Can migration make deadly recessions look healthy? Evidence from large-scale linked microdata*

Vellore Arthi
University of Essex

Brian Beach
William & Mary
and NBER

W. Walker Hanlon
NYU Stern
and NBER

March 15, 2018

Abstract

Are recessions good for health? A large existing literature using aggregate panel data suggests that the answer to this question is yes. However, this approach may be biased if people move in response to recessions. We offer direct evidence of this "migration bias" using a historical setting that provides both large-scale longitudinal microdata and an exogenous shock to local economic conditions. Using microdata to overcome migration bias, we find that the recession increased overall mortality. However, our estimates from aggregate data fail to recover this effect. Thus, in at least some settings, migration bias substantially affects results obtained from standard methodologies.

JEL Codes: I1, J60, N33

*Arthi: v.arthi@essex.ac.uk; Beach: bbbeach@wm.edu; Hanlon: whanlon@stern.nyu.edu. We thank James Fenske, Joe Ferrie, Marco Gonzalez-Navarro, Tim Hatton, Taylor Jaworski, Amir Jina, Shawn Kantor, Carl Kitchens, Adriana Lleras-Muney, Doug Miller, Grant Miller, Christopher Ruhm, William Strange; audiences at the 2017 ASSA Annual Meeting, 2017 NBER Cohort Studies Meeting, 2017 PAA Annual Meeting, and 2017 SDU Workshop on Applied Microeconomics; and seminar participants at University of California, Davis, Columbia, Cornell, Essex, Florida State, Michigan, Queen's University, Queen's University Belfast, RAND, Toronto, and Warwick; for helpful comments. For funding, we thank the UCLA Rosalinde and Arthur Gilbert Program in Real Estate, Finance and Urban Economics, the California Center for Population Research, the UCLA Academic Senate Faculty Research Grant Fund, and the National Science Foundation (CAREER Grant No. 1552692). This study builds on a previous NBER Working Paper (No. 23507), "Estimating the Recession-Mortality Relationship when Migration Matters."

1 Introduction

Following the seminal paper by Ruhm (2000), there is now a large literature in economics documenting that mortality rates are procyclical.[1] The papers in this literature tend to apply a panel data methodology that compares unemployment rates to aggregate mortality rates. The results from this literature are widely cited and regularly reported in the media.[2] Recently, however, a small number of studies have examined this relationship using longitudinal microdata, which allow researchers to track individuals over time (Gerdtham & Johannesson, 2005; Halliday, 2014). Despite covering similar locations and time periods as earlier studies, these microdata studies typically do not find evidence that mortality is procyclical. Why these two sets of results differ, and whether we should trust results derived from aggregate data, remain open questions.

In this paper, we offer one possible explanation for these diverging results: migration. A critical assumption embedded in the aggregate panel data approach employed by Ruhm (2000) and much of the literature that followed is either that business cycles do not induce a large migratory response, or that both the size and composition of any migratory flows are accurately captured in intercensal population estimates. If this assumption fails, then we may observe a spurious change in the mortality rate which we will misinterpret as reflecting the true impact of business cycles on health. We call this phenomenon *migration bias*. In contrast to the aggregate panel data approach, which infers health change by focusing on data anchored to specific geographic units,

[1] A review of leading papers is provided in Appendix A.1. Summarizing this literature, Ruhm (2015) writes, "Using data from a variety of countries and time periods, these investigations provide strong evidence of procyclical fluctuations in total mortality and several specific causes of death." Notable exceptions which find either mixed health effects or evidence of countercyclical mortality include Svensson (2007), Economou *et al.* (2008), and Miller & Urdinola (2010). Cutler *et al.* (2016) find mixed evidence depending on the size of the shock: large recessions are bad for health while small ones may improve health.

[2] A recent example is an article in the *New York Times' Upshot* blog by Austin Frakt on Oct. 16, 2017 titled, "How a Healthy Economy can Shorten Life Spans."

studies using microdata are able to analyze how health among a given *cohort* changes regardless of where their deaths ultimately occur. Accordingly, microdata studies are much less vulnerable to migration bias.[3]

This is the first study to directly examine the impact of migration on estimates of the recession-mortality relationship, and to compare results obtained using both longitudinal microdata and aggregate-data methods.[4] To do so, we draw on a historical setting—the recession in the cotton textile-producing regions of Britain caused by the U.S. Civil War (1861-1865)—that offers two features that are vital for examining these issues.

First, this setting offers plausibly exogenous variation in the timing and spatial distribution of a short-term economic shock, which allows us to more cleanly identify the impact of changes in local economic conditions on mortality. On the eve of the U.S. Civil War, the British cotton textile industry was heavily reliant on raw cotton inputs from the United States. The war sharply reduced these supplies, resulting in a deep but temporary downturn in British districts with economies dependent on cotton textile production.[5] The temporal component of the economic shock was short, sharp, and generated by outside forces that were largely unexpected. Meanwhile, its spatial incidence was determined by the pre-existing distribution of economic activity, which was in turn due to underlying natural endowments. This approach to identifying a temporary change in local economic conditions, which follows Miller & Urdinola (2010), helps us deal with the fact that unobserved migration may cause unemployment rates, the key explanatory variable used in most of the studies following Ruhm (2000), to become endogenous.[6] More importantly, by using an exogenous shock

[3]Aggregate-data studies implicitly assume that geographic units and cohorts are effectively equivalent. This is only the case if there is no migration, or if migration is perfectly observable.

[4]While we are not aware of a study assessing the extent to which migration may affect the results in this literature, we are not the first to suggest that migration may be a source of bias in this literature. The potential bias created by migration is mentioned in Halliday (2014), Ruhm (2007), Stuckler *et al.* (2012), and Lindo (2015).

[5]This setting has previously been studied by Hanlon (2015) and Hanlon (2017), which examined outcomes related to technological change and urban growth, respectively.

[6]Unemployment rates both affect and are affected by migration; a change in the unemployment

to local economic conditions, we can more cleanly identify and correct for spatial spillovers due to migration.

A second, and perhaps more important, feature of this historical setting is that it allows us to draw on comprehensive, individually-identified, and publicly available census and death records. Linking these sources allows us to construct a much larger sample of longitudinal microdata—covering 151,558 deaths—than has previously been possible in this literature.[7] Our linking of census and death data is informed by previous work in this area, including Ferrie (1996), Abramitzky *et al.* (2012, 2014), Feigenbaum (2015), and Bailey *et al.* (2017). Drawing on this linked microdata, we are able to generate accurate estimates of the recession-mortality relationship, even in the presence of migration. This is because, rather fortuitously in our setting, the Census of 1861 was taken in April of that year, just before the U.S. Civil War began affecting the British economy. Thus, we think of the population resident in the cotton-producing areas at that time as the group at risk of experiencing the effects of the recession. By linking deaths occurring during the cotton shortage (1861-1865) back to the census, we can accurately identify mortality patterns in this group, relative to residents of other locations, irrespective of where they may have subsequently migrated and died. This is impossible in more aggregated data.

The cotton shortage, then, offers us a prime setting in which to examine the sensitivity of the standard aggregate-data approach to migration bias by directly comparing estimates of the recession-mortality relationship derived from longitudinal microdata to those obtained from traditional aggregate data. Our main analysis applies a difference-in-difference methodology that compares deaths from 1861-1865 linked back to the 1861 census, relative to deaths from 1851-1855, linked back to the

rate may induce migration, and this migration will in turn affect both the numerator and denominator used to calculate unemployment rates in the next period. Foote *et al.* (2017) raise concerns about the endogeneity of unemployment rates.

[7]The small number of deaths has been a challenge in previous studies using microdata. In Gerdtham & Johannesson (2005) the data include 6,571 deaths. The authors opt for a time-series analysis with the national unemployment rate as the key explanatory variable. The PSID data used in Halliday (2014) include 1,932 deaths from 1984-1993.

1851 census, in cotton districts relative to other parts of the country. Importantly, our linked data allow us to generate a mortality rate for these populations regardless of whether some people migrated as a result of the recession and died outside of the cotton districts. Thus, these results, which do not suffer from migration bias, provide us a benchmark for the true mortality impact of the cotton shortage.[8] Next, we re-estimate our main specification after re-assigning the deaths in our linked data to the district in which they occurred, rather than to each person's initial district of residence.[9] Comparing these results to our benchmark findings directly identifies the impact of migration bias on estimates of the recession-mortality relationship. Finally, we set aside the linked data and apply our estimating approach to actual aggregate data. Comparing these results to those obtained from our linked data, we investigate whether an analysis based on aggregate data is likely to be able to reliably recover what we know from our preferred linked-data estimates to be the migration-unbiased impact of the cotton shortage on mortality.

This study generates four main results. First, we show that temporary changes in local economic conditions can induce substantial and systematic migration responses. Specifically, we find evidence that during the downturn, large numbers migrated out of the cotton districts and into nearby non-cotton districts, mostly within a 50 km radius. We also find that in-migration to cotton districts fell during the shortage. Migration of this sort is not only important because it affects the size and composition of the underlying at-risk population, but also because migrant spillovers may undermine inference in a differences-in-differences framework. To address this issue, we identify migrant-sending and migrant-receiving districts, and separately estimate the mortality effects of the cotton shortage on each of these sets of districts, relative to a third set of more distant control districts which offer a cleaner counterfactual.

[8] Additional graphical evidence on the relationship between the cotton shock and mortality can be found in Appendix D.1.

[9] This configuration of the linked data mimics the organization of standard aggregate data, but allows us to directly examine the migration channel, since both the place of origin at the onset of the shock, and the place of death, are known.

These results show that migration led to substantial spatial spillovers that would have biased our results if left unaddressed.

Second, we provide direct evidence that migration can lead us to draw meaningfully inaccurate conclusions about the impact of a given downturn on mortality. Our preferred linked microdata results, which assign deaths to the place of initial residence, indicate that the downturn raised the mortality rate among those residing in a cotton textile district at the onset of the cotton shortage, particularly for older adults. However, when we re-assign deaths to the location of death, as in traditional aggregate data, we find no evidence of a mortality increase during the cotton shortage.[10] If anything, these results suggest that mortality may have decreased. Thus, addressing migration bias completely alters the conclusions that we draw from the same data.

Third, we show that the aggregate panel data approach common in the literature is unlikely to reliably recover accurate estimates of the recession-mortality relationship in the presence of unobserved migration. In particular, we find that for most age groups, the aggregate-data results are both qualitatively and statistically equivalent to the results obtained from microdata when deaths are assigned to the location in which they occur, but differ sharply from our preferred results, in which deaths are assigned to the individual's initial location. Thus, we find that for most age groups, aggregate data fails to recover the true impact of the recession on mortality. Importantly, we can explicitly pinpoint migration as the source of this discrepancy.

Finally, we provide substantive evidence on the health effects of a specific historical recession. We show that the cotton shortage had an adverse impact on mortality in

[10]It is worth noting that, even though modern mortality data is often enumerated by the decedent's location of residence rather than where death occurred, this type of data is still subject to migration bias. For example, if a recession in one location causes a migrant to move to a new location where they establish residence and then subsequently die, the death will be assigned to the new residence location. As a result, when using aggregate data this death would not contribute to the estimated effect of the recession on the mortality rate of the population who experienced it. This is particularly true given that most U.S. states require people to change their residency to their new state fairly quickly, usually in the first month of living there.

the affected population, as well as in the populations of nearby districts that were the destination of migrants leaving the cotton textile areas. The latter findings are suggestive of possible congestion effects, such as contagion and overcrowding. Within cotton districts, the mortality effects were concentrated amongst the elderly. We also look at how individuals' occupations, and the occupations of their family members, influenced mortality. Specifically, we show that mortality increased among cotton workers and members of their households as a result of the cotton shortage, but that this direct impact explains only part of the adverse effect of the downturn on the mortality of cotton-district residents. These results are broadly consistent with the mechanisms outlined in studies such as Stevens *et al.* (2015) and Sullivan & von Wachter (2009).

Together, these results show that migration bias is important for estimates of the relationship between recessions and mortality in our setting. These results also offer a potential explanation for why recent results obtained from modern microdata diverge from those using more aggregate data. Of course, the question of whether migration also affects the results obtained in other, more modern, settings depends on the strength and selectivity of the migratory response to local economic shocks, as well as the extent to which these patterns are accurately reflected in interpolated intercensal population data. There is, however, evidence that migration remains a key margin of adjustment to local economic shocks in modern contexts (Blanchard & Katz, 1992; Bound & Holzer, 2000; Foote *et al.*, 2017; Monras, 2015, 2018), and that it may be imperfectly observed in even the best available modern aggregate data (see, e.g., Phipps *et al.* (2005)).[11] Moreover, unobserved migration has gained attention in other literatures as a plausibly large source of bias in modern settings. For instance, Borjas (2003, 2014) has argued that modern migratory responses can be large enough to bias estimates of the impact of foreign migrants on local labor market conditions.

[11] See Appendix A.2 for more on the quality of modern intercensal population data.

2 Business cycles and health

2.1 Related literature

There is a large existing literature estimating the impact of recessions and booms on mortality.[12] While work in this area was pioneered by Harvey Brenner (see, e.g., Brenner (1979)), the bulk of this literature follows the panel data approach introduced by Ruhm (2000), a paper that has been tremendously influential in academic, policy, and popular debates, and that has been cited over 1,600 times to date.[13] A review of leading papers in the extensive literature building on Ruhm (2000) is provided in Appendix A.1.

The existing literature highlights a number of channels through which economic conditions can affect mortality rates. For example, recessions may improve health by removing individuals from environmental and work-related hazards such as pollution, traffic accidents, and on-the-job injuries (Muller, 1989; Chay & Greenstone, 2003; Miller *et al.*, 2009); by freeing up time for breastfeeding, childcare, exercise, and other salutary activities (Dehejia & Lleras-Muney, 2004; Ruhm, 2000); by raising the quality of elder-care (Stevens *et al.*, 2015); and by limiting the capacity for unhealthy behaviors such as smoking and alcohol use (Ruhm & Black, 2002; Ruhm, 2005). On the other hand, these adverse income shocks can worsen health by compromising access to proper nutrition, shelter, and medical care (Griffith *et al.*, 2013; Painter, 2010) or by causing psychological stress, which may in turn raise rates of suicide and risky behavior (Eliason & Storrie, 2009; Sullivan & von Wachter, 2009).[14] Accordingly,

[12] We are using the terms "recession" and "boom" loosely, to refer to temporary changes in local economic conditions (i.e., changes in unemployment rates and individual incomes) rather than to more technical changes in national GDP. This is convention is consistent with existing literature; see, for instance, Ruhm (2000, p. 617).

[13] According to Google Scholar.

[14] Work such as Adda *et al.* (2009) has suggested that some of these same channels may operate for less transitory income shocks as well. However, they find less support for the idea that behavioral change feeds instantaneously into mortality outcomes.

the net effects of business cycles on mortality are ambiguous ex ante. Nevertheless, the bulk of the studies in this literature, which rely on aggregate panel data, consistently find that mortality is procyclical on net (see Appendix A.1 and Ruhm (2015)).

It is striking, then, that a handful of recent papers covering much the same periods and geographies, but using individual-level longitudinal data, have found at best mixed evidence of this result. For instance, Halliday (2014), using PSID data from the U.S., finds strong evidence of countercyclical mortality in working-age men, and little evidence of cyclical mortality in other groups.[15] Edwards (2008), also working on the U.S., but drawing on NLMS data, reports mixed findings that include countercyclical mortality among the less educated.[16] Similarly, Gerdtham & Johannesson (2005), who analyze modern Swedish microdata, find countercyclical mortality in men, and acyclical mortality in women.

Importantly, these papers use microdata to track treated cohorts over time, which renders them less vulnerable to the issues that arise from migration in response to business-cycle fluctuations—migration that typically goes unobserved in traditional aggregate data. The dissonance between their findings and those of the broader aggregate-data literature suggests that migration bias may not only exist, but may also have a meaningful impact on estimates of the recession-mortality relationship. However, it is difficult to draw strong conclusions from these results because the datasets these authors employ cover relatively small numbers of linked deaths: 6,571 deaths in Gerdtham & Johannesson (2005) and 1,932 deaths in Halliday (2014). One reason that we focus on a historical setting is that it allows us to work with a much larger sample of linked deaths, which helps us generate clearer results.

[15]The PSID has also yielded evidence of countercyclicality in other health outcomes. See, e.g., Wang *et al.* (2018); Halliday (2007). Another related study using microdata is Ruhm (2003), which uses individual-level survey data on reported health status and medical care use, but not mortality.

[16]Although Edwards (2008) uses microdata, this study differs from ones like Halliday (2014) and Gerdtham & Johannesson (2005) in that the location of individuals is observed only at the beginning of the study period but is nevertheless used to infer exposure to business-cycle shocks up to two decades later. As a result this study is more vulnerable to being influenced by migration that studies using data that track individual locations over time.

Underscoring the possibility that unobserved migration may play a role in the discrepancy between aggregate-data and microdata studies is a recent paper by Lindo (2015), which shows that results obtained from aggregate data in the modern U.S. are sensitive to the level of spatial aggregation. Lindo suggests that migration across smaller geographic units is a likely explanation for this pattern. In the following section, we discuss the ways in which unobserved migration can affect the standard estimating framework used in the broader aggregate-data literature.

2.2 Methodological issues introduced by migration

The starting point for our analysis is the standard estimating equation introduced by Ruhm (2000):

$$\ln(MR_{dt}) = \beta U_{dt} + X_{dt}\Gamma + \phi_d + \eta_t + \epsilon_{dt}\,, \tag{1}$$

where MR_{dt} is the mortality rate in a given location (e.g., district) d; η_t and ϕ_d are a full set of time-period and location fixed effects; U_{dt} is the location's unemployment rate (or another similar variable representing local economic conditions); and X_{dt} is a set of controls.

There are three channels through which migration can affect estimates generated using Eq. 1. First, migration may affect the key explanatory variable, U_{dt}. Second, migration may affect the dependent variable, MR_{dt}, by causing the size and/or the composition of the observed at-risk population to differ from that of the true at-risk population. Third, migration spillovers across locations may affect results through the comparison, implicit in Eq. 1, between treated and control locations. The goal of our analysis is to generate an estimating approach that deals with each of these three potential issues, and then to compare the results with the more standard approach shown in Eq. 1. Below we discuss each of these potential channels for bias, and how they are addressed in our analysis.

First, migration may cause unemployment rates, the key explanatory variable used in most of this literature, to become endogenous. This is particularly true if individuals migrate toward areas that offer better economic opportunities, since migration will not only respond to the local unemployment rate, but it will also directly affect the unemployment rate by changing the size and composition of the labor force.[17] In our analysis, we deal with this issue by replacing the unemployment rate U_{dt} with a plausibly exogenous measure of temporary changes in local economic conditions, $SHOCK_{dt}$. This yields the following modified specification,

$$\ln(MR_{dt}) = \beta\, SHOCK_{dt} + X_{dt}\Gamma + \phi_d + \eta_t + \epsilon_{dt}\,. \tag{2}$$

Conditional on $SHOCK_{dt}$ being exogenously determined, Eq. 2 removes concerns about the measure of local economic conditions being endogenous in the presence of migration.[18] In our setting, $SHOCK_{dt}$ is generated by interacting a time-varying industry-specific shock to the cotton textile industry with measures of the initial importance of the affected industry in each location.[19] Specifically, we interact an indicator for the U.S. Civil War period (1861-65) with another indicator for being a cotton textile district.[20]

Next, migration may affect estimates obtained from Equations 1 and 2 by changing the size and composition of the true at-risk population in ways that are unobservable to the researcher. This in turn will create a wedge between the true mortality

[17]Moreover, migration will affect the unemployment rate in both migrant-sending and migrant-receiving locations, potentially tainting our control group.

