Estimating the Recession-Mortality Relationship when Migration Matters*

Vellore Arthi[†]
University of Essex

Brian Beach[‡]
William & Mary
and NBER

W. Walker Hanlon[§]
NYU Stern
and NBER

June 5, 2017

Abstract

A large literature following Ruhm (2000) suggests that mortality falls during recessions and rises during booms. The panel-data approach used to generate these results assumes that either there is no substantial migration response to temporary changes in local economic conditions, or that any such response is accurately captured by intercensal population estimates. To assess the importance of these assumptions, we examine two natural experiments: the recession in cotton textile-producing districts of Britain during the U.S. Civil War, and the coal boom in Appalachian counties of the U.S. that followed the OPEC oil embargo in the 1970s. In both settings, we find evidence of a substantial migratory response. Moreover, we show that estimates of the relationship between business cycles and mortality are highly sensitive to assumptions related to migration. After adjusting for migration, we find that mortality increased during the cotton recession, but was largely unaffected by the coal boom. Overall, our results suggest that migration can meaningfully bias estimates of the impact of business-cycle fluctuations on mortality.

JEL Codes: I1, J60, N32, N33

*We thank James Fenske, Joe Ferrie, Marco Gonzalez-Navarro, Tim Hatton, Taylor Jaworski, Amir Jina, Shawn Kantor, Carl Kitchens, Adriana Lleras-Muney, 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, University of Essex, Florida State University, Queen's University, RAND, University of Toronto, and University of 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).

[†]Department of Economics, University of Essex; v.arthi@essex.ac.uk

[‡]Department of Economics, College of William & Mary; bbbeach@wm.edu

[§]Department of Economics, NYU Stern School of Business; whanlon@stern.nyu.edu

1 Introduction

How do business cycles affect mortality? This question has attracted substantial attention following an influential study by Ruhm (2000), which found that mortality rates in the U.S. fall during recessions and rise during booms.[1] Methodologically, Ruhm compared state-level unemployment rates to state-level total mortality rates. This panel-data approach has now been applied to a wide variety of settings in developed and developing countries, both modern and historical. The majority of studies in this literature yield consistent results: mortality appears to be pro-cyclical.[2]

One critical assumption embedded in this empirical approach is that either there is no large migration response to business-cycle fluctuations, or that short-run population flows are accurately captured by intercensal population estimates. If this assumption fails, say, because unobserved migration changes either the size or the composition of an affected location's at-risk population, then we may observe a spurious change in the observed mortality rate which we will misinterpret as reflecting the true impact of business cycles on health. We call this phenomenon *migration bias*.[3] The presence of a migratory response raises two further concerns. First, migration may cause unemployment rates, the key explanatory variable used in most of the studies following Ruhm (2000), to become endogenous.[4] Second, to the extent that

[1] We acknowledge that we are using the terms "recession", "boom", and "business cycles" 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. Nevertheless, we adopt this convention to remain consistent with the existing literature; see, for instance, Ruhm (2000, p. 617).

[2] 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 counter-cyclical mortality include Svensson (2007), Economou *et al.* (2008), and Miller & Urdinola (2010). Also, Ruhm (2015) suggests that this relationship may have changed in the last decade. We review this literature in Appendix A.1.

[3] This potential issue is mentioned in Ruhm (2007), Stuckler *et al.* (2012), and Lindo (2015), but we are not aware of a study that assess the extent to which migration may affect the results in this literature.

[4] This is because unemployment rates both affect and are affected by migration. That is, a change in the unemployment rate may induce migration, and this migration will in turn affect both

individuals migrate towards areas that offer better economic opportunities, we are likely to observe migration between treatment and control locations, which has the potential to bias coefficient estimates obtained in panel-data regressions.

In light of these issues, the aims of this study are twofold: first, to assess the migration response to short-run changes in economic conditions; and second, to examine the impact of such migration on estimates of the relationship between business cycles and mortality.

To do this, we draw on two empirical settings. The first setting is the large recession in cotton-textile producing regions of Britain caused by the U.S. Civil War (1861-1865), an event which sharply reduced the industry's supply of raw cotton inputs.[5] The second setting is the boom in coal-producing counties of the Appalachian U.S. over the period 1970-1977, an event precipitated by the OPEC oil embargo and national regulatory changes.[6] Both settings were chosen because they offer plausibly exogenous variation in the timing and spatial distribution of short-term economic shocks. This allows us to more cleanly identify the impact of changes in local economic conditions, as we can observe the local incidence of the economic shock without relying on unemployment rates, which drive and are endogenously affected by migration.[7] In both cases, the temporal component of the economic shock was short, sharp, and generated by outside forces that were largely unexpected. Meanwhile, the spatial incidence of each shock was determined by the pre-existing distribution of economic activity, which was in turn due to underlying natural endowments.[8]

---

the numerator and denominator used to calculate unemployment rates in the next period. Similarly, unemployment rates may affect and be affected by patterns in mortality and fertility, patterns which may also be influenced by migration. The endogeneity of unemployment rates has been raised as a concern in Miller & Urdinola (2010) and Foote *et al.* (2017).

[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]This setting has previously been studied by Black *et al.* (2002), Black *et al.* (2005), and Black *et al.* (2013), which investigated outcomes related to social welfare provision, local economic performance, and fertility, respectively.

[7]This approach follows work by Miller & Urdinola (2010).

[8]In the case of the coal boom, the key natural endowment was coal reserves. In the case of the cotton shortage, work by Crafts & Wolf (2014) shows that the spatial distribution of the cotton

We make three primary contributions in this study. First, we show that local economic shocks can induce substantial and systematic migratory responses. In the case of the cotton shortage, we observe evidence that workers left cotton textile-producing areas during the recession and migrated to nearby non-cotton locations. In the case of the coal boom, we observe that the pattern of out-migration from Appalachia in the decades prior to the 1970s saw a short-lived reversal during the boom decade of 1970-1980. In this setting in particular, migration appears to have been highly selective, with boom-time population changes driven by the return of healthy working-age adults.

Second, we provide evidence that the estimated relationship between business cycles and mortality is highly sensitive to modeling assumptions related to migration. In both settings, different sets of assumptions about the relationship between population, migration, and mortality lead to substantially different results—in some cases, with completely different signs. Moreover, as we describe below, the standard approach makes strong assumptions about both the accuracy of intercensal population estimates and the short-run relationship between population and mortality. These assumptions are at odds with the evidence on migration in the settings that we consider.

Accordingly, we use four strategies to diagnose and account for the impact of migration on estimates of the relationship between business cycles and mortality. First, we propose a set of alternative estimation strategies, each embedding different assumptions about the relationship between mortality and short-run changes in population. Comparing the results obtained from these alternative approaches shows that estimates of the business cycle-mortality relationship depend heavily on the choices made to model migration. Second, we use the available evidence on the migration

textile industry in the 1830s was driven by factors including the availability of water power sources, rugged terrain, and access to a natural port. There was strong persistence in industry location, so that the locations where the industry was concentrated in the 1830s remained the main centers of the industry in the 1860s.

response in each setting in order to choose between these alternative approaches and the underlying assumptions that they imply. Third, we validate our choice of modeling assumptions by generating additional results using windows around census years, when population estimates are more accurate and less subject to migration bias. Fourth, we identify migrant-sending and migrant-receiving locations, and separately estimate the effects in each in order to account for treatment spillovers.

Our final contribution is to provide substantive evidence on the relationship between business cycles and mortality in two very different environments. In the case of the cotton shortage, our results after adjusting for migration bias suggest that mortality *increased* during this recession. In the case of the Appalachian coal boom, there is evidence that migration was selective, with young, healthy workers migrating temporarily to coal areas during the boom. This had a substantial impact on the *observed* mortality rate by changing the size and composition of the at-risk population. We find little evidence, however, that the Appalachian coal boom had any *real* effect on underlying population mortality. Strikingly, neither our cotton nor our coal findings are consistent with the pro-cyclical results obtained in most of the existing literature on the net effects of business cycles on total mortality.

Together, our results suggest that migration bias is an important concern, and that it may cause the standard panel-data approach used in the literature to generate meaningfully inaccurate estimates of the relationship between business cycles and mortality. Although here we illustrate these issues in two specific empirical settings, there is evidence that migration may be a key margin of adjustment to local economic shocks in many other contexts (Blanchard & Katz, 1992; Bound & Holzer, 2000; Foote *et al.*, 2017).[9] This suggests that migration bias may be an important concern when assessing the relationship between business-cycle fluctuations and mortality in

[9]Two other recent papers documenting the dynamic migration response to local economic shocks are Monras (2015) and Monras (2017). Both of these suggest that the internal migration response in the modern U.S. is rapid, with much of the adjustment taking place within just a couple of years.

a variety of settings[10]

Given the capacity of migration bias to undermine inference, one approach to dealing with these effects is that proposed in this study: to focus on specific shocks, to assess the magnitude and direction of migration bias in the setting being considered, and to use additional information on those migratory responses to make empirical choices that will mitigate this bias. One alternative approach to dealing with migration bias—but one that is only possible in settings with unusually rich data—is to use individual-level panel data. This approach has been applied in Sweden by Gerdtham & Johannesson (2005) and in the U.S. by Edwards (2008). Notably, these papers find mixed or counter-cyclical effects. Together with our results, the findings in these studies suggest that accounting for migration bias may substantially change our understanding of the relationship between business cycles and mortality. Although studies using individual-level panel data may be less vulnerable to migration issues, current data constraints mean that this approach cannot be readily applied to developing-country settings or over longer periods of time. Another alternative approach is to analyze patterns at the national level, as has been done using life tables in recent work such as Cutler *et al.* (2016).[11] This can help deal with migration bias because of the barriers faced by international migrants. However, an analysis run at the national level in the contexts that we study either requires relying on time-series variation, which has been criticized in previous work (e.g., Ruhm (2000)), or requires conducting cross-country regressions.

In the next section, we briefly describe the mechanisms through which local eco-

[10]This finding complements existing work highlighting the potential for migration to bias estimates obtained when applying a fixed-effects estimation strategy to assess the impact of local economic shocks. For example, Borjas (2003) and Borjas (2014) provide evidence that migration responses can be an important source of bias in estimates of the impact of foreign migrants on local labor market conditions.

[11]Indeed, Cutler *et al.* (2016) find mixed evidence of pro-cyclicality as well: large recessions are bad for health, while small ones may improve health. Given that the cotton shortage we study was a very large shock, our results upon adjusting for migration appear to be consistent with their findings.

nomic shocks may affect health. Section 3 describes the impact that migration can have on observed mortality rates, while Section 4 describes the empirical approach that we will use to investigate these issues. We illustrate the impact of migration bias empirically in our cotton shortage and coal boom examples, which are presented in Sections 5 and 6, respectively. Section 7 concludes.

2 Business cycles and health

The existing literature highlights a number of channels through which booms or recessions 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).[12] Accordingly, the net effects of business cycles on mortality are ambiguous ex ante.

3 Migration and the standard approach

Despite the variety of channels through which business-cycle fluctuations might impact mortality, studies that apply the standard panel-data approach have consistently

[12]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.

found evidence that, on net, health improves during recessions, although there is some evidence that this relationship may have weakened in recent years (Ruhm, 2015). A review of leading papers in this literature is provided in Appendix A.1. Following the convention in this literature, we focus on the impact of temporary changes in local economic conditions on all-age mortality. Although related issues, such as the impact of business-cycle fluctuations on age- and cause-specific mortality, and on fertility, are certainly important for better understanding the underlying mechanisms, these fall beyond the scope of this particular study.

Equation 1 provides an example of the standard estimating equation applied in this literature,

$$\ln(MR_{it}) = \beta E_{it} + X_{it}\gamma + \phi_i + \eta_t + \epsilon_{it}\,, \tag{1}$$

where MR_{it} is the mortality rate in a given location (e.g., state) i; η_t and ϕ_i are a full set of time-period and location fixed effects; E_{it} is the location's unemployment rate (or another similar variable representing local economic conditions); and X_{it} 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, E_{it}. Second, migration can affect the dependent variable MR_{it}, which includes population in the denominator. Third, migration spillovers across locations may affect results through the comparison, implicit in Eq. 1, between treated and control locations. Below we discuss each of these potential channels in more detail. Before doing so, it is worth noting that some of these channels reflect a true effect of migration on mortality rates, while in other cases, migration will only affect estimates of the mortality rate.

One channel where migration can affect the results of Eq. 1 is through the E_{it} term. In particular, one may be concerned that unemployment rates are endogenous. This is particularly true if individuals migrate towards areas that offer better eco-

nomic opportunities, as migration will not only respond to the local unemployment rate, but migration will also directly affect the unemployment rate by changing the size and composition of the labor force.[13] Our solution is to study settings where we can observe plausibly exogenous variation in the timing and spatial incidence of temporary changes in local economic conditions. Specifically, we follow the work of Miller & Urdinola (2010), and interact global price shocks with the pre-exiting spatial distribution of economic activity.

Migration can also affect estimates obtained from Eq. 1 through the interpolated annual population variable, which appears in the denominator of the mortality rate term. Specifically, by taking the log of the mortality rate as the outcome variable, Eq. 1 implicitly assumes that mortality scales one-for-one with population.[14] Migration can cause this assumption to fail in several ways.

First, if short-run population changes are not accurately captured by intercensal population estimates, then we should not expect a one-for-one mortality response. This is a purely mechanical phenomenon that arises from the fact that while the mortality rate numerator (deaths) are observed annually, direct measures of the denominator (population) are only available in census years. Indeed, in 9 out of 10 years, population must be constructed, which introduces room for error.[15] This means that while deaths in the (annually-observed) numerator may increase or decrease as a function of changes in the true at-risk population, the denominator may not move in step.

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

[14]One way to see this point is to separate the log mortality rate into log mortality and log population. Moving log population to the right-hand side of the equation shows that the specification in Eq. 1 is assuming that population is reflected in mortality with a coefficient of one.

[15]Typically, population in a non-census year is 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. However, as we discuss in Appendix A.2.2, even in modern developed countries there remains some unobserved migration that must be allocated across intercensal years. Note also that this implies that all intercensal population values are interdependent and endogenously constructed by design.

If we expect that people migrate from places with worse economic conditions to those with better ones (as we find in both of our empirical examples), and if migration is not fully captured by intercensal population estimates, then this will mechanically bias results towards finding pro-cyclical mortality.[16]

A second channel that can cause the one-for-one assumption to fail is migrant selection. Indeed, even if population is perfectly observed, the one-for-one assumption may fail if migrants are not representative of the population as a whole. For example, if healthy young workers are more likely to migrate to areas offering better economic opportunities (as is the case in the two settings that we study), 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 can partially deal with the selective sorting of population, but it cannot account for selection of migrants within age groups.

While the first two channels illustrate ways in which migration can affect the *observed* mortality rate, it is also possible for migration to influence the *true* mortality rate. One such channel is through congestion costs (e.g., by straining fixed local resources), which will increase mortality in migrant-receiving areas while reducing mortality in migrant-sending areas. Another channel by which the act of migration itself can change underlying health is by relocating people across locations with different intrinsic conditions. If, for example, people move from less healthy to healthier locations, then migration will have a direct and beneficial impact on health.[17] As with the first two channels, the congestion effects and protective effects of migration

[16]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. For further discussion of this issue, see Appendix A.2.1.

[17]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.

will also undermine the one-for-one relationship between population and mortality.

To help address the potential issues generated by assuming a one-to-one relationship between population and mortality, we consider alternative specifications that do not impose this strict relationship. In particular, we consider a range of specifications embodying assumptions that vary from a strictly proportional relationship between population and mortality to no short-run relationship. Estimating results using this range of alternative specifications will help us assess the sensitivity of estimates to assumptions about the relationship between interpolated population and mortality. We also offer a second strategy to understanding the impact of migration bias that relies on the fact that intercensal population estimates are more accurate, and less vulnerable to migration-related error, closer to census years. Thus, by looking at treatment windows close to census years and expanding these windows iteratively, we can assess the stability of the results obtained from our range of specifications as the wedge between the true (unobserved) population and our best interpolated estimates of it becomes larger. Combining these strategies with additional information about the migration response in the settings under study allows us not only to diagnose migration bias, but also to mitigate it so that we can arrive at a more accurate understanding of the relationship between business cycles and health.

