Online Appendix

Recessions, Mortality, and Migration Bias: Evidence from the Lancashire Cotton Famine

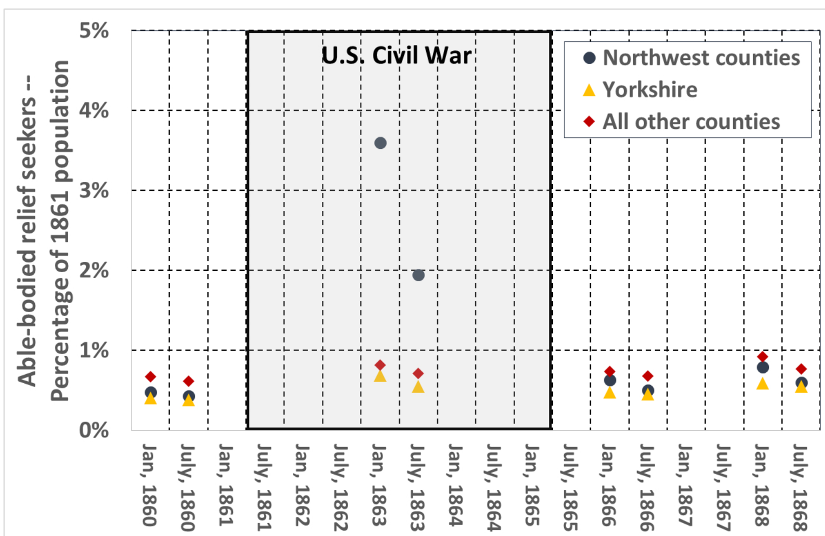
Vellore Arthi, Brian Beach, and W. Walker Hanlon

A Online Appendix: Empirical setting

A.1 Additional evidence of distress in Lancashire

Figure 5 describes the number of able-bodied relief-seekers who obtained aid from local Poor Law Boards, the main source of government support for the destitute in our setting. Consistent with the graph of Poor Law expenditures shown in the main text, during the downturn we see an increase in relief-seekers in the Northwest counties (Lancashire and Cheshire), where cotton textile production was concentrated. Non-cotton counties, however, were largely unaffected.

Figure 5: Evidence from able-bodied relief seekers



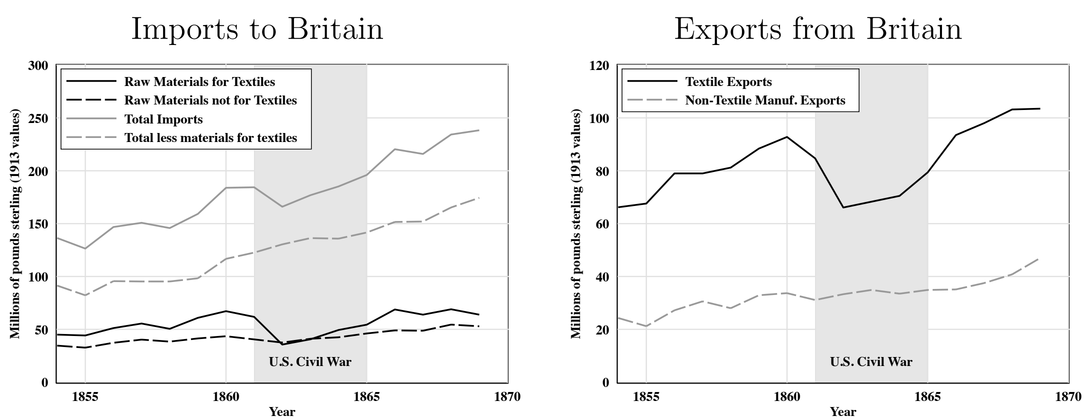
Data from Southall *et al.* (1998), Graph reproduced from Hanlon (2017).

A.2 Was the broader British economy affected?

To look for other effects of the U.S. Civil War on the British economy, a natural starting point is to examine imports and exports. The left-hand panel of Figure 6 shows that, aside from raw cotton, there does not appear to be a substantial change in total imports or raw material imports. This makes sense given that raw cotton made up 67% of total British imports from the U.S. in 1860. The right-hand panel examines exports. Aside from textiles, there is no evidence of a substantial change in British exports during the Civil War period.

One may expect that the U.S. Civil War would have had an impact on particular sectors of the British economy, such as arms or warship production. However, British producers were prohibited from selling arms to either side during the U.S. Civil War. While some producers were able to circumvent these restrictions, in general, these restrictions limited the impact that the conflict had on these industries.

Figure 6: British imports and exports, 1854-1869



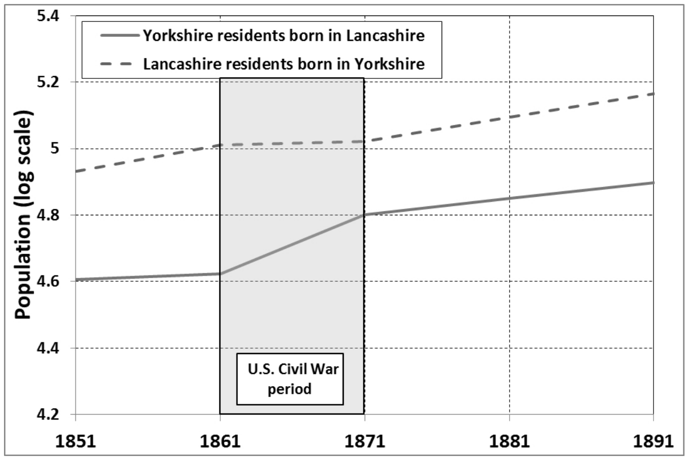
Data from Mitchell (1988).

A.3 Additional results on migration during the cotton shock

Evidence from birthplaces

Additional evidence on migration during the cotton shock can be gleaned from the location-of-birth information provided in the census. Specifically, changes in the share of the population born in one location who are resident in another can be used to provide evidence on net migration between locations. The location-of-birth data are

Figure 7: Evidence of migration for Yorkshire and Lancashire from birthplace data



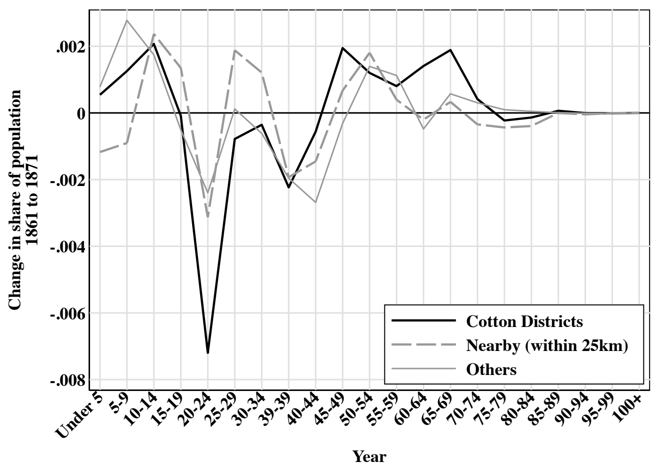
This graph, which is reproduced from Hanlon (2017), presents data on the birthplace of county residents from the Census of Population.

only available at the county level, so in Figure 7, which we reproduce from Hanlon (2017), we compare the largest cotton textile county, Lancashire, with the neighboring wool textile county of Yorkshire. The figure indicates that the number of Yorkshire residents who were born in Lancashire increased substantially from 1861-1871, while the number of Lancashire residents born in Yorkshire stagnated. This suggests an out-migration of Lancashire residents during the U.S. Civil War, as well as reduced in-migration to Lancashire.