[18]Note that we use a reduced-form approach rather than using $SHOCK_{dt}$ as an instrument for unemployment. This is because we do not have access to suitable unemployment data in our setting to enable an IV approach. The best proxy available to us is Poor Law relief-seekers, which is neither truly comparable to modern unemployment, nor consistently available for the entire period of our study.

[19]This approach closely follows Miller & Urdinola (2010) who use coffee price shocks and spatial variation in coffee cultivation as an exogenous shock to local economic conditions in Colombia.

[20]Cotton textile districts are defined as those with greater than 10% of employment in cotton textiles in 1851, a decade before the U.S. Civil War. Because the location of industry was relatively persistent, we also obtain similar results when using the spatial distribution of industry in 1861.

rate, and that which is observed, MR_{dt}. While mortality data are typically observed annually, population is only accurately observed in census years.[21] Thus, migration can introduce errors in interpolated population values which cause the size of the observed population in a given location to differ from that of the true at-risk population.[22] Migrant selection introduces another issue: if migrants are not representative of the relevant at-risk population, then their intercensal movements will cause the composition (and so, the average mortality risk) of the true at-risk population to differ from that which is observed.[23] For instance, if out-migrants are positively selected on health, then the observed population of their place of origin will appear less healthy. Meanwhile, the opposite will be true if they are negatively selected. In either case, however, the result of this selection is to bias estimates of the recession-mortality relationship.

Linked individual-level longitudinal data offers a solution to these issues by al-

[21] Population in non-census years is typically interpolated using observed population counts from two censuses as well as data on births and deaths occurring in the interim. In data from modern developed countries, additional sources are used to help capture that portion of migration which can be observed. For example, the U.S. Census uses tax information from the IRS as well as Medicare data to track migration among working-age and older adults, respectively. In Appendix A.2 we show that the amount of unobserved migration (termed the error of closure) that must be allocated across intercensal years in modern settings is relatively similar to that in our setting. Note also that this approach implies that all intercensal population values are interdependent and endogenously constructed by design.

[22] Assuming that our population measures on the eve of any business cycle-induced migration are reasonably accurate, if we expect people to migrate from places with worse economic conditions to those with better ones, and if migration is not fully captured by intercensal population estimates, then this will mechanically bias results towards finding procyclical mortality. To see this, note that unobserved out-migration will lead to an artificially high population denominator and fewer observed deaths because of a smaller at-risk population. Conversely, unobserved in-migration will lead to an artificially low population denominator but more observed deaths because the at-risk population has increased. Thus, even if migration is not selective, and even if there is no change in the true underlying mortality rate (say, due to either the direct effect of business-cycle fluctuations on health or to the effects on health of the act of migration itself), the relocation of individuals from one region to another can mechanically generate the false appearance of health change where there has been none.

[23] For example, if healthy young workers are more likely to migrate to areas offering better economic opportunities, then we should expect the observed mortality rate in the location experiencing worse economic conditions to rise even if the change in economic conditions itself has no causal impact on mortality. Focusing on age-standardized mortality rates can partially deal with the selective sorting of population, but it cannot account for selection of migrants within age groups.

lowing the researcher to isolate and trace cohorts, rather than geographies, that were exposed to the relevant shock.[24] Specifically, by fixing individuals to their location at the onset of the shock, their deaths can be correctly attributed to the relevant (e.g., treated) location irrespective of where they ultimately occur. Thus, this approach ensures that the population represented in the denominator of the mortality rate corresponds to the group of people whose deaths appear in the numerator. Thus, we modify our specification to,

$$\ln\left(\frac{MORT_{dt}}{POP_{dt}}\right) = \beta\, SHOCK_{dt} + X_{dt}\Gamma + \phi_d + \eta_t + \epsilon_{dt} \tag{3}$$

where POP_{dt} is the population in a district d at the beginning of period t (in our empirical setting, at the 1851 or 1861 census) and $MORT_{dt}$ is the number of deaths among that population during the period (i.e., from 1851-55 or 1861-65).

It is worth noting that, in addition to the purely statistical effects mentioned above, migration may also have very real effects on mortality. For example, migration may affect mortality in both migrant-sending and migrant-receiving areas through congestion effects (e.g., disease contagion or strain on fixed local resources). Alternatively, migration can change underlying population health by relocating people across locations with different intrinsic conditions. If, for example, people move to healthier locations, then migration will have a real and beneficial impact on health.[25] While estimates obtained from a linked-data approach will purge the spurious impact of migration on observed mortality patterns, they will capture—alongside the direct effects of recession on mortality—any *real* effects of migration on mortality. Put another way, our results reflect the true impact of a local economic shock on mortality in a setting in which migration is costly but nevertheless possible.

[24] While the standard aggregate-data panel approach effectively assumes these are one and the same, unobserved migration can create a wedge between these conceptualizations of the panel unit.

[25] This would also be an example of an indirect but true effect of business cycles on mortality—it is indirect in that it occurs as a result of the migration that business cycles induce.

Finally, migration can affect results obtained from Equations 1, 2, and 3 by generating spillovers from treated to putative control locations, thus violating the assumptions necessary for causal inference in a differences-in-differences approach. This issue is a concern in both aggregate and microdata-based analysis, but it can be addressed if migrant-sending and migrant-receiving locations can be identified and compared to a third set of locations that were not contaminated by substantial spillovers, and that can thus serve as a cleaner counterfactual.[26] To operationalize this intuition, we modify our specification to separately estimate the impact of the shock on migrant-receiving districts,

$$\ln\left(\frac{MORT_{dt}}{POP_{dt}}\right) = \beta\, SHOCK_{dt} + \gamma RECEIVING_{dt} + X_{dt}\Gamma + \phi_d + \eta_t + \epsilon_{dt} \qquad (4)$$

where $RECEIVING_{dt}$ is an indicator equal to one for districts receiving migrants from the treated districts during the treatment period. As we will show below, in the case of the cotton shortage, most migration occurred to nearby locations. Thus in our setting, $RECEIVING_{dt}$ will simply be an indicator variable (or variables) identifying districts within a specified radius of cotton districts. We estimate these effects out to 75 km because our evidence suggests that the effect on nearby areas disappears by that distance.

[26] An alternative approach is to aggregate to higher geographic levels. For instance, one could combine migrant-sending and migrant-receiving areas and, in essence, treat the two areas as a single unit. While this type of aggregation would solve the migration-related spillovers issue, it does so by ignoring the fact that the various local labor markets within the aggregated study area are likely experiencing dramatically different economic conditions. Indeed, if there are migrant-sending and migrant-receiving locations within the aggregated unit, it is likely that the difference in local economic conditions is precisely what drove this migration. Accordingly, such aggregation may undermine inference by assuming that the aggregated unit's full population was subject to the same change in local economic conditions. For this reason, we prefer to conduct our analysis at the local labor market level, and to correct for migration spillovers by analyzing migrant-sending and migrant-receiving locations separately. Nevertheless, in Appendix D.2 we present the results of a county-level analysis. The results are similar to those obtained at the district level, and suggest that in our setting, migrant-sending and migrant-receiving districts are not being captured within the same spatial unit when moving to the county or even the regional level.

With these modifications, our linked-data specifications should produce estimates that capture only the real effects of recessions on mortality: both directly, and indirectly through migration. Meanwhile, the results obtained from aggregate-data will reflect these real effects as well as the purely statistical effects generated by unobserved migration. To examine the importance of this migration bias in practice, we now turn to an empirical setting in which both approaches can be compared directly: the Lancashire "Cotton Famine," a sharp, severe, and short-lived downturn in the cotton textile-producing regions of 19th Century Britain.[27]

3 Empirical setting

The cotton textile industry was the largest and most important industrial sector of the British economy during the 19th century. For historical reasons dating to the 1700s, British cotton textile production was geographically concentrated in the Northwest counties of Lancashire and Cheshire, which held over 80% of the cotton textile workers in England & Wales in 1861.[28] The right-hand panel of Figure 1 depicts this spatial distribution by plotting the share of employment accounted for by the cotton textile industry in each district using data from the Census of Population of 1851.

The British cotton textile industry was entirely reliant on imported raw cotton and in the run-up to the U.S. Civil War, 70% of these inputs from the U.S. South (Mitchell, 1988). The war sharply reduced British imports of U.S. cotton, prompting a sudden and dramatic rise in cotton prices and a sharp drop production. These effects are depicted in the left-hand panel of Figure 1.[29] Additional information,

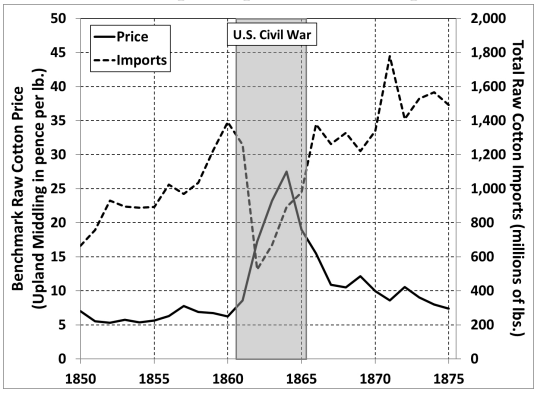
[27]Historians often refer to this event as the "Cotton Famine," where the term "famine" is used metaphorically to describe the dearth of cotton inputs. Henceforth in this paper, we avoid using this term because it can be misleading in a study focused on health.

[28]Calculation based on data collected by the authors from the 1861 Census of Population reports. Crafts & Wolf (2014) suggest the main factors determining the location of the cotton textile industry prior to 1830 was the location of rivers, which were used for power, access to the port of Liverpool, and a history of textile innovation in the 18th century.

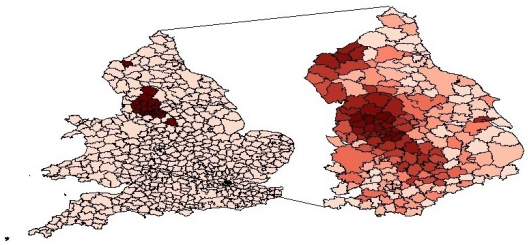
[29]Appendix Figure 5 shows British firms' raw cotton consumption and variable operating costs

Figure 1: Cotton prices, imports, and spatial distribution of cotton textile industry

British import quantities and prices



Spatial distribution of cotton textiles



Import data from Mitchell (1988). Price data, from Mitchell & Deane (1962), are for the benchmark Upland Middling variety. Data on the geography of the cotton textile industry the are calculated from the 1851 Census of Population. Shaded in the map of England & Wales are districts with over 10% of employment in cotton, while the inset shows the share of employment in cotton in the core cotton region, with darker colors indicating a greater share of employment in cotton.

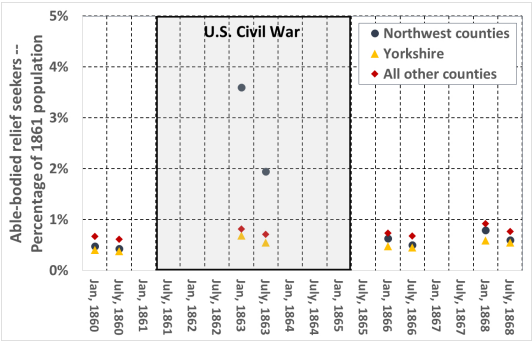
in Appendix B, shows that the direct impact of the U.S. Civil War on the British economy was largely confined to the cotton textile sector.

Both contemporary reports on public assistance and data on relief-seekers suggest the shock had a substantial and geographically-concentrated impact on workers. One reflection of these effects is presented in Figure 2. The left-hand panel of this figure 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. During the U.S. Civil War, 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. Contemporary reports suggest that at the nadir of the recession in 1862 and early 1863, roughly half a million people in the cotton-producing districts relied on relief from government sources or

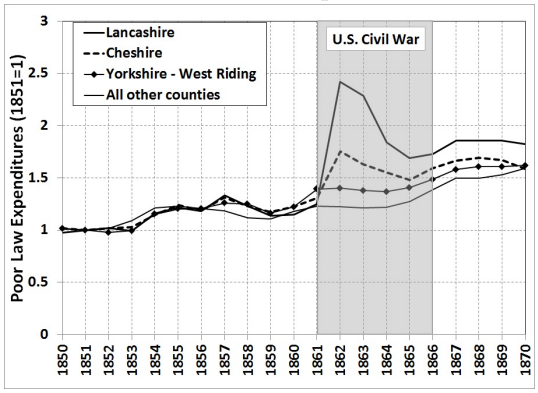
(excluding cotton), good proxies for industry output. These show a sharp drop in production and factor payments during the 1861-1865 period equal to roughly a 50% reduction compared to pre-war output levels. As described in Hanlon (2015), other cotton-producing countries such as India, Egypt, and Brazil rapidly increased their output during the U.S. Civil War period. However, these increases were not large enough to offset the lost U.S. supplies, though they did contribute to the rapid rebound in imports after 1865.

Figure 2: The spatial incidence of the cotton shock

Able-bodied relief-seekers



Poor Law Board expenditures



Expenditure data were collected by the authors from the annual reports of the Poor Law Board. Data on relief-seekers come from Southall *et al.* (1998) (left-hand graph reproduced from Hanlon (2017)).

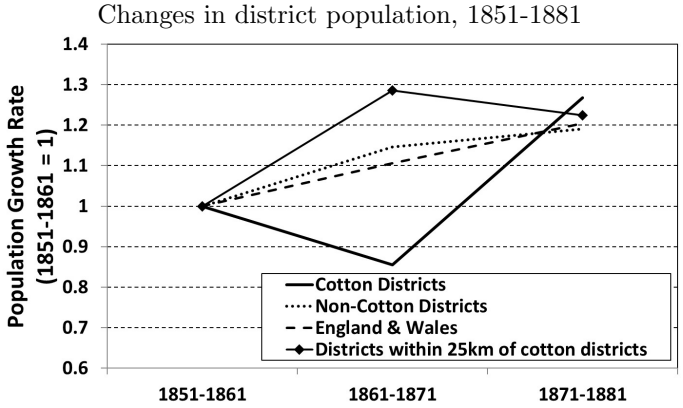
private charities (Arnold, 1864). The right-hand panel of Figure 2 describes expenditures by local Poor Law Boards. These spiked in Lancashire and Cheshire, the two main cotton textile counties, during the U.S. Civil War. In the face of the cotton shortage, workers in the affected areas employed a variety of coping mechanisms, including running down savings, pawning valuables, seeking government and charitable support, and migrating. In Appendix B.2, we review contemporary evidence on all these private and institutional responses to the cotton shortage, with the exception of migration, which we discuss in Section 4.

4 Migration during the cotton shortage

Contemporary accounts indicate that out-migration occurred in response to the cotton shortage, but the extent of these flows remains debated. To assess the magnitude of the migratory response, we look for changes in population growth patterns using data from the 1851-1881 censuses. These data are presented in Figure 3, which describes changes in district population across each decade, normalized by the change

in 1851-1861 (the decade preceding the downturn).[30] This figure reveals three important patterns. First, it presents evidence of a substantial slowdown in population growth in the cotton textile districts in the decade spanning the cotton shortage. This suggests either that the downturn coincided with greater out-migration from those districts or less in-migration. Appendix B.3 presents additional evidence consistent with both increased out-migration and decreased in-migration.[31] Second, we observe an acceleration in population growth in nearby districts, which in this figure we define as non-cotton districts that are within 25 km of a cotton district. Meanwhile, there is little change in the population growth trend in districts beyond 25 km. These patterns are consistent with short-distance migration from cotton textile districts during the downturn. Third, we see that these changes essentially disappear after 1871, a fact which highlights the temporary nature of the shock.

Figure 3: Migration response to the cotton shortage



This graph describes the change in population for all cotton districts, all non-cotton districts, all districts in England & Wales, and all non-cotton districts within 25km of a cotton district using Census data. Cotton districts are defined as those districts with more than 10% of employment in cotton textile production in 1851. The population growth rate for each group of districts is normalized to one in 1851-1861. Data are from the Census of Population.

[30]The 1861 Census was collected in April of that year, before the U.S. Civil War had any substantial impact. As a result, this should be thought of as a clean pre-war population observation.

[31]Specifically, we show 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. Full count census data are not available for 1871, and so the county is the smallest observation for which we can conduct this analysis. Evidence of a sharp and short-lived increase in emigration from Britain during this period is also presented in Appendix B.3.

These implied migration flows were meaningfully large. In terms of magnitude, had the population of the cotton districts grown from 1861-1871 at the same rate that it grew in 1851-1861, these districts would have had 54,000 additional residents in 1871, a figure equal to 2.2% of the districts' 1861 population.[32] Similarly, if nearby districts had grown in 1861-1871 at the rate they grew during 1851-1861, they would have had 61,000 fewer residents, which is equal to 4% of the districts' 1861 population. Note that these figures will understate the migration response if some migrants returned between 1865 and 1871.[33]

An alternative view of migration is provided by calculating implied net migration rates at the district level. These are calculated as the difference between the observed population count in a district in a given census year and the population that we would have expected in that district-year given the population in the previous census plus all births and less all deaths in the intervening years.[34] In the left-hand panel of Figure 4, we map implied net migration by district from 1851-1861. The cotton textile districts show a strong pattern of net in-migration (dark colors) in this decade. In the right-hand panel we plot the change in net migration in 1861-1871 compared to 1851-1861. This figure provides evidence of a large reduction in net in-migration in most of the cotton textile districts during the U.S. Civil War decade. We also see evidence of an increase in migration into districts surrounding the cotton areas, consistent with migration from the cotton districts into nearby locations. Overall, the results in Figures 3-4 show a substantial and short-distance migration response to the cotton shortage.

There is also some evidence of selective migration away from the cotton textile

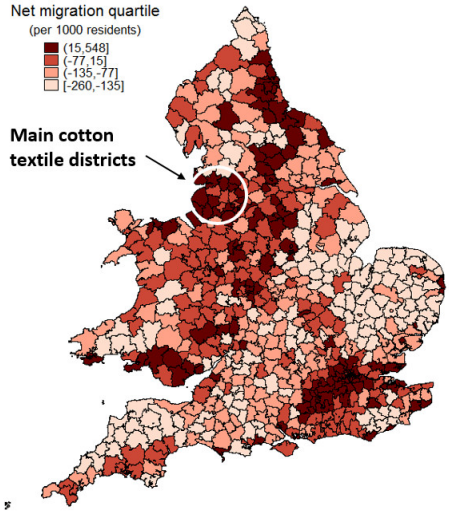
[32] As a point of comparison, during the Great Depression in the U.S., Fishback *et al.* (2006) report that from 1935-1940, 11% of the U.S. population moved, with 60% of the moves occurring within state. This suggests that the migration response observed in our setting may not be unusually large compared to other historical recessions.

[33] These patterns are consistent with the city-level experiences documented in Hanlon (2017).

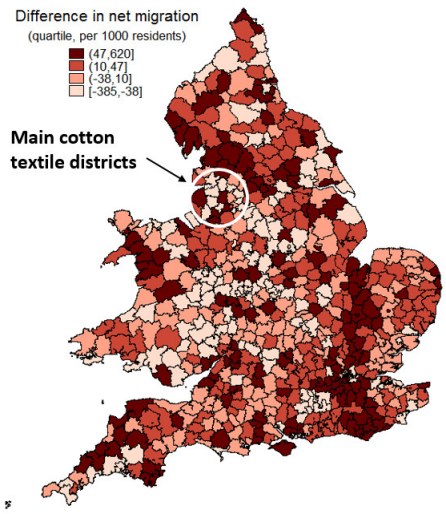
[34] Put another way, implied net migration is the difference between the census count and the postcensal estimate obtained via the components of change method. This conceptual approach has been used in studies of migration such as Fishback *et al.* (2006).

