amount							0.872*	0.068	0.854†	0.064
Pseudo-R^2	0.224		0.227		0.235		0.235		0.236	

307,914 observations where each observation is each veteran's choice of a ward. The coefficients are exponents from a conditional logit model. The symbols *, †, and ‡ indicate statistical significance at the 10, 5, and 1 percent level. Robust standard errors, clustered on the company level. Additional controls in regressions 4-5 are interactions of dummies for poor health and for missing health information with the fraction of new immigrants. Additional controls in regression 5 include interactions of dummies for poor health and for missing health information interacted with the fraction Irish and the fraction German.

that ward by only 5%. Although veterans were more likely to live in the same ward as fellow veteran from the same company and birth city, an extra veterans from the same company or birth city increased the probability of choosing that ward by only 6% (second column of Table 2).

We examined interactions between the number of veterans in a ward and measures of the strength of other ties and of the war-time network (see Table 3). We found that married veterans were more likely to live in a ward with fellow veterans suggesting that existing relationships did not weaken ties with veterans. We found no evidence that a more cohesive war-time company made veterans more likely to move to a ward with veterans.[15] We did not find that a higher company death rate, whether overall or from wounds, arguably a measure of the emotional intensity of wartime ties led veterans to move to wards with veterans.

We also examined interactions between the number of veterans in a ward and measures of economic status (see the last two columns of Table 3). We interacted the number of veterans in a ward with a veteran's occupational score at enlistment but the interaction term was statistically insignificant suggesting that information or assistance from fellow veterans was not more important for men in lower or higher occupational classes. We did not find that a higher pension increased the probability of living in a ward with more men from the same company implying that income and fellow company members were neither substitutes nor complements.

Any non-experimental networks study faces the concern that observed patterns are generated by correlated unobservables. We downplay this concern both because of our same birth city and all other veterans findings and because wards were not labor markets. Wards were residential communities. We have controlled for distance to the central business district, a center for jobs, as well as other ward attributes. In the period we study, jobs were more concentrated in the central business district than they are today. In Chicago, 80 percent of jobs were in the central business district in the nineteenth century (Fales and Moses, 1972).

4.2.2. *Ward characteristics*

A veteran's probability of choosing a ward was lower if it was a high mortality ward, if the fraction of "new" immigrants was higher, if the ward was further from the center city, and if the city was closer to the city of enlistment. We found no differential effect of distance from the center city by veterans' retirement status. When we interacted whether a veteran was Irish with the fraction of the ward born in Ireland we found that

[15] We measured cohesion by creating an index based on company heterogeneity in occupation, birth place, and age, where the weights on each variable were determined by the coefficients on a regression of each factor on the probability of desertion.

Table 3
City Ward Locational Choice Regressions: Interactions.

	(1) e^{β}	(2) e^{β}	(3) e^{β}	(4) e^{β}	(5) e^{β}	(6) e^{β}	(7) e^{β}	(8) e^{β}
Number of veterans from the same company and birth city	1.110							
	(0.096)							
Number of veterans in ward from same company	1.288‡	1.687‡	1.762‡	1.892‡	1.837‡	1.803‡	1.881‡	1.782‡
	(0.045)	(0.166)	(0.099)	(0.169)	(0.232)	(0.214)	(0.235)	(0.179)
Same city as city of enlistment	9.002‡	10.602‡	10.508‡	10.490‡	10.485‡	10.504‡	8.064‡	10.509‡
	(1.698)	(2.044)	(2.022)	(2.029)	(2.024)	(2.024)	(1.733)	(2.022)
Same city as city of enlistment ×								
Number of veterans in ward from same company		0.720‡	0.730‡	0.760‡	0.726‡	0.727‡	0.742‡	0.729‡
		(0.069)	(0.067)	(0.069)	(0.068)	(0.066)	(0.086)	(0.070)
Number of veterans in ward from same company ×								
Dummy = 1 if veteran married in 1900		1.093†						
		(0.041)						
Dummy = 1 if laborer at enlistment			1.087					
			(0.099)					
Company cohesion				0.887				
				(0.076)				
Fraction of company dying in war					0.767			
					(0.334)			
Fraction of company dying of wounds						0.793		
						(0.727)		
Occupational score in 1900							0.998	
							(0.002)	
Pension amount in 1900								0.999
								(0.003)
Dummy = 1 if in poor health								1.035
								(0.047)