Evidence on the age distribution of migrants

Next, we consider some results that help us think about how migration patterns varied across age groups. Figure 8 plots the change in population shares for several age categories between 1861 and 1871. We plot these changes separately for cotton districts, districts that were proximate to cotton districts, and all other districts. The most prominent feature in this graph is that there was a substantial decline in the share of 20-24 year-olds between 1861 and 1871. This suggests that the migration response to the shock was strongest among young adults.

Figure 8: Share of population in each age group in cotton districts



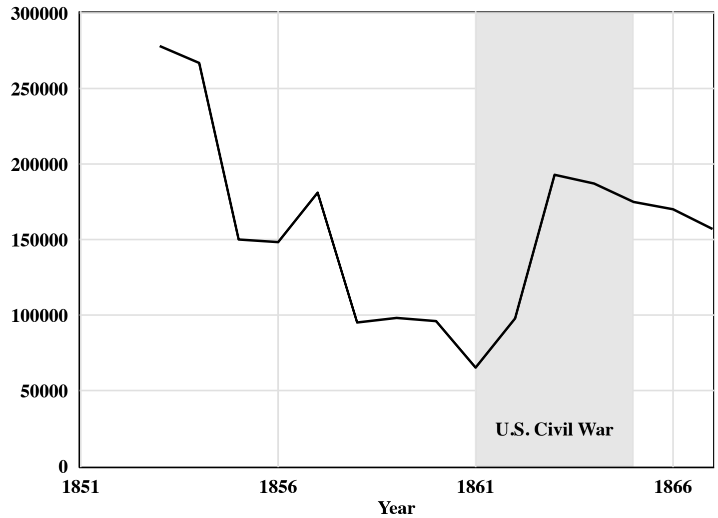
Population data are from the Census of Population for 1861 and 1871. Cotton districts are identified as those with over 10% of workers employed in cotton textile production in the 1851 Census. Nearby districts are those within 25 km of cotton districts.

Evidence on emigration

Tracking emigration from Britain in response to the cotton shock is more difficult than tracking internal migration. What information is available was collected at the ports of embarkation and reported in the British Parliamentary Papers.[44]. Figure 9 uses data from the 1868 report to the House of Commons, which provides total emigration numbers for 1853-1867. This graph shows that the total number of emigrants leaving Great Britain fell almost continuously from 1851-1861 and then increased substantially from 1861-1863. Unfortunately, we do not know what areas these emigrants were coming from, though we do know that most emigrants were Irish by birth. The English made up roughly one-third of emigrants across this period. However, by 1860 there were many Irish and Scottish living in cotton districts, so international emigrants from cotton districts need not be English.

[44]UK Parliamentary Archives (1868)

Figure 9: Emigration from Britain, 1852-1867



Data from the British Parliamentary Papers (1868, no. 045515).

A.4 Contemporary reports on health effects

Contemporary reports offer a mixed view of the impact that the cotton shortage had on health. Some 19th century observers, such as Arnold (1864), report that there was a "lessened death-rate throughout nearly the whole of the [cotton] district, and, generally speaking, the improved health of the people." In the words of the Registrar of Wigan, these gains were attributed primarily to "more freedom to breathe the fresh air, inability to indulge in spirituous liquors, and better nursing of children."[45]

The importance of childcare is highlighted in a number of reports, such as Dr. Buchanan's 1862 *Report on the Sanitary Conditions of the Cotton Towns* (Reports from Commissioners, British Parliamentary Papers, Feb-July 1863, p. 304), which discusses the importance of the "greater care bestowed on infants by their unemployed mothers than by the hired nursery keepers." This channel was likely to be particularly

[45]Quoted from the *Report of the Registrar General*, 1862.

important in the setting we study because female labor force participation rates were high, even among mothers. Using 1861 Census occupation data, we calculate that nationally, 41% of women over 20 were working, and they made up 31% of the labor force. This rate was much higher in major cotton textile areas. In districts with over 10% of employment in cotton textiles in 1861, the average female labor force participation rate for women over 20 was 55%, and women made up 38% of the labor force. For comparison, these are similar to the levels achieved in the U.S. in the 1970s and 1980s (Olivetti, 2014), though of course the nature of the work done by women was quite different.

On the other hand, there were also reports of negative health effects due to poor nutrition and crowded living conditions. Dr Buchanan, in his *Report on the Sanitary Conditions of the Cotton Towns*, states that "There is a wan and haggard look about the people..." (Reports from Commissioners, British Parliamentary Papers, Feb-July 1863, p. 301). Typhus and scurvy, diseases strongly associated with deprivation, made an appearance in Manchester and Preston in 1862 after being absent for many years, while the prevalence of measles, whooping cough, and scarlet fever may have also increased (*Report on the Sanitary Conditions of the Cotton Towns*, Reports from Commissioners, British Parliamentary Papers, Feb-July 1863). Seasonality features prominently in these reports, with conditions worsening during the winters, when the shortage of clothing, bedding, and coal for heating increased individuals' vulnerability to winter diseases such as influenza.

B Online Appendix: Data

B.1 Obtaining linkable death records

The death records used in our main analysis were obtained from the General Registrar's Office website for the years 1851-1855 and 1861-1865. This process involved several steps. First, we identify "linkable" records in the 1851 and 1861 censuses, where a linkable record is defined as one where the name is unique within a 5 year age band. This means that no two records in the same census can have the same first name, same last name, and a birth year within 5 years of each other. This yields 4,291,185 "linkable" records in 1851, and 5,228,528 "linkable" records in 1861. Of course, not all of the people underlying these linkable records will die in the subsequent five years. Thus, to help guide our search of the GRO index, we take these "linkable" records and identify whether there is at least one person in the Registrar's master death index with the same first name and same last name. Again, the master death index does not include age, so this process simply tells us who among the "linkable" names might have died in the next five years. This allows us to generate a list of surnames for each five year period (1851-1855 or 1861-1855), which we then fed into the GRO search query to obtain a new death index with age at time of death.

Note that we could have obtained all names, rather than only those likely to be linkable. However, the nature of the website meant that there are difficulties involved when searching for very common names, and moreover, this commonality means that such names are very unlikely to ultimately end up in any linked dataset. Thus, we decided to restrict our attention to only the set of death records associated with names that were likely to ultimately be linkable.

It is worth pointing out that in addition to having to provide a surname, sex, and year of death, the GRO website only displays up to 50 records per page, with a maximum of display of 250 records. We built an automated algorithm to feed in queries of surnames and death years, prioritizing surnames with fewer than 50 unique deaths by year, as identified by the master death index. Unfortunately, shortly after extracting the surnames with less than 40 unique queries, the GRO changed their terms of conditions, banning automated search algorithms. Nevertheless, we were able to extract 964,567 complete records for the 1851-1855 period and 790,623 complete records for the 1861-1865 period. To put this in perspective, the death master index lists 2,122,875 deaths over the 1851-1855 period and 2,196,602 deaths over the 1861-1865 period. Thus, we were able to extract complete information for roughly 45% of all deaths occurring between 1851-1855 and 36% of all deaths occurring between 1861-1865. Moreover, because our surnames were drawn from "linkable" census records, these percentages correspond to the records where we have the greatest change of establishing a valid link.