It should be noted that a key feature of several of the sources of bias listed above—the endogeneity of unemployment rates, the mechanical relocation of population, migrant selectivity, and congestion effects—is that the impacts of migration in migrant-sending locations will be matched by opposite responses in migrant-receiving locations. Such spillovers violate the assumptions behind the panel-data approach in Eq. 1. Accordingly, treatment spillovers reflect yet another channel through which migration may bias results generated using the standard approach. However, this issue 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. This intuition provides the foundation for our empirical approach to

addressing spillover concerns.

An alternative approach to dealing with spillovers is to simply 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 to be experiencing dramatically different economic conditions.[18] 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.

4 Empirical approach

Our starting point is the standard estimating equation introduced by Ruhm (2000), shown in Eq. 1 above. To deal with the possibility that measures of local economic conditions might be endogenous, we modify Eq. 1 by replacing the unemployment rate E_{it} with a plausibly exogenous measure of the incidence of local economic shocks, $SHOCK_{it}$. In our examples, this variable is generated by interacting time-varying industry-specific global price shocks with measures of the initial importance of the affected industry in each location. Our modified specification is

$$\ln(MR_{it}) = \beta_1 SHOCK_{it} + \phi_i + \eta_t + \epsilon_{it} \,. \tag{2}$$

Conditional on $SHOCK_{it}$ being exogenously determined, Eq. 2 removes concerns about the measure of local economic conditions being endogenous in the presence of

[18] 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.

migration.[19] Still implicit in Eq. 2, however, is the assumption that that mortality and population move one-for-one. As discussed in the previous section, migration is likely to cause this assumption to fail, either because the true at-risk population is not well measured by intercensal population estimates, because migration is selective, or because of congestion effects. Thus, the specification in Eq. 2 is only likely to be valid under ideal circumstances, i.e., when intercensal population estimates are accurate and neither migrant selection nor congestion effects are important.

Motivated by this concern, we consider two alternative regression specifications. The first is,

$$\ln(MORT_{it}) = \tilde{\beta}_1 SHOCK_{it} + \tilde{\beta}_2 \ln(P\tilde{O}P_{it}) + \phi_i + \eta_t + \epsilon_{it} \tag{3}$$

where $MORT_{it}$ is the number of deaths in location i and period t and $P\tilde{O}P_{it}$ is interpolated population in period t.[20] The key difference in Eq. 3, relative to Eq. 2, is that the relationship between population and mortality is no longer constrained to be one-to-one. Rather, it is estimated in the data.

The estimated values of $\tilde{\beta}_2$ obtained from Eq. 3 can be a useful tool for diagnosing estimation issues created by migration. In particular, the approach shown in Eq. 2 assumes that $\tilde{\beta}_2 = 1$. If instead we estimate a $\tilde{\beta}_2$ significantly different from one, then this signals that migration is likely to be affecting results obtained using Eq. 2, say, by affecting the size or composition of the true at-risk population. One reason that we

[19] Note that we use a reduced-form approach in Eq. 2, rather than using $SHOCK_{it}$ as an instrument for unemployment. This is because we do not observe sufficient unemployment data in either of our empirical examples to enable an IV approach. Specifically, in the cotton shortage, we only observe Poor Law relief-seekers, which is not comparable to modern unemployment, and then only for a subset of years. In the case of the coal boom, the U.S. Census does not report unemployment at the county level prior to 1970, which would leave an IV approach with no pre-boom period.

[20] Note that throughout this paper, in any part of our empirical analysis that incorporates population values (i.e., in any graph or regressions that includes mortality rates or population), the population values in census years are observed, and in non-census years are estimates that are constructed using the Das Gupta interpolation method. This is not to suggest that we believe that these intercensal estimates are necessarily reliable reflections of the true population at that point in time. To the contrary, we use these constructed values to illustrate the methodological problems arising from the use of population data that may be inaccurate due to unobserved migration.

might find $\tilde{\beta}_2 < 1$ is that intercensal population estimates are inaccurate reflections of the true population at that moment in time. In that case, and if migrants move from places with worse economic conditions to places with better ones, then $\tilde{\beta}_2$ will be systematically biased downwards. Attenuation bias due to random mis-measurement of population will also push $\tilde{\beta}_2$ towards zero. We should also expect to see $\tilde{\beta}_2 < 1$ if migrants tend to be healthier than the general population. This is because such population movements will cause less-than-proportional changes in receiving-location mortality.[21]

While we view Eq. 3 as an improvement on Eq. 2 because it does not require the assumption that mortality moves one-for-one with population, it still suffers from two potential problems. First, we may be concerned that the $\ln(POP_{it})$ term is not only systematically mis-measured, but is also endogenously affected by both mortality and migration. This is because the best-available interpolation methods use the observed population in the following census as an explicit input for interpolation.

A second, and ultimately more important issue, is that the estimated relationship between population and mortality reflects a combination of the long-run relationship and the short-run relationship between these variables. Over the long run, we expect that a location's population and number of deaths will move together; as population increases, there are more people at risk of mortality, and it is reasonable to expect the number of deaths to also rise. However, in the short-run, the relationship between population and mortality may be very different from the long-run relationship. For example, if short-run population changes are driven by young workers who have very low mortality risk, then there may be little or no short-run relationship between population and mortality.[22]

[21] On the other hand, if migrants are negatively selected on health, or if congestion forces are strong, then we may observe $\tilde{\beta}_2 > 1$. For example, if congestion is important, when population moves away from a location we should expect a more-than-proportional fall in deaths because the fall in deaths generated by a reduction in congestion is added to the reduction in deaths due to the mechanical removal of population. We do not emphasize this possibility here because it does not appear to be the dominant force in our empirical examples.

[22] This issue will be exacerbated by the use of interpolated population values in non-census years.

One way to partially address issues raised by the difference between the short-run and long-run relationship between population and mortality in Eq. 3 is to include location-specific time trends. Because population growth rates tend to be fairly persistent over the time periods that we study, the inclusion of time trends in Eq. 3 will serve to absorb differences in long-run population growth, allowing $\tilde{\beta}_2$ to pick up the short-run relationship between population and mortality. To allow this, we will incorporate linear location-specific time trends in some regression specifications.

An alternative approach to dealing with the concerns we have raised about Eqs. 2-3 is to consider a third specification:

$$\ln(MORT_{it}) = \bar{\beta}_1 SHOCK_{it} + \phi_i + \eta_t + \epsilon_{it} \,. \tag{4}$$

Note that Eq. 4 differs from Eq. 3 only in that we omit the log population variable. Implicitly, this specification assumes that in the short run, there is no relationship between population and mortality, reflecting the opposite extreme to the strictly proportional relationship assumed in Eq. 2.

To summarize, Eqs. 2, 3, and 4 reflect a range of assumptions about the short-run relationship between population and mortality, and therefore generate a range of estimates of the relationship between business cycles and mortality.[23] When these

In Appendix A.2 we provide a short example illustrating this concern.

[23]One minor issue that warrants discussion is whether we should weight our regressions. When using the death rate as the dependent variable, weighting accounts for the increased precision with which this value is measured when it represents a greater number of underlying observations. Accordingly, Ruhm (2000) weights his regressions, as do some of the other papers in this literature. In this study, we typically do not weight our regressions. One reason for this is that two of the three specifications that we consider use deaths as the outcome variable, where differential precision is less of a concern. Another reason is that weighting can place considerable weight on large outliers, where effects may be different than those experienced in the average location. For example, national or local government centers may experience recessions differently than the typical town, either because of the stability of government employment or because they are more likely to receive relief in a downturn. Finally, and perhaps most important, population is likely to be systematically mis-measured in the presence of unobserved migration, which will result in our giving too much weight to observations from locations that lose population and too little weight to districts that gain population. Nevertheless, we provide weighted results for both empirical settings in the appendix, and our results are largely unaffected by weighting.

estimates are close to agreement, we can be reasonably confident in the results, even in the presence of migration. When estimates diverge substantially, this signals that assumptions about how to model underlying population and its impact on mortality play an important role in determining the results. In those cases, carefully studying the available evidence on migration patterns in the setting under study can help us choose between alternative specifications.

The last issue we need to address before moving to our empirical results is the impact of spillovers across locations. The specifications in Eqs. 2-4 rely on a comparison between a set of treated locations and a set of control locations. However, if migration in response to local economic shocks in the treated locations affects mortality in the control locations, then the control locations cannot provide a valid counterfactual. We implement a relatively simple approach for dealing with this issue, which involves identifying those locations that are likely to send or receive migrants to or from the treated locations. Since migrant flows are often related to distance, we refer to these as "nearby" locations.[24] Once we have identified nearby locations that are likely to share migrant flows with the treated locations, we can estimate the impact of the shock on both the treated and nearby locations by comparing these to a third set of uncontaminated control locations.

5 The Lancashire cotton shortage

Having laid out our basic empirical approach, we now turn to the first of our two examples: a sharp, severe, and short-lived downturn in the cotton textile-producing region of Britain in the 19th Century.[25] This historical example is useful because

[24]Of course, in many, and particularly, in modern settings, geographic distance may not be the dimension over which migration takes place. For instance, migrant-sending regions may be identified on the basis of occupational similarity, social or transportation networks, etc.

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

it allows us to cleanly identify the spatial and temporal incidence of a short-term adverse shock without having to rely on potentially endogenous measures of local economic conditions, such as unemployment. In addition, high migration costs in this setting make it easier for us to track and demonstrate the impact of migrant spillovers, which occur mainly between geographically proximate locations. Below we provide a brief description of this setting and the data we use. Further details are available in Appendix B.

5.1 Background and Data

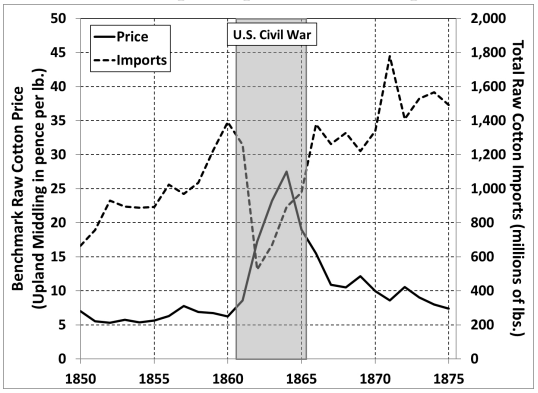
The cotton textile industry was the largest and most important industrial sector of the British economy during the 19th century. For historical reasons dating to the 1700s, British cotton textile production was geographically concentrated in the Northwest counties of Lancashire and Cheshire, which held over 80% of the cotton textile workers in England & Wales in 1861.[26] The industry was entirely reliant on imported raw cotton and, in the run-up to U.S. Civil War, 70% of these imports came 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.[27] The right-hand panel shows the spatial distribution of the British cotton textile industry on the eve of the U.S. Civil War. Additional information, in Appendix B.1, shows that the direct impact of the U.S. Civil War on the British economy was largely confined to

[26] Crafts & Wolf (2014) suggest that the main factor 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. Calculation based on data collected by the authors from the 1861 Census of Population reports.

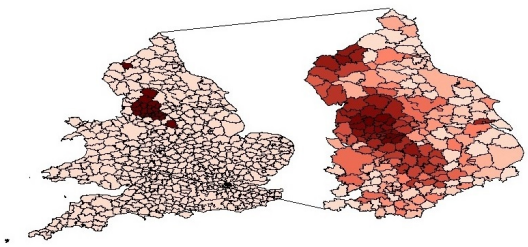
[27] Appendix Figure 11 shows British firms' raw cotton consumption and variable operating costs (excluding cotton), good proxies for industry output. These show a sharp drop in production and factor payments during the 1861-1865 period equal to roughly a 50% reduction compared to pre-war output levels. 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 1: Cotton prices, imports, and spatial distribution of cotton textile industry

British import quantities and prices



Spatial distribution of cotton textiles



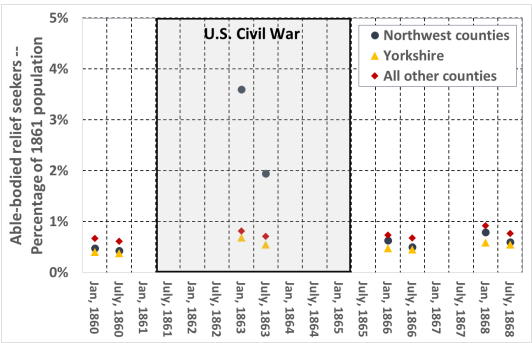
Import data from Mitchell (1988). Price data, from Mitchell & Deane (1962), are for the benchmark Upland Middling variety. Data on the geography of the cotton textile industry the are calculated from the 1851 Census of Population. Shaded in the map of England & Wales are districts with over 10% of employment 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.

the cotton textile sector.

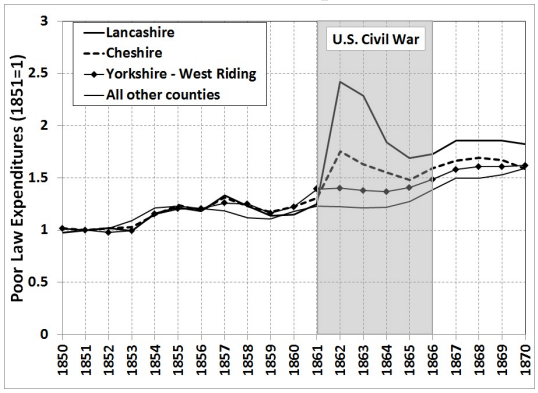
Both contemporary reports on public assistance and data on relief-seekers suggest that the shock had a substantial and geographically-concentrated impact on workers. One reflection of these effects is presented in Figure 2. The left-hand panel of this figure describes the number of able-bodied relief-seekers 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, 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 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, work-

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)).

ers in the affected areas employed a variety of coping mechanisms, including running down savings, pawning valuables and furniture, 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 5.2.

To analyze the impact of the U.S. Civil War on migration and mortality in the cotton-textile regions of Britain, we draw on population data from the British Census for every decade from 1851-1881, and mortality data over the period 1851-1865, the latter taken from reports produced annually by the Registrar General's office.[28] Our analysis is conducted at the district level using 539 geographically-consistent districts covering all of England & Wales over this period.[29]

To identify treated districts, we use information on the industrial composition of

[28]Further details on the data are available in Appendix B.3.

[29]The available data cover around 630 districts in each year, but some districts experienced boundary changes over time. To obtain geographically consistent districts, 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 539 consistent districts in the main analysis.

employment in each district on the eve of the U.S. Civil War, based on occupation data from the 1851 Census. In the main analysis, we define as cotton (i.e., "treated") districts those districts with over 10% of employment in cotton textile production.[30] Our study covers the years 1851-1865, with the U.S. Civil War years (1861-1865) defining the cotton shortage period. We focus only on the periods before and during the shortage because of concerns that the post-shortage years may have been influenced by persistent effects of the shortage.[31] Wherever intercensal population estimates are used, these are generated using Das Gupta interpolation, which accounts for annual changes in population due to observed births and deaths within each district, and distributes any residual population change at the decade level (the error of closure, a measure of implied net migration) smoothly across intercensal years.[32]

5.2 Migration

Contemporary accounts indicate that out-migration occurred in response to the cotton shortage, but the extent of these flows remains debated. Thus, 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

[30]In robustness checks we also generate results using a continuous measure of local cotton textile employment shares. We focus our main results on a discrete cutoff both in order to make interpretation easier, and in order to reduce the impact of measurement error. In particular, within the set of major cotton textile districts, variation in cotton textile employment share does not necessarily correspond to variation in shock intensity because the impact of the cotton shortage depended in part on the type of cotton textile products produced in each district. For example, districts that produced finer fabrics, where raw cotton was a smaller portion of total cost, were less affected by the cotton shortage.

[31]For example, we know that one form of relief during the cotton shortage was public works employment, which was mainly focused on projects improving health or transportation infrastructure. These projects were undertaken relatively late in the study period, so they are unlikely to have substantially affected mortality during the U.S. Civil War. Furthermore, we omit the post-1865 recovery period so as to avoid comparisons that generate spurious results driven by mean reversion in population, i.e., by large-scale return migration during the recovery period. See Appendix Section B.3 for more.

[32]The Das Gupta method is the same approach used by the U.S. Census Bureau. A more detailed discussion of this approach, and its limitations, appears in Appendix A.2.2.

change in 1851-1861 (the decade preceding the downturn).[33] 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 a rise in out-migration from those districts, or a reduction in in-migration. Indeed, additional evidence on the birthplaces of Northwest county residents, which we present in Appendix B.4, suggests that both an increase in out-migration and a reduction in in-migration occurred during this period.[34] Second, we observe an acceleration in population growth in nearby districts, which 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.