Figure 4: Maps of implied net migration

Net migration before the war (1851-61)



Implied change in net migration from 1851-1861 to 1861-1871



The left-hand panel maps implied net migration (per 1000 residents) for each district in the decade before the shock, 1851-1861. Darker colors indicate net in-migration. The right-hand panel plots the difference in net migration (per 1000 residents) between the 1851-61 and 1861-71 decade. Lighter colors indicate an increase in net out-migration from a district during the U.S. Civil War decade (1861-71). Implied net migration is calculated as the difference between the observed population count in a district in a given census year and the population that we would have expected in that district-year given the population in the previous census count plus all births and less all deaths in the intervening years.

districts during the U.S. Civil War period. In particular, in Appendix B.3 we provide evidence showing that young adults were somewhat more likely to migrate in response to the cotton shock than the elderly. However, the change in population in the 20-39 age group accounts for only about three-fifths of the overall change in population of the cotton districts between 1861 and 1871. Thus, a substantial amount of migration likely occurred among other segments of the population as well.

5 Mortality analysis

The U.S. Civil War-era downturn in the cotton textile-producing regions of Britain offers us the opportunity to directly test the vulnerability of the traditional aggregate-data approach to unobserved migration. To test this, we begin by generating estimates known to be free of migration bias: namely, results from linked data wherein deaths are assigned to the district in which the decedent was resident on the eve of the cotton shortage. We then compare these benchmark estimates to those obtained after re-assigning the deaths in our linked data to the district in which they occurred. This comparison directly illustrates the impact of migration bias on estimates of the recession-mortality relationship, since migration is known to be the wedge between these two sets of results. Finally, we compare both sets of these results to those obtained from actual aggregate data, and ask whether the standard methodological approach used in this literature can reliably recover the true impact of the cotton downturn on mortality.

5.1 Mortality analysis using linked individual-level data

An ideal way to estimate the relationship between recessions and mortality in the presence of unobserved migration is by using individual-level longitudinal data that identifies both an individual's place of residence at the beginning of the recession, and

whether that individual died within a specified period thereafter. The data available for our historical setting allow us to recover this information. Specifically, we have access to comprehensive individual-level census data and death records for England and Wales. Moreover, because census enumeration took place in April of 1861, just as the U.S. Civil War began, this means that with some effort, we can identify deaths in the cohort of individuals actually exposed to the cotton shortage.

5.1.1 Conceptual overview

Ideally we would like to run the regression shown in Eq. 4 in order to recover β, which reflects the impact of the cotton shortage on the mortality rate of the treated population, regardless of where they died. However, the linking of death records to individuals observed in the census complicates our analysis. Our data do not include unique individual identifiers (e.g., a social security number), so we are not able to link every death to a census record. To see how this affects our analysis, let $MORT_{dt}$ be the number of deaths of individuals initially resident in district d and let λ_{dt} be the share of deaths of these deaths that we are able to match back to census records. What we can observe in our linked data is $\tilde{MORT}_{dt} = MORT_{dt}\,\lambda_{dt}$. Substituting out $MORT_{dt}$ in Eq. 4 and reorganizing, we have,

$$\ln\left(\frac{\tilde{MORT}_{dt}}{POP_{dt}}\right) = \beta\, SHOCK_{dt} + \sum_{i\in\{25,50,75\}} \gamma_i NEAR^i_{dt} + X_{dt}\Gamma + \phi_d + \eta_t - \ln(\lambda_{dt}) + \epsilon_{dt}\,. \quad (5)$$

This equation tells us that we will be able to recover β under the assumption that any change in our linking rate λ_{dt} is uncorrelated with the shock. Note that our approach to analysis can handle variation in the linking rate across districts that is constant over time, which will be absorbed by the district fixed effects, as well as changes in the linking rate over time that are constant across districts, which

will be absorbed by the time effects. Before discussing threats to this identification assumption, it is useful to discuss our linking procedure.

5.1.2 Data and linking procedure

Appendix C provides a detailed summary of our data, linking procedure, and the representativeness of our sample. Below, we briefly summarize each of these topics.

To generate data wherein deaths can be assigned to the place of residence at the onset of the cotton shortage, we draw on two main sources: first, comprehensive individual death records for the entire population of England and Wales over the years 1851-55 and 1861-65; and second, the full-count British census for the years 1851 and 1861. The census data come from the UK Data Archive, and include individual names, location at the time of enumeration, age, and some additional information.[35] Our death records come from the Registrar General's master death index, which we obtain from freeBMD.org. In addition to the decedent's first and last name, these records include information on the district, year, and quarter of death.[36]

Using these two sources, we then find each decedent in the census immediately preceding their death, allowing us to recover from the census data their location at the time of enumeration, irrespective of where they eventually went on to die. Because in our setting, the cotton shortage spanned the years 1861-65, and enumeration occurred just before the onset of the shortage,[37] this information allows us to compare mortality patterns across two five-year periods while correctly identifying outcomes in the cohort actually at risk of exposure to the shock we wish to study.

[35]Note that individual-level microdata are not available from the next closest censuses, in 1871 or 1841.

[36]Cause of death is not available from the death index, and so we will not be able to conduct a cause-specific analysis to shed light on mechanisms.

[37]This census was taken in April, the same month that the U.S. Civil War began. Historical evidence makes it clear that people in both the U.S. and abroad failed throughout most of 1861 to anticipate the severity of the conflict that had begun, and there is little evidence that the British economy was substantially affected until late 1861 or early 1862.

Finally, we collapse these individual-level matched observations to the district of residence in the census year (i.e., at the onset of the shock), and generate mortality rates wherein the numerator is the total number of linked deaths occurring in each district over the relevant five-year period, and the denominator is the population reported in each district in the census year that begins the five-year period. We then compare our 1861-65 mortality rates to those in the 1851-55 period using a difference-in-difference approach. This differencing is important, since it helps address concerns about the quality of linking.[38]

As is standard in the literature, we begin by identifying potential links based on time-invariant variables.[39] In the case of our data, and arising from the limited information available in the death records, the only information fitting this description is the individual's name. As such, the matching problem we face differs from that faced in most existing studies using linked microdata. Given the nature of our data, we apply a very straightforward linking methodology: we restrict our sample to the set of deaths and census records with unique first and last names, and we link those that match perfectly as written. Unlike much of the existing literature, we do not allow links among names that are similar but do not match perfectly, or among names that sound alike. We find that doing so decreases the usable linked sample by increasing the number of individuals with multiple matches, particularly since we do not have additional information that can be used to break ties.

While we have less information to use in linking, we do have a couple of advantages in generating links relative to most of the existing literature. One advantage is that name changes and misspellings are likely to be less common in the British census data that we use, than in the U.S. census data used in much of the existing work.[40]

[38] This point is noted in Feigenbaum (2015).

[39] Important contributions to the literature on linking individual-level historical data include (Ferrie, 1996; Abramitzky *et al.*, 2012, 2014; Feigenbaum, 2015, 2016; Bailey *et al.*, 2017).

[40] One reason for this is that there were many foreign migrants in the U.S. who may have changed their name, or had their name misspelled by the census enumerator. Another reason is that in Britain, households typically filled out their own census form, rather than providing their information verbally to an enumerator.

Another difference, relative to existing work, is that we look for individuals over relatively short periods. One implication of this is that we lose relatively few women to name changes resulting from marriage. As such, women are well represented in our linked sample.

5.1.3 Assessing our linked sample

Our main linking procedure relies on direct matching of first and last names. This procedure yields a matched sample of 71,566 individuals who died between 1851 and 1855 (3.6% of all deaths) and 81,221 for the 1861-1865 period (3.8% of all deaths). Of the 539 districts in our dataset, we have linked deaths occurring in 538 districts, and these deaths are linked back to 537 initial districts of residence in the census. Our linking rate is lower than the rate obtained in many linking studies, reflecting the fact that we are unable to use additional information (e.g., age or birthplace) to help uniquely identify a death record. However, an alternative way of calculating the match rate is to identify, among the set of individuals we tried to link (i.e., those with unique names), the share that we are able to successfully link. This rate is about 17%, and is similar to the rate achieved in Ferrie (1996), which constructs a sample of unique name matches across censuses and links about 19% of the records.

Recent work by Bailey *et al.* (2017) highlights the prevalence of false positives in linked data. One way to assess the presence of false positives in our sample is to examine how often the middle initials in the census and death records agree. Middle initials were not universally reported, and as a result we do not use them for linking, but we can use them as a diagnostic tool. Using the sample for which middle names are available, we observe a false positive rate of 28% in the 1850s and 39.8% in the 1860s. This is comparable to the false match rates obtained in other papers using linked data (see Bailey *et al.* (2017)). However, so long as these false matches are roughly randomly distributed, then their effect will be to push our mortality coefficients toward zero, thus working against our findings.

As a check on whether our linked data appear reasonable, we can look at how the probability of finding a link declines as the distance between the death location and enumeration location increases. This analysis, presented in Appendix C.1.3, shows that deaths are much more likely to be matched to individuals previously enumerated in the same district, and the chances of observing a link fall off rapidly and fairly smoothly as the distance between the death district and the census enumeration district increases. This pattern, which one would expect given that migration declines with distance, and which cannot be a mechanical result of our linking procedure, suggests that our linking approach is performing well.

In Appendix C, we also provide evidence that our linked sample of deaths is representative along many dimensions of the aggregate deaths reported by the Registrar General. Looking over time, we find that our the change in the number of deaths over time is well reflected in our linked data: the ratio of deaths in the linked data in 1861 relative to 1851 is 1.17, while in the aggregate data it is 1.14. In terms of gender, women are slightly over-represented in our linked sample, where they account for 53.7% of deaths. In the aggregate data women account for 48.9% of deaths. This is most likely due to women having names that are more unique than those of men.[41] We can also analyze representativeness in terms of socioeconomic status using the occupation data in the census. We find that the shares of deaths among white- vs. blue-collar workers in the linked sample is very similar to the shares generated from aggregate mortality data. Thus, for the working population, our linked sample appears to be quite representative in terms of socioeconomic status.

The main dimension on which our linked deaths sample differs from aggregate mortality patterns is in the age distribution. In our linked sample, working-age adults are over-represented, while young children and, to a lesser extent, the elderly, are under-represented. The fact that young children are under-represented is a mechanical

[41] This is consistent with evidence from sociology that parents are more likely to give girls less traditional names (Rossi, 1965).

consequence of our procedure, since an infant death in, say, 1865, can never be linked to someone alive in the 1861 census. The source of the under-representation of the elderly is less clear, but it may be related to an increase in the likelihood of spelling mistakes in the census. We take two approaches to dealing with this issue. One approach is simply to analyze different age groups separately. Alternatively, when estimating effects across all age groups, we re-weight our linked sample so that its age distribution is representative of that in the corresponding aggregate deaths data, which covers all deaths in the population.

5.1.4 Estimation strategy

Our main identification assumption is that our probability of finding a link conditional on death, λ_{dt}, is uncorrelated with the cotton shock. The most plausible violation of this assumption is that migration generated by the shock may have made it more difficult to link cotton-district residents observed in the 1861 census to deaths over the period 1861-65. However, if individuals who emigrated are less likely to be linked (e.g., because they left the country), and emigration increased from the cotton districts during the shock, then this will bias the estimated effect of the shock downwards, since it will cause the linked deaths among those initially resident in the cotton districts to *understate* the true number of deaths. Thus, if anything, this form of bias will work against the countercyclical results that we present.

The preferred specification used in our analysis of the linked data assigns linked deaths to the initial location of residence as given in the census, and then aggregates these deaths to generate mortality at the district level ($M\tilde{O}RT_{dt}$).[42] We then divide

[42] An alternative approach used in previous studies drawing on microdata is, instead, to run logit or probit regressions at the individual level. There are three reasons why we have chosen to run our analysis at the district rather than the individual level. First, while running individual-level regressions offers the advantage of being able to include controls for individual covariates, this methodology is difficult to implement in a dataset as large as ours, which draws on every individual in two full British censuses. Second, aggregating to the district level makes it easier to compare the results obtained using the linked data to results obtained using traditional aggregate data. Third, it is likely that error terms will be correlated across individuals in the same location. Aggregating

by district population at the beginning of the period (POP_{dt}) obtained from aggregate census data to generate mortality rates.[43] We then take differences across the two time periods available in our data. This procedure leads to the specification shown in Eq. 6. We also standardize the dependent variables in order to deal with the fact that using our linked sample substantially increases the variation in district-level mortality rates. This will be useful later on, when comparing these results to those obtained from aggregate data. Finally, we include controls for initial district population, and, when looking at all-age mortality, controls for the share of different age groups in the district in 1851 and 1861.

$$\Delta \ln \left(\frac{M\tilde{O}RT_{dt}}{POP_{dt}} \right) = \beta \, SHOCK_d + \sum_{i \in \{25,50,75\}} \gamma_i NEAR_d^i + X_d \Gamma + e_{dt} \, . \tag{6}$$

5.1.5 Results

Table 1 presents our main findings. The results across all age groups, in Column 1, indicate that total mortality rates increased by about 0.8 s.d. in the cotton districts in the 1861-65 period relative to the 1851-55 period.[44] This suggests a large, but not unreasonable, effect of the cotton shortage on health. Column 2 presents results looking only at the 0-14 age group, Column 3 presents results looking at those aged 15-54, and Column 4 presents results looking at those aged 55 and over. We consider these three broad age categories so that each category includes a sufficient number

eliminates concerns related to this type of correlated error.

[43] An alternative is to use population denominators obtained by collapsing the individual-level census microdata. In theory these should be the same, but in practice the two series do not match up perfectly, though they are close: the correlation between them is 0.972 in 1851 and 0.991 in 1861. To ensure that this choice is not influencing our results, we have checked alternative estimates using denominators derived from the microdata. These regressions deliver similar findings.

[44] Note that throughout the linked analysis where deaths are assigned to the enumeration district, what we really mean when we say that mortality increased in cotton districts is that it increased in the population initially resident in cotton districts. The same is true when describing the nearby coefficients, which represent the effect on the cohorts initially resident in non-cotton districts nearby to cotton districts.

of linked deaths to keep the standard errors from growing too large.[45] These results indicate that the effects we observe are being driven primarily by deaths among older adults. This pattern fits with some existing results showing that older adults are responsible for much of the effect of business cycles on mortality (e.g., Stevens *et al.* (2015)).

For all age groups, we also observe an increase in mortality rates in nearby districts during the shock period, which disappears outside of about 50 km from cotton districts. There are several factors that may be behind the clear mortality effects that we observe in the population initially resident in nearby districts. One possible channel is through congestion effects caused by the migration of cotton workers into nearby areas. Similarly, cotton workers, who were coming from some of the least healthy districts in the country, may have also brought disease with them. Alternatively, it may be that the economic effects of the cotton shortage also negatively impacted economic activity in nearby non-cotton areas. Our methodology does not allow us to differentiate between these alternative channels. Notably, because the cohorts are fixed at the place of enumeration rather than the place of death, these nearby effects do not incorporate any spurious effects due to the mechanical relocation of deaths from cotton districts to nearby areas.

Overall these results provide evidence that during the cotton shortage, the mortality rate increased among those initially resident in the cotton textile districts, as well as among residents of non-cotton districts within 50 km of the cotton areas. When looking across all age groups, these results are statistically significant at the 90% confidence level. When focusing only on the most heavily affected group, older adults, our estimates are statistically significant at the 99% level.

Appendix Tables 13-14 present additional results using the linked data across all

[45]While 55 is a somewhat young age to begin our older category relative to, say, the results of Stevens *et al.* (2015) which show deaths concentrated in the 80s, it is worth remembering that life expectancy was substantially lower in our study period than in the modern U.S. Also, we start our working age category at 15 since most people in the 15-20 age group would have been working during the period we study.

Table 1: Estimated effects using linked data

DV: Std. change in log mortality rate from 1851-5 to 1861-5				
	All ages	**Under 15**	**Age 15-54**	**Over 54**
	(1)	(2)	(3)	(4)
Cotton district indicator	0.823*	0.370	0.642	0.858***
	(0.437)	(0.351)	(0.406)	(0.319)
Nearby (0-25km)	0.446***	0.254*	0.360***	0.416***
	(0.125)	(0.143)	(0.0907)	(0.128)
Nearby (25-50km)	0.384**	0.106	0.295**	0.340***
	(0.153)	(0.137)	(0.126)	(0.129)
Nearby (50-75km)	-0.00220	0.00550	0.0936	0.0122
	(0.153)	(0.153)	(0.129)	(0.158)
Log initial pop.	-0.0762	-0.0928*	-0.122**	-0.115**
	(0.0879)	(0.0510)	(0.0504)	(0.0566)
Age share controls	Yes			
Observations	537	534	535	533
R-squared	0.090	0.020	0.056	0.080
Linked deaths	152,787	55,723	69,978	27,086

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. Deaths are assigned to the district of initial residence (i.e., district of census enumeration). Regressions are weighted by district population in 1851. In the all-age regressions we include age group controls for the share of district residents under 5, the share 5-15, and the share over 50 for both the 1851 and 1861 census. Note that the number of observations varies because some small districts with no linked deaths in one period are dropped.

age groups. These exercises show that we obtain fairly similar results regardless of whether we weight regressions by district population, use a continuous treatment variable (district cotton employment share), drop outlier locations such as Manchester, Liverpool, or Leeds, or omit the foreign-born from our linked deaths sample. All of these results show evidence that over the 1861-65 period, mortality increased among the population initially resident in the cotton districts. In addition, we consistently find statistically significant evidence that mortality in nearby areas also increased during the downturn. In Appendix D.2 we consider age-group effects using smaller age bins. These estimates suggest that we begin observing statistically significant increases in mortality for those over age 45. We still see positive effects for younger age groups, but these are not statistically significant.

It is interesting to note that we see smaller adverse effects for children. Our review of evidence from contemporary reports, in Appendix B.2, suggests that one

factor that positively affected the health of young children during this period was that many working mothers lost their jobs and could spend more time on childcare, a finding also consistent with Miller & Urdinola (2010).[46]

Because the preceding analysis assigns deaths to the place of initial residence at the onset of the shock, it effectively tracks cohorts, and so provides us estimates of the real impact of the cotton shortage in the absence of migration bias. To directly assess the importance to these results of correctly accounting for migration—indeed, the sort of migration that typically goes unobserved in aggregate data—we next re-assign deaths to the location in which they occurred, rather than to individuals' initial, pre-shock location. This reconfiguration of the linked individual-level data brings it closer in structure to traditional aggregate data, wherein deaths are also effectively reported by location of occurrence.[47]

The results of this exercise are presented in Table 2. For comparability, this table follows the same format as Table 1. Re-assigning deaths to the location in which they occurred generates substantially different results: we no longer observe evidence that overall mortality increased in the cotton districts during the cotton shortage. Similarly, for children and working-age adults, this approach would lead us to conclude that mortality decreased during the cotton shortage, while for older adults we would conclude that mortality increased, but by much less than was indicated by the results in Table 1.[48]

In the bottom panel of Table 2 we test more formally for the impact of (traditionally unobserved) migration on our estimates. Specifically, we test whether the

[46]Female labor force participation was very high in Britain during this period, and particularly in the cotton districts, similar to the levels observed in the late 20th century.

[47]Note that at least nominally, deaths in typical aggregate datasets are reported by the location of residence. However, on a practical level, these are equivalent to the location of occurrence if an individual's death occurred in a location other than that in which they originally experienced the relevant change in local economic conditions.

[48]This simply tells us that the impact of the cotton shortage on mortality among older adults documented in Table 1 was strong enough to offset the bias created by migration away from the cotton textile districts, while this was not the case for children and adults.