B.2 Linked data

B.2.1 Assessing the quality of our linked dataset

A natural check on the accuracy of our linking procedure is to compare the distance between census district and death district in the linked sample. We would expect the share of matches to diminish rapidly with the distance between the census and death district, since migration should be less common between more distant locations. This provides an opportunity to test the reasonableness of our results.

Panel A of Figure 10 presents histograms showing the share of linked deaths by distance bins using data from both the 1851-1855 and 1861-1865 periods. Distance is calculated using latitude and longitude coordinates for the main town or administrative center for each district (or the geographic center for a few very rural districts). The left panel includes links within the same district, while these are dropped in the right panel in order to make it easier to view the pattern for links across districts. In the left graph, we can see that just under half of all links occurred within a district. In the right graph, we see that the share of links across districts declines rapidly with distance.[46]

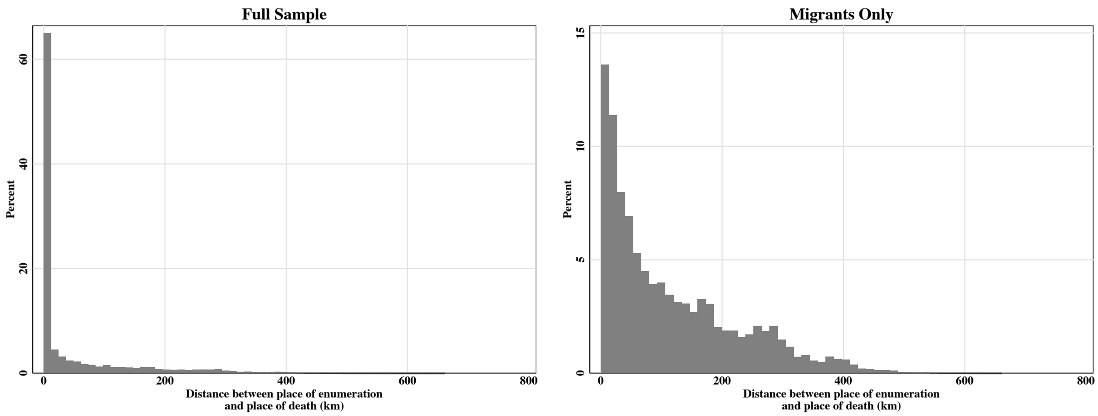
As a point of comparison, we can apply our linking algorithm to link the universe of living individuals between the 1851 census and the 1861 census. Note here that we

[46] The bump at about 250 km corresponds to the distance between the two major population centers in the country, London in the Southeast, and Manchester and Liverpool in the Northwest.

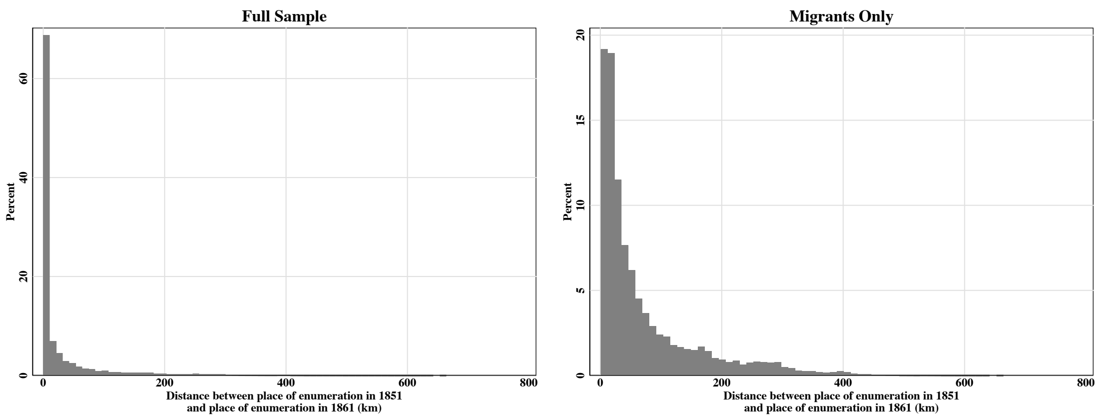
are linking individuals that we think survived to be enumerated in the 1861 census. After generating those links, we can then compute the distance between the district of enumeration in 1851 and that in 1861. The histogram of these distances is provided in Panel B. Here we see a pattern that is nearly identical to Panel A. One may worry that this is a statistical artifact. To address this, in Panel C we randomly link individuals from the 1851 census to other records in the 1851 census. In this sample, essentially all of the records are false positives. We then calculate the distance between the two districts of enumeration. Here, we see a distribution that has more of a hump shape, and is clearly different from both Panel A and Panel B. Thus, to summarize, the results of this exercise suggest that the patterns of migration in our linked dataset mimic the patterns of migration that we should expect to see during this time period, and do not appear to be an artifact of the linking procedure.

Figure 10: Share of links by distance between census and death districts

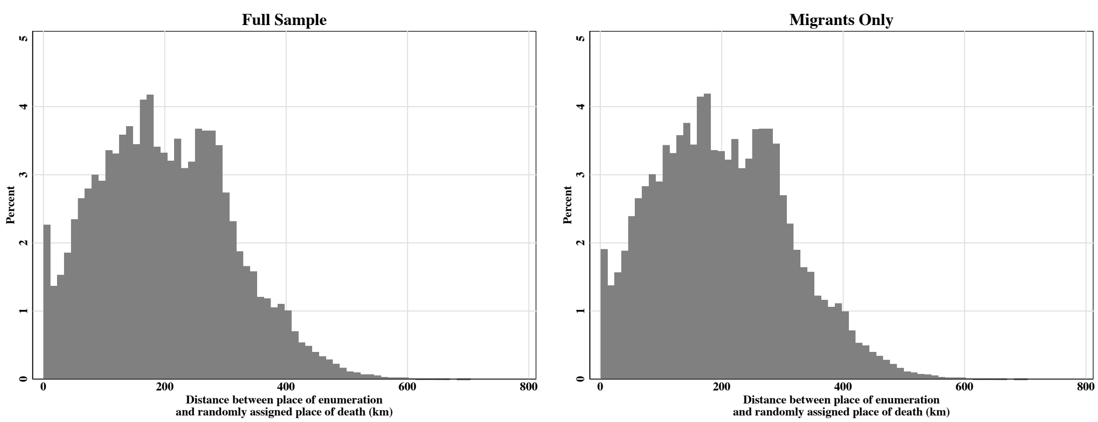
Panel A: Distance between district of death and district of enumeration



Panel B: Distribution of distances after applying our linking algorithm to link individuals between the 1851 and 1861 censuses



Panel C: Distribution of distances when randomly linking individuals from the 1851 census to other records in 1851



B.3 Representativeness: Comparing the linked sample to aggregate data

Representativeness by age

This subsection analyzes how representative our linked sample of deaths is of aggregate deaths, i.e., the comprehensive mortality aggregates published by the Registrar General. One dimension that we can compare is the age distribution in the two datasets. This is perhaps the most important dimension to consider, given the strong relationship between age and mortality risk. The top panels of Figure 11 compare the share of linked deaths in each age group and the share of aggregate deaths in each age group for the 1851 and 1861 periods. These graphs show that infant and young child deaths are substantially under-represented in the linked sample. The fact that our linked sample struggles to reflect deaths among young children is a mechanical result of our approach, since deaths among infants born after enumeration cannot be linked back to the corresponding census. It is worth noting that the linked distributions are nearly identical to each other (see the bottom panel of Figure 11), and so this potential bias is unlikely to be related to treatment. As discussed in the main text, we consider two approaches for dealing with differences in the age distribution of deaths in the linked sample relative to true distribution. In one approach we re-weight each linked death such that our linked sample mirrors the aggregate age distribution before collapsing to the district level. Alternatively, we analyze different age groups separately.