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.[35] 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.[36]

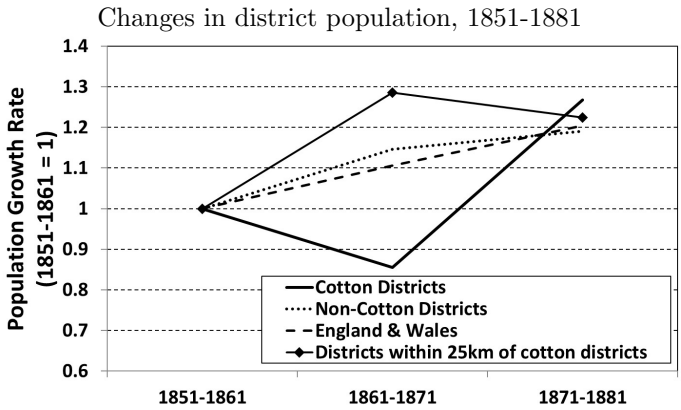
[33] The 1861 Census was collected in April of that year, before the U.S. Civil War had any substantial impact. As a result, this should be thought of as a clean pre-war population observation.

[34] Appendix B.4 also provides evidence of a sharp and short-lived increase in emigration from Britain during this period.

[35] 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.

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

Figure 3: Migration response to the cotton shortage

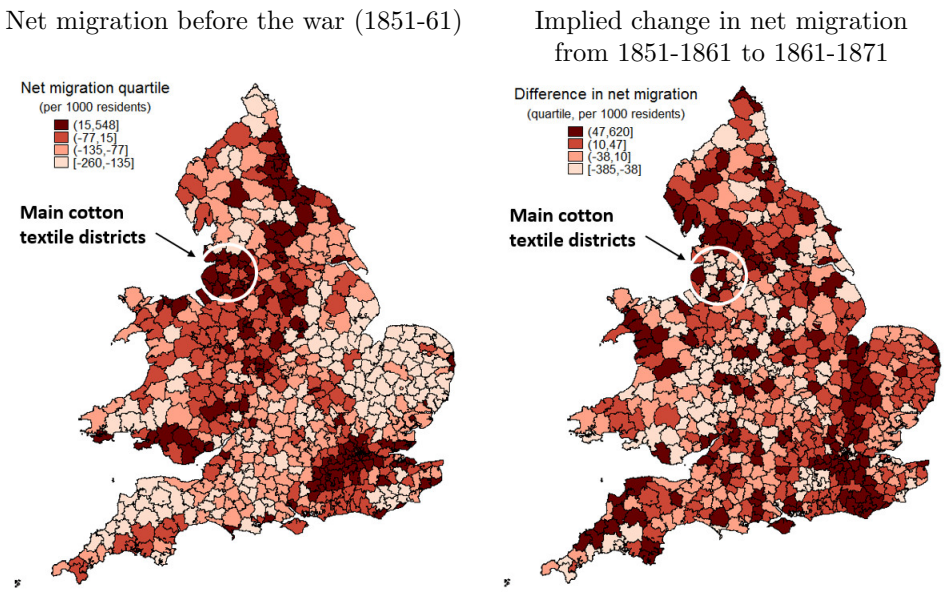


This graph describes the change in population for all cotton districts, all non-cotton districts, all districts in England & Wales, and all non-cotton districts within 25km of a cotton district using Census data. Cotton districts are defined as those districts with more than 10% of employment in cotton 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.

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.[37] 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,

[37]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



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.

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.

Although selectivity will become much more important in our second empirical example, there is also some evidence of selective migration away from the cotton textile districts during the U.S. Civil War period. In particular, in Appendix B.4 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, even though this

group was important, a substantial amount of migration likely occurred among other segments of the population as well.

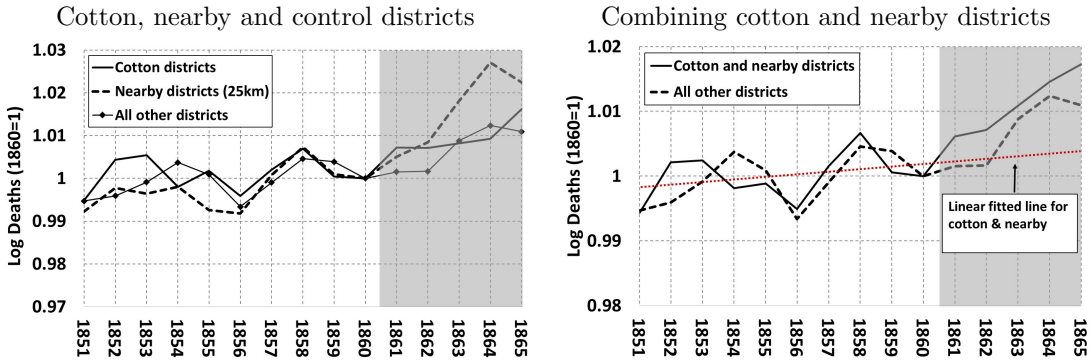
5.3 Mortality

Having established that the cotton shortage induced a substantial and spatially concentrated migratory response, we next analyze the impact of this event on mortality patterns. As a first step, Figure 5 describes the evolution of log mortality across the study period (normalized such that 1860=1).[38] In the left-hand panel we separate cotton districts, nearby districts (those within 25 km), and all other districts. There are three notable features here. First, in the pre-shortage period, mortality in the three groups track each other 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.

Since our evidence suggests that most migration out of cotton districts went to nearby areas, combining the migrant-sending and migrant-receiving districts can help us get a sense of the aggregate mortality impact of the downturn. Recall that most of the channels through which migration generates bias, discussed in Section 3, 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.

[38] Graphs of the mortality rate are available in Appendix B.5.

Figure 5: Mortality effects of 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.

Next, we analyze these patterns econometrically. Table 1 present our results, with coefficients in Columns 1-3 corresponding to the specifications in Eqs. 2-4, respectively. In these regressions, standard errors are clustered by district to adjust for serial correlation.[39] In Column 1, we present results using the log mortality rate as the dependent variable, as in the standard approach. Here we find evidence that the mortality rate fell during the recession, a finding that is consistent with the procyclical mortality results obtained in the existing literature. In Column 2, we move the population denominator to the right-hand side of the equation. This weakens the relationship between the cotton shock and mortality. Moreover, the estimated relationship between population and mortality is well below (and statistically different from) one, suggesting that the one-to-one relationship between population and mortality implied by the specification in Column 1 may be inaccurate. Column 3 presents results based on the specification in Eq. 4, which assumes no short-run rela-

[39]Clustering standard errors by district is similar to the approach used in most studies in this literature. In Appendix table 12, we present additional results including those clustered by county (a larger geographic unit), and those adjusting simultaneously for serial and spatial correlation across locations. Results are qualitatively unaffected as all three approaches produce similar standard errors.

tionship between population and mortality. Here the sign of the relationship between the cotton shock and mortality becomes positive.

Table 1: Preliminary estimates of the mortality effect of the cotton shortage

Dependent variable:	Ln(MORT. RATE)	Ln(MORTALITY)	Ln(MORTALITY)
	(1)	(2)	(3)
Cotton district × shortage	-0.0276**	-0.0221*	0.0155
	(0.0108)	(0.0114)	(0.0189)
Ln(Pop)		0.872***	
		(0.0283)	
Observations	8,085	8,085	8,085
R-squared	0.137	0.334	0.195

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by district. Data cover 539 districts from 1851-1865. All regressions include district fixed effects and year effects. Shock period is 1861-1865. Cotton districts are defined as those with a cotton employment share greater than 10%.

The results in Table 1 show that the estimated relationship between population and mortality is sensitive to assumptions about the relationship between short-run population changes and mortality. To provide substantive evidence on the true effect of the cotton shortage on mortality, we need a way to choose between these alternative sets of assumptions. Before we tackle that issue, however, it is useful to deal with a second concern that is not addressed in Table 1: the impact of spillovers across treated and control districts due to migration.

In order to deal with the issue of migrant spillovers, we include a term to capture the impact of migration from cotton textile districts to other nearby areas. Our primary approach to dealing with the impact of migration on nearby districts is to construct a variable, $NEARcot_j^{0-25}$, which reflects, for each non-cotton district, the amount of cotton textile employment in other districts within 25 km (or alternative distance windows). Specifically, we calculate,

$$NEARcot_j^{0-25} = \ln\left(\sum_{i \neq j} 1[d(i,j) < 25km] * (COT_i + 1) * 1[COTDIST_j = 0]\right) \quad (5)$$

where $1[d(i,j) < 25km]$ is an indicator variable that equals one when the distance between districts i and j is less than 25 km and COT_i is cotton textile employment in district i.[40] The $1[COTDIST_j = 0]$ component of this equation ensures that we only cacluate nearby cotton exposure for non-cotton districts, defined as those with 10% or less of employment in cotton textile production.[41] Finally, because nearby cotton textile employment is zero for districts that were far from cotton textile-producing areas, we add one before taking logs. We then include this employment-weighted distance term, interacted with an indicator for the shock period, in our regressions. Including this variable effectively removes indirectly-treated districts from the control group, allowing us to separately evaluate the direct and indirect mortality effects against a clean counterfactual.[42]

Table 2 presents results that account for the impact of cotton textile employment in nearby districts. Again, we display results that correspond to each of our three estimating equations in Columns 1-3. The results in Table 2 provide evidence that mortality increased in nearby districts during the cotton shortage, consistent with the relocation of migrants from the cotton to nearby districts. Once we account

[40] To calculate the distance between any pair of districts, we collect the latitude and longitude of the main town or district seat for each district, which we call the district center. For a small number of very rural districts, we use the geographic center of the district.

[41] We include the $1[COTDIST_j = 0]$ term because we expect that the impact of the recession in nearby districts will influence migration into non-cotton districts, but that this is unlikely to influence net migration into other cotton districts. Contemporary evidence consistently shows that those leaving cotton districts were not migrating to other major cotton-producing areas. See, e.g., Arnold (1864).

[42] In Appendix Table 8, we include employment-weighted distance terms for each of the following bands: 0-25 km, 25-50 km, 50-75 km, and 75-100 km. The inclusion of these variables highlights that the majority of systematic spillovers are captured within 25 km. We also consider alternative approaches to measuring "nearby" districts, for instance those that use unweighted (i.e., indicator) variables for districts within particular distance bands of the major cotton textile districts, or those that discount a continuous measure of distance. These alternatives deliver similar results.

for migration spillovers, the results in Column 3 suggest that the recession led to a statistically significant increase in mortality, while the results in Column 1 continue to suggest that the recession reduced mortality. Again, it is clear that the results are highly sensitive to assumptions about the relationship between population and mortality.[43]

Table 2: Accounting for spillovers to nearby districts

Dependent variable:	Ln(MORT. RATE)	Ln(MORTALITY)	Ln(MORTALITY)
	(1)	(2)	(3)
Cotton district × shortage	-0.0229**	-0.0125	0.0494**
	(0.0114)	(0.0123)	(0.0195)
Nearby cotton emp. × shortage	0.00154	0.00287**	0.0109***
	(0.00117)	(0.00122)	(0.00187)
Ln(Pop)		0.857***	
		(0.0293)	
Observations	8,085	8,085	8,085
R-squared	0.137	0.335	0.207

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by district. Data cover 539 districts from 1851-1865. All regressions include district fixed effects and year effects. Shock period is 1861-1865. Cotton districts are defined as those with a cotton employment share greater than 10%.

Given the sensitivity of these results, we consider two approaches to help choose among the alternative assumptions. First, we review the evidence on migration during the cotton shortage to assess the validity of the assumptions underlying each of these specifications. Second, we examine how results change when we use narrower shock windows that are closer to the census year, and so, less vulnerable to population mis-measurement due to migration. In this case, the census year (1861) falls at the beginning of the cotton shortage period.[44]

[43] As mentioned previously, we choose not to weight our main results because we are concerned that population is mis-measured. Nevertheless, weighted results are available in Appendix Table 10. Weighting does not qualitatively affect the results.

[44] Note that although our shock begins in the census year 1861, the usefulness of the census-window strategy does not in fact rely on the shock being timed close to a census year. Rather, it exploits

The evidence in Section 5.2 shows that, relative to the previous decade, cotton textile districts experienced a substantial reduction in population growth during the the downturn. Despite this reduction, however, population in most cotton districts increased over the 1861-1871 decade. This fact is important because it means that intercensal population estimates will imply positive growth across all years, which is at odds with the available evidence. The available evidence suggests that people left the cotton districts during the cotton shortage, but that population in these areas rebounded strongly from 1866-1871, as the cotton industry recovered.[45] When this pattern is combined with the one-to-one relationship between population and mortality embedded in the specification in Column 1 of Table 2, the result will be to bias results towards finding that the mortality rate decreased during the shortage, simply because the interpolated population denominator overstates the true population. Under these population growth conditions (a point that we discuss further in Appendix A.2), the specification in Column 3 of Table 2 offers a more suitable approach to modeling mortality. Accordingly, the results in Column 3 are our preferred results.

Another way to differentiate between the alternative results in Table 2 is to study their robustness. In particular, in Appendix Table 9 we consider how these specifications react to the inclusion of district time trends, which are included in some specifications in the literature following Ruhm. The results in Appendix Table 9 show that once district time trends are included, all three of our alternative specifications deliver results that are very similar to those produced without trends in Column 3 of Table 2.[46] This reinforces our choice of Eq. 4 as our preferred specification.

---

the declining accuracy of population measures as we move further and further away from a census year, irrespective of when the shock occurs during the intercensal period. As such, this approach is suitable for use even when shocks do not align closely with census years.

[45]For example, in the 1871 Census, the Registrar for Salford, near Manchester, an important cotton textile area, wrote that the population of the district experienced an "Increase due to the revival of trade since the conclusion of the American war..."

[46]That all three specifications deliver very similar results once time trends are included suggests that most of the differences across specifications in this context are due to long-run population trends.

As a final test, we exploit the fact that population estimates will be more accurate, and less subject to migration-related mis-measurement, closer to census years. Table 3 presents additional results showing how the estimates obtained from each specification change as we iteratively expand the shock period window to include years further from the census year 1861. The top panel presents results using the log mortality rate as the dependent variable, as in Column 1 of Table 2. The middle panel corresponds to Column 2 of Table 2, where the outcome variable is log mortality, and log population is on the right-hand side. The bottom panel presents results with log mortality, as in Column 3 of Table 2.

There are two important features to note in Table 3. First, near the census year, when population estimates are more accurate, all three specifications deliver very similar results: the downturn results in higher mortality. Second, as we consider larger shock windows which include years further from the census year, the results in Panel C consistently show that the cotton shortage increased cotton-district mortality. In contrast, the results in Panels A and B are unstable: in Panel A, they switch from positive and statistically significant to negative and statistically significant, while in Panel B, they switch from positive and statistically significant to negative and insignificant. Both of these patterns provide support for our preferred specification, that in Panel C, as it is the the least vulnerable to population mis-measurement.