Table 2: Linked results assigning mortality to district of occurrence

DV: Std. change in log mortality rate from 1851-5 to 1861-5				
	All ages	**Under 15**	**Age 15-54**	**Over 54**
	(1)	(2)	(3)	(4)
Cotton district indicator	0.00148	-0.266*	-0.180*	0.513***
	(0.131)	(0.160)	(0.101)	(0.123)
Nearby (0-25km)	0.247**	0.105	0.118	0.380***
	(0.104)	(0.116)	(0.114)	(0.112)
Nearby (25-50km)	0.100	-0.161	0.0697	0.393***
	(0.133)	(0.137)	(0.0969)	(0.136)
Nearby (50-75km)	-0.0366	0.237*	0.0275	-0.0226
	(0.203)	(0.134)	(0.199)	(0.143)
Log initial pop.	0.0252	-0.0347	-0.0768**	-0.0492
	(0.0519)	(0.0431)	(0.0360)	(0.0473)
Age controls	Yes			
Observations	538	531	534	529
R-squared	0.038	0.027	0.020	0.049
Linked deaths	152,787	55,723	69,978	27,086
Testing difference relative to results in Table 1				
Chi-sq. stat.	3.30	3.30	3.63	8.57
P-value	0.0693	0.0695	0.0569	0.0034

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. Deaths are assigned to the district of death. Regressions are weighted by district population in 1851. In the all-age regressions we include age group controls for the share of district residents under 5, the share 5-15, and the share over 50 for both the 1851 and 1861 census. Note that the number of observations varies because some small districts with no linked deaths in one period are dropped. At the bottom of the table we present tests comparing the cotton-district coefficients estimated in this table to those obtained when deaths are assigned to each person's initial place of residence, as in Table 1. This test is implemented by running seemingly unrelated regressions on non-standardized variables with robust standard errors. The sample used in the tests at the bottom of this table differs slightly from the sample used for the regression results because the test of coefficient equality uses regressions run only on the 531 districts where mortality rates are observed when assigning deaths by both census district and death district.

estimated coefficients here are statistically different from the coefficients in Table 1. In all cases we can reject at a 90% confidence level the equality between these estimates, wherein deaths accrue to the district of occurrence, and those in Table 1, in which they accrue to the initial district of residence. This indicates that without being able to identify and track the relevant at-risk population, we would not be able to recover the true effect of this recession on mortality.

The linked individual-level census data can also shed light on migration patterns. For example, of those resident in the cotton textile districts in 1851, 59.7% also died in one of the cotton districts. In 1861, this fell to 55.2%, indicating that individuals

initially resident in the cotton textile areas were more likely to die elsewhere during the cotton shortage than in the early 1850s. At the same time, the share of linked deaths in the nearby districts (within 50 km) accounted for by those initially resident in the cotton districts rose from 9.5% to 9.9% and the share of linked death in all other districts accounted for by those initially resident in the cotton districts rose from 2.5% to 3.3%.

The increased share of cotton-district residents dying in other locations is not simply a function of increased out-migration from these locations. Rather, there is evidence that while migration from the cotton districts increased, migration into the cotton districts also decreased. One way to see this is to look at the share of residents from other locations that died in the cotton districts. For example, the share of individuals initially resident in the nearby districts who died in the cotton textile districts fell from 11.0% in 1851-55 to 10.4% in 1861-65, which suggests that there were also fewer migrants flowing into the cotton areas and then dying.

Overall, the share of all linked deaths accounted for by individuals initially living in the cotton textile districts rose from 8.27% in the 1851-55 period to 8.98% in the period 1861-65. This represents an increase of 8.6% (or 0.7 percentage points). The share of linked deaths that actually occurred in the cotton districts also rose, but only by 1.8% (or 0.17 percentage points). Thus, people initially resident in the cotton textile districts made up a larger share of all deaths in 1861-65 than in 1851-55, while at the same time, the share of these deaths occurring outside of the cotton districts increased.

5.1.6 Additional evidence

Before moving on to direct comparisons of our linked results with those drawn from aggregate data, we use our linked data to shed some additional light on how different populations were impacted by the cotton shortage. In particular, the occupation

Table 3: Share of deaths accounted for by cotton households

	Cotton workers	Households with at least one cotton worker	Non-cotton workers living in a household with a cotton worker
1851	0.0085	0.0356	0.0271
1861	0.0149	0.0449	0.0300
Increase (p.p.)	0.00638	0.009304	0.002925
Percent increase	75.21%	26.16%	10.80%

information in the linked census data allow us to identify cotton workers and those in their household, and to look at mortality patterns in these groups. The share of deaths accounted for by cotton workers or those in their household are described in Table 3.[49] These results show that cotton workers accounted for 0.85% of the deaths in our linked sample in 1851-55, but this increased to 1.49% of deaths in the sample from 1861-65. This is a 75.2% increase in the share of all deaths accounted for by cotton textile workers, indicating that the cotton shortage had a substantial direct effect on mortality among cotton workers. In the second column, we see that deaths among all members of cotton households increased from 3.56% of all deaths to 4.49%. Thus, the share of overall deaths accounted for by cotton textile households increased by 26.2%. In the last column we can see that the share of deaths accounted for by those not listed as cotton workers but who lived in a household with at least one cotton worker increased from 2.7% of all deaths to 3.0%. Together, these results suggest that the cotton shortage period was characterized by a substantial increase in the share of deaths accounted for by cotton households. This effect appears both for cotton workers themselves, as well as for other members of cotton households.

To get a sense of how important these cotton-worker and cotton-household deaths were to the overall increase in mortality that we observed in the population initially resident in cotton textile districts, we can run regressions dropping these deaths from

[49]Note that we cannot analyze cotton worker deaths using the difference-in-difference framework from our main analysis since cotton workers and their families were heavily concentrated in the cotton textile districts.

our linked deaths sample. These results are presented in Table 4. As a point of comparison, Column 1 presents the results for the full linked sample, replicating our preferred specification from Table 1. In Column 2, we drop all cotton workers from the sample. We can see that this reduces the estimated impact of the cotton shock on mortality among residents of the cotton districts by about 0.1 s.d., or about 12%. In Column 3 we drop all members of cotton households, i.e., households with at least one cotton worker. Relative to Column 1, this reduces the estimated impact of the cotton shock by almost one-third. As expected, dropping cotton workers or cotton households has little impact on nearby districts, where there were relatively few cotton workers. Overall, the direct impact of the shock on mortality through cotton workers and their families appears to be important, but there is also evidence that the effects spilled over onto workers in other occupations in the cotton areas. Thus, our findings support both the notion that personal experiences of job loss are important to health (see, e.g., Sullivan & von Wachter (2009)), as well as the idea that local employment shocks can have a broader social impact on individuals whose income or labor force status is not personally affected (see, e.g., Ruhm (2007); Stevens *et al.* (2015)).

5.2 Mortality analysis using aggregate data

Our analysis of linked microdata make it clear that the cotton shortage had deleterious effects on the health not only of those initially resident in cotton districts, but also on the health of those residing nearby, who were indirectly affected by the downturn through migration. Next, we ask whether the more standard aggregate- and panel-data approaches commonly used in this literature can recover the true effect of the cotton shortage on mortality. Our main approach follows the methodology applied to the microdata, comparing total deaths in 1861-65 to deaths in 1851-55, but uses data taken directly from aggregate district death counts. This allows an apples-to-apples comparison between results obtained using the linked data and those generated from aggregate data. In Appendix D.4, we also consider results obtained using aggregate

Table 4: Estimated impact of cotton households on our main results

DV: Std. change in log mortality rate from 1851-5 to 1861-5			
	Full linked sample	Dropping cotton workers	Dropping all members of cotton households
	(1)	(2)	(3)
Cotton district indicator	0.823*	0.726	0.586
	(0.437)	(0.443)	(0.591)
Nearby (0-25km)	0.446***	0.436***	0.425***
	(0.125)	(0.127)	(0.160)
Nearby (25-50km)	0.384**	0.380**	0.424***
	(0.153)	(0.154)	(0.141)
Nearby (50-75km)	-0.00220	-0.00283	-0.0130
	(0.153)	(0.155)	(0.161)
Log initial pop.	-0.0762	-0.0813	-0.115
	(0.0879)	(0.0887)	(0.0949)
Age share controls	Yes	Yes	Yes
Observations	537	537	537
R-squared	0.090	0.079	0.058

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. Deaths are assigned to the initial district of residence (i.e., census enumeration district). Regressions are weighted by district population in 1851. Age group controls include the share of district residents under 5, the share 5-15, and the share over 50 for both the 1851 and 1861 census.

annual panel data and a methodology closer to the standard panel-data approach from Ruhm (2000).

The aggregate mortality data used in this analysis are taken from reports produced annually by the Registrar General's office. Further details on the data are available in Appendix C. These data cover the same districts used in the linked data analysis. As in the linked analysis, population data come from the census and are available every ten years starting in 1851.

5.2.1 Results

As a first step, it is useful to compare the results obtained from the aggregate data to the results we obtain from analogous linked data (i.e., linked data wherein deaths are assigned to the location in which the death occurred). Because in the aggregate data, deaths are assigned to the location in which they were registered, we should expect these results to be similar to the linked results in Table 2. This comparison, therefore, provides a check on how representative our linked deaths data are of all deaths.

The results of this comparison for each age group are described in Table 5. For each age group, we present results obtained from both aggregate data and linked data in which deaths are assigned to the location in which they were registered. At the bottom of the table, we report tests looking at whether the estimates using the two alternative data types are statistically distinguishable. Our results show that for adults aged 15-54 and those 55 and over, the effects estimated using the place-of-death linked data are almost identical to those estimates obtained from the traditional aggregate data. This tells us that, for these populations, the linked data sample appears to almost perfectly recover the results obtained from aggregate data, i.e., the linked deaths appear to be representative of aggregate deaths. For the population under age 15, we can see that the aggregate and linked data give qualitatively similar

results, but the magnitude of the effects are different. This difference is clearly due to the fact that infant deaths are under-represented in our linked sample. Thus, when comparing the linked results to the aggregate results, our main focus should be on the older populations, where our linked sample is representative.

Table 5: Comparing aggregate-data results to linked-data results using district of death

DV: Std. change in log mortality rate from 1851-5 to 1861-5						
	Ages under 15		**Ages 15 to 54**		**Ages 55 and up**	
	Agg. data	Linked data	Agg. data	Linked data	Agg. data	Linked data
	(1)	(2)	(3)	(4)	(5)	(6)
Cotton dist. ind.	-0.604***	-0.266*	-0.231	-0.180*	0.508***	0.513***
	(0.136)	(0.160)	(0.179)	(0.101)	(0.141)	(0.123)
Nearby dist.	Yes	Yes	Yes	Yes	Yes	Yes
Observations	531	531	534	534	529	529
R-squared	0.089	0.027	0.120	0.020	0.133	0.049
Testing difference between linked and aggregate cotton dist. coef.						
Chi-sq stat.	2.46		0.10		0.00	
P-value	0.1166		0.7547		0.9728	

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. Linked results come from data in which deaths are assigned to the district of death. Regressions are weighted by district population in 1851. All regressions include controls for nearby districts 0-25 km, 25-50 km and 50-75 km away from the cotton districts. The sample size changes slightly across specifications if there are small districts where we do not have links when cutting by age group. At the bottom of the table we present tests comparing the cotton district coefficients estimated in this table to those obtained in linked data when deaths are assigned to each person's place of death as in Table 2. This test is implemented by running seemingly unrelated regressions with robust standard errors.

Having established that the results from aggregate data are essentially equivalent to those from the version of the linked data that intentionally fails to account for migration, in Table 6 we compare the results obtained from the aggregate data to estimates from our preferred linked approach, which avoids migration bias by assigning deaths to each person's district of residence at the beginning of each period. This provides our most direct test of the ability of results obtained from aggregate data to correctly capture the impact of the cotton shortage on the relevant population at risk of exposure. While we include all three age groups in the table, attention should be focused on the groups over age 15, where we know from Table 5 that our linked sample is representative.

For adults aged 15-54, the evidence in Table 6 leads us to reject the hypothesis that the aggregate data deliver results that are the same as those obtained from our preferred linked approach. The fact that this is true for the adults, where, when deaths are assigned by death district, the linked sample can almost perfectly replicate results obtained from the aggregate data (Table 5 Columns 3 and 4), is particularly telling. This means that the difference is due to migration rather than to differences between the linked sample and the aggregate data. Thus, for adults it is clear that (1) at least using this approach, estimates obtained from aggregate data cannot capture the real impact of the cotton shock on population health, and (2) the source of this discrepancy is migration. For the elderly, the analysis based on aggregate data comes closer to the result obtained with our linked data, and the estimates cannot be statistically distinguished. This indicates that migration bias is less likely to be a problem for older age groups, a pattern that is reasonable given that we would expect the elderly to be less mobile in response to changes in local economic conditions.

As a final check, we also generate results using annual panel data on aggregate mortality. While less comparable to the approach used above in the linked data analysis, this panel-data approach is closer to that used in most of the literature. These results, presented in Appendix D.4, show that, across a variety of specifications, results obtained using aggregate data typically fail to recover what we know from our preferred linked data to be the true impact of the cotton shock on mortality. Notably, while our aggregate data's largely procyclical results are consistent with most of the previous studies using aggregate panel data (as well as with the results presented above for aggregate and place-of-death linked data, where we know that migration is the source of the apparent bias), they are contradicted by the preferred results that we obtained using place-of-origin linked data. Thus, it seems unlikely that an analysis based on aggregate panel data, and following the standard approach used in this literature, will be able to correctly and reliably recover the impact of this downturn on mortality.

Table 6: Comparing aggregate data results to our preferred linked data estimates

DV: Std. change in log mortality rate from 1851-5 to 1861-5						
	Ages under 15		**Ages 15 to 54**		**Ages 55 and up**	
	Agg. data	Linked data	Agg. data	Linked data	Agg. data	Linked data
	(1)	(2)	(3)	(4)	(5)	(6)
Cotton district ind.	-0.605***	0.370	-0.231	0.642	0.510***	0.858***
	(0.136)	(0.351)	(0.178)	(0.406)	(0.138)	(0.319)
Nearby dist.	Yes	Yes	Yes	Yes	Yes	Yes
Observations	534	534	535	535	533	533
R-squared	0.087	0.020	0.120	0.056	0.137	0.080
Testing difference between linked and aggregate cotton dist. coef.						
Chi-sq stat.	7.45		4.20		0.85	
P-value	0.0064		0.0404		0.3559	

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. Linked results come from data in which deaths are assigned to the individual's district at the beginning of the period (i.e., census enumeration district). Regressions are weighted by district population in 1851. All regressions include controls for nearby districts 0-25 km, 25-50 km and 50-75 km away from the cotton districts. The sample size changes slightly across specifications if there are small districts where we do not have links when cutting by age group. At the bottom of the table we present tests comparing the cotton district coefficients estimated in this table to those obtained when deaths are assigned to each person's initial place of residence as in Table 1. This test is implemented by running seemingly unrelated regressions with robust standard errors. Note that each pair of comparisons uses only districts for which death rates from the linked data are available, which causes the aggregate results to differ slightly from those shown in Table 5.

6 Conclusion

There is reason to believe that, in at least some settings, migration may be an important part of how people deal with local economic shocks (see, e.g., Blanchard & Katz (1992)). Drawing on a historical setting that allows us to use large-scale linked individual-level census and death records, we show that this type of migration response can substantially undermine estimates of the impact of economic fluctuations on mortality. In our setting, we find that an analysis based on aggregated data in a fixed-effects framework is unlikely to be able to accurately and reliably recover the impact of the recession on mortality.

A large body of existing work uses aggregate data and panel data methods to study the relationship between business cycles and mortality in modern settings. The vast majority of these studies find that mortality is procyclical. We do not interpret

our results as implying that these findings are incorrect. However, our results do suggest that the methodology behind these results is vulnerable to migration bias if the migration response is sufficiently strong, and if it is not perfectly reflected in interpolated population values. Together with the evidence of contemporary cross-labor market migration responses (see, e.g., Blanchard & Katz (1992); Foote *et al.* (2017)), the evidence on population data quality presented in Appendix A.2, and the dissonance between modern microdata and aggregate-data estimates (see, e.g., Halliday (2014); Gerdtham & Johannesson (2005)), it seems plausible that unobserved migration may pose concerns not just in our setting, but also in a variety of others.

The solution to dealing with migration bias is to track individuals over time, so that when people migrate away from areas experiencing severe recessions and then die elsewhere, we can tie their deaths back to their recession experience. Currently there are only a small number of studies that use individual-level data to assess the impact of economic fluctuations on mortality. The fact that these studies often contradict the procyclical findings obtained from aggregate data suggests that migration bias may indeed be important in modern settings. However, existing microdata studies are limited by small sample sizes and the fact that, without observing exogenous sources of local economic shocks, it is difficult to directly study migration responses. This calls for further work combining larger samples of individual-level panel data with plausibly exogenous local economic shocks.

References

Abramitzky, R, Boustan, LP, & Eriksson, K. 2012. Europes tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration. *American Economic Review*, **102**(5), 1832–1856.

Abramitzky, R, Boustan, LP, & Eriksson, K. 2014. A nation of immigrants: Assimilation and economic outcomes in the age of mass migration. *Journal of Political Economy*, **122**(3), 467–506.

Adda, Jérôme, Banks, James, & von Gaudecker, Hans-Martin. 2009. The Impact of Income Shocks on Health: Evidence from Cohort Data. *Journal of the European Economic Association*, **7**(6), 1361–1399.

Arnold, Arthur. 1864. *The History of The Cotton Famine: From the Fall of Sumter to the Passing of The Public Works Act*. London: Saunders, Otley, and Co.

Bailey, M, Cole, C, Henderson, M, & Massey, C. 2017. *How Well Do Automated Linking Methods Perform in Historical Samples? Evidence from New Ground Truth*. NBER Working Paper No. 24019.

Blanchard, Olivier Jean, & Katz, Lawrence F. 1992. Regional Evolutions. *Brookings Papers on Economic Activity*, 1 – 61.

Borjas, G. J. 2014. *Immigration Economics*. Cambridge, MA: Harvard University Press.

Borjas, George J. 2003. The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The quarterly journal of economics*, **118**(4), 1335–1374.

Bound, John, & Holzer, Harry J. 2000. Demand Shifts, Population Adjustments, and Labor Market Outcomes during the 1980s. *Journal of Labor Economics*, **18**(1), pp. 20–54.

Boyer, George R. 1997. Poor Relief, Informal Assistance, and Short Time during the Lancashire Cotton Famine. *Explorations in Economic History*, **34**(1), 56 – 76.

Brenner, M. H. 1979. Mortality and the National Economy: A Review and the experience of England & Wales, 1936-76. *Lancet*, **314**, 568–573.

Chay, Kenneth Y., & Greenstone, Michael. 2003. The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession. *The Quarterly Journal of Economics*, **118**(3), pp. 1121–1167.

Crafts, Nicholas, & Wolf, Nikolaus. 2014. The Location of the UK Cotton Textiles Industry in 1838: a Quantitative Analysis. *Journal of Economic History*, **74**(4), 1103–1139.

Cutler, D., Huang, W., & Lleras-Muney, A. 2016 (September). *Economic Conditions and Mortality: Evidence from 200 Years of Data*. NBER Working Paper No. 22690.

Dehejia, Rajeev, & Lleras-Muney, Adriana. 2004. Booms, Busts, and Babies' Health. *The Quarterly Journal of Economics*, **119**(3), 1091–1130.