307,914 observations where each observation is each veteran's choice of a ward. The coefficients are exponents from a conditional logit model. The symbols *, †, and ‡ indicate statistical significance at the 10, 5, and 1 percent level. Robust standard errors, clustered on the company level. For additional controls, see Table 2.

the non-Irish avoid areas with a high fraction of Irish but that the Irish were more likely to be in areas with a high fraction of Irish. An increase of 0.1 in the fraction of the ward that was Irish decreased the probability of the non-Irish moving there by 9% but increased the probability of the Irish moving there by more than 11,000%, illustrating how highly segregated immigrants were in ethnic enclaves. We found similar effects for the German-born and the fraction of the ward that was German (see the third specification).

The probability that a veteran would locate in a ward fell by 9% if the fraction of the ward that was composed of "new" immigrants was 0.1 greater. Our results imply that the veterans were 50% less likely to live in the ward with the greatest fraction of "new" immigrants (0.56) than in the ward with the lowest fraction (0.00). Veterans who received a larger pension were less likely to live in "new" immigrant areas (see the fourth and fifth specifications). Each additional dollar increase in monthly pension amount decreased the probability of living in a ward where the fraction of new immigrants was 0.1 greater by an additional 2%. The fifth specification also shows that veterans who received a larger pension also were less likely to live in areas with a high fraction of Irish. (There was no statistically significant effect of pension on the probability of living in a German ward.) We found no evidence that pensions affected the probability of living in a higher mortality ward. We also found no evidence that pensions affected the probability of living further from the city center (results not shown).

4.3. Benefits of Social Networks

What benefits did a veteran derive from having a war-time company member living nearby? We can provide a suggestive answer using our urban sample. Veterans were not randomly assigned to locations and a veteran's death could be indicative of a common ward-level mortality shock. But common mortality shocks should affect both all veterans from the same company in a ward and all other veterans in ward. We thus can compare the mortality benefits of veterans from the same company with those of all other veteran to provide suggestive evidence that veterans chose to participate in the wartime network and to test theories of the benefits of social networks. We estimate a Gompertz hazard model of the form

$$h(t) = h_0(t)\exp(\beta_x X) \tag{2}$$

where $h_0(t)$ is the baseline hazard and where X is a vector of characteristics specific to each veteran, including a time-varying covariate of living near a fellow veteran from the same company, a time-varying covariate of living near a fellow veteran from a different company, a time-varying covariate of living near a fellow veteran from the same city but not from the same company, a time-varying covariate of having a living spouse, and various control variables, including city fixed effects. We cluster the standard errors on the ward. We also estimate a random effects (shared frailty) model in which we assume that there is an unobserved multiplicative effect, α, such that

$$h_j(t|\alpha) = \alpha_j h_j(t) \tag{3}$$

where α is shared by all veterans in ward j and is assumed to be distributed as a gamma distribution.[16] An estimate of the frailty variance, θ, that is statistically significantly different from 0 suggests the presence of a common ward-level mortality effect.

Table 4 shows that having a man from the same war-time company in the same ward in 1900 decreased a veteran's probability of dying by 6%. In our first specification (without shared frailty), the coefficient on the number of men from the same wartime company is statistically significantly different from 1 at the 5 percent level. The coefficient is statistically significantly different from 1 at the 10% level in our shared frailty specification but our estimate of the shared frailty parameter suggests that there are no common ward-level mortality effects. We therefore prefer the first specification. In both specifications, the impact of the odds ratio on the number of veterans from the same company is statistically significantly different from the odds ratio on the number of other veterans and the number of other veterans from the same city but not from the same company, suggesting that no common ward-level mortality were operating. The odds ratio on the number of veterans, other than

[16] We have too few observations to control for common ward-level mortality effects using ward fixed effects.