Overall, the results in Tables 1-3 illustrate the sensitivity of the panel-data approach to assumptions about the relationship between population and mortality in the short run. However, by using additional information on migration patterns, and by looking closer to census years, when population is less vulnerable to mis-measurement, we can select a specification that delivers consistent results that more accurately reflect the business cycle-mortality relationship. In our setting, that specification is Eq. 4, and under this specification, results indicate that mortality in the cotton districts increased during the downturn.[47] Importantly, this finding is corroborated by the

[47]The robustness of this result is assessed in Appendix 11. There we make several sample restric-

Table 3: Results for various windows starting in 1861

Shock years:	1861	1861-62	1861-63	1861-64	1861-65
	(1)	(2)	(3)	(4)	(5)
			Panel A – DV: Log Mortality Rate		
Cotton × shortage	0.0246*	0.0334***	0.000559	-0.0212*	-0.0229**
	(0.0146)	(0.0116)	(0.00980)	(0.0115)	(0.0114)
Nearby × shortage	-2.77e-05	0.00152	0.00134	0.00205*	0.00154
	(0.00155)	(0.00139)	(0.00133)	(0.00123)	(0.00117)
			Panel B – DV: Log Mortality		
Cotton × shortage	0.0405**	0.0482***	0.0154	-0.00887	-0.0125
	(0.0158)	(0.0122)	(0.0100)	(0.0120)	(0.0123)
Nearby × shortage	0.00188	0.00333**	0.00319**	0.00361***	0.00287**
	(0.00156)	(0.00140)	(0.00135)	(0.00128)	(0.00122)
Ln(Pop)	0.727***	0.761***	0.774***	0.822***	0.857***
	(0.0398)	(0.0359)	(0.0334)	(0.0308)	(0.0293)
			Panel C – DV: Log Mortality		
Cotton × shortage	0.0829***	0.0953***	0.0660***	0.0478***	0.0494**
	(0.0203)	(0.0168)	(0.0152)	(0.0178)	(0.0195)
Nearby × shortage	0.00695***	0.00909***	0.00949***	0.0108***	0.0109***
	(0.00176)	(0.00168)	(0.00173)	(0.00182)	(0.00187)

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by district. All regressions include district fixed effects and year effects. In each column, the sample runs from 1851 to the last year of the specified shock window. N=5,929 in Column 1; 6,468 in Column 2; 7,007 in Column 3; 7,546 in Column 4; 8,085 in Column 5.

patterns shown in Figure 5. In addition to the downturn's direct effects on cotton districts, we also find evidence that mortality increased in districts proximate to cotton textile areas, consistent with the impact of the relocation of population to these areas. It is important to recognize, however, that the increase in mortality in nearby districts does not necessarily imply that true mortality rates rose in those areas; this result may simply reflect an increase in the size of the at-risk population.

6 The Appalachian coal boom

Next we consider a more recent setting in which we can identify a spatially concentrated shock to local economic conditions. In particular, following Black *et al.* (2005) (hereafter BMS), we study the impact of the commodity price boom that affected the Appalachian coal-mining region of the U.S. between 1970 and 1977.[48] This example allows us to consider the impact of migration bias on the relationship between temporary local economic shocks and mortality in a setting that is more similar to most of the previous literature.[49] As before, we only briefly describe the empirical setting; further details can be found in Black *et al.* (2005).

6.1 Background and Data

The Appalachian coal boom of the 1970s resulted from a combination of regulatory changes that took place in 1969, and the rise in oil prices due to the OPEC oil embargo

tions to ensure that the treatment and control districts are as similar as possible. We also consider alternative measures of vulnerability to the cotton price shock. Furthermore, we test whether our findings on total mortality are driven by fertility responses to the downturn. We a positive but statistically insignificant fertility response. Moreover, the magnitude of the response accounts for less than 10% of the total increase in deaths. For more, see Appendix Table 13 and the related discussion.

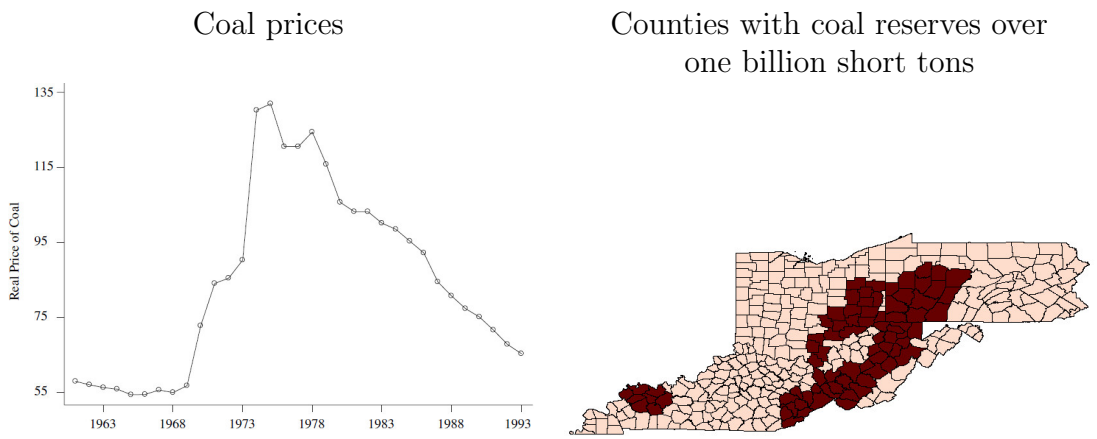
[48]In a related paper, Black *et al.* (2002) study the impact of the coal boom and bust on participation in Disability Insurance. They find that disability program participation fell during the boom and then increased during the bust.

[49]In fact, the time period covered by this example partially overlaps with the period covered in the original Ruhm (2000) paper.

in 1973-74.[50] The left-hand panel of Figure 6 presents data from BMS showing the evolution of coal prices across this period. Coal prices rose sharply starting in 1970 and then fell after 1982 as competition from western coal mines increased. BMS show that these price increases were matched by a similar increase in coal industry earnings. Accordingly, BMS divide their study into three periods—the boom (1970-1977), a stable peak period (1978-1982), and the bust (1983-1989). For consistency, we adopt the same definitions.

As in our first example, the spatial distribution of the economic impact was highly concentrated. In this case, the presence of coal reserves within a county determined the impact of the shock. The right-hand panel of Figure 6 provides a map showing the distribution of coal reserves across the four states studied by BMS: Kentucky, Ohio, Pennsylvania, and West Virginia. This figure shows that the main coal reserves were located in a band of counties stretching from southern Pennsylvania, across West Virginia and Ohio, and into Kentucky.

Figure 6: Coal prices and the spatial distribution of coal reserves



Left-hand panel is from Black *et al.* (2005). The right-hand panel is produced using the USCOAL database provided by the U.S. Geological Survey using data from surveys conducted mainly in the 1960s.

[50]See Black *et al.* (2005). The Appalachian region also benefited from new roads during this period as a result of the Appalachian Development Highway Program (Jaworkski & Kitchens, 2016), but this affected the region as a whole rather than just those areas with coal resources.

The main data used in our analysis come from the the decennial Census of Population, and from the annual records of deaths and births given by Bailey *et al.* (2016) and the Centers for Disease Control. Focusing on the four states studied by BMS, we draw on data from the U.S. Geological Survey's USCOAL database in order to define coal counties as those with over one billion short tons of reserves.[51] There are 72 counties that satisfy this criterion in the four analysis states. Our control counties are those with coal reserves below the cutoff, and following BMS, we limit the set of control counties to those counties in the four analysis states with populations ranging from 8,000 to 225,000. This population criterion is applied to control counties in order to obtain a set that is similar to the fairly rural coal counties, and results in our dropping the main urban counties in these states from the control group. This yields 202 non-coal control counties. In robustness exercises, we consider a larger set of control counties including all non-coal counties in the four states studied by BMS as well as additional counties in a set of four further neighboring states.

We define the coal boom as spanning the period 1970-1977, and we limit our analysis to the pre-boom (1950-1969) and boom (1970-1977) periods. This parallels the approach used in our first example, and also avoids questions about whether the period after 1977 should be treated as a downturn or merely the end of a boom.[52] As in the cotton shortage example, intercensal population estimates are generated using the Das Gupta interpolation method.[53]

[51]This definition differs slightly from the approach used by BMS. BMS identify coal counties as those where more than 10% of earnings came from the coal industry in 1969, a criterion that yields 32 coal counties. We prefer to use coal reserves rather than earnings because reserves are less likely to reflect other factors affecting the counties in the period just before the coal boom. However, in the end, both approaches yield fairly similar results.

[52]In Appendix C.3, however, we do explore alternative definitions of the boom period.

[53]These account for births and deaths in each county-year, and distribute decade-wide implied net migration across intercensal years. It is worth noting that for reasons of data availability, these interpolations use slightly less information than estimates produced by the census today, the latter of which take advantage of additional information on migration from IRS and Medicare data. However, in Appendix Figure 18 we provide evidence that in the 1970s, when we have both our own intercensal estimates and the published census interpolations (the latter of which are partially adjusted for migration), these two series are quite similar. Note that the Census does not provide annual county-level population estimates for years prior to 1970. There is some suggestion that county-level SEER

6.2 Migration

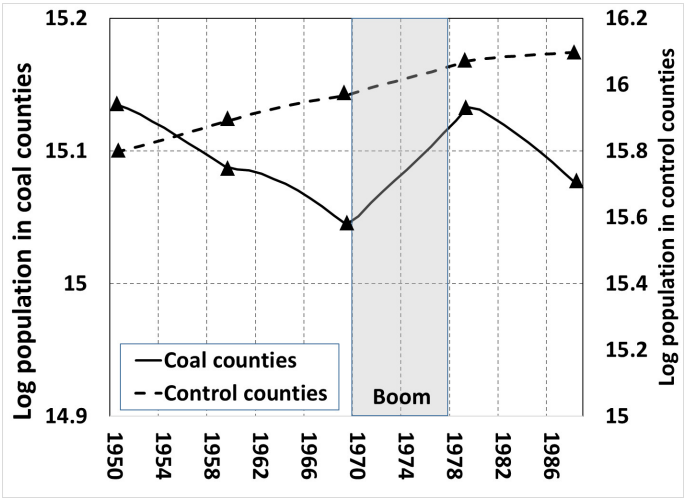
Figure 7 describes the migration response generated by the coal boom and bust in the coal and control counties. Population in the coal counties was shrinking in the decades before the coal boom, but increased substantially from 1970-1980. Once the boom ended, population again started to decline. These patterns suggest that how we deal with differential initial population growth trends and longer-run migration flows will play an important role in our results.[54]

Population trends in control counties do not appear to respond to the coal boom. This suggests that either the control counties do not share migration flows with the coal counties, or that these flows are so small relative to the size of the migrant-sending regions that they have little impact in these locations. There are several likely explanations for this pattern. First, our control counties do not include major urban areas, which may have been the preferred destination of Appalachian migrants before the coal boom, and a source of return migrants during the boom. Second, coal counties were relatively small in terms of population, so even if the migration inflows had a large effect in those counties, they may not have had any meaningful impact elsewhere. Finally, the low costs of migration in this setting may have caused many of those returning to the coal counties during the boom decade to have originated form a set of locations that was much more spatially diffuse than in our cotton

(Surveillance, Epidemiology, and End Results Program) population data published by the National Cancer Institute and used in studies like Phipps *et al.* (2005) may be slightly more reliable than published U.S. Census estimates. However, these are only available beginning in 1969, and so are unsuitable for our analysis. What's more, even if this data were available further back in time, it is unlikely that the sorts of small-scale improvements it makes on Census data would give us substantially better estimates of population, given that there is already negligible difference between the population estimates that we produce and those published by the U.S. Census Bureau, despite the large difference in the quality of inputs (ours make no observed-migration corrections, while the Census makes many such adjustments).

[54]These migration patterns are consistent with the results of BMS, who also find evidence of an increase in employment and population growth in the coal counties from 1970-1980, followed by a reduction from 1980-1990. Similar results were obtained by Carrington (1996) in a study of the impact of the temporary boom in the Alaskan economy that occurred during the construction of the Trans-Alaska Pipeline.

Figure 7: Migration response to the coal boom



This figure describes the change in population for all coal counties and the control counties in the four states in the original BMS study. Coal counties are defined as those in which more than 10% of income came from coal in 1969. Black triangles indicate census observations. The remaining intercensal population estimates are constructed using the Das Gupta interpolation method.

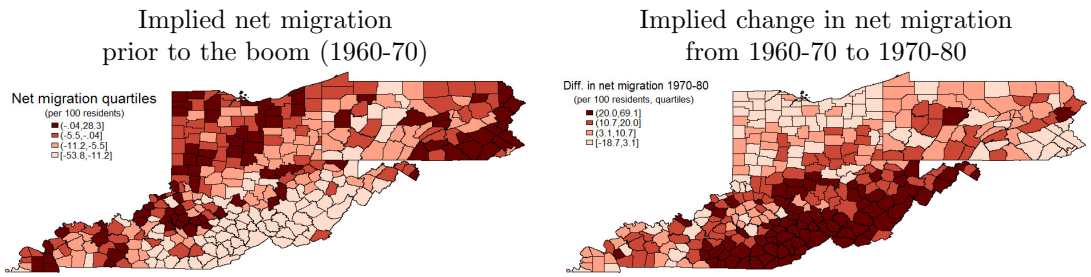
example. Regardless of cause, the fact that we find no large population response in the control counties has important implications for our analysis, since it implies that spillovers between migrant-sending and migrant-receiving counties are not likely to be an important source of bias in this setting.

Figure 8 presents maps of implied net migration by county in the decade before the coal boom, and the difference in implied net migration during the coal boom decade relative to the preceding one.[55] In the left-hand panel we see that coal districts were experiencing strong net out-migration in the decade before the coal boom. Similar patterns are also observed in 1950-60, as well as in the decade following the coal boom (see Appendix C.2). The right-hand panel, however, indicates a reversal in this trend during the 1970s boom decade. Furthermore, in Figure 8 we see no evidence that counties proximate to coal counties experienced a stark reversal in migration patterns

[55]Maps including four additional adjacent states are provided in Appendix Figures 19 and 7.

during the boom period, as we had observed for nearby districts in the cotton shortage example. This suggests that spillovers between coal and nearby counties are unlikely to be an important concern in this particular setting, consistent with the results in Figure 7.

Figure 8: Maps of migration before and during the coal boom



The left-hand panel describes implied net migration over the decade from 1960-70. Lighter colors indicate counties experiencing net out-migration. The right-hand panel describes the change in net migration from the 1960-70 decade to the 1970-80 decade. Darker colors indicate districts that experienced an increase in net in-migration in the 1970s relative to the previous decade. 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.

In Appendix Figure 21, we present evidence suggesting that migration to coal counties was selective. During the boom, population in coal counties increased by about 342,000, with the number of residents aged 25-34 increasing by 210,000 and the number of residents aged 15-24 increasing by 101,000. These increases, especially the increase among 25-34 year-olds, were much larger in percentage terms than the change in population in the control counties.[56] That the vast majority of new residents were young adults has important implications for our study because young adults have low

[56]Similar results are also reported by BMS, who find that most of the increase in population in the coal counties from 1970-1980 was concentrated among those aged 20-29, particularly men. In addition, Black *et al.* (2013) provide evidence that fertility increased in the coal counties during the boom, a pattern which Kearney & Wilson (2017) suggest was confined to marital fertility. Despite this, Figure 21 suggests that the impact of fertility changes was small relative to the population changes among young workers.

mortality rates relative to those in other age groups. Thus, although the population in-flows were large, they may contribute little to total mortality.[57]

6.3 Mortality

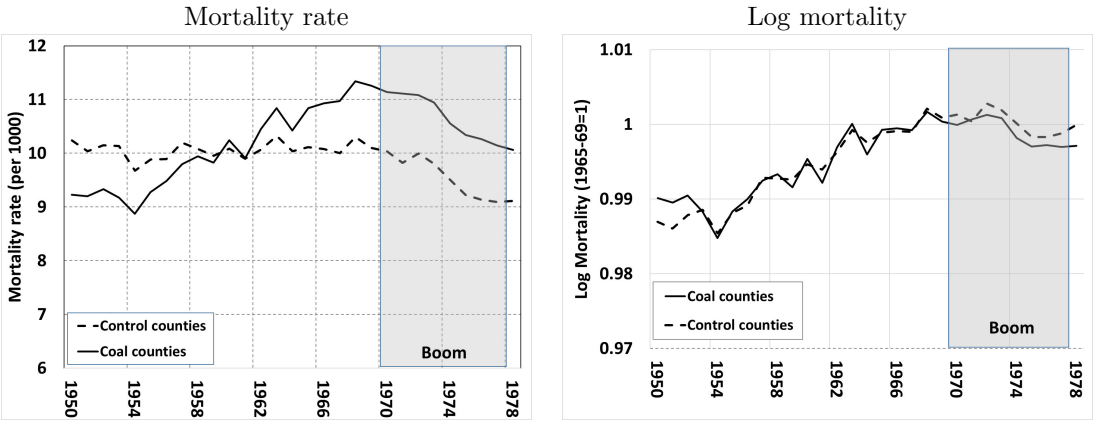
As in our cotton example, we begin by providing graphical evidence on the impact of the boom—and the migration it induced—on mortality. The left-hand panel of Figure 9 shows that mortality rates in the coal counties were rising prior to the boom, and fell rapidly during the coal boom period. While the mortality rate was higher in the coal counties during the coal boom, this figure does not give the impression that health became worse in the coal counties during this period. Rather, the mortality rate in the coal counties appears to evolve roughly in parallel to the mortality rate in control counties during the boom.