Economou, Athina, Nikolaou, Agelike, & Theodossiou, Ioannis. 2008. Are recessions harmful to health after all?: Evidence from the European Union. *Journal of Economic Studies*, **35**(5), 368–384.

Edwards, Ryan. 2008. Who is hurt by procyclical mortality? *Social Science & Medicine*, **67**(12), 2051–2058.

Eliason, Marcus, & Storrie, Donald. 2009. Does Job Loss Shorten Life? *Journal of Human Resources*, **44**(2), 277–302.

Ellison, T. 1886. *The Cotton Trade of Great Britain.* London: Effingham Wilson, Royal Exchange.

Feigenbaum, J. 2015. *Intergenerational Mobility during the Great Depression.* Mimeo.

Feigenbaum, J. 2016. *A Machine Learning Approach to Census Record Linking.* Mimeo.

Ferrie, JP. 1996. A New Sample of American Males Linked From the 1850 Public Use Micro Sample to the Manuscript Schedules of the 1860 Federal Census of Population. *Historical Methods*, **16**.

Fishback, P. V., Horrace, W.C., & Kantor, S. 2006. The impact of New Deal expenditures on mobility during the Great Depression. *Explorations in Economic History*, **43**, 179–222.

Foote, Andrew, Grosz, Michel, & Stevens, Ann. 2017 (March). *Locate Your Nearest Exit: Mass Layoffs and Local Labor Market Response.*

Forwood, WB. 1870. The Influence of Price upon the Cultivation and Consumption of Cotton During the Ten Years 1860-70. *Journal of the Statistical Society of London*, **33**(3), 366–383.

Gerdtham, Ulf-G, & Johannesson, Magnus. 2005. Business cycles and mortality: results from Swedish microdata. *Social Science & Medicine*, **60**(1), 205–218.

Griffith, Rachel, O'Connell, Martin, & Smith, Kate. 2013. *Food Expenditure and Nutritional Quality over the Great Recession.* Tech. rept.

Halliday, Timothy J. 2007. Business cycles, migration and health. *Social Science & Medicine*, **64**(7), 1420–1424.

Halliday, Timothy J. 2014. Unemployment and Mortality: Evidence from the PSID. *Social Science & Medicine*, **113**, 15–22.

Hanlon, W. Walker. 2015. Necessity is the Mother of Invention: Input Supplies and Directed Technical Change. *Econometrica*, **83**(1), 67–100.

Hanlon, W. Walker. 2017. Temporary Shocks and Persistent Effects in the Urban System: Evidence from British Cities after the U.S. Civil War. *Review of Economics and Statistics*, **99**(1), 67–79.

Henderson, W.O. 1969. *The Lancashire Cotton Famine 1861-1865.* New York: Augustus M. Kelley Publishers.

Hund, Lauren Brooke. 2012. *Survey Designs and Spatio-Temporal Methods for Disease Surveillance*. Ph.D. thesis.

Kiesling, L.Lynne. 1996. Institutional Choice Matters: The Poor Law and Implicit Labor Contracts in Victorian Lancashire. *Explorations in Economic History*, **33**(1), 65 – 85.

Lindo, Jason M. 2015. Aggregation and the estimated effects of economic conditions on health. *Journal of Health Economics*, **40**, 83–96.

Long, Jason, & Ferrie, Joseph. 2013. Intergenerational Occupational Mobility in Britain and the US since 1850. *American Economic Review*, **103**(4), 1109–1137.

Miller, Douglas L., Page, Marianne E., Stevens, Ann Huff, & Filipski, Mateusz. 2009. Why Are Recessions Good for Your Health? *The American Economic Review, Papers & Proceedings*, **99**(2), pp. 122–127.

Miller, Grant, & Urdinola, B Piedad. 2010. Cyclicality, Mortality, and the Value of Time: The Case of Coffee Price Fluctuations and Child Survival in Colombia. *The Jurnal of Political Economy*, **118**(1), 113.

Mitchell, Brian R. 1988. *British Historical Statistics*. Cambridge, UK: Cambridge University Press.

Mitchell, Brian R, & Deane, Phyllis. 1962. *Abstract of British Historical Statistics*. London: Cambridge University Press.

Monras, J. 2015 (November). *Immigration and Wage Dynamics: Evidence from the Mexican Peso Crises*. Working paper.

Monras, J. 2018 (February). *Economic Shocks and Internal Migration*. Working paper.

Muller, Andreas. 1989. Business Recession, Alcohol Consumption, Drinking and Driving Laws: Impact on Oklahoma Motor Vehicle Fatalities and Fatal Crashes. *American Journal of Public Health*, **79**(10), 1366?1370.

Olivetti, C. 2013 (June). *Human Capital in History: The American Record*. NBER Working Paper No. 19131.

Painter, Gary. 2010. What Happens to Household Formation in a Recession? *46th Annual AREUEA Conference Paper*.

Phipps, Amanda I., Clarke, Christina A., & Ereman, Rochelle R. 2005. Impact of Intercensal Population Projections and Error of Closure on Breast Cancer Surveillance: Examples from 10 California Counties. *Breast Cancer Research*, **7**(5), R655 – R660.

Robinson, Gregory, & West, Kirsten. 2005. *Understanding Factors that Contributed to the Large Error of Closure in Census 2000*.

Rossi, Alice S. 1965. Naming children in middle-class families. *American sociological review*, 499–513.

Ruhm, Christopher J. 2000. Are Recessions Good for Your Health? *The Quarterly Journal of*

Economics, **115**(2), 617–650.

Ruhm, Christopher J. 2003. Good times make you sick. *Journal of Health Economics*, **22**(4), 637–658.

Ruhm, Christopher J. 2005. Healthy living in hard times. *Journal of Health Economics*, **24**(2), 341–363.

Ruhm, Christopher J. 2007. A Healthy Economy Can Break Your Heart. *Demography*, **44**(4), pp. 829–848.

Ruhm, Christopher J. 2015. Recessions, healthy no more? *Journal of Health Economics*, **42**, 17–28.

Ruhm, Christopher J, & Black, William E. 2002. Does Drinking Really Decrease in Bad Times? *Journal of Health Economics*, **21**(4), 659–678.

Southall, Humphrey R, Gilbert, David R, & Gregory, Ian. 1998 (Jan.). *Great Britain Historical Database : Labour Markets Database, Poor Law Statistics, 1859-1939.* [computer file]. UK Data Archive [distributor] SN: 3713.

Stevens, Ann H., Miller, Douglas L., Page, Marianne E., & Filipski, Mateusz. 2015. The Best of Times, the Worst of Times: Understanding Pro-cyclical Mortality. *American Economic Journal: Economic Policy*, **7**(4), 279–311.

Stuckler, David, Meissner, Christopher, Fishback, Price, Basu, Sanjay, & McKee, Martin. 2012. Banking Crises and Mortality During the Great Depression: Evidence from US Urban Populations, 1929–1937. *Journal of Epidemiology and Community Health*, **66**(5), 410–419.

Sullivan, Daniel, & von Wachter, Till. 2009. Job Displacement and Mortality: An Analysis Using Administrative Data. *The Quarterly Journal of Economics*, **124**(3), 1265–1306.

Svensson, M. 2007. Do Not Go Breaking Your Heart: Do Economic Upturns Really Increase Heart Attack Mortality? *Social Science & Medicine*, **65**, 833–841.

Wang, Huixia, Wang, Chenggang, & Halliday, Timothy J. 2018. Health and Health Inequality during the Great Recession: Evidence from the PSID. *Economics & Human Biology*, **29**, 17–30.

Watts, John. 1866. *The Facts of the Cotton Famine.* London: Simpkin, Marshall, & Co.

Woods, R. 1997 (March). *Causes of Death in England and Wales, 1851-60 to 1891-1900 : The Decennial Supplements.* [computer file].

Woods, Robert. 2000. *The Demography of Victorian England and Wales.* Cambridge, UK: Cambridge University Press.

A Appendix: Literature, concepts and methods

A.1 Review of selected related literature

Below, we provide a review of select leading studies on the relationship between recessions and public health. In particular, we highlight the methodological approaches used, the main findings, and the setting in which these results are found.

	Study	Data	Dependent variable	Specification	Standard errors	Result
1	Ruhm (2000) QJE	50 states, 20 years	Ln(mortality rate), Ln(mortality)	Fixed effects	Robust, weighted by population	Procyclical mortality
2	Ruhm & Black (2002) J. Health Ec.	13 years, 15-45 states (repeated cross-sections of individual-level data)	Alcohol use	Linear probability mode with state fixed effects and time trends	Clustered by state-month	Procyclical alcohol use
3	Ruhm (2003) J. Health Ec.	20 states (31 MSAs), 10 years (individual-level data)	Various health indicators	Linear probability model with state FEs	Clustered by state	Countercyclical health
4	Chay & Greenstone (2003) QJE	3 years, 1200 counties	Infant mortality rate	Fixed effects at the county level with state time trends	Robust, weighted by births	Recessions reduce mortality
5	Dehijia & Lleras-Muney (2004) QJE	Individual data, state level explanatory variables, 50 states, 25 years	Mothers characteristics, infant health indicators, prenatal care	Fixed effects at state level, with some state time trends	Clustered at state level, weighted and unweighted	Improved infant health during recessions
6	Neumayer (2004) Soc. Sci. & Medicine	20 years, 11-16 German states	Ln(mortality rate), mortality by cause	Fixed effects at state level with lagged dependent variable (Arellano-Bond)	Robust, weighted by state population	Procyclical mortality
7	Ruhm (2005) J. Health Ec.	34-45 states, 14 years (repeated cross-sections of individual-level data)	Smoking, overweight	Probit regressions	Robust, with correlation within state-month or by state	Smoking and obesity are procyclical
8	Gerdtham & Johannesson (2005) Soc. Sci. Med.	Individual-level panel data, 10-16 years	Prob. of death	Probit model, individual level with time-series explanatory variable	Robust, clustered by individual	Mortality risk countercyclical for men, unclear for women
9	Tapia Granados (2005) European J. of Pop.	18 years, 50 provinces (Spain)	Ln(mortality rate)	Fixed effects with some province time trends	Weighted by population	Procyclical mortality
10	Gerdtham & Ruhm (2006) Ec. and Human Bio.	23 OECD countries, 37 years	Ln(mortality rate)	Fixed effects at country level	Robust and AR1, weighted by country pop.	Procyclical mortality
11	Svensson (2007) Soc. Sci. & Medicine	21 Swedish regions, 17 years	Heart disease	Fixed effects	Robust	Mixed results
12	Ruhm (2007) Demography	50 states +DC, 20 years	Coronary heart death rates, all heart-related death rates	Fixed effects	Robust, AR1, weighted by population	Recessions decrease coronary mortality
13	Fishback et al. (2007) Review of Ec. and Stat.	114 U.S. cities, 1929-1940	Infant mortality rate, overall death rate	Fixed effects	Robust	Procyclical mortality

14	Edwards (2008) Soc. Sci. & Medicine	Individual level data panel data by state & year	Mortality rate	Logit regressions		Procyclical mortality
15	Economou et al. (2008) J. of Economic Studies	13 EU countries, 20 years 1977-1996	Mortality rate	Fixed effects	Robust	Countercyclical mortality
16	Miller et al (2009) AER P&P	50 states + DC, 1978-2004	Ln(mortality rate) by group	Fixed effects Poisson at state level, some time trends	Clustered by state, weighted by state population	Procyclical mortality
17	Lin (2009) Applied Econ.	8 Asia-Pacific countries, 1976-2003	Ln(mortality rate)	Fixed effects, with some country time trends	Robust, weighted by population	Procyclical mortality
18	Stuckler, et al. (2009) Lancet	26 EU countries, 1970-2007	Mortality rate by cause	Fixed effects in differences	Clustered by country	Mixed
19	Gonzalez & Quast (2011) Empir. Econ.	32 Mexican states, 1993-2004	Ln(mortality rate)	Fixed effects	Clustered by state	Procyclical mortality
20	Stuckler et al (2012) J. of Epid. & Community Health	114 cities in 36 states, 9 years	Ln(mortality rate)	Fixed effect and distributed lag	Clustered by state	Procyclical mortality
21	Ariizumi & Schirle (2012)	10 Canadian provinces, 33 years 1977-2009	Ln(mortality rate)	Fixed effects, with provincen time trends	Clustered by province or bootstrapped	Procyclical mortality
22	McInerney & Mellor (2012), J. Health Econ.	50 US states from 1976-2008, Individual-level data repeated cross-sections from 1994-2008	Ln(mortality rate) and other senior health indicators	Fixed effects with location time trends	Unclear	Countercyclical health among seniors in recent decades
23	Tekin et al. (2013) NBER Working Paper No. 19234	Repeated cross-sections of individual-level data, 2005-2005-2011	Variety of health indicators (reported health, smoking, etc.)	Fixed effects at the state level with some state time trends	Clustered by state-month	Zero recession-mortality relationship
24	Ruhm (2015) J. Health Ec.	50 states, 35 years 1976-2010 (using different time windows)	Ln(mortality rate)	Fixed effects, with some state time trends	Clustered by state	Mortality becoming less procyclical recently
25	Stevens et al (2015) AEJ: Policy	50 US states, 1978-2006	Ln(mortality rate) by group	Fixed effects with location time trends	Clustered by state	Procyclical mortality
26	Ruhm (2015) NBER Working Paper No. 21604	50 US states or 3,142 US counties, 1976-2013	Ln(mortality rate)	Fixed effects with location time trends	Clustered by state	Procyclical mortality

A.2 Errors of closure and Das Gupta interpolation in modern data

The results presented in this study suggest that migration bias has an important impact on the estimated business cycle-mortality relationship. Should we be concerned that migration may cause similar issues in other contexts? The answer to this question depends in large part on the accuracy of intercensal population estimates.

The standard interpolation approach, termed the Das Gupta method, calculates

annual estimates of population using a base and terminal year census count combined with annual information on births, deaths, and net internal and international migration.[50] Specifically, for each intercensal year, they calculate a cumulative population value that adds to observed population in a base census year all the net change in population (up to that year) due to natural increase (i.e., births minus deaths) and migration (i.e., internal and international migrant inflows minus internal and international migrant outflows). Working in this manner until they reach the next census year, they obtain what is called a "postcensal estimate" for the entire 10-year period. This represents the population we would expect to find in the terminal census, based on the observed initial census count and all known entries and exits from the population. The difference between this postcensal estimate and the terminal year's observed census count is termed the *error of closure*, and is equivalent to -1 times the residual implied net migration that has gone unobserved in the intercensal period. Thus, as a final step, intercensal estimates are obtained by taking this error of closure and then distributing it across intercensal years geometrically, thus "topping up" the year-specific value obtained on the way to the postcensal estimate (see *Methodology for the Intercensal Population and Housing Unit Estimates: 2000 to 2010,* U.S. Census Bureau (2012) for further details of the procedure).

The error of closure can be used as a diagnostic of data quality. Specifically, the error of closure can be interpreted as the (negative of) implied net migration that has gone unobserved: the higher the magnitude of the error of closure, the less well observed is migration.[51] Furthermore, the larger the error of closure, the more important are assumptions about how to distribute growth across intercensal years.

To get a sense of the quality of best-available intercensal population estimates and

[50]To compute the latter two measures, census officials rely on a wealth of information, such as Social Security and Medicare records and records of troops stationed abroad.

[51]This is, of course, not a perfect measure of the potential for migration bias. The measure only reflects unobserved *net* migration, and so it would miss migration that occurred in one part of the census period but was revered before the next census. It will also miss selective migration. Furthermore, other factors may contribute to this residual, including unreported births and deaths, as well as census under-enumeration.

the likelihood of bias related to unobserved migration, we now compare the errors in the intercensal population estimates available in our settings to the errors found in several modern developed and developing countries. In particular, we calculate the error of closure at several levels of geographic aggregation for three wealthy countries (England and Wales, the United States, and Canada), and for one developing country (India). We compare these to similar figures obtained from our historical data for England and Wales. These errors of closure, reported in absolute value as a percentage of final-year population, are presented in Table 7.

Here, we find evidence that migration is imperfectly observed—and so, may be imperfectly reflected in intercensal population—even in modern rich-country settings. Panel A shows that migration-adjusted errors of closure at the national level are non-negligible, ranging from 0.10% to 2.42% in magnitude. Here, our historical estimates for England & Wales (which, crucially, cannot adjust for migration), compare favorably. For contrast, modern errors of closure from India are several times higher, ranging from 4.61-9.94%.

Naturally, errors of closure are likely to be larger at lower levels of spatial aggregation, where internal migration is more likely to go unobserved. In Column 5 of Panel B, we present estimates for modern England and Wales at the Local Authority District (LAD) level. Using best-available methods (i.e., those that can adjust for observed migration), LAD-level errors of closure are on average roughly one quarter the size of those we obtain (without adjusting for observed migration) from historical England and Wales at the Registration District (RD) level.[52] Comparing measures on a more consistent basis (that is, where neither error of closure adjusts for migration) reduces this accuracy gap substantially: average historical errors of closure at the RD level are only half that for modern LADs. Indeed, comparing migration-adjusted to

[52]Errors of closure look similar for other small, sub-provincial units in modern rich countries. See, e.g., errors of closure reported for census metropolitan areas and economic regions of Canada, which in 2011 were as high in magnitude as 3.0% and 4.4%, respectively based on data from Statistics Canada.

non-migration-adjusted errors of closure for those modern developed-country units for which it is available, it is clear that even while partially adjusting for migration, best-available modern errors of closure still leave a substantial portion of migration unaccounted for.

Furthermore, there is considerable cross-unit heterogeneity in the error of closure: even adjusting for migration, 65% of modern English and Welsh LADs have errors of closure above 1.00%. Indeed, for the City of London, the rate is 52.71%. More importantly from an estimation standpoint, it should be noted that although this table presents the *absolute value* of the error of closure, inaccuracy is liable to be compounded where errors of closure move in opposite directions in recession-stricken and migrant-receiving spatial units—say, because of unobserved migration from one unit to the other.

What influences the size of the error of closure? Stratifying further, and using the U.S. over 1990-2000 as an example, we find that at its smallest, the error of closure is 0.37%, among the relatively immobile elderly (age 85+). Meanwhile, the modern U.S. error of closure is understandably largest in the lower end of the prime working-age population (25-34 years old), where it is -5.92% (see *Intercensal National Methodology*, U.S. Census Bureau, 2010). This pattern is consistent with labor mobility and other age-related factors that may make this age group particularly hard to track as they move, given the sort of tax and Medicare data the U.S. Census Bureau uses to make these migration adjustments. Indeed, the 1990-2000 census period in the U.S. has been widely acknowledged as having an unusually large error of closure, and this large error in turn has been attributed to precisely those issues we highlight as concerns in this analysis: the mis-measurement of highly mobile populations—in the case of the U.S. in the 1990s, the growing Hispanic population (Robinson & West, 2005).[53] Recent studies in the medical and biostatistics literatures have also raised

[53]Further undermining the accuracy of such analyses is the fact that those populations that may be especially vulnerable to mis-measurement (for reasons including but not limited to their higher geographic mobility) may also face different mortality risk profiles and access to medical and social

the issue of denominator measurement error, highlighting in particular the issue of measurement error in small demographic strata, such as those by age and race (see Phipps *et al.* (2005) for a discussion of intercensal interpolation-driven bias in estimates of breast cancer incidence rates, and Hund (2012, Ch. 4-5) for a broader methodological discussion of related denominator issues).[54]

Together, these findings suggest that the bias resulting from poor intercensal population denominators is likely to be a significant factor in a variety of settings, even beyond the obvious historical and developing-country ones. From a practical perspective, as these error-of-closure comparisons imply, even modern rich-country studies should be concerned about denominator-interpolation problems that could arise when studying granular, sparsely populated,[55] or rapidly growing geographic regions; small strata (for instance, by age, gender, race, and the interactions of these, especially where the sub-population in question may also experience mortality risk that is different from that of the general population); and populations that are vulnerable, mobile, and/or prone to under-registration (e.g. in the modern U.S., Hispanics and immigrant populations and young working-age people more generally). What's more, in the presence of migration spillovers, this bias will be further magnified when using standard panel data approaches to estimation. Finally, it should be noted that errors of closure only reflect the potential mis-measurement of the size of the at-risk

welfare services. Note also that these errors of closure reflect nothing of migrant selectivity—indeed, one may mis-measure the composition of the population even where the error of closure appears to be small in magnitude.