Table 4
Mortality Regressions.

	e^{β}	Std. Err	e^{β}	Std. Err
Time-varying covariates:				
Number of veterans from company in 1900 ward	0.943†	0.027	0.947*	0.031
Number of other veterans in 1900 ward	1.004	0.013	1.011	0.014
Number of veterans born in the same city but not in the same company	1.025	0.027	1.026	0.025
Dummy = 1 if living wife	0.974	0.057	0.974	0.062
γ	0.007‡	0.000	0.007‡	0.000
θ (shared ward-level frailty)			0.010	0.017
Test of $\theta = 0$				
$\chi^2(1) =$			2.17	
$p > \chi^2 =$			0.070	
Test of equality odds ratios of number of veterans from the same company and number of other veterans				
$\chi^2(1) =$	3.90†		3.26*	
$p > \chi^2 =$	0.048		0.071	
Test of equality odds ratios of number of veterans from the same company and number of veterans born in the same city but not in the same company				
$\chi^2(1) =$	3.62†		3.38*	
$p > \chi^2 =$	0.057		0.066	
Test of equality odds ratios of veterans from the same company and wife living				
$\chi^2(1) =$	0.25		0.15	
$p > \chi^2 =$	0.616		0.696	

1213 veterans. The coefficients are hazard ratios from a gompertz parametric survival model of months lived with both time-varying (by the month) and time-invariant covariates. The symbols *, †, and ‡ indicate statistical significance at the 10, 5, and 1 percent level. Robust standard errors, clustered on the ward level, in the first regression. The regressions include a constant and controls for age in 1900, the logarithm of ward-level death rates, a dummy variable equal to one if in poor health in 1900, a dummy equal to 1 if information on poor health was missing (poor health was set equal to 0 if this information was missing), a dummy variable equal to one if wounded, a dummy variable equal to one if in the same city as city of enlistment, a dummy variable equal to one if born in Ireland, a dummy variable equal to one if a laborer in 1900 and a dummy variable equal to one if this information was missing, a dummy variable equal to one if a laborer at enlistment, a dummy variable equal to one if enlisted in a city of 50,000 plus inhabitants (one of the 13 largest city in the US in 1860), an interaction term between laborer at enlistment and enlisted in a city of 50,000 plus, and dummy variables for city of enlistment.

those in the same company, is not statistically significantly different from 1 and the magnitude is small and greater than 1. Social networks that arguably are not as strong have a smaller impact on mortality. The impact of having a living wife was not statistically significantly in either specification, but we could not reject the hypothesis that the odds ratios on the number of veterans from the same company and having a living wife were statistically significantly different.

How large is a mortality effect of 6%? Our control variables (not shown) included whether the veteran was in poor health in 1900, whether he had been wounded during the war, and whether he was Irish-born. The coefficients on these variables indicate that mortality was 13, 12, and 33 percent higher among those veterans than among their more fortunate counter-parts. Our results are consistent with a 20 year follow-up of older Dutch respondents which found that the probability of death was 1% lower among those embedded in large networks and 5% lower among those embedded in diverse networks (Ellwardt et al., 2015).

As a robustness test, we instrumented for the number of veterans from the same company within the ward using the total number of veterans in the same company and the mean distance from place of enlistment to the city for all men in the company using a control function approach. The total number of veterans from the same company is arguably the maximum number of men at risk to be within a single ward. The further the distance from their place of enlistment to the city of residence in 1900, the less likely they would be to live there. We are assuming that a small total number of men from the same company (which could arise from high wartime mortality) is not measuring the strength of network ties and that distance from place of enlistment to city of residence in 1900 does not measure company cohesion. We obtained a statistically insignificantly different from 1 odds ratio of 0.926 ($\sigma = \hat{}$ 0.039), similar in magnitude to our favorite specification. Because a Hausman test revealed no evidence of endogeneity, we suspect that our estimated odds ratio is not statistically significant because of sample size but also acknowledge that our instruments might not be valid.