The right-hand panel of Figure 9 plots log mortality in the coal and control counties, normalized for comparability. Here we see that log mortality in both sets of counties track each other very closely from about 1955 through the main coal boom period. If anything, it appears that mortality in the coal counties may have dropped relative to the control counties during the boom. This is somewhat surprising, given evidence that population was flowing into the coal counties during the boom.

The striking difference between the left- and right-hand panels of Figure 9 tells us that *essentially all* of the mortality-rate differences observed in these two groups of counties from the mid-1950s through 1977 are due to population movements. Comparing the mortality-rate movements illustrated in Figure 9 to the population movements in Figure 7 reinforces this result: as population in the coal counties falls in the 1950s and 1960s, the mortality rate increases, and as population rebounds in the 1970s, the

[57]In particular, in the four states used in our main analysis, the mortality rate for those aged 20-24 was 1.40 deaths per thousand in 1970 and for people aged 25-34 it was 1.44 per thousand. In contrast, the average death rate across the entire population in that year was 10.24 deaths per thousand, with the vast majority of deaths concentrated among those over age 55. See Appendix Table 15 for more.

Figure 9: Mortality effects of the coal boom and bust



Mortality data from Bailey *et al.* (2016). Population data for years ending in zero are observed, and are taken from the Census. Population values in the intervening years are annualized using Das Gupta interpolation.

mortality rate falls. Despite all this, deaths in the coal and control counties trend together throughout.[58]

That the increase in population in the coal counties during the boom was not matched by an increase in deaths suggests that the migrants entering these counties during the boom were less likely to die than the average resident of these counties on the eve of the boom. This pattern is consistent with the previously discussed evidence that most migrants were young adults. Because these migrants had very low mortality risk, they had little effect on the number of deaths, but their presence in the population did make it *appear* that the overall mortality *rate* had declined. Overall, the patterns in Figure 9 suggest that the economic shock had little impact on total mortality, and that changes in the observed mortality rate primarily reflect changes in the size and composition of the population.

Next, we examine the results obtained when applying the three alternative econo-

[58]In Appendix Figure 22 we present additional evidence comparing the evolution of mortality within different age groups in coal and control counties. These results are similar to the all-age results shown in Figure 9—that is, the evolution of mortality looks similar in the treatment and control counties. Note that we cannot run our analysis reliably using age-specific data, since this data is only available disaggregated by age starting in 1968.

metric approaches discussed in Section 4. To illustrate the problems caused by the migration-driven violation of the parallel trends assumption when using the log mortality rate as the dependent variable (as evidenced in the left panel of Figure 9), we show results both without and with county time trends. Our results are presented in Table 4. The top panel presents estimates from our three alternative specifications without county time trends, while the bottom panel presents estimates with time trends. Standard errors are clustered by county.[59]

As in the cotton example, the results in Table 4 are highly sensitive to the assumptions made about how to model the relationship between population and mortality in the short run. Results change substantially both as we move across Columns 1 to 3, and as we include or exclude county time trends. The results obtained using the mortality rate as the dependent variable, in Columns 1a and 1b, are particularly sensitive to the inclusion of time trends. In contrast, results in Columns 3a and 3b are less sensitive to the inclusion of time trends. It is also worth noting that in the results in Columns 2a and 2b, the coefficient on the population term is well below one, suggesting that the one-to-one relationship between population and mortality assumed by the specification used in Columns 1a and 1b is likely to be violated in this setting. Indeed, this coefficient is consistent with the boom-time in-flows of healthy workers who contributed less than proportionally to coal-county deaths.

Given the patterns described in Figure 9 and the evidence on migration, our assessment is that the results in Columns 1a and 1b of Table 4 are likely to provide misleading results that are driven primarily by changes in the size and composition of the at-risk population. The results in Columns 2 and 3 are likely to be more reliable. These estimates suggest that the coal boom had little effect on total mortality.

As in our cotton example, our shock begins in a census year, which provides a convenient way to examine the stability of the results as population measures be-

[59]Results using spatial standard errors are available in Appendix Table 16. These deliver similar results.

Table 4: Estimated mortality effects of the coal boom

Dependent variable:	Ln(MR)	Ln(MORT)	Ln(MORT)
Panel A: Without county time trends			
	(1a)	(2a)	(3a)
Coal county × boom	0.096***	0.010	-0.023*
	(0.017)	(0.011)	(0.012)
Ln(Pop)		0.278***	
		(0.0230)	
Panel B: With county time trends			
	(1b)	(2b)	(3b)
Coal county × boom	-0.038***	-0.003	0.009
	(0.010)	(0.009)	(0.010)
Ln(Pop)		0.267***	
		(0.043)	

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by county. All regressions include county fixed effects and year effects. N=7,672.

come more vulnerable to migration-related error. Specifically, examining how results change as we expand the shock window away from the census year provides additional evidence on how to choose between specifications. This is done in Table 5 for various windows starting in 1970. Panel A presents results using the log mortality rate as the dependent variable as per the standard approach, Panel B presents results using log mortality as the dependent variable while still controlling for log population, and Panel C presents results with log mortality as the dependent variable, without including log population.

There are two important patterns to note in Table 5. First, for windows closer to the census year, all three approaches yield similar results: the coal boom has very little impact on mortality. Second, as we expand the window, the results in Panel B and Panel C continue to show evidence that the boom had little impact on mortality, while the results in Panel A change substantially, becoming negative and statistically significant. Both of these patterns justify our using the third, log

Table 5: Results for various windows starting in 1970

Shock years:	1970-1971 (1)	1970-73 (2)	1970-75 (3)	1970-77 (4)
	Panel A – DV: Log Mortality Rate			
Coal county × boom	-0.009	-0.021*	-0.030***	-0.035***
	(0.010)	(0.011)	(0.010)	(0.011)
	Panel B – DV: Log Mortality			
Coal county × boom	0.004	-0.002	-0.007	-0.007
	(0.010)	(0.011)	(0.010)	(0.010)
Ln(Pop)	0.155**	0.224***	0.280***	0.267***
	(0.077)	(0.062)	(0.052)	(0.044)
	Panel C – DV: Log Mortality			
Coal county × boom	0.006	0.004	0.002	0.003
	(0.010)	(0.011)	(0.011)	(0.011)

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by county. All regressions include county fixed effects, year effects, and linear county time trends. In each column, the sample runs from 1950 to the last year of the specified shock window. N=6,028 in Column 1; 6,576 in Column 2; 7,124 in Column 3; 7,672 in Column 4.

mortality, specification as our preferred approach. Moreover, the fact that the results in Panel A become negative and statistically significant as we move away from the census year is consistent with the expected effect of migration in this setting, as the selective inflow of young, healthy workers pushes the population denominator up while hardly affecting the total number of deaths.

The overriding message from Figure 9 and Tables 4-5 is that the boom had little impact on mortality, but substantially affected *observed* mortality rates. This finding is robust to alternative sample definitions, alternative boom periods, and alternative shock measures (See Appendix Tables 17 and 18.) Our interpretation of these results is that the coal boom induced a large migration response, but one that was concentrated among young workers who carried very little mortality risk. For those who were at a higher risk of mortality (i.e., older residents), the changing economic conditions seem to have had little effect, such that we see no evidence of a change

in their patterns of mortality. This is most likely because older residents, who bore higher mortality risk, were insured against economic shocks by retirements savings and programs such as Social Security, Medicare, and Medicaid, and because they were unlikely to have been the recipients of the rise in mining incomes.

In summary, we draw two main conclusions from the experience of the Appalachian coal boom. First, as in the cotton downturn, we find evidence that the coal boom caused a substantial change in migration patterns. Second, we observe changes in the observed mortality rate across locations, but little effect on the number of deaths. Thus, it appears that changes in mortality rates were largely statistical artifacts driven by changes in the population denominator. Consistent with this pattern, we find that much of the migration that took place was selective and concentrated among young workers, a population with very little mortality risk. Overall, we conclude that the coal boom had little effect on mortality, but that an analysis focused on the overall mortality rates may find large and spurious results driven by the unobserved short-term migration of young, healthy workers.

7 Conclusion

Drawing on two natural experiments in very different settings, this paper assesses the sensitivity of estimates of the relationship between business cycles and mortality to unobserved migration. In both settings, we find that changes in local economic conditions generated a substantial migration response, and that this migration has the potential to systematically bias estimates of the impact of business cycles on mortality. The standard approach used in the literature following Ruhm (2000), which takes the mortality rate as the dependent variable, appears to be especially vulnerable to bias generated by migration because it imposes strong assumptions on the relationship between population and mortality in the short run. These assumptions are likely to be violated in the presence of migration, resulting in misleading estimates.

If migration bias may pose serious problems in estimating the impact of business-cycle fluctuations on mortality, how can we correct for it? An ideal solution is to use individual-level panel data, as has been done by Gerdtham & Johannesson (2005), Edwards (2008) and Dehejia & Lleras-Muney (2004). Unfortunately, this type of data is only available in a very small number of modern developed countries. In settings where such detailed data are not available, one alternative is to focus only on infant mortality, where migration is less of a concern.[60] However, the impact of business-cycle fluctuations on infant mortality may be qualitatively different from the impact on other segments of the population. Thus, to focus on infant outcomes alone may yield an incomplete picture of the relationship between business cycles and health more generally. A second alternative is to focus on patterns at the national level, as is done by Cutler *et al.* (2016), whose cross-country approach abstracts from difficult-to-track internal migration flows. Finally, a third approach, proposed in this paper, is to carefully study individual events that provide plausibly exogenous variation in the incidence of temporary economic shocks, and to use the available evidence in the empirical context to diagnose and mitigate migration bias.

Recent work by Ruhm (2015) shows that the pro-cyclical relationship between business cycles and mortality documented in the U.S. using the standard approach has substantially weakened since about 1980. This trend is concentrated among the young and those in middle age. By highlighting the important influence of migration in estimates of the relationship between business cycles and mortality, our study offers a potential explanation for this trend. In particular, recent work (e.g., Malloy *et al.* (2016)) documents a fall in the rate of internal migration that also began around 1980, particularly among young and middle-aged workers. If rising internal migration was partially responsible for the pro-cyclical mortality estimated by Ruhm (2015) before

[60]Recent studies in this literature include Paxson & Schady (2005), Ferreira & Schady (2009), Bhalotra (2010), Miller & Urdinola (2010), Baird *et al.* (2011), Cruces *et al.* (2012), Friedman & Schady (2013), and Bozzoli & Quintana-Domeque (2014). Many, though not all, of these studies find that negative economic shocks increase infant mortality.

1980, then the fall in internal migration rates after 1980 may explain the weakening of the pro-cyclical relationship between recessions and mortality. More work is needed to assess whether this is indeed the case.

References

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., Clay, K., Fishback, P., Haines, M., Kantor, S., Severnini, E., & Wentz, A. 2016 (September). *U.S. County-Level Natality and Mortality Data, 1915-2007*. Inter-university Consortium of Political and Social Research.

Baird, Sarah, Friedman, Jed, & Schady, Norbert. 2011. Aggregate income shocks and infant mortality in the developing world. *Review of Economics and Statistics*, **93**(3), 847–856.

Bhalotra, Sonia. 2010. Fatal fluctuations? Cyclicality in infant mortality in India. *Journal of Development Economics*, **93**(1), 7–19.

Black, D.A., Kolesnikova, N., Sanders, S.G., & Taylor, L.J. 2013. Are Children "Normal"? *Review of Economics and Statistics*, **95**(1), 21–33.

Black, Dan, Daniel, Kermit, & Sanders, Seth. 2002. The impact of economic conditions on participation in disability programs: Evidence from the coal boom and bust. *The American Economic Review*, **92**(1), 27–50.

Black, Dan, McKinnish, Terra, & Sanders, Seth. 2005. The Economic Impact Of The Coal Boom And Bust. *The Economic Journal*, **115**(503), 449–476.

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.

Bozzoli, Carlos, & Quintana-Domeque, Climent. 2014. The weight of the crisis: Evidence from newborns in Argentina. *Review of Economics and Statistics*, **96**(3), 550–562.

Carrington, William J. 1996. The Alaskan Labor Market during the Pipeline Era. *Journal of Political*

Economy, **104**(1), pp. 186–218.

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.

Conley, Timothy G. 1999. GMM Estimation with Cross Sectional Dependence. *Journal of Econometrics*, **92**(1), 1 – 45.

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.

Cruces, Guillermo, Glüzmann, Pablo, & Calva, Luis Felipe López. 2012. Economic crises, maternal and infant mortality, low birth weight and enrollment rates: evidence from Argentinas downturns. *World Development*, **40**(2), 303–314.

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.

Ferreira, Francisco HG, & Schady, Norbert. 2009. Aggregate economic shocks, child schooling, and child health. *The World Bank Research Observer*, **24**(2), 147–181.

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.

Friedman, Jed, & Schady, Norbert. 2013. How many infants likely died in Africa as a result of the 2008–2009 global financial crisis? *Health Economics*, **22**(5), 611–622.

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.

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.

Jaworkski, T., & Kitchens, C.T. 2016 (March). *National Policy For Regional Development: Evidence from Appalachian Highways.* NBER Working Paper No. 22073.

Kearney, Melissa S., & Wilson, Riley. 2017 (May). *Male Earnings, Marriageable Men, and Nonmarital Fertility: Evidence from the Fracking Boom.* NBER Working Paper No. 23408.

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.

Malloy, R., Smith, C.S., & Wozniak, A.K. 2016. Job Changing and the Decline of Long-Distance Migration in the United States. *Demography*, **54**(2).

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. 2017 (April). *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.

Newey, Whitney K., & West, Kenneth D. 1987. A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, **55**(3), pp. 703–708.

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.*

Paxson, Christina, & Schady, Norbert. 2005. Child health and economic crisis in Peru. *The World Bank Economic Review*, **19**(2), 203–223.

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.*

Ruhm, Christopher J. 2000. Are Recessions Good for Your Health? *The Quarterly Journal of Economics*, **115**(2), 617–650.

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.

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 Intercensal population estimation

A.2.1 Assumptions about population growth in the presence of migration

A key issue in this literature has to do with the fact that in most intercensal years (indeed, in 9 out of every 10), the available population measure is an interpolated variable that does not perfectly reflect the true underlying population. This can seriously bias estimates of the true relationship between business cycles and health, because in the presence of short-term migration flows, this interpolation error is systematic

rather than random.

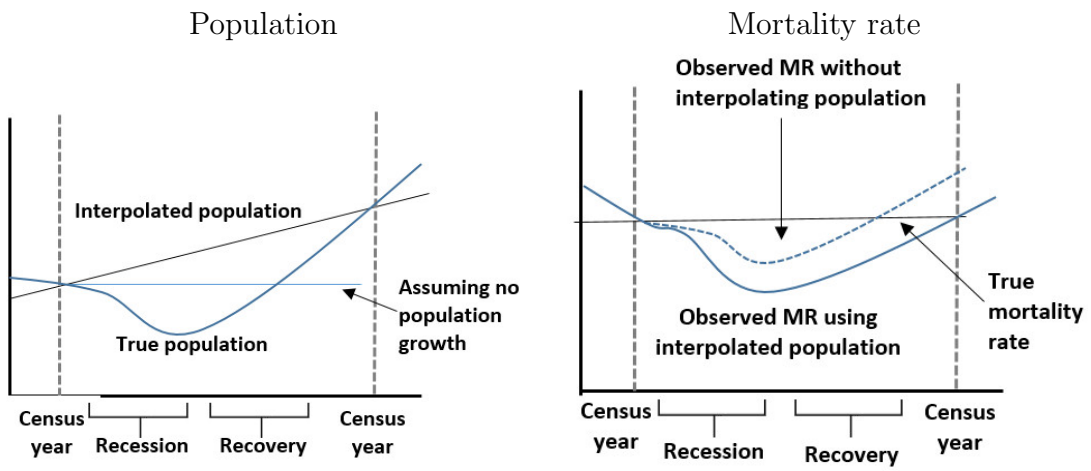
Consider a hypothetical situation in which a district experiencing long-run population growth is hit by a short-run recession followed by a recovery. Further, assume that the recession causes out-migration which is reversed during the recovery period. These are essentially the conditions observed in the case of the cotton shortage. To keep things simple, assume that the recession has no impact on the mortality rate.

The left-hand panel of Figure 10 describes the evolution of population in this hypothetical setting. In this example, the interpolated population builds in population growth during the recession, even though population was falling during that period. If that interpolated population is used in the denominator when calculating the mortality rate, as shown in the right-hand panel, then it will make it appear that the mortality rate fell during the recession even when it did not.[61] If instead we do not use the interpolated population information (as in Eq. 4), then in this case, we will obtain a mortality rate estimate that is closer to the true value. Thus, including the interpolated population value in the regression does not necessarily lead to a more accurate counterfactual.

The main conclusion to draw from this discussion is that when there is a migration response to economic shocks, assumptions made about the evolution of population and the relationship between population and mortality in the short versus long run are important. Moreover, in the absence of accurate annual population estimates, it is not necessarily true that using interpolated values will lead to more accurate results. Indeed, erroneous assumptions about population change may *introduce* new sources of bias into our calculations.

[61]Note also that this interpolation does not merely result in population mis-measurement in the recession period. Rather, by forcing part of the growth that occurred at one point in the intercensal period into another, this approach also leads population in the recovery sub-period to be underestimated, which will further contribute to the spurious impression that the recession was good for health.

Figure 10: Example



A.2.2 Errors of closure and Das Gupta interpolation

The results presented in this study suggest that migration bias has an important impact on the estimated business cycle-mortality relationship in both of our empirical settings.[62] 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.[63] 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

[62]Indeed, despite the fact that we might expect this bias to be most problematic in settings—such as historical and developing-country ones—with mobile labor forces and a poor capacity to track such movements intercensally, our coal boom analysis nevertheless indicates that migration bias can be an issue even in a modern rich-country setting.

[63]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.

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.[64] 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 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

[64]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.

England and Wales. These errors of closure, reported in absolute value as a percentage of final-year population, are presented in Table 6.

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.[65] 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 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

[65] 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.

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).[66] Recent studies in the medical and biostatistics literatures have also raised 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

[66]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 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.

methodological discussion of related denominator issues).[67]

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 both our coal boom analysis and 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,[68] 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.

[67]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.

[68]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).

Table 6: 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

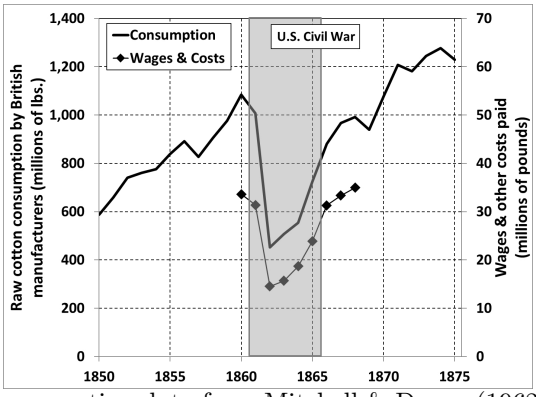
We report the absolute value (magnitude) of the error of closure in the total population, calculated as a percentage of the final-year census count of population. Figures are given for England & Wales, the United States, Canada, and India, as well as for smaller geographic units within these countries (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: The Lancashire cotton shortage

B.1 Additional evidence on the effects of the cotton shortage

Figure 11 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 11: 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 look at imports and exports. The left-hand panel of Figure 12 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 12 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.

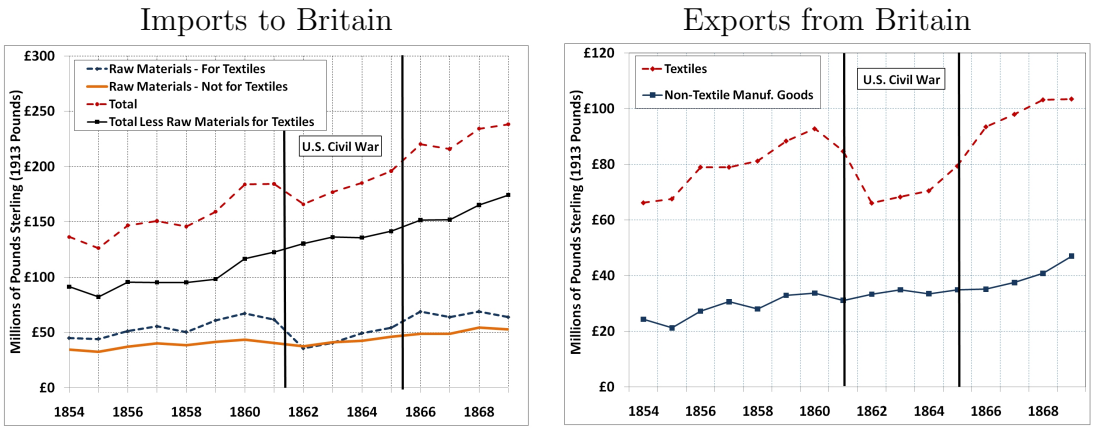
One may expect that the U.S. Civil War would have had an impact on particular sectors of the British economy, such as arms or warship production. However, British producers were prohibited from selling arms to either side during the Civil War. While it is well known that some producers were able to circumvent 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 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 and London, the two most important British ports, from our data.

B.2 Responses to the cotton shortage

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

Figure 12: British imports and exports, 1854-1869



Data from Mitchell (1988).

fresh air, inability to indulge in spirituous liquors, and better nursing of children."[69] On the other hand, there were also reports of negative health effects due to poor nutrition and crowded living conditions.[70] 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.

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

[70]Dr Buchanan, in his *Report on the Sanitary Conditions of the Cotton Towns*, states that "There is a wan and haggard look about the people..." (Reports from Commissioners, British Parliamentary Papers, Feb-July 1863, p. 301). Typhus and scurvy, diseases strongly associated with deprivation, made an appearance in Manchester and Preston in 1862 after being absent for many years, while the prevalence of measles, whooping cough, and scarlet fever may have also increased (*Report on the Sanitary Conditions of the Cotton Towns*, Reports from Commissioners, British Parliamentary Papers, Feb-July 1863).

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) 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 5.2, 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.[71] 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 period.[72] 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

[71]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."

[72]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.

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.[73]

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.

B.3 Data used in the cotton shock analysis

To assess the health consequences of the cotton shortage, we construct a new panel of annual district-level mortality spanning 1851-1865. These detailed data, which we digitized from original reports of the Registrar General, include information on both the age and cause of death for over 600 registration districts covering all of England

[73]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.

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 and births. The population data are based on information from the census years 1851, 1861 and 1871, while the births data were collected annually. 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.2.2).

In our main analysis, we take the entire U.S. Civil War, 1861-1865, as shock period. As noted above, contemporary reports suggest that most of the adverse impacts of the U.S. Civil War were concentrated in the first three years of the event, but we focus primarily on the entire Civil War period so as to avoid concerns that our main results may be dependent on the choice of shock years that we consider. In particular, as part of the government response to the cotton shortage, a number of public works projects focused on sanitary improvements were undertaken during the recession. Most of these did not come into operation until late in the 1861-1865 period. In addition, migration occurring during the shock period was also likely to have affected mortality patterns in the post-shock period.[74] Focusing on the pre-shock and shock periods avoids concerns about how to treat these factors.

[74]For instance, if, as contemporary sources suggest, people returned to cotton districts shortly after the end of the cotton shortage, then there will be mean reversion in the population. Hence, comparisons of the cotton shortage period with the recovery (i.e., post-1865) period will suffer from some of the same population mis-measurement issues we highlight in our analysis: namely, the number of deaths will appear higher during recovery than during the shortage, because population has rebounded. Thus, this issue amounts to contamination of observations over time rather than over space. As such, it parallels the problems we raise in making comparisons between cotton and nearby (i.e., indirectly treated) districts during the cotton shortage period itself.

In order 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. Nevertheless, we provide robustness results using occupation data from the 1861 Census. 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 539 consistent districts in the main analysis.[75] Summary statistics for these 539 districts appear in Table 7.

[75] 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 7: Summary Statistics

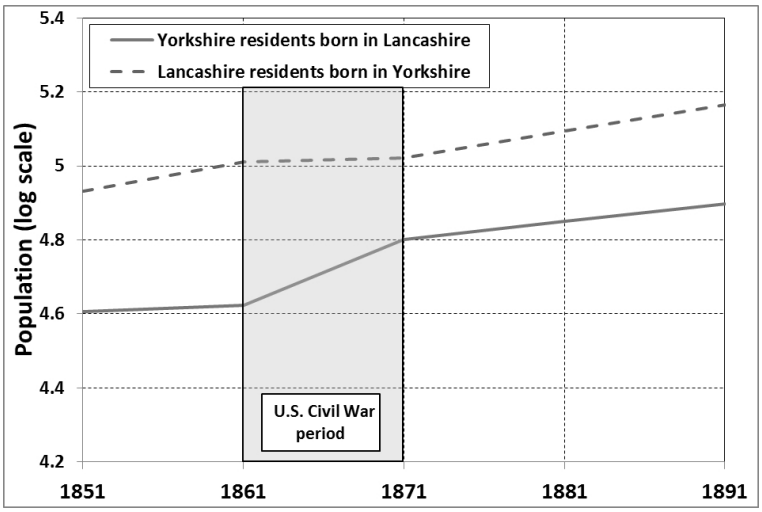
Panel A: Full sample of districts	(1) Mean	(2) Standard deviation	(3) Min	(4) Max
Average annual deaths (full sample)	809.23	1050.19	34	11,256
Cotton employment share (1851 census)	0.017	0.07	0	0.51
Nearby cotton employment (1851 census)	4569.01	20,175.54	0	158,490
Population (1851 census)	33,260.87	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

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 the total workforce (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 percent). Pre-shock period is 1851-1860 while the shock period is 1861-1865.

B.4 Additional results on migration during the cotton shock

B.4.0.1 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 13, 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.

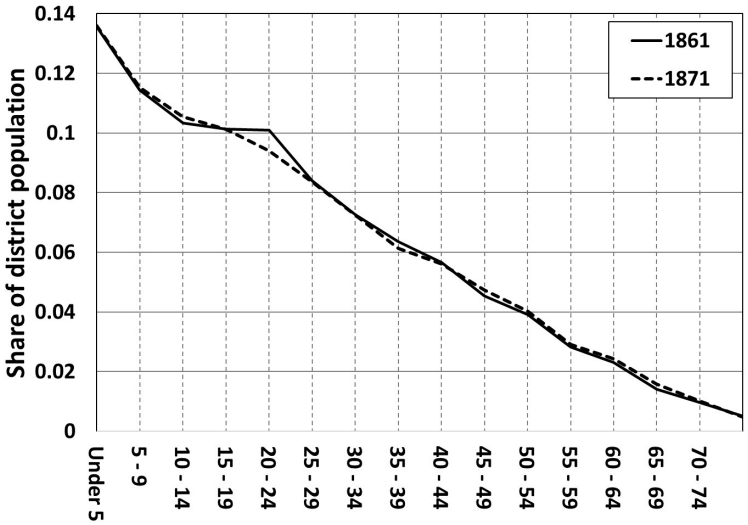
Figure 13: Evidence of migration for Yorkshire and Lancashire from birthplace data



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 14 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.

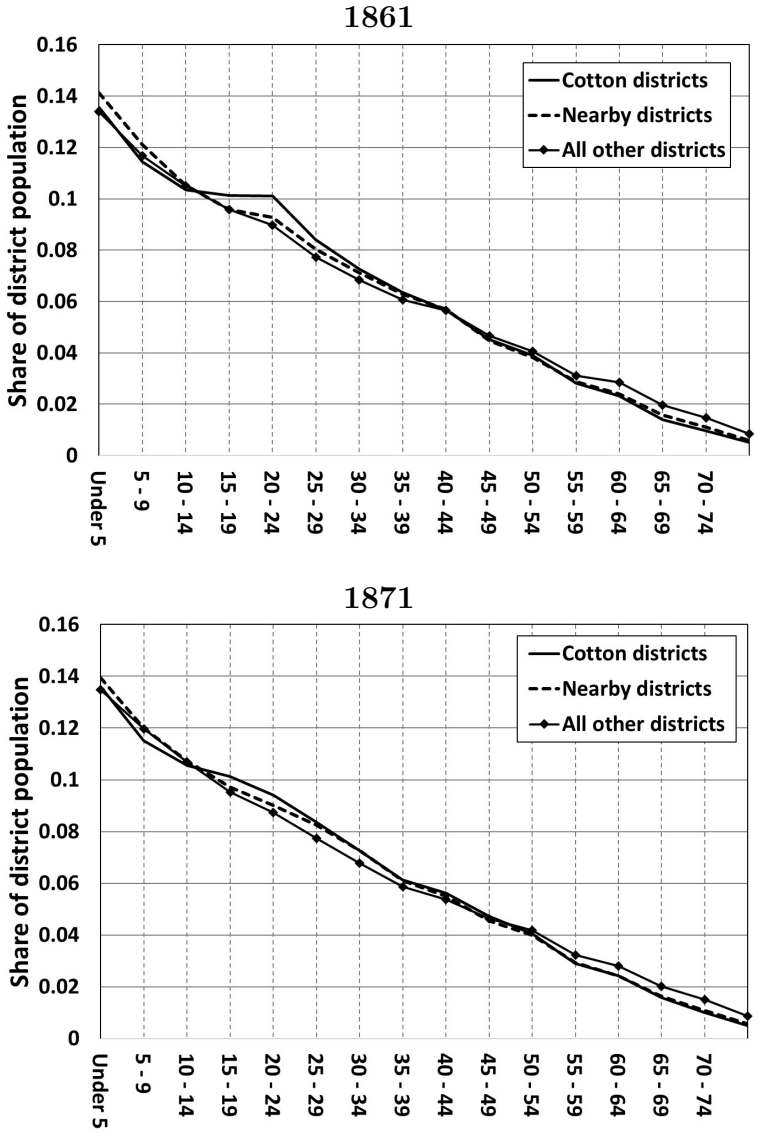
Figure 14: Share of population in each age group in cotton districts



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 15, 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 15: Share of population in each age group by type of district

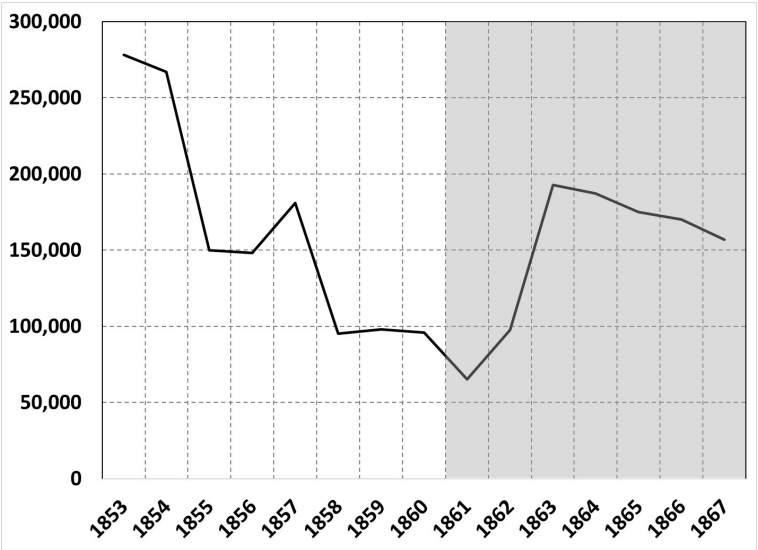


Population data are from the Census of Population for 1861 and 1871. Cotton districts are identified as those with over 10 percent of workers employed in cotton textile production in the 1851 Census. Nearby districts are those within 25 km of cotton districts.

B.4.0.2 Emigration from Britain Tracking emigration from Britain in response to the cotton shock is more difficult that tracking internal migration. What information is available was collected at the ports of embarkation and reported in the

British Parliamentary Papers. Figure 16 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.