Representativeness by gender

Table 6 breaks down gender shares of deaths in the linked and aggregate data by time period. We can see that women are over-represented in the sample of linked deaths relative to their share of aggregate deaths. It is worth noting that this feature appears in both the 1851 and 1861 data. Also, both the linked and aggregate data

Figure 11: Histogram of deaths by age from linked and aggregate data

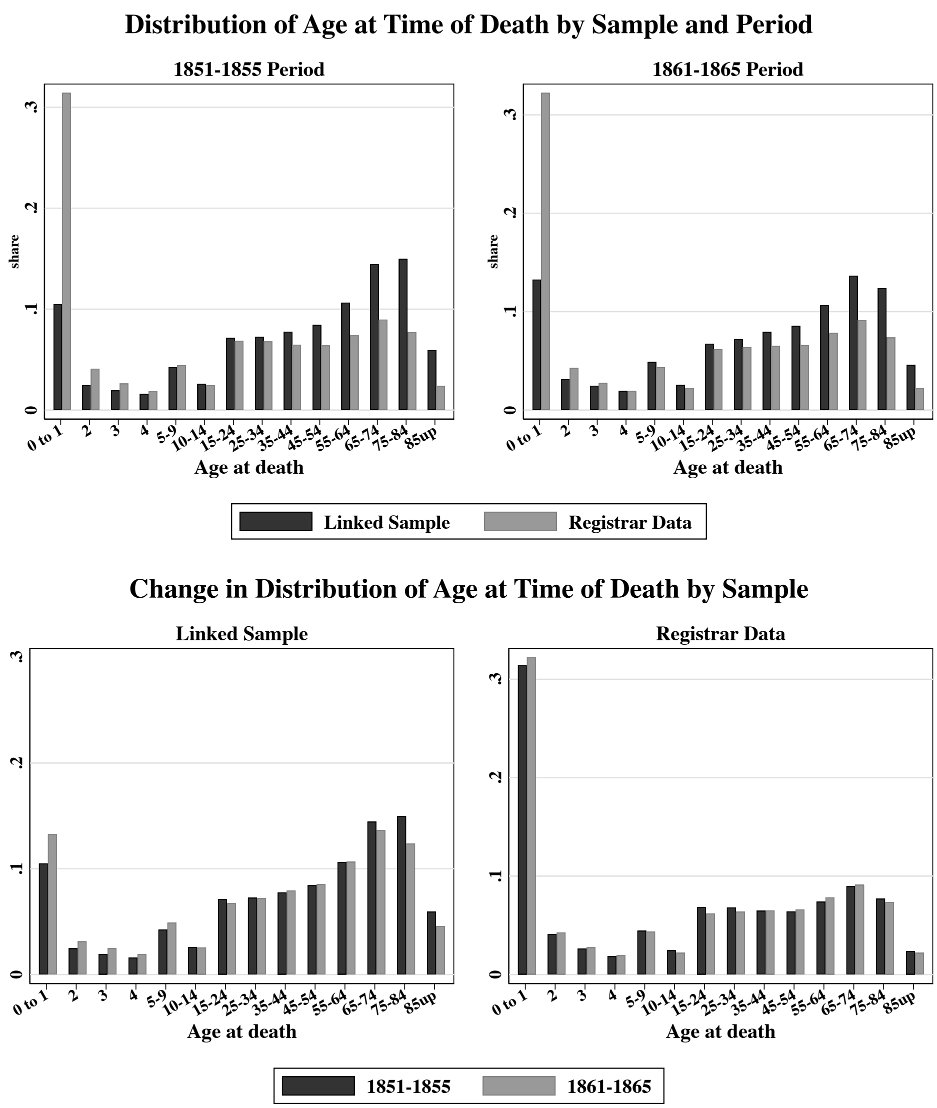


Table 6: Female shares of linked and aggregate deaths by time period

	Linked deaths		Aggregate deaths	
	1851	1861	1851	1861
Women	51.4%	50.5%	49.2%	48.8%

show very similar declines in the female share of deaths between the two periods. The most likely explanation for the higher share of female deaths in the linked sample is that women had more unique names, allowing us to generate more unique matches. This is consistent with evidence from the sociology literature suggesting that parents are more likely to give male children traditional names (Rossi, 1965).

Representativeness by SES

Next we examine the representativeness of our sample by socioeconomic status (SES). To do so, we take advantage of the occupation data available in the linked census data, which has been classified by HISCO score. We can compare this to aggregate data that we have gathered from the Registrar General's report for 1851, which lists the number of deaths among people in each occupation in that year. In the analysis below, we focus on comparing shares of white-collar and blue-collar workers among the linked sample for which occupations are given in the census. We define white-collar workers as HISCO groups 1-3, which includes professional and technical workers, administrative and managerial workers, and clerical workers. We focus on the white-collar vs. blue-collar comparisons because an analysis at more detailed occupation levels is made difficult by the fact that the HISCO classes used in the individual-level census data are not a great fit for the groupings used in the aggregate data, so it is difficult to do this comparison at a detailed occupation level. For example, those working in sales are classed with their industry in the aggregate data (e.g., "Other workers, dealers in flax and cotton") while in the HISCO classifications, a cotton dealer would be categorized differently from a cotton worker. As similar issue exists

Table 7: White-collar share of deaths in linked and aggregate data

	Linked deaths in 1851	Linked deaths in 1861	Aggregate deaths in 1851
White-collar share	6.0%	6.4%	7.6%

for foremen and managers, who are classed with their industry in the aggregate data but not in the HISCO classifications.

The results, in Table 7, show that the shares of deaths among white-collar vs. blue-collar workers in the linked sample are very similar to the shares observed in the aggregate data for 1851. Thus, our linked sample appears to be fairly representative of the aggregate data in terms of SES.

C Online Appendix: Methods

C.1 Description of the permutation exercise

One potential worry in our analysis is that the spatially concentrated nature of our set of treated districts may be influencing the statistical significance of our results. This section describes a permutation test designed to provide an alternative approach to evaluating the statistical significance of our results, while still respecting the spatial structure of our data. The basic approach here is to construct alternative sets of spatially concentrated placebo treatment districts, surrounded by rings of placebo "nearby" districts, apply our standard estimating procedure to each of these sets of placebo treated and nearby districts, and then compare the distribution of estimated results to the coefficients obtained from our true treated and nearby districts.

To implement this approach, we start each permutation with a different "anchor" district. Since there are 538 districts in our main linked data analysis, we run 538 different permutations using each district as an anchor district. For each anchor district, we identify the 23 nearest districts and call them our placebo treated districts. This gives us 24 treated districts (23 plus the anchor district), matching the number of cotton districts that were actually treated. We then identify the next 26 nearest districts and call them the first set of nearby districts, matching the 26 districts within 25 km of the cotton districts in our main analysis. The next 32 districts are called the second set of nearby districts (matching the 32 districts 25-50 km from the cotton districts), while the next 36 nearest districts are the third set of nearby districts (matching the 36 districts 50-75 km from the cotton districts). Thus, we end up with a set of placebo treated and placebo nearby districts which are both spatially concentrated and the same, in terms of number, as the true treated and nearby districts used in the main analysis. Note that, as is standard in permutation exercises, we apply this approach to all districts in the data, including the districts that were actually treated (the cotton districts).