We found mixed evidence that those living near a former comrade experienced better economic outcomes (see Table 5). The impact of a former comrade was statistically insignificantly different from 0 in determining labor force participation. Conditional on being in the labor force, men living near a former comrade were statistically significantly less likely to have a higher occupational income score but the coefficient was statistically indistinguishable from that on the number of other veterans and the number of other veterans born in the same city but not the same company. In addition, nearby veterans of any type did not change the odds of being a laborer in 1900 (conditional on being in the labor force). We also found that nearby veterans did not change the odds of being the head of the household. However, men having a nearby comrade were more likely to own their homes and the impact was statistically significantly different from that of all other veterans and all other veterans born in the same city but not from the same company.[17]

Although we cannot pinpoint why social networks may have had positive mortality effects, we find no evidence that living near a fellow veteran was beneficial because of the information and assistance he could provide. The impact of a having a former comrade nearby was statistically indistinguishable between migrants and non-migrants (results not shown). We also find no evidence that the strength and intensity of the tie was an important determinant of survival – the interaction term between number of fellow veterans living nearby and proxies for the strength and intensity of the tie such as company cohesion and the fraction of the company dying was statistically insignificant (results not shown). One plausible explanation for the beneficial effect of social networks is biological processes such as improved cellular immune responses and neuroendocrine functioning (Seeman, 1996). Feelings of social isolation may even be linked to alterations in the activity of genes that drive inflammation, the first response of the immune system (Cole et al., 2007).

5. Conclusion

According to Oliver Wendell Holmes, "the generation that carried on the war has been set apart by its experience."[18] At least locationally, veterans were set apart. They selected to be with fellow veterans, preferably with veterans from the same company, with whom they would have had stronger ties. These locational preferences were strongest for migrants, suggesting that veteran networks were a source of material support such as job information. Veterans avoided the South and areas populated by the new immigrants who arrived after the Civil War and an income transfer made them even more likely to avoid immigrant areas. Veterans who lived near former comrades from the same company faced a lower mortality risk.

Were there other effects of veterans' locational choices? Geographic clustering leads to knowledge spillovers and thus has economic

[17] All homeowners were household heads. When we examined household heads only, we obtained a similar and statistically significant coefficient on the number of veterans from the same company of 0.029. There were no housing benefits for veterans.

[18] 1895 Memorial Day Speech, http://www.people.virginia.edu/mmd5f/holmesfa.htm.

Table 5
Occupational Outcome Regressions.

	Out of the Labor Force $\frac{\partial P}{\partial X}$	Occupational score Coef.	Laborer $\frac{\partial P}{\partial X}$	Household head $\frac{\partial P}{\partial X}$	Home owner $\frac{\partial P}{\partial X}$
Mean of dependent variable	0.272	27.541	0.341	0.252	0.844
Number of veterans from company in 1900 ward	−0.023	−0.467†	0.011	0.004	0.024‡
	(0.016)	(0.210)	(0.015)	(0.009)	(0.008)
Number of other veterans in 1900 ward	−0.002	0.053	0.000	0.001	−0.003
	(0.001)	(0.041)	(0.002)	(0.001)	(0.002)
Number of veterans born in the same city but not in the same company	0.020	0.431	0.027	0.008	−0.030*
	(0.034)	(0.719)	(0.043)	(0.021)	(0.016)
Test of equality coefficients of number of veterans from the same company and number of other veterans					
$\chi^2(1) =$	1.69		0.46	0.07	9.44
$p > \chi^2 =$	0.194		0.498	0.788	0.002
$F(1,179) =$		2.40			
$p > F =$		0.123			
Test of equality coefficients number of veterans rom the same company and number of veterans born in the same city but not in the same company					
$\chi^2(1) =$	0.92		0.08	0.02	6.47
$p > \chi^2 =$	0.336		0.774	0.891	0.011
$F(1,179) =$		0.00			
$p > F =$		0.978			
Observations	1213	873	850	1213	1213

Estimates are the derivatives from a probit, with the exception of the occupational score regression which is a simple linear regression. The symbols *, †, and ‡ indicate statistical significance at the 10, 5, and 1 percent level. Robust standard errors, clustered on the ward level, in the first regression. The regressions include the variables listed in the footnote to Table 4, as well as pension amount.

implications. It also has political implications because spatial clustering can reinforce political narratives. Union Army veterans intentionally avoided the anti-War areas and the South, thus leading to their agglomeration in specific areas. This spatial clustering provided Union Army veterans with voting power (which favored the Republican party) and with the members for organizations that could campaign for keeping the Civil War in public memory (Logue, 2007; McConnell, 2007). Thus both selection (of the types of people who moved to different areas) and treatment effects (the social interactions that take place among people who live in a given place) meant that the Tiebourt sorting of veterans re-inforced a pro-Union narrative of the Civil War. A similar ideological sorting among Confederate veterans, as documented by Eli et al. (2016), would have reinforced ideological divisions in the country.

Acknowledgements

Dora Costa, Christopher Roudiez, and Sven Wilson gratefully acknowledge the support of NIH grant P01 AG10120. Dora Costa also acknowledges the use of facilities and resources at the California Center for Population Research, UCLA, which is supported in part by NICDH grant P2C R24HD041022.

Appendix A. Data Appendix

A.1. County choice regressions

The data set is the original Union Army sample, available at uadata.org. Our explanatory variables are

1. *The number of veterans from the same company, the number of veterans in the sample not from the same company, the number of veterans from the same 1860 town, and the number of veterans from the same enlistment town*. These variables are specific to the veteran who is excluded from the calculation of the variables. The variables were generated from our primary sample. The number of veterans and the number of veterans from the same birth city in the ward are thus underestimated.
2. *Logarithm of county population, the fraction of workers in manufacturing, and the fraction of "new" immigrants*. These variables were generated from Haines (2010). "New" immigrants are those from Eastern and Southern Europe.
3. *Distance from the origin county to the destination county*. Measured in miles, from county centroids.
4. *Dummy variables indicating southern and coastal counties*.
5. *Mean February temperature*. Obtained from https://www.ncdc.noaa.gov/cdo-web/.
6. *The percentage voting for McClellan in 1864, the percentage voting for McKinley in 1900, and dummy variables for each year indicating that no data were available*. Obtained from Clubb et al. (2006).
7. *The interactions between pension amount in 1900 and the fraction of "new" immigrants*. Pension amount (in dollars per month) was arguably an exogenous income transfer which depended upon health status and whether the veteran could claim his disability was related to the war. Veterans who could argue that their rheumatism was caused by being out in the damp during a march received more money than veterans whose rheumatism could not be related to the war according to the medical theories of the time.
8. *The interaction between poor health status and the fraction of "new" immigrants*. Health status is a dummy variable indicating that a veteran's Body Mass Index (BMI, weight in kilograms divided by height in meters squared) was either too low (below 18.5) or too high (greater than or equal to 25).
9. *The interaction between a dummy equal to one if health status was unknown and the fraction of "new" immigrants*.

A.2. Ward choice regressions

The data set consists of every veteran in either the original Union Army data or the urban sample over-sample whom we could place in a ward in 1900 in Baltimore, Boston, Chicago, Cincinnati, New York City (all boroughs), and Philadelphia and for whom we knew city of enlistment. To this sample we merged ward-level characteristics obtained from uadata.org and known as Historical Urban Ecological (HUE) data.