[54]For instance, Phipps *et al.* (2005) suggest that "areas with a high growth rate, a large population of retirees, or a large population of foreign-born individuals are likely to be underestimated, while areas with high poverty, and areas with a negative growth rate are likely to be overestimated." Using alternative denominators, they find "the DOF-based [breast cancer incidence] rates for Marin County were approximately 22% higher than census-based rates based on the same numerators." It should be noted that this 22% bias does not include the additional bias that could occur in the presence of spillovers, which in a differences-in-differences framework could compound denominator error if undercounts in one area lead to overcounts in the region used as a comparator.

[55]In sparsely populated or difficult-to-monitor regions, errors can be even higher: for instance, average absolute errors of closure for sub-provincial units are highest in rural provinces/territories, such as the Northwest Territories, whose census divisions had an average absolute error of closure of 3.5% in 2011 (data from Statistics Canada).

population—they say nothing regarding the composition of the population, or potential unobserved changes therein. Since selective sorting can also contributes to bias, it is possible that even the errors of closure shown here underestimate the potential for migration bias in modern settings.

Table 7: Absolute error of closure (%)

Panel A: National

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Country	E&W	E&W	E&W	E&W	US	US	Canada	Canada	Canada	India	India
Period	1851-1861	1861-1871	1991-2001	2001-2011	1990-2000	2000-2010	1996-2001	2001-2006	2006-2011	1991-2001	2001-2011
EOC (Natural increase only)	0.61	0.25	2.21	4.41		3.29				9.94	4.61
EOC (Migration-adjusted)			0.92	0.40	2.42	0.16	0.20	0.10	0.50		

Panel B: Sub-national

Country	E&W	E&W	E&W	E&W	E&W	E&W	US	Canada	Canada	Canada	India
Geo Unit	RD	County	RD	County	LAD	County	State + DC	Province	Province	Province	State
Period	1851-1861	1851-1861	1861-1871	1861-1871	2001-2011	2001-2011	2000-2010	1996-2001	2001-2006	2006-2011	2001-2011
Mean EOC (Natural increase only)	10.68	7.23	10.44	6.56	4.97	4.93	4.26				6.31
Min	0.02	0.32	0.01	0.20	0.01	0.82	0.16				0.59
Max	31.91	18.68	68.44	19.55	16.02	8.85	19.28				21.73
Share $< 1.00\%$	0.03	0.13	0.05	0.13	0.10	0.02	0.20				0.11
Share $> 5.00\%$	0.78	0.69	0.76	0.60	0.44	0.49	0.33				0.46
Mean EOC (Migration-adjusted)					2.26	1.10	0.96	0.73	0.95	0.97	
Min					0.01	0.02	0.00	0.00	0.00	0.00	
Max					52.71	2.69	4.80	2.20	3.20	2.10	
Share $< 1.00\%$					0.35	0.53	0.73	0.54	0.62	0.54	
Share $> 5.00\%$					0.10	0.00	0.00	0.00	0.00	0.00	
Observations	539	55	539	55	347	55	51	13	13	13	35

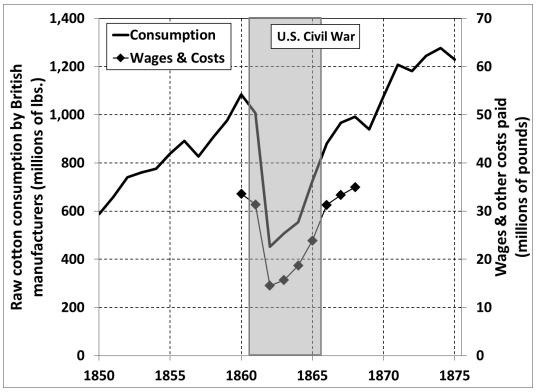
We report the absolute value (magnitude) of the error of closure in the total population, calculated as a percentage of the final-year census count of population. Figures are given for England & Wales, the United States, Canada, and India, as well as for smaller geographic units within these countries (i.e. Registration Districts, Local Authority Districts, counties, and states/provinces). "EOC (natural increase only)" refers to the error of closure based on accounting for births and deaths only. "EOC (migration-adjusted)" refers to the error of closure based on accounting for births, deaths, internal and international migration, and other similar adjustments. With the exception of the figures for all of Canada and for the US in 1990-2000, which are taken directly from data provided by Statistics Canada and the U.S. Census Bureau, respectively, the table reports the authors' calculations based on data from the U.S. Census, the U.K. Office for National Statistics, and the Office of the Registrar General & Census Commissioner of India.

B Appendix: Empirical setting

B.1 Additional background on the cotton shortage

Figure 5 describes domestic raw cotton consumption in Britain from 1850-1875. This is the best available measure of the change in production in the industry across this period. The graph also describes the evolution of payments for wages and other variable costs, other than cotton, in the industry, from 1860-1868. Both of these statistics suggest that the shock period was characterized by a large reduction, equal to roughly half of pre-war production, in both industry production and wage payments.

Figure 5: British domestic cotton consumption and input payments, 1850-1875



Domestic raw cotton consumption data from Mitchell & Deane (1962). Wage and cost data from Forwood (1870).

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 focuses on imports. This figure shows that, once imports of raw cotton are excluded, there do not appear to be any substantial changes in either total imports or raw material imports to Britain. This makes sense given that raw cotton made up 67% of total British imports from the U.S. in 1860. Of the other major U.S. exports to Britain, only tobacco was heavily sourced from the U.S. South, and that made up

only 2.6% of imports from the U.S.

The right-hand panel of Figure 6 shows the behavior of exports from Britain over the study period. There was a substantial drop in exports of textiles during the U.S. Civil War period, which was due entirely to exports of cotton (these were relatively good years for other textile industries such as wool and linen). However, once textile exports are removed, 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 it is well known that some producers were able to circumvent these restrictions, in general they limited the impact that the conflict had on these industries.

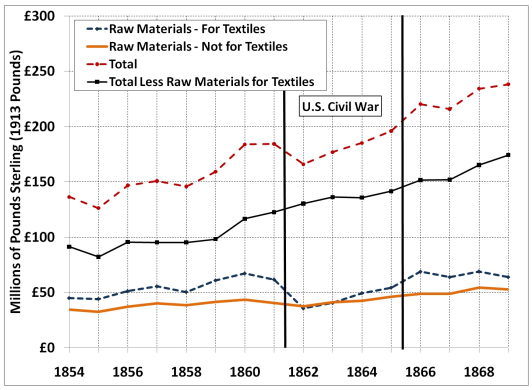
One sector of the British economy that was substantially affected, however, was the shipping industry. To avoid the risk of capture by Southern privateers, many U.S. merchant ships, which were primarily owned by Northern shipping interests, were transferred to British ownership during the U.S. Civil War. This resulted in a substantial expansion of the British merchant fleet, which had impacts in major shipping centers, particularly Liverpool. To account for the potential effect of these changes, in robustness exercises we explore the impact of dropping Liverpool, the main beneficiary of the transfer of shipping capacity to Britain, from our data.

B.2 Responses to the cotton shortage

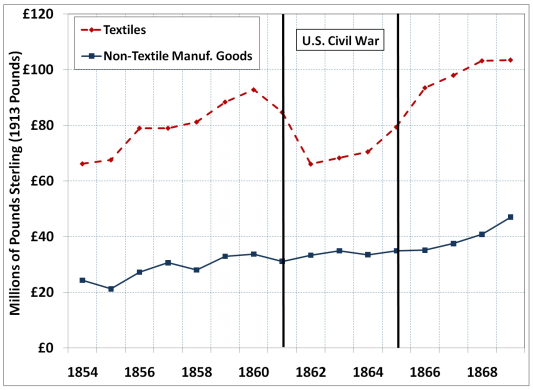
The response of both individuals and institutions to the recession caused by the cotton shortage played an important role in influencing health outcomes during this period. Workers who found themselves unemployed responded, first, by reducing costs and dipping into any available savings, and later, by pawning or selling items of value, including furniture, household goods, clothing and bedding (see Watts (1866, p. 214)

Figure 6: British imports and exports, 1854-1869

Imports to Britain



Exports from Britain



Data from Mitchell (1988).

and Arnold (1864)). Evidence suggests that many workers exhausted these private resources before turning to public relief—indeed, some previously proud workers were even found begging or busking on the streets (Henderson, 1969, p. 98-99). Even those who remained employed generally suffered substantial reductions in income, due to working short-time or to the substitution of Indian for U.S. cotton, a practice which slowed down production and reduced pay, which was largely based on piece rates. Finally, as discussed briefly in Section 4, many left cotton districts in search of work in other areas.

The recession also generated an unprecedented institutional response aimed at relieving the suffering in cotton districts. Contemporary reports largely credit public relief efforts for the fact that no widespread famine occurred during the recession.[56] Relief funds came in two main forms. First, funds were provided at the local level through the Poor Law Boards, the primary system for poor relief in Britain during this

[56]For example, the Registrar General's report of 1864 states that (p. xv), "that famine did not bear the fruit which in the history of nations it has too often borne, the spectacle of thousands struck by fever and death,—is mainly due to that legal provision for the poor which Christian civilization has established, and to the spontaneous munificence of a people amongst whom the seeds of charity have been liberally scattered."

period.[57] However, because Poor Law funds were associated with pauperism, provided funds for only the barest level of subsistence, and required "labour tests" such as rock-breaking, which cotton workers found demeaning, there is evidence that workers tried to avoid drawing on this stigmatized source of support (Kiesling, 1996; Boyer, 1997). The second source of funds was a large number of charitable contributions. These funds could take the form of cash, vouchers, and in-kind assistance, and came from voluntary subscriptions from across the country and even as far away as Australia (Watts, 1866). Direct relief was not the only institutional response. Additional relief programs included schools for children and adults, such as girl's sewing schools, as well as public works employment for unemployed cotton workers, though most public works employment began in 1863, after the worst of the crisis had passed.[58]

At the height of the recession in the winter of 1862, reports indicate that roughly 500,000 persons depended on public relief funds, with over 270,000 of these supported by the local Poor Law boards and an additional 230,000 reliant on the voluntary relief funds (Arnold, 1864, p. 296). The number of persons supported by public sources would fall to 264,014 by mid-summer 1863, and by 1865, the number of persons on relief fell back to where it had been at the beginning of the crisis (Arnold, 1864; Ellison, 1886).

Despite the best attempts of institutions and individuals to cope with the crisis—for instance, through short-time work, public relief funds, in-kind transfers, and public works employment—these efforts were insufficient in the face of such an intense and unexpected shock. Accordingly, migration became a popular means of adjustment as many erstwhile cotton operatives left cotton districts in search of work in other areas.

[57] These funds were provided by taxes levied on local property owners. See Watts (1866) for a description of the workings of the Poor Law Boards during the Cotton Famine.

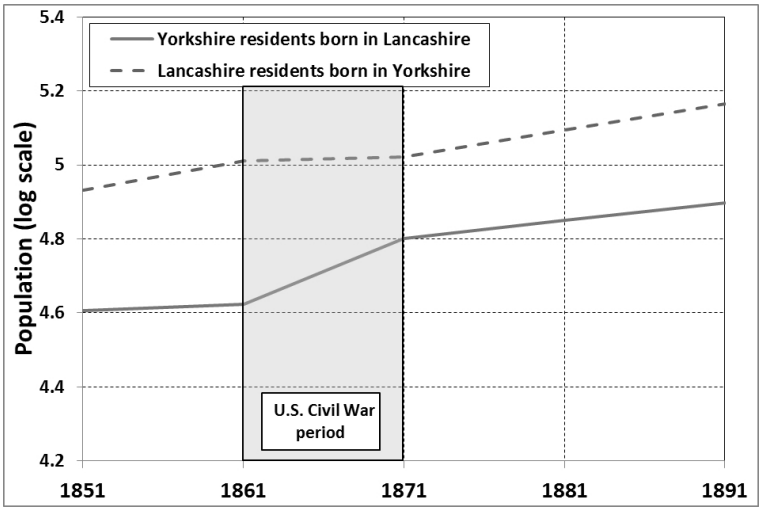
[58] See Arnold (1864) for a discussion of public works. The availability of public works expanded substantially starting in the summer of 1863, when Parliament passed the Public Works (Manufacturing Districts) Act. This Act used the central government's borrowing authority to provide long-term low interest rate loans to municipal governments so that they could undertake needed public works projects using unemployed cotton operatives. Most of these projects were aimed at improvements to roads and water or sewer systems.

B.3 Additional results on migration during the cotton shock

B.3.1 Internal migration

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 only available at the county level, so in Figure 7, which is reproduced 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.

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

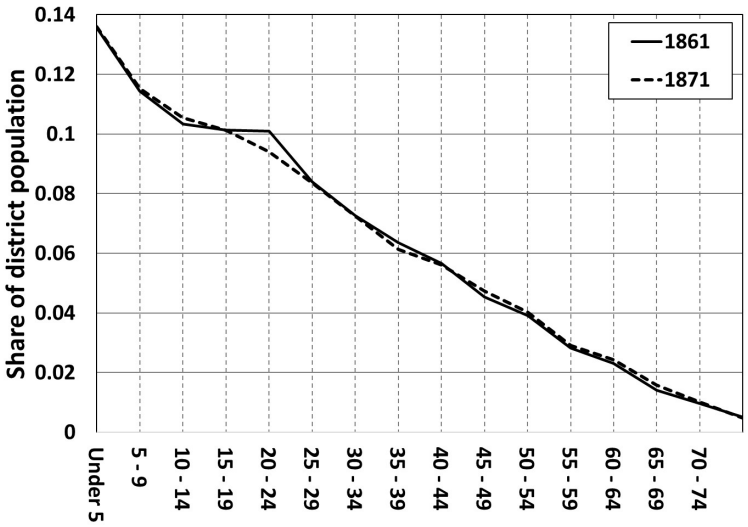


This graph, which is reproduced from Hanlon (2017), presents data on the birthplace of county residents from the Census of Population.

Next, we consider some results that help us think about how migration patterns

varied across age groups. Figure 8 describes the share of the population in each age category up to 79 in the cotton districts. The most prominent feature in this graph is that there was a substantial excess of young workers in the 20-24 age group in cotton districts in 1861, which had largely disappeared by 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

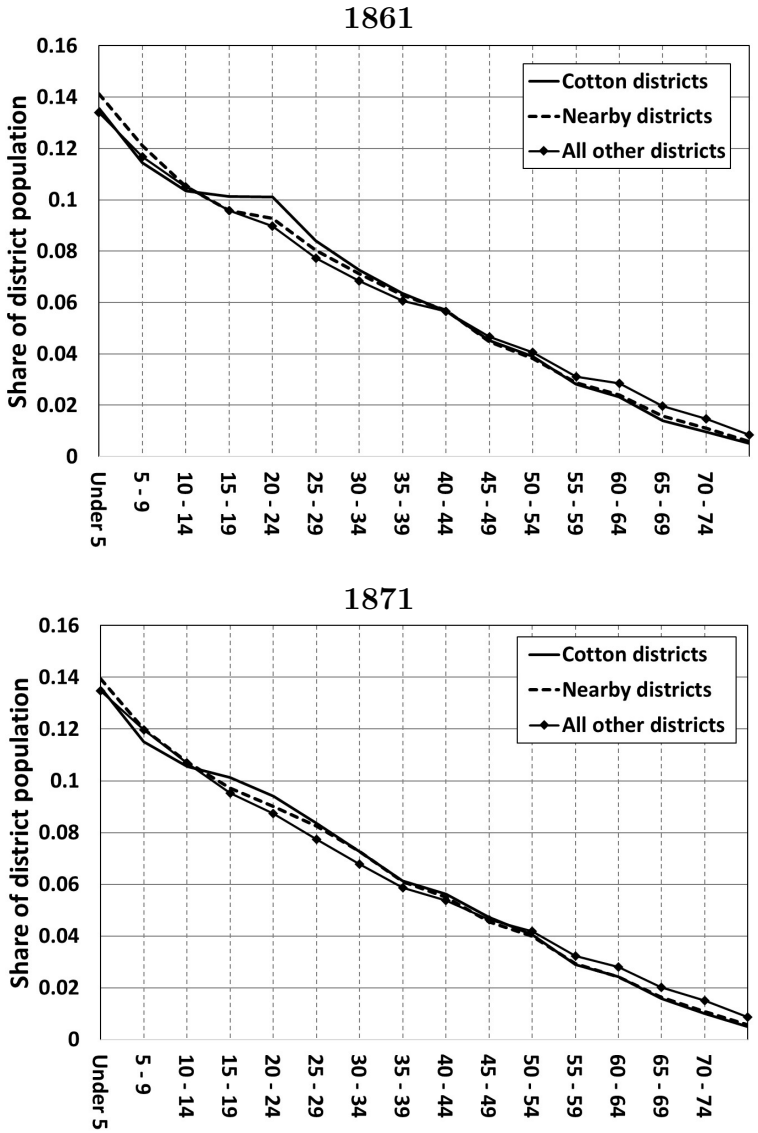


Population data are from the Census of Population for 1861 and 1871. Cotton districts are identified as those with over 10% percent of workers employed in cotton textile production in the 1851 Census, as in the main analysis.

An alternative view of the same pattern is provided in Figure 9, which compares the share of population by age group in cotton districts, nearby districts, and all other districts. This is done for 1861 in the top panel and for 1871 in the bottom panel. In 1861, cotton districts had a much larger share of young workers, particularly in the 15-19 and 20-24 age groups, than the other districts. By 1871, most of that difference had disappeared. It is also worth noting that in 1871, the nearby districts had substantially more population in the 25-29 and 30-34 age groups than the "all-other" districts. This pattern is consistent with the migration of workers who were in the 15-19 and 20-24 age group and living in cotton districts in 1861, into nearby

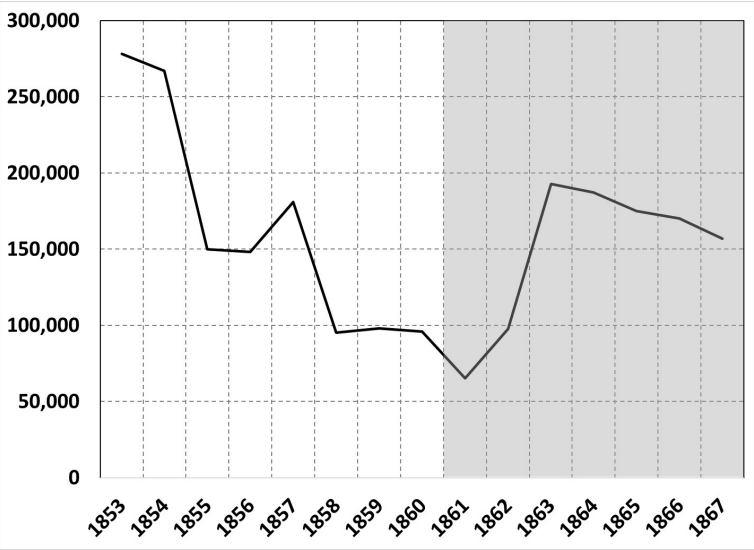
districts where, by 1871, they appear in the 25-29 and 30-34 age groups.