Given each set of placebo treated and nearby districts, we then apply our standard estimation procedure and recover estimated coefficients for the change in mortality in each group of placebo treated districts comparing the 1861-1865 period to the 1851-1855 period. This provides a distribution of coefficients which can then be compared to the coefficients obtained when running the analysis on the actual cotton districts. The permutation test p-values reported in the main text reflect the share of coefficients for the placebo treated districts which exceed the coefficient obtained from the actual cotton districts.

Intuitively, the idea behind this exercise is that if having spatially clustered treated districts leads to understated standard errors in our main analysis, then applying our analysis to spatially clustered sets of placebo districts should generate a more spread-out distribution of placebo coefficients than we would expect given our standard errors, and as a result, the p-values from the permutation exercise should be larger than the standard p-values obtained from our main regressions. Of course, there are a number of reasonable alternative ways to implement a permutation exercise. For example, we could have used districts within 25 km of the placebo treated districts as our first set of nearby districts, rather than the 26 nearest districts. Ultimately, we think it is unlikely that variations like this will make any substantial difference in the results.

There are two natural alternative approaches to dealing with spatial correlation in our data. One approach is to cluster standard errors at some higher geographic level, such as the county. However, we think this approach is undesirable for two reasons. First, many of the cotton districts are in Lancashire, which is a large and diverse county with different areas that have starkly different economic structures. For example, Barrow-in-Furness was a major steel and shipbuilding center with an economy that was starkly different than that of the cotton textile districts, while Liverpool was a major trading center. As such, we do not think it is reasonable to include these together. Second, if we cluster by county, the cotton districts fall into only two counties. Clustering data in this way is likely to cause statistical issues (Donald & Lang, 2007). Despite these concerns, we have generated results clustering standard errors by county. We find that these tend to deliver smaller standard errors (more statistically significant results) than those reported in the main text, suggesting that there may be negative spatial correlation across districts within the same county. This is not surprising given that other studies looking across British districts during this time period have found evidence of negative spatial correlation (Hanlon, 2017). See that paper for a discussion of why negative spatial correlation is not surprising in this context.

A more promising alternative to clustering is to implement spatial standard errors following Conley (1999). However, we are also hesitant to take this approach because the statistical properties when the treated districts are spatially concentrated are not well-studied. Despite these concerns, we have also implemented this approach on our main linked data sample, and we find that it delivers smaller standard errors (more statistically significant results) than those reported in the main text. Again, this finding is consistent with negative spatial correlation across districts.

D Online Appendix: Analysis

D.1 Additional analysis of the linked data

Various robustness checks

Here, we investigate the robustness of the all-age result in Table 8. For comparison purposes, in Column 1 we reproduce our preferred specification. In Column 2, we present results where we do not weight the regressions by initial district population. In Column 3 we drop Manchester, which is an outlier among the cotton districts in terms of city size and because it was the commercial center of the industry. In Column 4 we also drop both Liverpool, which was not a cotton district, but which was the main port serving the industry; and Leeds, which was an important nearby wool-producing center. In Column 5, we use the birthplace information in our linked census data to confine our linked sample to only those workers born in England or Wales ("native-born") in order to assess whether deaths among immigrant workers are driving our results.

Additional linking criteria/handling of false positives

Next, we consider how our main results are affected when we modify our underlying linking procedure. In particular, we consider two modifications that eliminate from our linked data any deaths where there are other potential links that sound similar to the name in question when spoken. In the first modification, we eliminate a linked individual if, for that individual, the first and last names are unique in the death records, but in the census record, there is another record where the first name matches exactly and there is a similar-sounding last name, i.e., the last name is the

Table 8: Robustness of linked-data results on all-age mortality rates

DV: Death per 1,000 Individuals (per year)	Baseline	No weights	Without Manchester	Drop Manchester Liverpool & Leeds	Only native-born
	(1)	(2)	(3)	(4)	(5)
Cotton District × Shortage	2.534***	2.715***	2.609***	2.424***	2.291***
	(0.605)	(0.769)	(0.657)	(0.661)	(0.564)
Nearby (0-25 km) × Shortage	1.054*	1.447	1.058*	0.853	0.609
	(0.597)	(0.925)	(0.596)	(0.602)	(0.591)
Nearby (25-50 km) × Shortage	0.191	0.735	0.194	0.729	0.233
	(0.623)	(0.686)	(0.624)	(0.559)	(0.507)
Nearby (50-75 km) × Shortage	0.586	0.143	0.583	0.482	0.245
	(0.656)	(0.646)	(0.656)	(0.645)	(0.596)
Observations	1,076	1,076	1,074	1,070	1,076
R-squared	0.398	0.333	0.396	0.401	0.404

same after Soundex cleaning. Results obtained using these two alternative linked samples, and assigning deaths to individuals' initial district of residence (as in Table 3), are presented in Columns 1-2 of Table 9. This table shows that our main results continue to hold, even when using these severely limited samples.

As a second way to assess the impact of false positives on our results, we limit our analysis to samples where false positives are less common. In particular, we take advantage of the fact that, mechanically, false positives are more likely for links where the death district is further from the census district. This is because people are less likely to migrate over longer distances, but false matches are just as likely to occur between distant districts as they are between nearby districts. As a result, the share of false matches to true matches will rise as the distance between the census district and the death district increases.

Columns 3-6 of Table 9 present results obtained when limiting our linked sample to those where the census district and death district are proximate to one another.

In Column 3 we consider only links under 200km and we further reduce this distance in Columns 4 and 5. In Column 6 we consider only results where the census district and the death district match. At the bottom of the table we can see that applying these restrictions progressively reduces the rate of false positives in our sample (with these rates coming from the test described in in section 3.3). If bias generated by false positives was affecting our results, then we should expect our estimates to fall as we limit the sample to observations with fewer false positive links. Instead, the results remain very stable across Columns 3-6.

Table 9: Linked-data analysis using alternative linking restrictions

	DV: Deaths per 1,000 Individuals (per year)					
	Requiring links to have a unique-sounding name		Omitting links where death and enumeration district are far apart			
	Last Name (1)	First and Last (2)	0-200 km only (3)	0-100 km only (4)	0-50 km only (5)	No Migration (6)
Cotton District × Shortage	2.592***	2.375**	2.826***	2.533***	2.352***	5.545***
	(0.900)	(1.027)	(0.607)	(0.685)	(0.735)	(1.316)
Nearby (0-25 km) × Shortage	0.342	0.605	1.064*	0.307	0.077	1.039
	(0.835)	(0.926)	(0.612)	(0.728)	(0.798)	(1.013)
Nearby (25-50 km) × Shortage	0.193	0.731	0.137	-0.741	-0.943	-1.741**
	(0.887)	(1.002)	(0.605)	(0.641)	(0.676)	(0.853)
Nearby (50-75 km) × Shortage	0.435	0.642	0.484	-0.056	0.101	0.572
	(0.950)	(0.993)	(0.644)	(0.709)	(0.819)	(0.937)
Observations	1,074	1,074	1,076	1,076	1,074	1,066
R-squared	0.342	0.315	0.414	0.434	0.435	0.460
Lower Bound False Positive Rate	7.72%	7.50%	8.19%	7.57%	6.88%	5.31%
Upper Bound False Positive Rate	13.16%	12.38%	16.35%	15.23%	14.15%	11.80%
Linked Deaths	127,676	112,297	252,667	227,660	208,109	166,406

D.2 Additional results by employment and location

In the results breaking effects down both by location and sector of employment in main text Table 4, we classify individuals based on the occupation of the household head. That is a useful approach because of the important role of household head earnings, as well as the fact that it allows us to study the mortality effect on nonworkers, such as children. However, an alternative approach that can be used for workers is to classify them by their own occupation, rather than the household head's occupation. Results obtained using this alternative approach, for those over age 15 with an occupation provided in the census, are presented in Table 10. These results show that being located in a cotton textile district had a strong influence on workers, while we see weaker effects if the workers were themselves employed in the cotton textile industry. Note that the magnitude of the results are larger than those shown in Table 4 (except the results for older adults) because these data do not include those under 15, where we saw no increase in mortality associated with the shock.

Table 10: Effects by employment sector and location using own occupation

	DV: Deaths per 1,000 persons (per year)		
	(1)	(2)	(3)
Cotton District × Cotton Shortage	2.306***		2.254***
	(0.699)		(0.643)
Employed in Cotton × Cotton Shortage		1.090	0.204
		(0.886)	(0.813)
Observations	33,445	33,445	33,445
R-squared	0.0203	0.0195	0.0203

*** p<0.01, ** p<0.05, * p<0.1. Standard errors, clustered at the district level, in parentheses. All regressions include period and district fixed effects, a series of indicators for whether the district is 0-25, 25-50, or 50-75 km from a cotton textile district in the shortage period, region-by-period fixed effects, ln(population density), the share of the population that is under 15, and the share of the population that is over 54. Deaths are assigned to the initial district of residence (i.e., census enumeration district). Regressions are weighted by district population.

Next, we exploit the occupation-by-location data to study changes in mortality patterns across all occupation groups. As in the main analysis, we begin by classifying each individual in the data based on the occupation of their household head. We then run regressions using the following specification:

$$\left(\frac{\widetilde{MORT}_{dst}}{POP_{dst}}\right) = \sum_{s \neq COT} b_{1s} OCC_s * POST_t$$

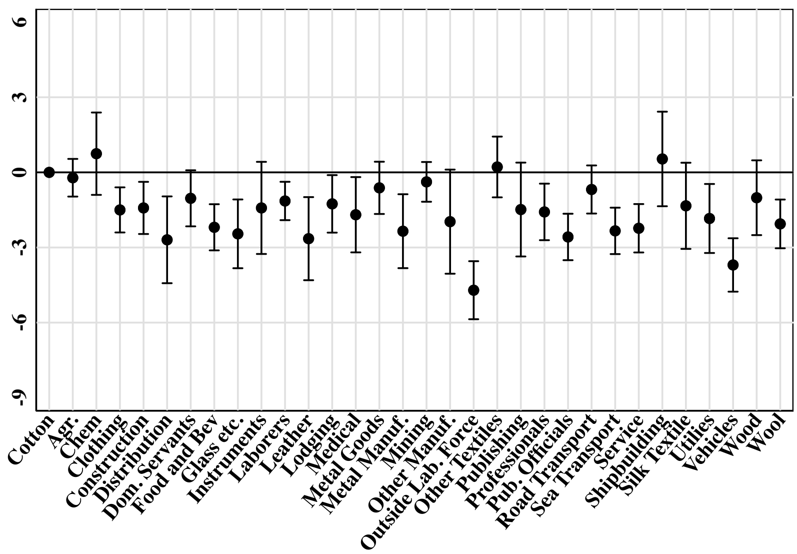
$$+ \sum_{s \neq COT} b_{2s} OCC_s * COTDIST_d * POST_t + X_{dt}\Gamma + \phi_d + \eta_t + \theta_s + \epsilon_{dot} \,,$$

where OCC_s is an indicator for households classified in occupation s. In this specification, the b_{1s} coefficients allow us to study how mortality changed in each occupation in the 1861-1865 period compared to 1851-1855. The omitted industry here is cotton textiles, so that the b_{1s} coefficients reflect how mortality for households associated with each non-cotton industry changed compared to the mortality change observed in the cotton textile industry. The b_{2s} coefficients allow us to study how mortality for households associated with each non-cotton industry changed in the cotton textile districts. These coefficients allow us to learn about how mortality changed for households not directly reliant on cotton textile employment, but living in the areas affected by cotton shock.

Figure 12 presents the estimated b_{1s} coefficients and confidence intervals. The main pattern to note here is that the vast majority of coefficient estimates are negative, and many are also statistically significant. Since these estimates reflect mortality rates relative to cotton textile households, this is telling us that cotton textile households experienced a greater increase in mortality during the shock period than households reliant on almost any other occupation.

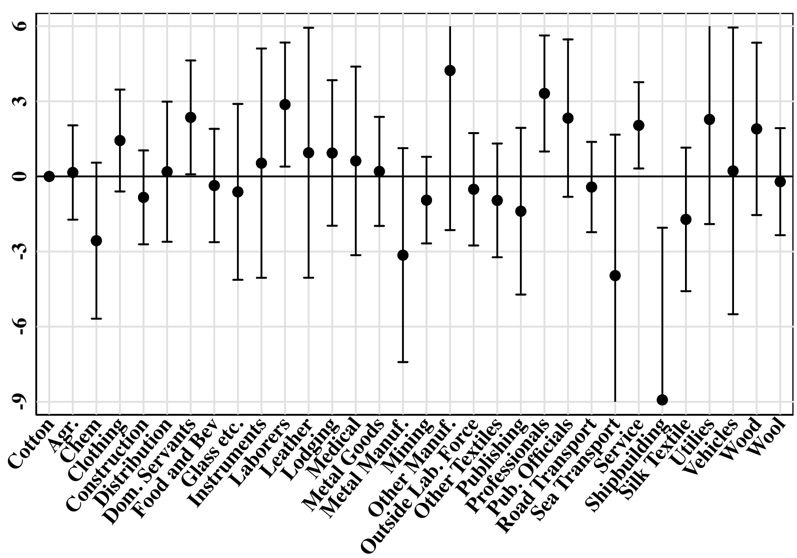
Figure 13 presents the estimated b_2s coefficients and confidence intervals. These reflect how mortality among households reliant on non-cotton industries, but resident in the cotton textile districts, fared during the shock, relative to mortality in those industries in other locations. The households that experienced particularly large increases in mortality in the cotton districts during the shock are those with household heads classified as domestic servants, laborers, service workers, and professional services. There is a common thread here; these are all providers of non-traded local services. In addition, all except professionals are low-wage occupations where households would likely have already been near poverty. It is worth contrasting this

Figure 12: Estimated change in mortality in 1861-1865 compared by 1851-1855 by occupation



list with the set of households that experienced relatively smaller mortality effects (negative coefficients). In order, these are: shipbuilding, sea transport, metal manufacturing, chemicals, silk textiles, and publishing. Other than sea transport, these are all producers of tradable goods. While we should be careful not to draw overly-strong conclusions from these results, the patterns certainly suggest that being in an area negatively impacted by the cotton shock hurt those employed in providing local services more than those reliant on manufacturing traded goods.

Figure 13: Estimated change in mortality in 1861-1865 compared by 1851-1865 by occupation in the cotton textile districts



D.3 Results using an alternative linked dataset

This subsection presents results from a different linked dataset which uses "BMD" deaths data. This alternative set of death records comes from freeBMD.org, which has separately digitized the underlying death records from the General Registrar's Office. These records include the decedent's first and last name, as well as district, year, and quarter of death. These data are available for the full 1851-1870 period.

Since the BMD data come from a different source and are linked using a different procedure, they can provide a useful robustness check on our main results using the GRO deaths data. The BMD data are inferior to the GRO data in that they do not include information on age at death, which is why the GRO data are preferred for our main analysis. However, the BMD data do have one advantage, which is that they are available for 1856-1860 and 1866-70, which allows us to undertake some additional placebo exercises as well as to look at effects post-shock.

The linking procedure applied to this dataset involves matching unique first and last name combinations between the census and death records. We begin by restricting our set of potential links to the set of deaths and census records with unique first and last names. We then link those that match perfectly as written. This procedure yields a matched sample of 71,566 individuals who died between 1851 and 1855, and 81,221 individuals for the 1861-1865 period, representing 3.6% and 3.8% of all deaths over the respective periods. Note that, without the age information used in our main GRO death data, we obtain lower matching rates and a substantially smaller matched dataset. We also link deaths from 1856-1860 back to the 1851 census and deaths from 1866-70 back to the 1861 census.

Below, we begin by examining the representativeness of the BMD data and comparing it to the GRO data. We then replicate our main results using the BMD data. That is followed by a placebo exercise and then an examination of longer-run outcomes after 1865.

Representativeness

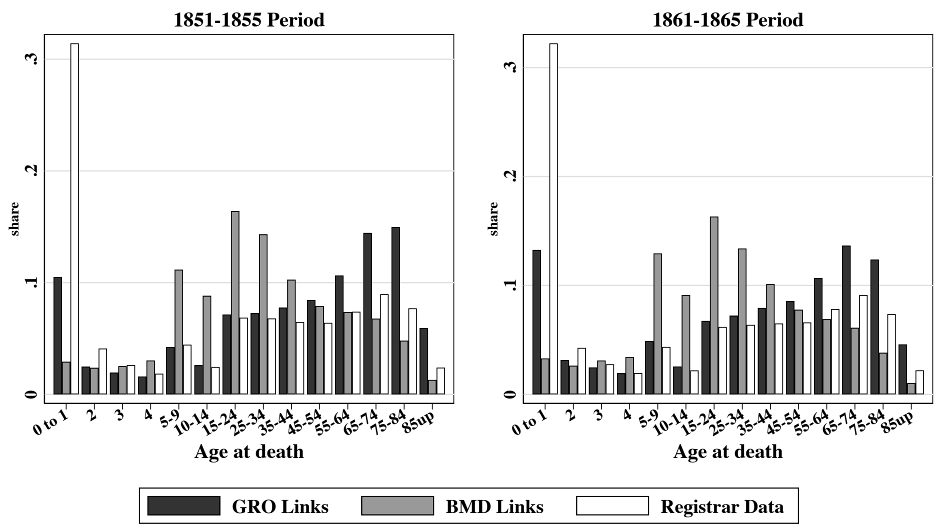
As with the GRO data, the representativeness of the linked BMD data can be assessed along several dimensions. One of these is the breakdown by gender. In the BMD data, the female share of deaths in 1851-1855 is 54.3 percent and in the 1861-1865 period it is 53.4 percent. Thus, females are overrepresented relative to the universe of deaths, reported in the aggregate Registrar General's reports, where the female share is 49.2 in 1851-55 and 48.8 in 1861-65. The BMD data is also less representative than the GRO data in this dimension; the GRO shares, 51.4 in 1851-55 and 50.5 in 1861-65 are substantially closer to the shares reported in the aggregate statistics. As discussed in the main text, females are most likely overrepresented because female names are more likely to be unique than male names. This has a bigger effect in the BMD data, which rely only on unique names to make links, than in the GRO data, where name uniqueness is less important because we are also able to use age information in linking. This is a potentially important issue since females may have been affected by the shock in a way that differed from males.

Another dimension along which we can assess representativeness is age. Figure 14 provides histograms comparing the share of deaths in the aggregate Registrar's data to both the GRO and BMD data. As discussed in the main text, our data will not be representative of deaths at young ages for mechanical reasons. However, as Figure 14 shows, this issue is more severe in the BMD data than in the GRO data. In addition young adults are substantially overrepresented in the BMD data while older adults are underrepresented.

Replicating our main results

In this section we replicate our main analysis results using the alternative BMD data set. We follow a procedure that is essentially identical to that used in our main analysis of the GRO data. Specifically, we link deaths from 1851-1855 back to their corresponding 1851 census record, and deaths from 1861-1865 back to their 1861 census entry and then conduct a difference-in-difference analysis.

Figure 14: Histogram of deaths by age at time of death from linked and aggregate data



Results obtained using the BMD data are presented in Table 11. Column 1 presents results calculated across all age groups while Columns 2-4 break results down by age group. Note that, because the linked BMD data set is substantially smaller than the GRO data set, it is necessary to use more aggregated age groups in the results in Columns 2-4 than those used in the main text.

As in our main analysis, the BMD results show evidence of a substantial increase in mortality among older adults, but no measurable effect among the young or prime-aged adults. The magnitude of the effect among the elderly is similar to what we obtained for older ages in our main GRO analysis. The estimated all-age effect presented in Column 1 is smaller than that observed in the GRO data. This is clearly driven by negative estimated effect among younger populations. It is notable that these are populations where the BMD data set is less representative than the GRO data. In general, the broad patterns observed in the BMD data are generally in-line with those presented in our main analysis of the GRO data, despite the fact that substantial differences exist between these two alternative linked datasets.

Table 11: Baseline regressions from BMD to Census links

DV: Deaths per 1,000 Individuals (per year)	**All Ages**	**Under 15**	**Age 15-54**	**Over 54**
	(1)	(2)	(3)	(4)
Cotton District × Cotton Shortage	0.869	-1.026	-0.244	11.255***
	(0.835)	(1.408)	(0.469)	(3.497)
Nearby (0-25 km) × Cotton Shortage	1.766	4.250**	0.488	8.124**
	(1.188)	(1.848)	(0.442)	(3.434)
Nearby (25-50 km) × Cotton Shortage	0.819	0.937	0.343	3.872
	(1.164)	(1.690)	(0.483)	(3.294)
Nearby (50-75 km) × Cotton Shortage	-1.229	-1.507	-0.048	0.224
	(0.932)	(1.327)	(0.467)	(3.269)
Observations	1,074	1,074	1,074	1,074
Linked deaths	147,434	53,927	67,284	26,223

*** p<0.01, ** p<0.05, * p<0.1. Standard errors, clustered at the district level, in parentheses. Deaths are assigned to the district of initial residence (i.e., district of census enumeration). Regressions are weighted by population. All regressions include district fixed effects, period fixed effects, region-by-period fixed effects, and controls for ln(population density), and the share of the population (in each district-by-period cell) that has a linkable name. The all-age regression in Column 1 also includes controls for the share of the population in each of the following age categories: under 15, 15-54, and over 54 (with 15-54 as the omitted category).