Our explanatory variables are

1. *Number of veterans in the company, number of veterans, number of veterans from the same birth city*. These variables are specific to the veteran and exclude him. The variables were generated from our primary sample. The number of veterans and the number of veterans from the same birth city in the ward are thus underestimated.
2. *Adjusted death rate*. The ward death rate divided by the mean city death rate, all multiplied by 100.
3. *Logarithm of distance to the city center*. Calculated from the ward centroid to City Hall (in meters).
4. *Population density*. Ward square footage divided by ward population where ward population was obtained from the published 1900 census.
5. *Fraction of "new" immigrants, blacks, Irish, and Germans*. Calculated from the complete count census indices available from the Minnesota Population Center and Ancestry (2013). "New" immigrants are immigrants born in Eastern or Southern Europe. We thank Carlos Villarreal for providing us with a mapping of ward numbers to enumeration districts for New York City where the census manuscript schedules do not provide ward numbers.
6. *Dummy equal to one if current city of residence was the same as the city of enlistment*.
7. *The distance from the city of enlistment to the current city*. In kilometers and estimated from the city center.
8. *City population in 1900* from the 1900 published census. See http://www.census.gov/prod/www/decennial.html
9. *The interaction term between a dummy equal to one if the veteran was a professional, proprietor, or artisan at enlistment and city size*.
10. *Interaction terms between dummies equal to one if the veteran was born in Ireland or Germany and the fraction of Irish or Germans in the ward*.
11. *The interactions between pension amount in 1900 and the fraction of "new" immigrants, the fraction Irish, and the fraction German*. Pension amount (in dollars per month) was arguably an exogenous income transfer which depended upon health status and whether the veteran could claim his disability was related to the war. Veterans who could argue that their rheumatism was caused by being out in the damp during a march received more money than veterans whose rheumatism could not be related to the war according to the medical theories of the time.
12. *The interaction between poor health status and the fraction of "new" immigrants, the fraction Irish, and the fraction German*. Health status is a dummy variable indicating that a veteran's Body Mass Index (BMI, weight in kilograms divided by height in meters squared) was either too low (below 18.5) or too high (greater than or equal to 25).
13. *The interaction between a dummy equal to one if health status was unknown and the fraction of "new" immigrants, the fraction Irish, and the fraction German*.

A.3. Mortality Regressions

The data set consists of every veteran in either the original Union Army data or the urban sample over-sample whom we could place in a ward in 1900 in Baltimore, Boston, Chicago, Cincinnati, New York City (all boroughs), and Philadelphia and for whom we knew city of enlistment. To this sample we merged ward-level characteristics obtained from uadata.org and known as Historical Urban Ecological (HUE) data. We restricted the sample to men with a known date of death.

The main variables are

1. *The number of veterans from the same wartime company in the ward*. A time-varying covariate affected by deaths. Unfortunately, we cannot observe yearly moves so the variable may either over- or underestiamte the number of men in the same ward.
2. *The number of other veterans in the ward*. A time-varying covariate affected by deaths.
3. *A dummy variable indicating if the wife is alive*. A time-varying covariate, affected by deaths and remarriage.

References

Bailey, Martha, April 15-16, 2016. LIFE-M Longitudinal, intergenerational family electronic microdata. Presentation at the NBER Cohort Studies Working Group Meeting, Los Angeles, CA.

Burnham, W.Dean, 1955. Presidential Ballots: 1836–1892. Johns Hopkins University Press, Baltimore.

Clubb, Jerome M., Flanigan, William H., Zingale, Nancy H., 2006. Electoral data for counties in the United States: presidential and congressional races, 1840–1972. ICPSR08611-v1.Inter-university Consortium for Political and Social Research (distributor), Ann Arbor, MI. 2006-11-13. https://doi.org/10.3886/ICPSR08611.v1.

Cole, Steve W., Hawkley, Louise C., Arevalo, Jesus M., Sung, Caroline Y., Rose, Robert M., Cacioppo, John T., 2007. Social regulation of gene expression in human leukocytes. Genome Biol. 8 (9) (R189.1-R189) ⟨http://genomebiology.com/2007/8/9/R189⟩.).

Costa, Dora L, Kahn, Matthew E., 2007. Surviving andersonville: the benefits of social networks in POW Camps. Am. Econ. Rev. 97 (4), 1467–1487.

Costa, Dora L., 1995. Pensions and retirement: evidence from union army veterans. Q. J. Econ. 110 (2), 297–320.

Costa, Dora L., Kahn, E.Matthew, 2007. Deserters, social norms, and migration. J. Law Econ. 50 (2), 323–353.

Costa, Dora L., Kahn, Matthew E., 2003. Cowards and heroes: group loyalty in the American civil war. Q. J. Econ. 118 (2), 519–548.

Costa, Dora L., Kahn, Matthew E., 2008. Heroes and Cowards: the Social Face of War. NJ and Oxford: Princeton University Press, Princeton.

Costa, Dora L., Kahn, Matthew E., 2010. Health, wartime stress, and unit cohesion: evidence from union army veterans. Demography 47 (1), 45–66.

Costa, Dora L., DeSomer, Heather, Hanss, Eric, Roudiez, Christopher, Wilson, Sven E., Yetter, Noelle, 2017. Union army veterans, all grown up. Hist. Methods 50 (2), 79–95.

Cutler, David M, Glaeser, Edward L., Vigdor, Jacob L., 2008. When are ghettos bad? Lessons from immigrant segregation in the United States. J. Urban Econ. 63 (3), 759–774.

Durlauf, Steven N., Ioannides, Yannis M., 2010. Social interactions. Ann. Rev. Econ. 2 (1), 451–478.

Eli, Shari, Salisbury, Laura, Shertzer. Allison, 2016. Migration Responses to Conflict: Evidence from the Border of the American Civil War. NBER Working Paper No. 22591.

Ellwardt, Lea, Tilburg, Theo van, Aarlsen, Marja, Wittek, Rafael, Steverink, Nardi, 2015. Personal networks and mortality risk in older adults: a twenty-year longitudinal study. PLos One 10 (3), e0116731.

Fales, Raymond L., Moses, Leon Nathan, 1972. Land-use theory and the spatial structure of the nineteenth-century city. Pap. Reg. Sci. Assoc. 28 (1), 49–80.

Fogel, Robert W., 1993. New sources and new techniques for the study of secular trends in nutritional status, health, mortality, and the process of aging. Hist. Methods: J. Quan. Interdiscip. Hist. 26 (1), 5–43.

Gin, Alan, Sonstelie, John, 1992. The streetcar and residential location in nineteenth century Philadelphia. J. Urban Econ. 32 (1), 92–107.

Haines, Michael R., 2010. Inter-university Consortium for Political and Social Research. Historical, Demographic, Economic, and Social Data: The United States, 1790–2002. ICPSR02896-v3. Ann Arbor, MI: Inter-university Consortium for Political and Social Research (distributor), 2010-05-21. ⟨http://doi.org/10.3886/ICPSR02896.v3⟩.

Helsley, Robert W., Zenou, Yves, 2014. Social networks and interactions in cities. J. Econ. Theory 150, 426–466.

Ioannides, Yannis M., 2013. From Neighborhoods to Nations: the Economics of Social Interactions. Princeton University Press, Princeton, NJ and Oxford, UK.

LeRoy, S.F., Sonstelie, J., 1983. Paradise lost and regained: transportation innovation, income, and residential location. J. Urban Econ. 13 (1), 67–89.

Logue, Larry M., 2007. The reality of veterans' voting. in the civil war veteran: a historical reader, Larry M. Logue, Michael Barton (Eds.). New York University Press, NY and London. pp 310–16.

Manski, Charles F., 1993. Identification of endogenous social effects: the reflection problem. Rev. Econ. Stud. 60 (3), 531–542.

Margo, Robert A., 1992. Explaining the postwar suburbanization of population in the United States: the role of income. J. Urban Econ. 31 (3), 301–310.

McConnell, Stuart, C., 2007. The patriotic boom. in the civil war veteran: A historical reader. In: Larry M. Logue, Michael Barton (Eds.). New York University Press, NY and London. pp. 365–375.

Minnesota Population Center and Ancestry.com., 2013. IPUMS Restricted Complete Count Data: Version 1.0 [Machine-readable database]. University of Minnesota, Minneapolis.

Ruggles, Steven, Genadek, Katie, Goeken, Ronald, Grover, Josiah, Sobek, Matthew, 2015. Integrated Public Use Microdata Series: Version 6.0 [dataset]. University of Minnesota, Minneapolis. https://doi.org/10.18128/D010.V6.0.

Seeman, Teresa E., 1996. Social ties and health: the benefits of social integration. Ann. Epidemiol. 6, 442–451.

Selod, Harris, Zenou, Yves, 2006. City structure, job search and labour discrimination: theory and policy implications. Econ. J. 116 (514), 1057–1087.

Shertzer, Allison, Walsh, Randall, P.,2016. Racial sorting and the emergence of segregation in American Cities. NBER Working Paper No. 22077.