Figure 16: Emigration from Britain, 1852-1867

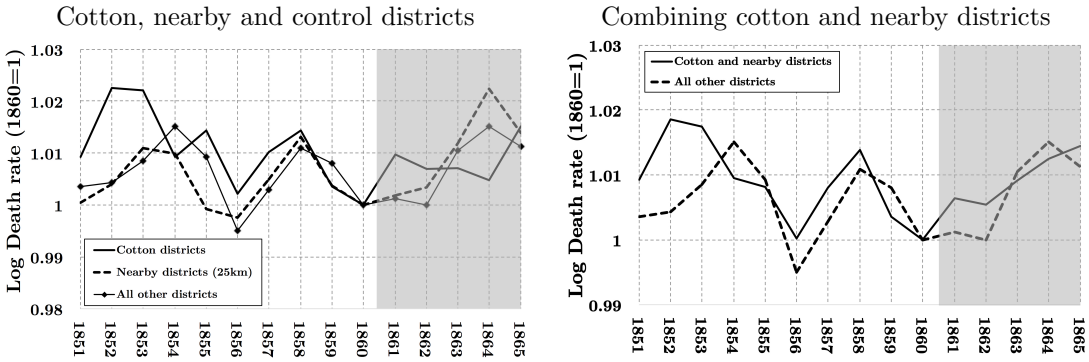


Data from the British Parliamentary Papers (1868, no. 045515).

B.5 Additional results on mortality during the cotton shock

B.5.0.1 Graphical evidence on mortality rates Figure 17 presents results similar to those shown in Figure 5 of the main text, but using mortality rates rather than mortality.

Figure 17: Mortality rates during the cotton shortage



Mortality data from the reports of the Registrar General. Data cover all of England & Wales. Population denominators are based on Census data and iterpolated using information on births and deaths from the Registrar General's reports using the Das Gupta method. 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.

B.5.0.2 Alternative measures of nearby cotton exposure Table 8 considers several alternatives to identifying nearby cotton districts. In Column 1 we add to our baseline specification variables capturing the effect of cotton employment in the 25-50 km, 50-75km, and 75-100 km bins. There we see that the effects are concentrated in the 0-25 km bin. In Column 2 we replicate this specification; however, instead of using log(nearby cotton employment) we use a series of indicators for whether the district was within 0-25 km, 25-50 km, 50-75 km, or 75-100 km of a cotton district. Further, in Columns 3 and 4 we consider continuous measures of cotton exposure. Specifically, for each non-cotton district, we calculate the distance to all other districts and then discount cotton employment for further away districts. In aggregate, these results suggest that those more proximate to cotton areas saw an increase in mortality, which is consistent with the graphical and contemporary evidence on the patterns of migration.

Table 8: Alternative measures of nearby cotton exposure

Dependent variable is ln(MORT)				
	(1)	(2)	(3)	(4)
1[Cotton district] ×	0.059***	0.031	0.241***	0.050**
Cotton shortage	(0.022)	(0.019)	(0.056)	(0.020)
0-25 km exposure ×	0.012***			
Cotton shortage	(0.002)			
25-50 km exposure ×	-0.004			
Cotton shortage	(0.003)			
50-75 km exposure ×	-0.001			
Cotton shortage	(0.003)			
75-100 km exposure ×	0.004			
Cotton shortage	(0.003)			
1[0-25 km of cot. dist.] ×		0.077***		
Cotton shortage		(0.024)		
1[25-50 km of cot. dist.] ×		0.054***		
Cotton shortage		(0.018)		
1[50-75 km of cot. dist.] ×		0.039**		
Cotton shortage		(0.018)		
1[75-100 km of cot. dist.] ×		0.051***		
Cotton shortage		(0.015)		
Nearby cot. emp. discounted as			1.221***	
exp(-distance/10000) × Cotton shortage			(0.287)	
Nearby cot. emp. discounted as				2.378***
1/(distance×10) × Cotton shortage				(0.461)
District effects	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes
Observations	8,085	8,085	8,085	8,085
R-squared	0.209	0.206	0.201	0.205

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered at the district level in parentheses. Data cover 539 districts from 1851-1865. Downturn period is 1861-1865. Cotton districts are defined as those with a cotton employment share greater than 10%. Nearby cotton exposure is calculated as the log of (1 + total cotton employment in other districts that lie within 25km). This variable is set to zero for cotton districts. Mortality data are from annual reports of the Registrar General.

B.5.0.3 Adding district-level time trends Table 9 presents results including district-specific time trends. Relative to Table 2, we see that all three specifications reveal similar patterns, and in fact, these patterns are most consistent with the specification that omits population in Table 2.

Table 9: Accounting for spillovers to nearby districts with time trends

Dependent variable:	Ln(MORT. RATE)	Ln(MORTALITY)	Ln(MORTALITY)
	(1)	(2)	(3)
Cotton district × shortage	0.079***	0.069***	0.072***
	(0.017)	(0.016)	(0.016)
Nearby cotton emp. × shortage	0.007***	0.007***	0.007***
	(0.002)	(0.002)	(0.002)
Ln(Pop)		-0.385**	
		(0.170)	
Observations	8,085	8,085	8,085
R-squared	0.239	0.421	0.420

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by district. Data cover 539 districts from 1851-1865. All regressions include district fixed effects, year fixed effects, and district-specific time trends. Shock period is 1861-1865. Cotton districts are defined as those with a cotton employment share greater than 10%.

B.5.0.4 Results weighted by population A number of papers in this literature weight their results by population, largely because the outcome of interest is given in rates. Given our concerns regarding the mis-measurement of intercensal population, our main specifications do not weight by population. However, in Table 10, we present results weighted by population, and show that although point estimates are slightly smaller here than in our main results, the qualitative story remains the same.

Table 10: Weighted results accounting for spillovers to nearby districts

Dependent variable:	Ln(MORT. RATE)	Ln(MORTALITY)	Ln(MORTALITY)
	(1)	(2)	(3)
Cotton district × shortage	-0.018*	-0.009	0.038*
	(0.011)	(0.012)	(0.022)
Nearby cotton emp. × shortage	0.003*	0.004**	0.011***
	(0.002)	(0.002)	(0.003)
Ln(Pop)		0.836***	
		(0.048)	
Observations	8,085	8,085	8,085
R-squared	0.239	0.421	0.420

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by district. Data cover 539 districts from 1851-1865. All regressions include district fixed effects, year fixed effects, and are weighted by population. Shock period is 1861-1865. Cotton districts are defined as those with a cotton employment share greater than 10%.

B.5.0.5 Robustness of the counter-cyclical result Table 11 presents some additional robustness results for the cotton shock example. We focus on results that include variables capturing the impact of the cotton shock in nearby districts, which is our preferred specification. In the first column we impose a population density cutoff which limits the set of comparison districts to be more similar to the cotton districts, which tended to be relatively dense urban areas. Specifically, we use a population density cutoff of 0.89, which is the minimum density among cotton districts. In the second column we drop London, Liverpool, Leeds and Manchester from the data. London is a clear outlier because of its size and status as the capital. Liverpool is dropped because as the main port for cotton shipments it may have also been affected by the shock. Leeds is dropped because it experienced several border changes during the 1861-1871 decade. Manchester is dropped because it was an outlier relative to other cotton districts; it was the largest cotton town and the main market for the industry. In Column 3, we use a shorter pre-shock period. In Column 4 we define cotton districts as those with an employment share greater than 5%, and in Column 5

we define cotton districts based on 1861 Census data. Both columns show an increase in mortality for both cotton and nearby districts, however, the increase for cotton districts is not statistically significant when we define cotton districts as those with greater than 5% employment share.

Table 11: Assessing the robustness of the cotton shortage results

	Dependent variable is ln(MORT)				
	(1) Population density >0.089	(2) No London, Liverpool, Leeds, or Manchester	(3) Data from 1856-65 only	(4) Cotton dist defined as emp. share>5%	(5) Cotton emp. based on 1861 census
1[Cotton district] × Cotton shortage	0.048**	0.039**	0.067***	0.020	0.051***
	(0.020)	(0.018)	(0.017)	(0.018)	(0.019)
0-25 km exposure × Cotton shortage	0.009***	0.009***	0.011***	0.010***	0.011***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
District effects	Yes	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes	Yes
Observations	7,905	7,680	5,390	8,085	8,085
R-squared	0.212	0.197	0.272	0.207	0.205

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered at the district-level in parentheses. Data cover 539 districts from 1851-1865. Downturn period is 1861-1865. Cotton districts are defined as those with a cotton employment share greater than 10%. Nearby cotton exposure is calculated as the log of (1 + total cotton employment in other districts that lie within 25km). This variable is set to zero for cotton districts. Mortality data are from annual reports of the Registrar General.

B.5.0.6 Alternative standard error adjustments Table 12 presents estimated results corresponding to Table 1 in the main text, but using two additional approaches to dealing with standard errors. In this table, the parentheses contain standard errors clustered by district, which is the approach used in the main text. This approach is the most similar to the one taken in most existing studies in this literature. In square brackets, we report standard errors clustered by county, which is the next largest geographic unit. There are 55 counties in our data. In the curly brackets, we present standard errors that allow for serial correlation using the approach from Conley (1999), up to a cutoff distance of 25 km. All three approaches yield fairly similar results, though allowing spatial correlation does increase the confidence intervals

somewhat.

Table 12: Assessing the importance of standard error adjustments

Dependent variable:	**Ln(MR)**	**Ln(MORT)**	**Ln(MORT)**
	(1)	(2)	(3)
Cotton district × shortage	-0.023	-0.013	0.049
	(0.011)	(0.012)	(0.019)
	[0.010]	[0.012]	[0.026]
	{0.020}	{0.020}	{0.020}
Cotton emp. within 0-25km × shortage	0.002	0.003	0.011
	(0.001)	(0.001)	(0.002)
	[0.002]	[0.002]	[0.003]
	{0.001}	{0.001}	{0.002}
Ln(Pop)		-0.857**	
		(0.029)	
		[0.028]	
		{0.032}	
District effects	Yes	Yes	Yes
Year effects	Yes	Yes	Yes

Standard errors clustered by district in parentheses. Standard errors clustered by county in brackets. 25 km spatially corrected standard errors in curly brackets. Data cover 539 districts from 1851-1865. Downturn period is 1861-1865. Cotton districts are defined as those with a cotton employment share greater than 10%.

B.5.0.7 Contribution of changes in fertility and infant mortality to total mortality Although we treat fertility responses and age-specific mortality effects as beyond the scope of this paper, we report these two results here so that we may comment on their potential contribution to our overall finding of counter-cyclical total mortality in the cotton shortage setting. Specifically, we might wonder if changes in fertility and infant mortality rates changed the size and composition of the population enough to raise the average mortality risk of the population, and so drive the observed rise in mortality during the cotton shortage.

Table 13 presents results on the effects of the cotton shortage on the log of births (which are observed annually), and on the infant mortality rate (or IMR, given as

the number of infant deaths, observed annually, per 1,000 live births). Crucially, the infant mortality rate abstracts from migration bias due to mis-measurement of the denominator, since in every year, the denominator is observed rather than constructed. We find that the cotton shortage resulted in an insignificant increase in births in cotton districts, and a significant increase in births in nearby districts. This is consistent with layoffs in a heavily female industry, and with the short-distance migration of young families from cotton districts to nearby areas. We also find that the downturn had a beneficial impact on infant health, significantly lowering the infant mortality rate in cotton districts. This, too, is consistent with the substitution effects of maternal time dominating any adverse income effects due to job loss, since in this period infant health interventions were more time-consuming than they were financially costly.

These results suggest that a cotton district saw a roughly exp(0.0247) = 2.50% increase in the number of infants born during the downturn. Taking into account the regression's constant (6.656), this is equivalent to roughly 19.44 additional births each year. Our preferred results in Table 2 suggest that a cotton district saw a exp(0.0494) = 5.06% increase in total mortality during the downturn, equivalent to roughly 26.56 additional deaths each year (based on a control-district baseline of exp(6.162) deaths). Given the estimated infant mortality rate in cotton districts during the downturn (135.5 - 5.141 = 130.359 infant deaths per 1,000 births), this would mean that the increase in cotton-district fertility implies roughly 2.53 extra deaths per year, or less than 10% of the total annual increase in cotton-district mortality during the downturn.

What's more, since infant mortality rates in this period were much higher than mortality rates for young children, it is unlikely that this cohort, whether through their larger initial cohort size or their higher rate of cohort survivorship, would have been large enough to drive the total mortality result as they moved through the age distribution in the few years that followed.

Thus, to sum up, while endogenous fertility responses may contribute to our main

results on the adverse mortality effects of the cotton shortage, they explain very little of this increase in total mortality.

Table 13: Assessing the effect on fertility and the infant mortality rate

Dependent variable:	**Ln(Births)**	**IMR**
	(1)	(2)
Cotton district × shortage	0.0247	-5.141***
	(0.0168)	(1.778)
Cotton emp. within 0-25km × shortage	0.00919***	0.351
	(0.00214)	(0.251)
Constant	6.656***	135.5***
	(0.00487)	(0.837)
District effects	Yes	Yes
Year effects	Yes	Yes
Observations	8,085	8,085
R-squared	0.264	0.066

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered at the district level in parentheses. Data cover 539 districts from 1851-1865. Downturn period is 1861-1865. Cotton districts are defined as those with a cotton employment share greater than 10%. The infant mortality rate (IMR) is given as the number of infant deaths per 1,000 live births.

C Appendix: The Appalachian coal boom

C.1 Data used in the coal boom analysis

The mortality data used in our analysis of the coal boom come from the Bailey *et al.* (2016) files deposited at ICPSR. These files contain annual county-level data on mortality.[76] In addition to deaths, the Bailey *et al.* (2016) data also provide the number of births, which we use when constructing intercensal population estimates.

The second type of data used in our study is a set of population data from the Census. These data are available at the county level for every decade. For intercensal

[76]Our analysis uses the location of residence mortality data.

years, we generate estimated population values using the Das Gupta method, which uses the previous and next census population values for each county as well as birth and death information.[77]

In the main analysis we use data starting covering 1950-1977. Summary statistics for the main variables used in the analysis across these years are presented in Table 14.

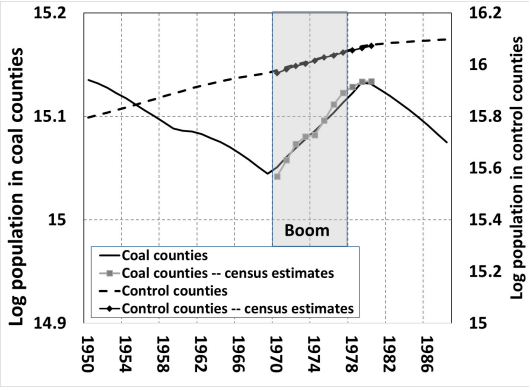
Table 14: Summary statistics for variables used in the main coal boom analysis

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Mortality rate (per 1000)	10.239	1.93	3.568	19.568
Log mortality rate	2.307	0.201	1.272	2.974
Mortality	432.16	411.531	43	2681
Log mortality	5.734	0.793	3.761	7.894
Population	43,189	40,276	8,008	231,122
Log population	10.335	0.803	8.988	12.351
N		7,672		

The interpolated population values used in the main analysis were generated by the authors using census population data together with births and deaths; no corrections were made for observed migration, since we do not have access to this data. The Census Bureau starts to provide annual county-level population estimates (i.e., adjusting for observed migration) starting in 1970. In Figure 18, we compare our intercensal population estimates for the coal and control counties to those published by the Census Bureau. This figure shows that our estimates (without adjusting for migration) are quite similar to those produced by the Census Bureau (which use tax and similar records to adjust for migration in the manner described in Appendix A.2.2).

[77]In the few instances in our sample where mortality or population data are missing in Bailey *et al.* (2016), we supplement this with data from the Centers for Disease Control, such that we have fully balanced temporal and spatial coverage within our sample.

Figure 18: Comparing our population estimates to the best-available estimates produced by the U.S. Census Bureau



In addition to the data used in the main analysis, we have also collected mortality data by age group. Unfortunately, however, we have only found these data at the county level starting in 1968, so we do not have a long enough pre-shock period to replicate the analysis that we applied to the total mortality data. Instead, in Appendix C.3 we present graphs describing the evolution of mortality within age categories from 1968 through the coal boom period.

Table 15 describes mortality rates by age group. We focus on mortality rates in 1970 in order to take advantage of the census population figures in that year and we use only data from the counties used in the main analysis to construct these figures. The main take-away from Table 15 is that mortality rates among those in their 20s and 30s were much lower than for older age groups. This is an important fact because, as discussed in the next appendix section, this group was the most likely to migrate in response to the coal boom.