Placebo exercise

One advantages of the BMD database is that we also have linked deaths for a longer period than in our main analysis. This allows us to undertake other tests to validate and extend our main findings. For one, the BMD data contain linked deaths for 1856-1860. In Table 12 we use these links to conduct a placebo exercise in which we treat the 1856-1860 period as a placebo "treated" period and look at whether we observe any differential mortality patterns in the cotton textile districts in that period relative to the 1851-1855. These results show no evidence of differential mortality patterns in the cotton textile districts in the 1856-1860 period. This tells us that there do not appear to be differential pre-trends in mortality that might be behind our main results.

Table 12: Placebo test: 1856-1860 (placebo) against 1851-1855 (control)

DV: Deaths per 1,000 Individuals (per year)	**Under 15**	**Age 15-54**	**Over 54**
	(1)	(2)	(3)
Cotton District × Placebo Period	-0.804	-0.736	-3.087
	(1.403)	(0.450)	(3.206)
Nearby (0-25 km) × Placebo Period	1.703	-1.175**	1.957
	(1.566)	(0.498)	(3.138)
Nearby (25-50 km) × Placebo Period	-0.212	0.002	-1.412
	(1.462)	(0.337)	(2.466)
Nearby (50-75 km) × Placebo Period	0.313	0.053	-6.507*
	(1.592)	(0.427)	(3.363)
Observations	1,074	1,074	1,074
R-squared	0.067	0.038	0.040
Linked deaths	40,376	57,809	22,424

*** p<0.01, ** p<0.05, * p<0.1. Standard errors, clustered at the district level, in parentheses. Deaths are assigned to the district of initial residence (i.e., district of census enumeration). Regressions are weighted by population. All regressions include district fixed effects, period fixed effects, region-by-period fixed effects, and controls for ln(population density, and the share of the population (in each district-by-period cell) that has a linkable name.

Examining harvesting versus scarring effects

For another, the linked BMD dataset also includes data from 1866-1870, which can be used to assess whether the cotton shock had persistent impacts on mortality—that is, we investigate whether our results reflect "harvesting," (i.e., the possibility that the cotton shortage merely hastened the deaths of those who would have anyway died within a short period thereafter), a phenomenon which has important implications for the overall mortality cost of the downturn. This is a particular concern given that our strongest adverse effects appear among older adults. To evaluate harvesting, we look at mortality rates among the treated populations in the period after the shock, 1866-1870, as compared to those in the 1856-1860 period, which is the most comparable available control period.

These results, presented in Table 13, indicate that mortality levels remained elevated among the adults exposed to the cotton shock (i.e., those resident in the cotton districts in 1861) in the years after 1865, especially older adults. This tells us that either no substantial harvesting occurred, or that any harvesting effect was dominated by the persistent effect of the recession on health. This sustained effect could be due to a number of factors including "scarring" i.e., a reduction in health capital during the shock that increased mortality risk in the next period; and persistent economic effects that continued even after cotton supplies resumed, for instance, through ongoing congestion. Regardless of channel, the fact that we do not find strong evidence of harvesting indicates that the pattern observed during the cotton shock was not merely confined to populations that would have died imminently in the absence of the downturn. These results represent an important contribution to the literature on business cycles and mortality, which has engaged relatively little with the mortality dynamics of local economic shocks. Here, scarring and harvesting are not only of substantive interest as phenomena affecting health, but they may also confound inference in traditional approaches using annual aggregate panel data.

Table 13: Harvesting test: 1866-70 (post-shortage) against 1856-60 (control)

	DV: Deaths per 1,000 Individuals (per year)		
	Under 15	**Age 15-54**	**Over 54**
	(1)	(2)	(3)
Cotton District × Post-Shortage Period	-0.726	0.851*	10.026**
	(1.356)	(0.488)	(4.328)
Nearby (0-25 km) × Post-Shortage Period	3.099	1.761***	5.466
	(1.989)	(0.609)	(4.182)
Nearby (25-50 km) × Post-Shortage Period	2.051*	0.208	-4.083
	(1.223)	(0.399)	(3.409)
Nearby (50-75 km) × Post-Shortage Period	0.366	-0.636	2.088
	(1.883)	(0.444)	(3.702)
Observations	1,076	1,076	1,076
R-squared	0.064	0.115	0.058
Linked deaths	46,580	66,235	23,003

*** p<0.01, ** p<0.05, * p<0.1. Standard errors, clustered at the district level, in parentheses. Deaths are assigned to the district of initial residence (i.e., district of census enumeration). Regressions are weighted by population. All regressions include district fixed effects, period fixed effects, region-by-period fixed effects, and controls for ln(population density, and the share of the population (in each district-by-period cell) that has a linkable name.

D.4 Additional results using aggregate data

In Table 14, we examine the results on infants in more detail.[47] We find evidence that the infant mortality rate fell in cotton districts during the crisis, as did the birth rate. The former results are consistent with the possibility that infant health improved as a function of maternal time reallocation (a story consistent with contemporary reports discussed in Section 3.6, as well as with the results in Miller & Urdinola (2010), which suggest that in poor settings, child health may be more intensive in maternal time than in income), the possibility of negatively-selected out-migration (thus raising the health "quality" of remaining mothers), or the possibility of positive selection into childbearing due to the downturn (consistent with work like Dehejia & Lleras-Muney (2004), which finds low- but not high-SES mothers deferring childbearing in a modern recession). Given the concurrent reduction in birth rates, the last explanation of the infant health results appears most plausible, though it is difficult to draw strong conclusions: because birth rates are taken from aggregate data, these rates may have fallen in cotton districts during the shortage at least in part as a mechanical result of out-migration.[48]

[47] When considering these results, there is one important caveat: namely, because this analysis uses aggregate data on births and infant deaths rather than linked data, these results may still be subject to migration bias, and so cannot be directly compared to those presented elsewhere in our age-specific analysis. Specifically, because these outcomes are reported by place of occurrence, we will only capture the births and infant deaths to accruing to those individuals actually living in a given district in the year in question. For instance, the births and infant deaths we observe in the cotton district will be just those coming from a dynamic subset of stayers. Practically speaking, this simply means that because of unobserved/imperfectly observed migration, the number of births may be affected by the size of the remaining population, and the infant mortality rate, which takes annually-reported births as its denominator, may be affected by the composition of the remaining population.

[48] Similarly, if out-migration was highly selective, leaving behind women with preferences for relatively low fertility, this too could contribute to a reduction in the birth rate.

Table 14: Examining Fertility and Infant Health (Aggregate Data)

	Births per 1,000 persons (1)	**Infant Deaths per 1,000 births** (2)
Cotton District × Shortage	-5.019***	-13.748***
	(0.648)	(4.369)
Nearby (0-25 km) × Shortage	-2.316**	-1.207
	(0.936)	(4.315)
Nearby (25-50 km) × Shortage	-0.311	-1.270
	(0.502)	(4.107)
Nearby (50-75 km) × Shortage	0.700	1.950
	(0.474)	(3.698)
Observations	1,076	1,076
R-squared	0.926	0.940

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors, clustered at the district level, in parentheses. Regressions are weighted by population. All regressions include district fixed effects, period fixed effects, region-by-period fixed effects, and controls for ln(population density, the share of the population (in each district-by-period cell) that has a linkable name, the under 15 share, and the over 54 population share.