Figure 9: Share of population in each age group by type of district



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.

Figure 10: Emigration from Britain, 1852-1867



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

B.3.2 Emigration from Britain

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. Figure 10 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.

B.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."[59]

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 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, 2013), 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,

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

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.

C Appendix: Data

C.1 Linked data

C.1.1 Linked data overview

To assess the extent to which migration may bias estimates of the recession-mortality relationship, we construct an individual-level longitudinal dataset. Specifically, we link comprehensive death records back to full-count census records, which allows us to identify not only where individuals lived at the time of death, but also where they lived at the time of enumeration—a point in time which in our setting corresponds to the onset of the shock under study.

To construct this dataset we draw on comprehensive and publicly available death records and census records. Our death records come from The Free Births, Marriages, and Deaths project (FreeBMD). FreeBMD's objective is to transcribe and provide free internet access to the General Register Office's Civilian Registration Index. The General Register Office was set up following the Births and Deaths Registration Act of 1836, which instituted mandatory reporting of births, deaths, and marriages. Of the three indices, the death index is thought to be the most accurate and comprehensive, as registration was required before the body could be legally disposed of Woods

(2000).

FreeBMD has currently transcribed the entire death index for the years spanning 1837 to 1978. For our linking exercise, we begin by extracting all death records occurring in the years 1851-55 and 1861-65.[60] During this period, the index reports the following information consistently: first and last name, the registration district where the death occurred, and the year and quarter that the death occurred. As to the accuracy of the transcription, entries in the index were independently transcribed by two or more volunteers. Further, transcribers were encouraged to identify whether an entry was illegible or whether they could not clearly identify the letter (e.g., o vs a). Because names are crucial for our linking, we drop any illegible entries. Roughly 0.5% of the index is dropped because either the first or last name could not be accurately transcribed. After dropping illegible entries, we are left with 1.97 million deaths during the 1851 to 1855 period and 2.13 million deaths between 1861 and 1865.[61]

We attempt to link each of these deaths to a unique census record. This is possible through the efforts of the Integrated Census Microdata Project (I-CeM), which partnered with findmypast.org to transcribe and standardize the 1851, 1861, 1881, 1891, 1901, and 1911 British censuses. Our application in particular requires access to the corresponding names of each entry in the entire 1851 and 1861 censuses, which we obtain from the UK Data Archive.

C.1.2 The linking process

With these two datasets—the full-count census and the entire death index—we now know where every individual was living at the time of census enumeration, and we know when and where each death occurred in the five years following enumeration. Our goal is to construct a district-level mortality rate based on where individuals

[60] Including the 1851-55 period allows us to analyze the impact of the cotton shock on mortality within a difference-in-differences framework.

[61] Because we are looking to link back to the preceding census, we omit any deaths in the first quarter of 1851 and 1861, as those deaths occurred prior to census enumeration.

lived (not where they died), which is complicated by the fact that these two sources of data were collected independently. Accordingly, there is no unique identifier linking records between the two datasets. Thus, we set about constructing our own linked sample based on the available data.

Linking across datasets to track individuals over time has become increasingly common in economic history. Our application shares many similarities to the existing literature, but it also offers some unique challenges and benefits. Given this, it is perhaps useful to provide an overview of a typical linking application.

A typical application might go something like this. A researcher is interested in studying intergenerational mobility, but the researcher faces a problem: few datasets allow us to simultaneously observe a child's adult outcomes and the characteristics of their parents. To overcome this challenge, a researcher might begin with a set of adults in one census and try to find those adults as children in an earlier census. In the earlier census, the adult is observed in the same household as their parent, which allows the researcher to recover the parental characteristics that are necessary to study mobility.

While the precise mechanical steps of constructing a linked dataset vary from paper to paper, the general approach can be briefly described in two steps. The first step is to use time-invariant variables to identify a set of potential links. The second step is to see if, among those potential links, a record exists that is clearly the best link. Important methodological contributions relating to these steps include: Ferrie (1996); Abramitzky *et al.* (2012, 2014); Feigenbaum (2015, 2016); Bailey *et al.* (2017).

A researcher typically identifies a set of potential links by specifying criteria that both observations must share. For instance, the researcher might require that the names be reasonably close in spelling or pronunciation, that the ages are similar, and that the birth places match. The reason researchers do not require an exact name is because, in the US, information was spoken to the census enumerator who then recorded the information by hand. Thus, requiring an exact string match for

names could be problematic if one enumerator wrote "Anderson" while the second enumerator wrote "Andersen".

One challenge we face is that by trying to link death records back to census records, the only identifying information we can leverage in our setting is the individual's name. Age was not recorded in the death index until 1866, and place of birth was never recorded. However, one thing that works in our favor is that unlike the U.S. setting, where information was spoken to enumerators, the British census was filled out by the household. British enumerators were only responsible for making sure that transcription sheets were complete at the time of collection. This suggests that information obtained from British enumerators is more accurate than in the U.S. Further, in contrast to the issues faced in a typical U.S. application of record linking, there were many fewer foreign immigrants residing in England and Wales during our time period. This, combined with the fact that we link records over a very short time horizon (at most five years, rather than the typical 20-40 years) suggests that the assimilation of immigrant names is likely not a problem in our setting. Another advantage gained by linking over a short time horizon is that, in contrast to much of the broader linking literature, we are able retain women in our sample.[62]

In light of this, we attempt to link death records back to census records by looking for unique first and last name string matches. To do this, we clean both the death and census records to remove titles (e.g. Lord, Duke, etc.), as the presence of titles may vary across the two datasets. We also remove spaces in names, which may also appear inconsistently (e.g., McArthur vs Mc Arthur). Aside from this minor cleaning, the first and last name records are largely untouched.

Next we attempt to link the death records back to the census records. We begin by restricting to the set of death records with a unique first and last name across each time period, so unique amongst all other deaths occurring between 1851-55 or

[62]Because most linking studies examine a period of 20 or more years it is standard practice to drop women, since they tend to change their last name at the time of marriage, which complicates the linking process.

unique amongst all other deaths occurring between 1861-65. Of the 1,974,147 deaths that occurred between 1851 and 1855, 437,307 of the records are uniquely identified by their first and last name. For the 1861-65 period 470,751 of the 2,132,246 deaths are uniquely identified by their first and last name.

With these unique records in hand, we turn to the census in order to attach a location of residence in either 1851 (for deaths occurring 1851-55) or 1861 (for deaths occurring 1861-65). Of course, having a unique name in the death records does not necessarily mean that the name will be unique in the entire enumerated population. For the 1851-55 period, 71,566 of our unique death records map to a unique census record in 1851, while 81,221 of the unique 1861-65 death records map to a unique census record in 1861. This yields a final match rate of about 3.7%. While this number is much lower than other linking studies—Abramitzky *et al.* (2012), for instance have a match rate of 29% while Long & Ferrie (2013), who link between the 1851 and 1881 British censuses, have a match rate of about 20%—the difference stems from the fact that we do not have additional information to help break ties. However, another useful point of comparison is Ferrie (1996), who begins with a sample of unique names in the census and attempts to link these back to an earlier census. In his study, Ferrie matches 19% of all unique names. Using the number of unique records (rather than all records) as our baseline, our match rate rises to a comparable 16.2% over the 1851-55 period, and 17.1% for the 1861-65 period.

Recent work by Bailey *et al.* (2017) has raised concerns about the presence of false positives in linking studies. One way to examine the presence of false positives in our study is to examine the rate of agreement between time-invariant variables that were not used to link records. In our study, we use information about middle names as this diagnostic tool. Because middle initials were not consistently collected in the census, this is not a variable that we can link on. However, for the set of records that have a middle initial (or name) in both the death and census records, we can examine how often the middle initial in each data source concurs within a given linked record

in order to shed light on the presence of false positives in our sample. This exercise reveals a false positive rate of 28% in the 1851 sample and 39.8% in the 1861 sample. While high, these rates are comparable to those in existing linking studies (see Bailey *et al.* (2017)).

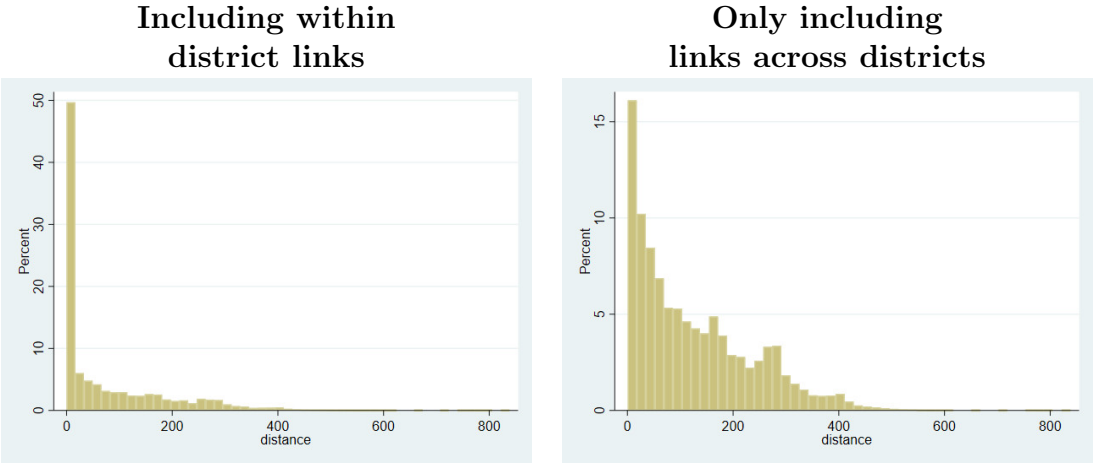
It is worth noting, however, that false positives are much less problematic in our setting than in traditional linking studies. The reason for this is twofold. First, many researchers use linked data to study intergenerational mobility. In these types of studies, false positives mean that a child's outcomes will be independent of the parent's characteristics, since the two records are randomly assigned. This is problematic because it will suggest perfect mobility. In our setting, however, deaths will be independent of the location of residence, which simply works against our finding by attenuating our coefficients, but otherwise should not affect the interpretation of our results. Second, and most important, in our analysis we follow Feigenbaum (2015) and examine our linked data within a differences-in-differences framework. This means that as long as errors are random, our estimating strategy will difference out the effect of any bad matches.

C.1.3 A check: linking against distance

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 chances of a match occurring 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.

Figure 11 presents histograms showing the share of linked deaths by distance bins using data from both the 1851-55 and 1861-65 periods. Distance is calculated using latitude and longitude coordinates for the main town or administrative center for

Figure 11: Share of links by distance between census and death districts



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.[63] If we split the data by period, similar patterns are observed in 1851 and 1861. Overall, these patterns suggest that our linking procedure is producing reasonable matches.

C.2 Aggregate data

C.2.1 Aggregate data overview

To assess the ability of traditional aggregate data to recover accurate estimates of the health consequences of the cotton shortage, we construct a new panel of annual district-level mortality spanning 1851-1871. These detailed data, which we digitized from original reports of the Registrar General, include information on both the age

[63]Note that the bump in links at about 250 km corresponds to the distance between the two major population centers in the country, London in the Southeast, and the area around Manchester and Liverpool in the Northwest.

and cause of death for over 600 registration districts covering all of England and Wales. The registration district-level tabulations are the finest geographic level covering the demography of all of England and Wales annually in this period. Previously available data from the Registrar General's reports, digitized by Woods (1997), are reported only at the decade level, and so is insufficiently detailed for our analysis. For an in-depth discussion of the Registrar General's data, see Woods (2000).

We also collect information from the Registrar General's reports on district population from the census years 1851, 1861 and 1871. When calculating mortality rates, we interpolate intercensal population using the Das Gupta method, per U.S. Census Bureau best practices (see *Methodology for the Intercensal Population and Housing Unit Estimates: 2000 to 2010*, U.S. Census Bureau, 2012, as well as the discussion in Appendix A.2).

To establish the spatial distribution of the shock, we measure the importance of the cotton textile industry in each registration district prior to the U.S. Civil War. This is done using data from the full-count 1851 Census of Population, which includes information on occupation, by district, for every person in England and Wales. Since the location of industries is highly persistent over time, we use use data from 1851 rather than from 1861 to avoid the possibility that our measure may be influenced by events occurring in cotton districts at the time of enumeration. Using these occupation data, we calculate the number of cotton textile workers as a share of the total working population for each district, which provides us with a cross-sectional measure of the importance of the cotton textile industry in each district on the eve of the shortage.

One factor complicating the use of these data is the change in district boundaries over time. To deal with this issue, we manually review the boundary changes for every district over our study period and combine any pair of districts experiencing a boundary change that resulted in the movement of over 100 people from one to the

other. This leaves us with 538 consistent districts in the main analysis.[64] Summary statistics for these 538 districts appear in Table 8.

Table 8: Summary statistics for aggregate data

	(1) Mean	(2) Standard deviation	(3) Min	(4) Max
Panel A: Full sample of districts				
Average annual deaths (full sample)	809.23	1050.19	34	11,256
Cotton employment share (1851 census)	0.017	0.07	0	0.51
Nearby cotton employment (1851 census)	4569.01	20,175.54	0	158,490
Population (1851 census)	33,260.87	35,019.68	2493	284,126
Panel B: Cotton districts only				
Average annual deaths prior to shock	1943.89	1546.46	207	7957
Average annual deaths during shock	2133.1	1684.46	199	8900

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

This subsection analyzes how representative our linked sample of deaths is of the aggregate deaths data. When comparing these two datasets, we use the version of the linked data in which deaths are assigned to the district of death, since this corresponds with the structure of the Registrar General's aggregate data.

wFirst, we present summary statistics for the linked sample over our two time periods (Table 9).

[64] One area where boundary changes create major issues is in a set of districts around Leeds. Ultimately, to obtain a consistent series we combine several neighboring districts into a single "Greater Leeds" district.

Table 9: Summary statistics for linked data

	Mean	Standard deviation	Min	Max	N
Ln(MR) in 1851	-5.641	0.536	-9.031	-4.639	538
Ln(MR) in 1861	-5.591	0.525	-9.268	-4.661	538
Δ Ln(MR)	0.046	0.341	-1.919	2.242	537
Cotton dist. ind.	0.045	0.206	0	1	539
Nearby (0-25km) ind.	0.048	0.214	0	1	539
Nearby (25-50km) ind.	0.059	0.237	0	1	539
Nearby (50-75km) ind.	0.067	0.25	0	1	539
Dist. pop. in 1851	33,284	35,047	2,493	284,126	538
Under 5 pop. shr 1851	0.131	0.009	0.098	0.163	538
Age 5-15 pop. shr. 1851	0.326	0.016	0.233	0.373	538
Over 50 pop. shr. 1851	0.154	0.021	0.093	0.213	538
Under 5 pop. shr 1861	0.133	0.01	0.1	0.166	538
Age 5-15 pop. shr. 1861	0.323	0.014	0.249	0.372	538
Over 50 pop. shr. 1861	0.161	0.024	0.098	0.236	538

Full sample includes 8,085 district-year observations spanning 1851-1865 for 539 unique districts. For the statistics that only draw on district-level data, there are District-level annual death data transcribed from annual reports of the Registrar General. Cotton employment share is simply the share of total employment (in 1851) that was employed in the cotton industry. Nearby cotton employment refers to the total number of workers in the 0-25 km radius of each district that were employed in the cotton industry in 1851. Nearby cotton employment is set to 0 for cotton districts (those with an 1851 cotton employment share greater than 10%). The pre-shock period is 1851-1860, while the shock period is 1861-1865.

Next, we examine representativeness by comparing linked deaths to deaths reported in the aggregate data. One dimension along which we can compare the linked sample to aggregate deaths is the in the number of deaths over the two time periods. In the aggregate data, the ratio of deaths in 1861 to deaths in 1851 is 1.14. In the linked sample the ratio is 1.17. Thus, the linked data appears to capture reasonably well the increase in the overall number of deaths across these two periods.

Another 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. Figure 12 provides histograms showing the share of linked deaths in each age group and the share of aggregate deaths in each age group. This graph shows that infant and young child deaths are substantially under-

represented in the linked sample. Deaths in ages 6-54 tend to be over-represented in the linked sample, while deaths among those over age 65 are also under-represented, though not as badly as for infants. The fact that our linked sample struggles to reflect deaths among young children is a mechanical result of our approach, since deaths among infants after 1862 cannot be linked back to people who were alive in the 1861 census. Similarly, the death of a child under 2 after 1863 cannot appear in our linked data. The explanation for why deaths in ages 65-85 are somewhat under-represented in the linked sample is not clear. The most likely explanation for this is that there are more spelling mistakes or legibility issues in the names among this age group, which could be a result of old age or the fact that older age groups were more likely to be illiterate. 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.

Figure 13 compares the age shares in the linked sample across the two time periods. This tells us that at least within the linked sample, our data are comparable over time.

We can also compare representativeness by gender. Table 10 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 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).

We can also look at the representativeness of our sample by socioeconomic status

Figure 12: Histogram of deaths by age from linked and aggregate data

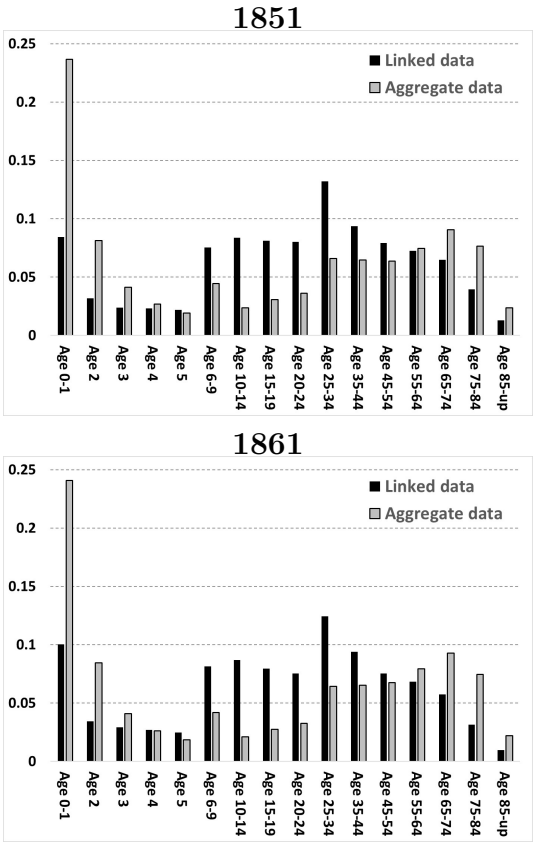


Figure 13: Histogram of deaths by age comparing linked samples from 1851 and 1861

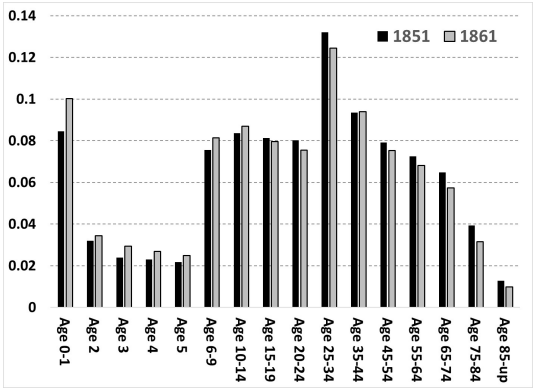


Table 10: Gender shares of linked and aggregate deaths by time period

	Linked deaths		Aggregate deaths	
	1851	1861	1851	1861
Women	0.542	0.533	0.492	0.487
Men	0.458	0.467	0.508	0.513

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

The results, in Table 11, 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.