Table 15: Mortality rate by age group in 1970 (per thousand)

0-4	5-9	10-14	15-19	20-24	25-34	35-44	45-54	55-64	65-74	Over 75
4.821	0.43	0.387	0.994	1.405	1.475	3.054	7.385	17.454	38.554	102.823

C.2 Additional results on migration during the coal boom

Figure 19 presents a map showing the difference in net migration by county in the 1970-80 decade compared to the 1960-70 decade including more surrounding states than the map in Figure 7. The additional states we include here are Maryland, Tennessee, and Virginia. Dark colors indicate counties that experienced an increase in in-migration in the coal boom decade, 1970-1980, relative to the decade before.

Figure 19: Difference in net migration with additional states

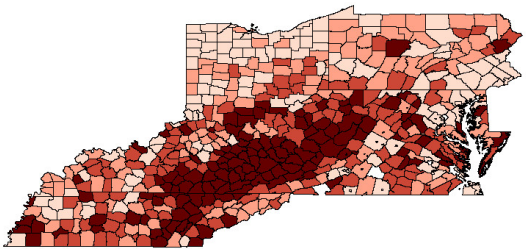
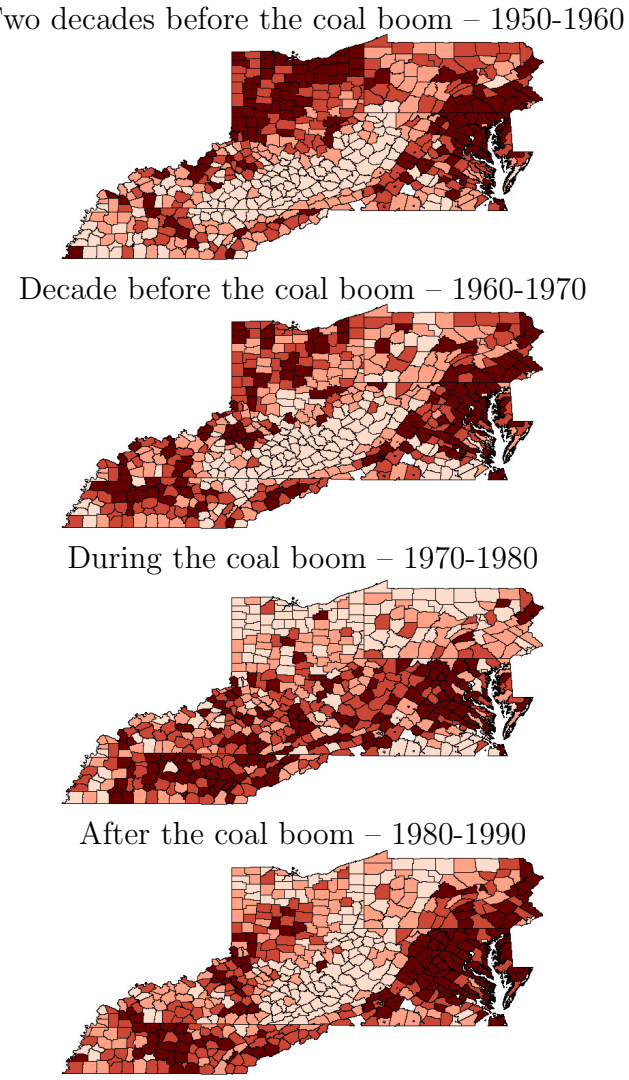


Figure 20 presents maps showing net migration based on the error of closure in the decade before the coal boom, the decade covering most of the boom, and the decade after the boom. Light colors indicate counties experiencing net out-migration. We can see that the coal counties experienced a clear pattern of out-migration during the decades before and after the boom, but that this pattern of out-migration completely disappeared during the 1970-1980 period.

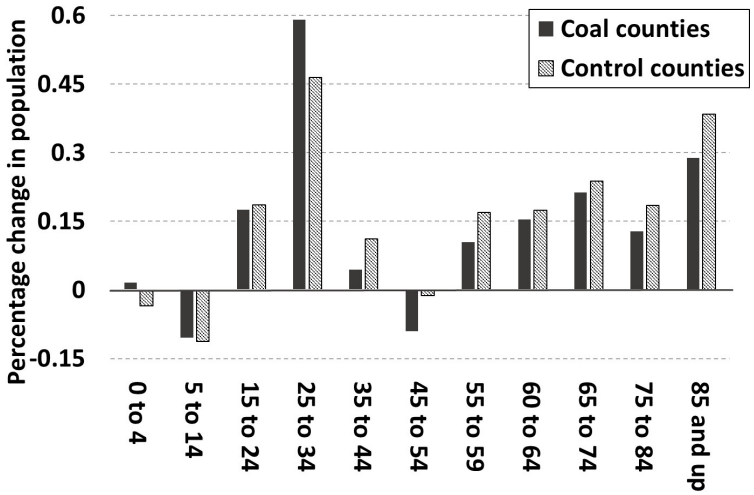
Figure 20: Estimated net migration before, during and after the coal boom



Next, we look at the distribution of migrants into the coal counties by age groups. Figure 21 plots the percentage change in population in the coal and control counties between 1970 and 1980. This shows that population growth in the coal counties during this period was concentrated in the 25-34 age group. While this age group also experienced growth in the control counties, this was less rapid than in the coal counties. Moreover, in terms of numbers, the change in population aged 25-34 in the

coal counties accounts for 100,259 people out of a total population increase of 180,903 during the 1970-1980 decade. The 15-24 year old age group accounted for another 43,870 of this increase. Thus, a substantial majority of the population increase in the coal counties from 1970-1980 was concentrated among young adults.

Figure 21: Percentage change in population by age group



C.3 Additional results on mortality during the coal boom

C.3.0.1 Alternative standard error adjustments In the main text we present results with standard errors that are clustered by county to allow for serial correlation. In Table 16 we present alternative standard errors that allow for spatial correlation between counties within 100 km of each other, following Conley (1999), and serial correlation for up to two years on each side of the observation year, following Newey & West (1987). We focus only on results with county time trends because of evidence that the parallel trends assumption is violated when time trends are not included. These alternative standard errors are presented in square brackets. In general these standard errors are smaller than those obtained when clustering by county, so we present the more conservative clustered standard errors in the main text.

Table 16: Estimated effects of the coal boom with alternative standard errors

Dependent variable:	Ln(MR) (1)	Ln(MORT) (2)	Ln(MORT) (3)
Coal county × boom	-0.0375	-0.00306	0.00947
	(0.00958)	(0.00906)	(0.0104)
	[0.00858]	[0.00712]	[0.00780]
Ln(Pop)		0.267	
		(0.0427)	
		[0.0332]	

Standard errors clustered by county in parenthesis. Spatial standard errors allowing correlation for counties within 100 km and serial correlation up to two lags in square brackets. All regressions include county fixed effects, year effects and county time trends. N=7,672.

Table 17: Coal boom regression results with additional control counties

Dependent variable:	Ln(MR) (1)	Ln(MORT) (2)	Ln(MORT) (3)
Coal county × boom	-0.0288***	0.00253	0.00947
	(0.00958)	(0.0102)	(0.0103)
Ln(Pop)		0.181	
		(0.164)	

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by county. All regressions include county fixed effects, year effects and county time trends. N=12,330.

C.3.0.2 Results with additional control counties Table 17 presents results with additional control counties. Specifically, we add to the original BMS sample set of counties in three neighboring states, Maryland, Tennessee, and Virginia. We focus only on results with county time trends because of evidence that the parallel trends assumption is violated when time trends are not included.

C.3.0.3 Alternative definitions of coal counties Table 18 explores the robustness of our results to alternative definitions of coal counties. We consider two alternative definitions—those with over 2.5 billion tons of reserves, and those with over

Table 18: Coal boom regression results with alternative coal-county definitions

	Coal counties are those with more than 2.5 billion tons of reserves (42 coal counties)			Coal counties are those with 10% of earnings from coal in 1969 (32 coal counties)		
Dependent variable:	Ln(MR) (1)	Ln(MORT) (2)	Ln(MORT) (3)	Ln(MR) (4)	Ln(MORT) (5)	Ln(MORT) (6)
Coal county × boom	-0.0347***	-0.00739	0.00258	-0.0478***	0.0234	0.0469***
	(0.0106)	(0.00988)	(0.0108)	(0.0124)	(0.0151)	(0.0156)
Ln(Pop)		0.267***			0.248***	
		(0.0437)			(0.0468)	

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by county. All regressions include county fixed effects, year effects and county time trends. N=7,672.

10% of earnings coming from coal in 1969, as in BMS. The results in columns 1-3 are qualitatively identical to the results in the main text. When we use an earnings-based definition, (Columns 4-6) the main change is in results obtained using log mortality as the outcome without including log population as a control, in Column 6, where we see evidence that mortality may have increased during the recession. However, it is worth noting that this increase is sensitive to the inclusion of time trends; without time trends the coefficient obtained from the specification in Column 6 is much smaller and not statistically significant.

C.3.0.4 Weighted results Table 19 presents results from regressions in which each observation is weighted by county population. Though weighting is typical in many of the studies in this literature, we do not weight our main regression results due to reasons discussed in Section 4. However, as a comparison between Table 19 and Table 4 in the main text shows, our decision not to weight our main results does not substantially alter our findings.

C.3.0.5 Alternative boom periods The next set of results looks at alternative definitions of the coal boom period. While in our main analysis we follow Black

Table 19: Weighted regression results for the coal boom

Dependent variable:	Ln(MR)	Ln(MORT)	Ln(MORT)
Panel A: Without county time trends			
	(1a)	(2a)	(3a)
Coal county × boom	0.106***	0.0191*	-0.0274*
	(0.0163)	(0.0107)	(0.0140)
Ln(Pop)		0.347***	
		(0.0374)	
Panel B: With county time trends			
	(1b)	(2b)	(3b)
Coal county × boom	-0.0406***	-0.00233	0.0122
	(0.0106)	(0.00882)	(0.0108)
Ln(Pop)		0.275***	
		(0.0339)	

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by county. All regressions include county fixed effects and year effects. N=7,672. Regressions are weighted by interpolated county population.

et al. (2005) in treating 1970-1977 as the main coal boom period, a look at the price data shown in Figure 6 suggests that one could alternatively think of the boom as lasting well into the 1980s. In Table 20 we look at results using two alternative definitions of the boom period, one stretching to 1982, when prices begin to fall, and the second lasting until 1988, when coal prices had returned near to their pre-boom level. In general, we can see that these extended periods deliver results that are similar to those obtained in the main text. In particular, we continue to find that results are highly sensitive to assumptions about population. The main differences in these results is that, when the period is extended to 1988, our preferred specification in Column 6b provides some evidence that mortality increased in the coal counties. It is interesting to note that we don't see any similar pattern when we only include years 1970-1982, which suggests that this effect is being driven entirely by the "bust" period from 1982-1988.

Table 20: Coal boom results using alternative boom periods

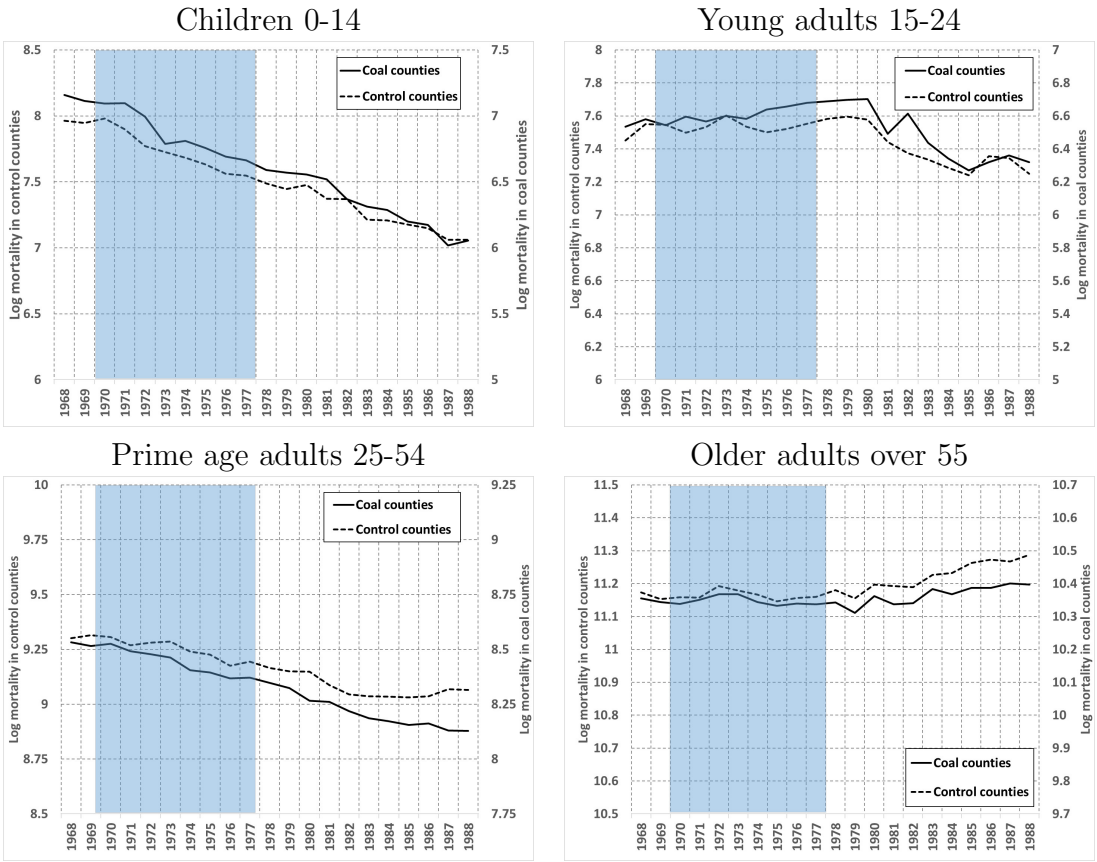
Boom period:		**1970-1982**			**1970-1988**	
Dep. variable:	Ln(MR)	Ln(MORT)	Ln(MORT)	Ln(MR)	Ln(MORT)	Ln(MORT)
			Panel A: Without county time trends			
	(1a)	(2a)	(3a)	(4a)	(5a)	(6a)
Coal county×boom	0.0909***	0.0120	-0.0306**	0.0932***	0.0134	-0.0445***
	(0.0169)	(0.0122)	(0.0132)	(0.0178)	(0.0137)	(0.0152)
Ln(Pop)		0.351***			0.421***	
		(0.0293)			(0.0277)	
			Panel B: With county time trends			
	(1b)	(2b)	(3b)	(4b)	(5b)	(6b)
Coal county×boom	-0.0360***	-0.00023	0.0096	-0.0147*	0.0107	0.0173*
	(0.0094)	(0.0092)	(0.0103)	(0.0085)	(0.0088)	(0.0094)
Ln(Pop)		0.216***			0.205***	
		(0.0389)			(0.0385)	

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by county. All regressions include county fixed effects and year effects. In Columns 1-3, N=9,042. In Column 4-6, N=10,686.

C.3.0.6 Mortality by age group Figure 22 presents graphs showing the evolution of log mortality for different age groups in the coal counties and control counties. In each graph we plot log mortality for the two series on separate axis, since there are more control counties with a larger overall population, but we use scales covering the same range for comparability.

For children, we can see that the number of deaths are falling in both locations although there is an increase in deaths around 1971 that is more pronounced in the coal counties. One likely explanation for this is that births in the coal counties increased in the early 1970s relative to the nearby counties. For young adults mortality is quite flat and similar across the two groups of counties, though there is some evidence of higher mortality in the coal counties towards the end of the boom and in the following years. Among working age adults over 24, the number of deaths was smoothly decreasing in both sets of counties across the study period, though this decline was somewhat more rapid in the coal counties. Among older adults, both locations experienced a

Figure 22: Mortality by age categories in the coal and control counties



brief increase in mortality from 1972-1974. This timing corresponds to the recession of 1973-75 caused by the oil embargo and provides further suggestive evidence that mortality may increase, rather than decrease, during recessions, particularly for the elderly. The number of deaths also increased in both locations starting in the 1980s. Overall, there does not appear to be any clear changes in relative mortality between the treatment and control counties during the coal boom period.[78]

[78] As mentioned earlier in the appendix, it is not possible to run reliable econometric analysis with age-specific data on the coal boom, since age-disaggregated mortality data are only available beginning in 1968, leaving us with a very short pre-boom period.