It is not possible to look at whether our linked sample is representative with respect

Table 11: Shares of white and blue-collar deaths in linked and aggregate data

	Linked deaths in 1851	Linked deaths in 1861	Aggregate deaths in 1851
White-collar	0.07599	0.07511	0.07601
Blue-collar	0.9240	0.9249	0.9240

to immigrant status because we do not have aggregate mortality data reported by birth location. However, we can get a sense of whether changes in immigrant deaths are influencing our results by looking at whether there was a differential change in the share of linked deaths made up by immigrants in 1861-65 compared to 1851-55. This is done in Table 12, where the dependent variable is the share of linked deaths among those born outside of England and Wales. We can see that the estimated coefficients on the cotton district indicator are close to zero and not statistically significant. Thus, there does not appear to have been any substantial change in the share of immigrant deaths in the cotton districts associated with the shock, unless these deaths occurred outside of England and Wales.

D Appendix: Analysis

D.1 Graphical evidence of the impact on mortality

In addition to our main results, we present here some simple time-series evidence in the tradition of early work in this area (e.g., Brenner (1979)). Figure 14 describes the evolution of log mortality across the study period (normalized such that 1860=1). Note that we focus on log mortality rather than the log mortality rate in these graphs. This is because the mortality rate includes population in the denominator. As previously discussed, measurement issues in the population variable can lead to misleading results when studying mortality rates in the presence of unobserved migration.

Table 12: Changes in immigrant share of deaths in cotton districts

DV: Change in immigrant share of linked deaths from 1851-5 to 1861-5				
	(1)	(2)	(3)	(4)
Cotton district indicator	-0.00193	-0.00319	0.000120	0.000181
	(0.0123)	(0.0123)	(0.0116)	(0.0116)
Nearby (0-25km)		-0.00784	-0.00731	-0.00718
		(0.00628)	(0.00662)	(0.00663)
Nearby (25-50km)		-0.0124**	-0.0118**	-0.0114**
		(0.00619)	(0.00514)	(0.00517)
Nearby (50-75km)		0.00723	0.00313	0.00327
		(0.00793)	(0.00813)	(0.00792)
Log initial pop.				-0.000887
				(0.00278)
Age share controls			Yes	Yes
Observations	537	537	537	537
R-squared	0.000	0.017	0.079	0.079

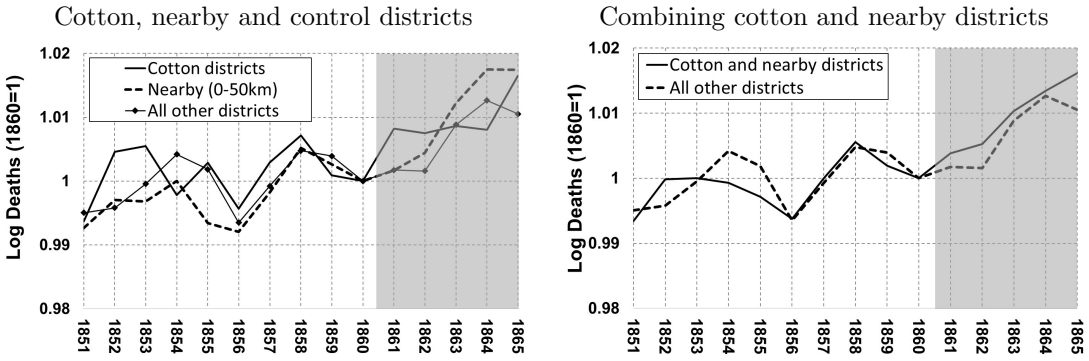
*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. Regressions are weighted by district population in 1851. Age group controls include the share of district residents under 5, the share 5-15, and the share over 50 for both the 1851 and 1861 census.

The left-hand panel separates cotton districts, nearby districts (non-cotton districts within 0-50 km of the cotton districts), and all other districts. There are three notable features here. First, in the pre-shortage period, mortality in the three groups track each other reasonably well. Second, in the first two years of the U.S. Civil War, there is evidence of elevated mortality in the cotton districts, though this increase disappears after 1862. Third, in the nearby districts we observe a substantial increase in mortality, particularly after 1862. One potential cause of this increase in mortality in nearby districts is migration from cotton to nearby districts. Migration can affect mortality in nearby districts either directly, if some migrants move and then die, or indirectly, if migrants bring disease or generate other congestion effects.

Since our evidence suggests that most migration out of cotton districts went to nearby areas, combining the migrant-sending and migrant-receiving districts into a single group can help give us a sense of the aggregate mortality impact of the downturn. Most of the channels through which migration generates bias, discussed above, suggest that we should observe offsetting effects in migrant-receiving and migrant-

sending districts. Given this, aggregating migrant-sending and migrant-receiving regions will allow us to assess the net impact of the recession on mortality. In the right-hand graph we combine the cotton and nearby districts. These results provide evidence that overall mortality in the cotton-and-nearby-districts category increased during the U.S. Civil War period.

Figure 14: Mortality during the cotton shortage



Mortality data from the reports of the Registrar General. Data cover all of England & Wales. Cotton districts are those with more than 10% of employment in cotton textile production in 1851. Nearby districts are non-cotton districts within 25 km of the cotton districts. The linear trend line in the right-hand panel is fitted to data from the cotton and nearby districts from 1851-1860.

D.2 Additional analysis of the linked data

This section presents some additional results obtained from the linked data. The first set of results, in Table 13, look at how our main all-age results change as we add control variables. We can see that adding controls for the shares of different age groups in each district has an important effect on our results, while other controls do not play a major role. Note that the results in Column 4 correspond to the specification presented in our main linked results table in the main text.

Of the control variables, the most important for predicting mortality are the share of very young and very old residents in a district. This makes sense given that these populations faced the greatest mortality risk.

Table 13: All-age results using linked data and adding controls

DV: Std. change in log mortality rate from 1851-5 to 1861-5				
	(1)	(2)	(3)	(4)
Cotton district indicator	0.620	0.665	0.818*	0.823*
	(0.442)	(0.442)	(0.441)	(0.437)
Nearby (0-25km)		0.323***	0.434***	0.446***
		(0.103)	(0.124)	(0.125)
Nearby (25-50km)		0.214	0.357**	0.384**
		(0.142)	(0.147)	(0.153)
Nearby (50-75km)		0.00791	-0.0138	-0.00220
		(0.136)	(0.143)	(0.153)
Under 5 shr. in 1851			21.36**	20.09*
			(10.63)	(10.56)
Under 5 shr. in 1861			-5.221	-5.174
			(10.93)	(10.92)
Age 15-54 shr. in 1851			-1.491	-1.713
			(3.677)	(3.713)
Age 15-54 shr. in 1861			-4.630	-5.115
			(5.185)	(5.438)
Over 54 shr. in 1851			-8.125	-9.493
			(6.414)	(6.913)
Over 54 shr. in 1861			15.38**	14.93**
			(6.370)	(6.263)
Log initial pop.				-0.0762
				(0.0879)
Observations	537	537	537	537
R-squared	0.037	0.049	0.087	0.090

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. Deaths are assigned to the district of initial residence (i.e., district of census enumeration). Regressions are weighted by district population in 1851. Age group controls include the share of district residents under 5, the share 5-15, and the share over 50 for both the 1851 and 1861 census.

Next, the results in Table 14 investigate the robustness of the all-age results to changes in the estimation methodology, underlying data set, or treatment variable. In Column 1 we present results where we do not weight the regressions by initial district population. In Column 2 we consider an alternative, continuous, measure of treatment, the share of employment in cotton textiles in each district in 1851. 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 Liverpool, which was not a cotton district but which was the main port

serving the industry, as well as 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.

Table 14: Additional linked-data results on all-age mortality rates

DV: Std. change in log mortality rate from 1851-5 to 1861-5					
	No weights	Continuous treatment	Without Manchester	Drop Manchester Liverpool & Leeds	Only native-born
	(1)	(2)	(3)	(4)	(5)
Cotton district ind.	0.847**		1.090**	1.071**	0.795*
	(0.383)		(0.436)	(0.429)	(0.427)
Cotton employment share		1.381**			
		(0.608)			
Nearby (0-25km)	0.274*	0.292*	0.443***	0.432***	0.419***
	(0.166)	(0.174)	(0.125)	(0.125)	(0.123)
Nearby (25-50km)	0.397***	0.296	0.345**	0.299**	0.400***
	(0.142)	(0.182)	(0.150)	(0.127)	(0.152)
Nearby (50-75km)	0.0917	-0.0203	-0.0139	-0.00208	-0.0375
	(0.206)	(0.152)	(0.150)	(0.155)	(0.152)
Log initial pop.	-0.142*	-0.0726	-0.0362	-0.0729	-0.0760
	(0.0759)	(0.0853)	(0.0774)	(0.0777)	(0.0860)
Age share controls	Yes	Yes	Yes	Yes	Yes
Observations	537	537	537	537	537
R-squared	0.058	0.059	0.117	0.118	0.090

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. Deaths are assigned to the district of initial residence (i.e., district of census enumeration). Regressions in Column 2-5 are weighted by district population in 1851. Age group controls include the share of district residents under 5, the share 5-15, and the share over 50 for both the 1851 and 1861 census.

The next set of results, in Table 15, looks across more detailed age categories. In these results, we find positive effects across all age groups, but these effects only become large and statistically significant only for age groups starting at 45, suggesting that the main effect of the shock on mortality was concentrated among older adults.

Table 15: Linked-data results for additional age categories

	DV: Std. change in log mortality rate from 1851-5 to 1861-5					
Age group:	15-24	25-34	35-44	45-54	55-64	65+
	(1)	(2)	(3)	(4)	(5)	(6)
Cotton district ind.	0.415	0.459	0.218	0.643**	0.540**	0.608***
	(0.258)	(0.309)	(0.304)	(0.276)	(0.227)	(0.210)
Nearby (0-25km)	0.265***	0.276**	0.231*	0.165	0.410**	0.384***
	(0.0840)	(0.115)	(0.128)	(0.114)	(0.162)	(0.137)
Nearby (25-50km)	0.135	0.150	0.171	0.356***	0.368***	0.223
	(0.129)	(0.122)	(0.111)	(0.132)	(0.122)	(0.139)
Nearby (50-75km)	0.00905	0.00214	0.301*	0.283*	0.273	-0.162
	(0.146)	(0.134)	(0.164)	(0.156)	(0.183)	(0.161)
Log initial pop.	-0.116***	-0.0806*	-0.0583	-0.0576	-0.124**	-0.0489
	(0.0438)	(0.0431)	(0.0420)	(0.0471)	(0.0507)	(0.0590)
Observations	521	521	512	507	500	527
R-squared	0.034	0.032	0.017	0.054	0.053	0.045
Linked deaths	24,272	19,636	14,305	11,765	10,709	16,377

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. Deaths are assigned to the district of initial residence (i.e., district of census enumeration). Regressions are weighted by district population in 1851. The sample size changes across specifications if there are small districts where we do not have links when cutting by age group.

One potential solution to issues associated with migration is to analyze larger geographic units. If most migration takes place within these larger units, then aggregation can potentially deal with migration bias. However, moving to larger geographic units comes with a cost. Larger units are more heterogeneous and may result in combining areas experiencing poor economic conditions with other areas that are doing well, making it more difficult to identify the effect of economic fluctuations on mortality. In the setting we consider the natural larger geographic unit to consider is the county, of which there were 55 in England and Wales.

In Table 16, we present the results obtained from a county-level analysis. In the first three columns of this table, linked deaths are assigned back to the individual's location in the census at the beginning of each period. For comparison, in the last three columns linked deaths are assigned to the death district. In Columns 1-2 we do not consider spillovers into nearby areas. Here we see very weak evidence that mortality may have increased in the cotton districts during the shock period. These results become much stronger in Column 3, where we separately estimate the impact

of the shock on mortality in the nearby West Riding, the main destination for migrants leaving the cotton districts. It appears that in our setting, even at the county level, accounting for spillovers is important. Mainly this is telling us that the county level in our setting is just not high enough to capture both migrant sending and receiving districts within the same unit. It is interesting to note that the estimated coefficient on the West Riding is larger than what we estimate for the cotton districts. The most likely explanation for this is that, even though the cotton districts were more directly affected by the shortage, they were in larger and more diverse counties which included areas that did well during this period.[65]

In Columns 4-6 we conduct a similar exercise, but assigning deaths to death locations rather than individual's location at the start of each period. We can see that, by failing to account for migration, this pushes results towards finding that the cotton shortage may have reduced mortality, though once we account for spillovers the effect is close to zero. Thus, in our setting, aggregation to the county level is not sufficient to deal with the issues created by migration. In fact, even aggregating to region does not solve the problem. This is likely because given the typically defined regions in Britain during this period, the migrant-sending cotton districts end up in the Northwest region while the migrant-receiving districts in the West Riding are in the Yorkshire region.

In summary, when assigning deaths to individual's initial location and controlling for spillovers in nearby areas we find evidence of a statistically significant increase in mortality in the cotton districts. However, when deaths are assigned to the location in which they occurred we find either negative effects or estimates that are very close to zero. Thus, aggregation to the county level does not solve the problems created by migration bias in our setting. The main reason for this is that most of the migration from the cotton districts crossed the county border from Lancashire and Cheshire

[65]For example, in the far North of Lancashire, the main cotton producing county, the area around Barrow-in-Furness was booming at this time due to iron ore discoveries that led to the opening of new steel works and a shipyard.

into Yorkshire. Even aggregating to the traditional regions of Britain is not likely to solve migration bias problems in our setting, since the affected areas span different administrative regions as well.

Table 16: All-age linked-data results at the county level

	DV: Std. change in log mortality rate from 1851-5 to 1861-5					
	Assigning deaths by census location			Assigning deaths to death location		
	(1)	(2)	(3)	(4)	(5)	(6)
Cotton counties	0.152	0.0914	0.687**	-0.286**	-0.452	0.203
	(0.176)	(0.334)	(0.274)	(0.125)	(0.359)	(0.287)
Nearby			1.085***			1.192***
(West Riding)			(0.263)			(0.268)
Log initial pop.		0.145	-0.0331		0.0157	-0.180
		(0.170)	(0.165)		(0.168)	(0.158)
Age controls		Yes	Yes		Yes	Yes
Observations	55	55	55	55	55	55
R-squared	0.007	0.164	0.266	0.026	0.095	0.219

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. Deaths are assigned to the district of initial residence (i.e., district of census enumeration). Regressions are weighted by county population in 1851. Age group controls include the share of county residents under 5, the share 5-15, and the share over 50 for both the 1851 and 1861 census.

D.3 Additional analysis of the aggregate data

In Table 17, we present results analogous to those presented in Table 13, but obtained from aggregate data. Here we can see that irrespective of the specification, standard aggregate data provides strong evidence that mortality is procyclical, a finding which is contrary to our graphical, qualitative, and preferred linked evidence. Here we can also see the importance of accounting for spatial spillovers, since the exclusion of the nearby terms in Column 1 serves to exaggerate the extent of procyclicality relative to what we see in Columns 2-4.

Table 17: All-age results using aggregate data and adding controls

DV: Std. change in log mortality rate from 1851-5 to 1861-5				
	(1)	(2)	(3)	(4)
Cotton district indicator	-0.732***	-0.630***	-0.503***	-0.508***
	(0.148)	(0.149)	(0.172)	(0.171)
Nearby (0-25km)		0.466**	0.624***	0.613***
		(0.214)	(0.229)	(0.227)
Nearby (25-50km)		0.368**	0.605***	0.579***
		(0.185)	(0.181)	(0.183)
Nearby (50-75km)		0.570***	0.602***	0.591***
		(0.154)	(0.155)	(0.149)
Under 5 shr. in 1851			29.10**	30.29**
			(12.32)	(12.17)
Under 5 shr. in 1861			-27.74**	-27.78**
			(10.82)	(10.76)
Age 15-54 shr. in 1851			-1.103	-0.895
			(5.219)	(5.156)
Age 15-54 shr. in 1861			-0.747	-0.294
			(5.972)	(5.966)
Over 54 shr. in 1851			-10.66	-9.375
			(7.149)	(7.182)
Over 54 shr. in 1861			16.23**	16.65**
			(6.657)	(6.633)
Log initial pop.				0.0713
				(0.0646)
Observations	537	537	537	537
R-squared	0.053	0.094	0.187	0.189

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. Regressions are weighted by district population in 1851. Age group controls include the share of district residents under 5, the share 5-15, and the share over 50 for both the 1851 and 1861 census.

D.4 Results from a more standard panel-data approach

In the main text, for purposes of easy comparability, our aggregate-data approach mirrored the first differences approach necessitated by the structure of our linked data. This section presents results also using aggregate mortality data, but instead applying an estimating approach more similar to that used in Ruhm (2000). In particular, these results are based on Eq. 2, which is as close to the standard specification as we can get in our setting, given that we do not have measures of district-level unemployment.

We also check the robustness of our panel-data results by considering alternative time periods and by including time trends in some annual specifications.[66] As is relatively standard in the literature, we weight the regressions by district population.

Since this analytical approach uses annual data, it is necessary to construct population estimates for non-census years. This is done using the Das Gupta interpolation method, which is the same method used by the U.S. Census today. Starting with population in a census year, this method adds births and subtracts deaths in each year until the next census year is reached. The difference between the district population in a census year based on the previous census, plus births less deaths, and the population observed in the current census—called the error of closure—is then allocated proportionally across the intervening years (see Section A.2 for further detail).

Table 18 presents the results. We consider a variety of different specifications that we think a reasonable empirical economist might consider. In the first four columns we use data from 1851-1871, which cover the period over which the population data are interpolated. This gives us both pre-shock and post-shock observations which we compare to the treatment period, 1861-65. We consider results with and without accounting for spillovers into nearby districts, and with and without time trends. These results consistently indicate that mortality fell as a result of the cotton shortage. It is worth noting that results are fairly consistent regardless of whether time trends are included.

In an attempt to obtain results that come closer to recovering the effects documented in our main analysis, Columns 5-8 compare just the pre-shock and cotton shortage periods, dropping data after 1865. When using this time window and including time trends, we can see that the aggregate approach does indicate that mortality increased during the cotton shortage, and we also begin to see evidence that mortality increased in nearby areas. The time trends clearly have a large effect on the results when there is no post period.

[66] Including time trends is a common practice in this literature starting with Ruhm (2000).

Overall, however, our conclusion from this panel-data exercise is that an analysis based on aggregate data is unlikely to reliably and consistently recover accurate estimates of the impact of recessions on mortality in the presence of unobserved migration. While certain specifications do generate results similar to what we obtain using the linked data, this relies on the researcher choosing just the right analysis period and specification.

Table 18: Results from aggregate panel data regressions

	Dependent variable is std. log mortality rate							
Sample:	**1851-1871**				**1851-1865**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cotton district ind.	-0.273***	-0.288***	-0.265***	-0.293***	-0.191***	0.337***	-0.160**	0.371***
	(0.0491)	(0.0448)	(0.0494)	(0.0453)	(0.0645)	(0.0697)	(0.0652)	(0.0706)
Nearby (0-25km)			0.0197	-0.0303			0.116	0.310**
			(0.0702)	(0.0546)			(0.104)	(0.140)
Nearby (25-50km)			0.00715	-0.0749			0.137	0.0958
			(0.0876)	(0.0683)			(0.132)	(0.184)
Nearby (50-75km)			0.110**	0.0893*			0.133*	-0.0345
			(0.0517)	(0.0472)			(0.0683)	(0.0900)
Time Trends	N	Y	N	Y	N	Y	N	Y
Observations	10,780	10,780	10,780	10,780	8,085	8,085	8,085	8,085
R-squared	0.139	0.256	0.140	0.256	0.170	0.278	0.172	0.280
No. districts	539	539	539	539	539	539	539	539

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses