CAMBRIDGE
UNIVERSITY PRESS

SPECIAL ISSUE ARTICLE

Fatal Places? Contextual Effects on Infant and Child Mortality in Early Twentieth Century England and Wales

Alice Reid[1] (iD), Eilidh Garrett[2] (iD), Hannaliis Jaadla[1] (iD), Kevin Schürer[1] (iD) and Sarah Rafferty[1] (iD)

[1]University of Cambridge, UK and [2]University of Edinburgh, Scotland, UK
Corresponding author: Alice Reid; Email: amr1001@cam.ac.uk

Abstract

This paper takes, as its starting point, Preston and Haines' observation in *Fatal Years* that social class was the most important influence on infant and child mortality in England and Wales in the early twentieth century. A subsequent study suggested that this could in part be due to the spatial distribution of the different classes across different types of place, and that some of the mortality differences by social class might actually reflect the contextual effects of healthy and unhealthy places. Although this line of argument has received a considerable amount of attention in health geography literature, it has rarely been examined for a specific historic period, and then only within particular urban areas. In this paper, we apply multi-level models to a complete count individual-level dataset of the 1911 census of England and Wales, comparing influences on infant and child mortality at the level of the individual couple and for two spatial levels. We find that although most variation in infant and child mortality operates at the individual level, there is also important variation at the two spatial levels and part of the mortality differences between social classes is better explained by the areas in which people lived rather than by their social class. A consideration of independent variables at all three levels suggests that different spatial scales capture different sorts of influences on early age mortality.

Introduction

Fatal Years, published in 1991 by Samuel Preston and Michael Haines, rapidly became a classic text on the influences affecting early age mortality during the demographic transition (Preston and Haines 1991). It was the first systematic study of nineteenth century early age mortality in the USA and the first large scale study of historic early age mortality using individual-level data that enabled the authors to compare different influences on child survival. The volume therefore quickly established itself as an important point of reference for studies in historic infant and child mortality. As one of the first studies to make use of census data (in this case the

1900 census of the USA) at an individual level, large-scale format, the authors were able to demonstrate the potential of the census as a source for this type of investigation, paving the way for a plethora of other studies. Furthermore, the book illustrated the applicability of indirect estimation techniques (originally developed by William Brass for use in lower- and middle-income countries lacking reliable vital registration) to historical data, thus broadening the range of methodological tools available for the study of historic mortality. Preston and Haines applied the principles of indirect estimation, designed for aggregate data, to generate an individual-level measure of mortality: the *Mortality Index (MI)*. This allowed multivariable regression to be used with the individual data, permitting a finer analysis of the influence of a range of factors thought to have impacted upon infant mortality and establishing an approach that has been used in a variety of subsequent analyses (Connor 2017; Dribe et al. 2020; Garrett et al. 2001; Reid 1997).

Although Preston and Haines's main focus was the American experience, they also provided a comparative element by undertaking a separate analysis of data from the 1911 census of England and Wales. This allowed them to argue that while race was perhaps the most powerful factor influencing infant and child mortality in the USA, in England and Wales that role was taken by social class, defined by father's occupation. Their analysis of England and Wales was, however, limited to the aggregate tables published in the census reports (HMSO 1917, 1923): they could not directly compare influences nor explore as wide a range of factors as they were able to for the USA.

Several years after the publication of *Fatal Years*, some of the authors of the current paper obtained specially negotiated early access to the individual-level data for the 1911 census for a set of 53 relatively homogenous areas spread across thirteen locales (Garrett et al. 2001). They took this opportunity to reassess Preston and Haines's results for England and Wales (Garrett et al. 2001; Reid 1997). Although their data set was comparatively small, it confirmed the strong social class differentials in early age mortality that Preston and Haines had shown in their work with the published tables. However, there were equally strong mortality differentials between agricultural, white collar, industrial, and other urban types of place. Both cross-classifications and multivariable regressions of the individual data demonstrated that the type of place was, apparently, a stronger influence on mortality than social class, at least in the communities covered by the data. All social classes enjoyed low mortality in agricultural places, while the opposite was broadly true in industrial areas. We concluded that the overall social class result was produced by the differential sorting of the various classes into different types of place, with the implication that the higher classes did not have lower mortality because they knew more or could purchase better food or medical services, but because they could afford to live in areas with better infrastructure and a more salutogenic environment. Another way of putting this is to say that agricultural places did not have better health because they were inhabited mainly by better-off people, but because of characteristics of the places themselves: it was not the composition of places in terms of their inhabitants, but aspects of the context, that affected health.

These composition-versus-context and class-versus-place debates have received considerable attention in the health geography literature, aided by the increasing availability of large individual-level datasets and the development of powerful

multi-level regression techniques (Owen et al. 2016; Smyth 2008). The literature has drawn attention to the complexity of operationalizing an analysis of contextual effects on health, and in particular to the fact that influences on health are not simply divided into those that operate at an individual level versus those that operate at a geographical or community level, but that different contextual influences may operate at different spatial scales. Although historical demographers have adopted multi-level modelling, it has more often been used simply to control for unobserved variation at the place level, rather than to explicitly examine such variation and address the composition versus context debate. This is perhaps a missed opportunity as multilevel modelling not only offers a better opportunity to compare the influences of class and place but, in a historic setting with a paucity of independent variables capturing particular influences on mortality, it affords the possibility of identifying different sorts of influences by detecting variation at particular scales.

In this paper, therefore, we take advantage of the full count data of the 1911 census of England and Wales now available through the Integrated Census Microdata (I-CeM) project (Schürer and Higgs 2014) to re-examine the extent and nature of contextual effects on infant and child mortality. We find that although most variation in infant and child mortality operates at the individual level, there is also important variation at a place level and part of the mortality differences between social classes is better explained by the areas in which people lived rather than by their social class. We also conclude that different spatial scales capture different sorts of influences on early age mortality.

In the next section, we outline some of the previous research on infant mortality differentials and introduce the class-versus-place and composition-versus-context debates. Section 3 describes our data and the methods we use. Section 4 presents and discusses our results, Section 5 recognizes the limitations of our research, and Section 6 offers some tentative conclusions.

Previous research

Using the published results from the 1911 census of England and Wales, Preston and Haines (1991) identified social class as the most important factor influencing infant and child mortality in England and Wales (see also Haines 1995). The existence of this gradient in the late nineteenth and early twentieth century had already been established using both this source and subsequent Registrar General's official reports (HMSO 1923; Morris and Heady 1955; Pamuk 1985; Watterson 1988; Woods et al. 1988, 1989).

Link and Phelan's fundamental cause theory argues that social class gradients in health have always been present, working through different pathways and mechanisms in different eras (Clouston et al. 2016; Link and Phelan 1995). According to this theory, the better-off in society will always be able to use their resources or knowledge to avoid hazards or exposure, to improve resistance to disease by purchasing higher quality and quantity of food, or to better aid recovery from illness through superior access to curative healthcare. The precise mechanisms through which the rich gain an advantage will vary according to the disease context and scientific knowledge at the time.

While the conclusion that social class differences have always been present is not well supported by the evidence (Antonovsky 1967; Reid 2021; Woods and Williams 1995), the focus on mechanisms and routes to better or worse health in different groups has chimed with the way that social differences in infant and child health have been conceptualized and investigated by (historical) demographers. Mosley and Chen (1984) expressed these mechanisms as 'proximate determinants': the intermediary factors through which the more distal (societal) causes have to operate. They divided the proximate determinants into various categories: maternal (age at birth, parity, birth interval), environmental (household crowding, water or food contamination), nutrient deficiency (in the child or mother), injury, and personal control of illness.

In historic Europe, social class could affect several of the proximate determinants. Firstly, fertility decline tended to start among the better off, so lower fertility could affect mortality through the maternal route, although the evidence for this is slim (Fernihough and McGovern 2014). Secondly, poorer people tended to live in households that were more crowded (Cage and Foster 2002) and had deficient ventilation (Barker and Osmond 1987), increasing the potential for transmission of infectious and respiratory diseases. Such housing may also have lacked reliable access to clean water (Jaadla and Puur 2016) or efficient sewage disposal (Morgan 2002), increasing the possibility of diarrheal disease through directly infected water or cross-contamination. Poor maternal nutrition has been linked to lower birth weight and survival chances (da Silva Lopes et al. 2017) and less well-off women may have been unable to afford a diet that was both sufficient and nutritious. Child diets may also have been suboptimal both in terms of quality and quantity, particularly for children who were artificially fed (Fildes 1998).[1] Finally, it has been suggested that the children of better-off parents may have had better survival because such couples could afford superior health care, either at delivery or while their children were growing up, or because their education rendered them more likely to seek healthcare and better able to communicate with health professionals (Pamuk 1985).

These mechanisms all focus on individual or household attributes, but another layer of influences operates at a community level (Williams and Galley 1995), examples of which could include environmental pollution (Hanlon 2022; Morgan 2002), social cohesion, and social capital (Szreter and Woolcock 2004). Many of the individual and household influences related to social class could also wholly or partly reflect resources and structures operating at a community or local level rather than the individual or household level. For example, a house cannot be connected to a mains water supply unless a water main has been laid in the area. Health service usage may similarly reflect local provision as much as individual constraints on access. These structural or local influences are not primarily dependent on individual social class or resources, although class may be instrumental in sorting people into more or less salubrious environments (Reid 1997).

---

[1]Breastfeeding itself was strongly linked to better infant survival, but this was associated with an inverse class gradient in early twentieth century Britain, with lower class women more likely to breastfeed their infants (Reid 2017).

When examining differences in mortality or health between different places or types of place, it can be difficult to identify whether poor survival is because a place is disproportionately inhabited by less well-off people, whose health is affected by their own characteristics or resources, or whether there is an additional and independent effect of place. The 'class-versus-place' debate asks whether the social class differences in mortality visible in aggregate figures are the consequence of individual socioeconomic characteristics or due to elements of the neighborhood in which people live. The closely related, but rather broader, 'composition-versus-context' debate recognizes that most variation in mortality operates at an individual level, and asks whether, once this is controlled, there is any additional place-level variation.

These debates have been hard to resolve because much of the data, particularly for historic health and mortality, have only been available at an aggregate level (Reid 2021). Aggregate data was used not only for all the early analyses of early age mortality by social class mentioned above, but also for many geographical analyses of early age mortality. There is a long tradition of investigating geographical differences in health in Britain, arguably starting with Farr's 'healthy districts': certain rural areas that had far lower mortality than contemporaneous towns and cities (Farr 1859).

Nineteenth century miasmatic and contagion theories of disease causation supported the notion that places could be injurious to health, and contemporary British observers focused on the deleterious effects of urban disamenities on the survival chances of young children (Gregory 2008). In the same tradition, Lee (1991) used county-level data to argue that employment structure, particularly the presence of mining, and housing density were key to explaining the geographical patterns in infant mortality in England and Wales. Using registration districts (RDs), Gregory (2008) demonstrated not only an urban–rural contrast in mortality, but argued for a core–periphery pattern, with places more distant from London having higher mortality and slower improvements in life expectancy. He argued that more attention should be paid to rural districts, particularly in terms of their early mortality decline, and this theme was taken up by Atkinson et al. (2017) who examined influences on the spatial pattern and declines among rural RDs. Using more detailed registration sub-district (RSD)-level data for the whole of England and Wales, Jaadla and Reid (2017) examined the factors affecting child mortality (ages 1–4) patterns over and above those affecting infant mortality. They concluded that aspects of the local disease environments, including overcrowding which governed transmission of air-borne diseases and urban disamenities such as poor sanitation, were more important than measures of human capital such as numbers of health service workers or teachers.

Spatial analyses of historical early age mortality outside of England and Wales have also focused on describing subnational patterns (Edvinsson et al. 2001; Ramiro-Fariñas and Sanz-Gimeno 2000; Thorvaldsen 2002; van den Boomen and Ekamper 2015), with a more limited amount of cross-country analysis (Edvinsson et al. 2008; Klüsener et al. 2014). The ecological nature of many of these studies, however, means they cannot address the place-versus-class and composition-versus-context questions.

The availability of individual-level data offers an avenue for disentangling place and class using regression techniques that include both individual socioeconomic status (SES) as well as indicators of the broader environment that would ideally capture aspects of the environment thought to influence the health outcome being measured. There are various methodological and conceptual issues that need to be considered, however. Firstly, if the multilevel structure associated with variables measured at different levels is not properly accounted for, standard errors for area-level variables can be under-estimated, leading to spuriously small confidence intervals: this can be easily dealt with through the use of multilevel models. Secondly, spatially structured data also pose methodological issues as values of a particular variable for neighboring areas may not be independent, for example if people in one area use services and interact with people in a contiguous area or further afield. Failure to take these sorts of dependencies into account can lead to mis-specification of models (Manley et al. 2006; Xu et al. 2014). This can be overcome by the use of spatial models, but these depend on the availability of boundary data and are highly resource intensive, limiting their application.

Related to these geographical effects, there is an issue regarding the most appropriate geographical unit to choose for analysis. The modifiable areal unit problem is very pertinent here. This recognizes that the results for variables calculated for geographical units are highly dependent on where boundaries are drawn (the aggregation or zonation issue) and the scale at which the units are created or the underlying data aggregated (the scale issue) (Openshaw 1983). Choosing units of an inappropriate size for the analysis, or which may be a suitable size but do not represent the areas in which the spatial effects operate, can therefore obscure results. This can pose a problem as the administrative boundaries for which data are available may not coincide with the areas for which influences on health operate, or with people's own conception of their neighborhood.

A strand of research in health geography has investigated the importance of these effects and, in relation to the aggregation issue, there is some evidence that units of similar population size but covering different areas could make a big difference to results (Flowerdew et al. 2008). Stafford et al. (2008), however, also comparing areas of similar size, concluded that administrative boundaries were reasonably good approximations of the areas in which health effects operate. The scale issue has attracted more investigation with respect to health outcomes. Although Duncan et al. (1993) suggested that the size of the geographical unit used does not make much difference to health outcomes, a larger body of work has indicated that scale is important, and that smaller units are generally better at capturing health effects (Boyle and Willms 1999; Flowerdew et al. 2008; Manley et al. 2006; Oliver and Hayes 2007). Haynes et al. (2007) also found that the neighborhoods that have meaning for residents are smaller than the administrative districts that are often used for research into health outcomes.

Many of these studies also demonstrated that the size of the appropriate unit varied according to the independent variable considered (Boyle and Willms 1999; Flowerdew et al. 2008; Xu et al. 2014). For example, automobile traffic affects air pollution within a radius of about 220 m (Wang et al. 2021), which suggests that small geographic units would best capture any health effects of traffic concentration. In contrast, the effect of variations in health service provision might be best captured

by health service delivery areas, which tend to be quite large. Tarkiainen et al. (2010) examined spatial variation in adult mortality in Helsinki and found that there was independent variation at both the sub-district and district levels (see also Meijer et al. 2012). Xu et al. (2014) found that egocentrically defined neighborhoods were better at capturing spatial variation in historic infant mortality than even small administrative units were, however such an approach requires precise geo-coding of individuals which is difficult to achieve on a very large scale.

Review articles have concluded that the spatial context has a relatively small, although usually significant, effect on health outcomes (Arcaya et al. 2016; Diez Roux and Mair 2010; Macintyre et al. 2002; Meijer et al. 2012; Pickett and Pearl 2001). However, spatial effects tend to be stronger for outcomes related to physical health compared to those related to subjective measures of well-being (Boyle and Willms 1999), and a systematic review found that place effects are often stronger for younger ages (Meijer et al. 2012). Xu et al. (2014) argued that child health outcomes may be susceptible to very local effects because children have a more limited daily activity space around their home.

Some of the relatively few studies that explicitly examine spatial influences, independent of individual-level influences, on historic health outcomes, have failed to take proper account of the multi-level structure of the data (Heady et al. 1955; Reid 1997). Others have concentrated on specific urban case studies that, although allowing valuable insights into intra-urban variation, do not allow wider aspects of rural versus urban areas to be examined, nor the different ways that intra-urban effects operate in different cities (Connor 2017; Thornton and Olson 2011; Xu et al. 2014). This paper investigates the independent roles of context and composition, and class and place, for the whole of England and Wales, using multi-level models. The considerations outlined above suggest that although individual influences are still expected to account for the most variation in mortality outcomes, place-level effects may well be important, particularly for the risk of death during infancy and early childhood.

Ideally, we would use independent variables that accurately measure each of the proximate determinants of mortality at the levels on which they are thought to operate: at an individual level these might include education, income, aspects of housing quality and sanitation, and at a contextual level community cohesion, area-level water, sanitation and health provision, pollution indicators, and so on. For many historic settings these are simply not available. It is common to use other indicators as proxy variables, but these cannot usually identify particular proximate determinants of mortality. The observation made above, that different variables operate at different spatial levels, means that variation at different spatial levels can provide clues about the contextual influences on mortality outcomes.

Data and methods

Our dataset is derived from the individual-level data for the 36 million people enumerated in the 1911 census of England and Wales.[2] We use a mortality outcome

---

[2]This work is based on I-CeM, a standardized, integrated data set of most of the censuses of Great Britain for the period 1851–1911; see K. Schürer and E. Higgs, Integrated Census Microdata (I-CeM); 1851–1911 [computer file]. Colchester, Essex: UK Data Archive [distributor], April 2014. SN: 7481,

based on responses to the special questions asked of married women in relation to their current marriage: about duration of marriage, children ever born and children who had died before the census was taken. Following previous work, the dependent variable in our analysis is the Mortality Index (MI) calculated for each individual woman (Connor 2017; Dribe et al. 2020; Preston and Haines 1991; Reid 1997; Garrett et al. 2001). This measures the ratio between the actual number of child deaths experienced by a woman and the expected number, where the latter is based on her fertility and marital duration: a value of one indicates that a woman had lost exactly the number of children expected given the number of children she had borne, how long she had been married, and overall mortality levels.[3] We use women married for less than 15 years with their husband present in the household and with valid data, who had given birth to at least one child.[4] Our measures of mortality are therefore essentially couple-level measures, although in our regressions we weight them by the number of children ever born so that they represent the risk to an individual child. Our regressions include a mortality reference date, which is an estimate of the date to which the mortality estimate refers, but this is still a family-level estimate (Dribe et al. 2020).[5]

It is important to realize that we do not have data on the sex, birth order, or age at death of children who died, and these important sources of variation at the child level cannot be controlled for. Our measure of mortality relates to children from birth up to the age of 15, but is heavily weighted towards infancy and early childhood. This is partly because mortality is highest soon after birth, but it is also a consequence of the fact that we use information from women married for up to 15 years. Only women married for 15 years could have had a child who died at age 15, but women of all marital durations could have had a child who died in infancy. Our measure is therefore neither a pure measure of mortality during infancy nor at any

[3]Our calculations use the England and Wales lifetable for 1911 as a standard. See Garrett et al. (2001: 459-65) for more details on the calculation of the Mortality Index.

[4]We excluded married women whose husband could not be identified as being present in the household partly because we are interested in social class and need the occupation of a woman's husband to identify this, and partly to exclude women who had separated from their husband or who incorrectly reported themselves to be married. We also excluded the answers of widowed and unmarried men and women who answered these questions, people with missing marital duration, implausible age at marriage, or inconsistent data about children born, surviving and died. Where answers were mistakenly written against a married man instead of his co-resident wife we transferred them to the latter. Of women married for less than 15 years, half of one percent were excluded due to invalid data, and a further 3.5% because they could not be linked to a husband in their household. Infant mortality rates calculated using indirect estimation techniques and the numbers of children born and died in I-CeM are very similar to those calculated using the numbers published in the official census report (HMSO 1923), indicating that the I-CeM data are of high quality.

[5]The children of women married for 0–4 years will have been younger at death, on average, than the children of women married for 10–15 years. The UN Manual X (1983) provides an equation and multipliers, to be used with indicators of children ever born, to allow the calculation of the average date to which the mortality estimates for each marital duration group of women apply. Here, these dates are calculated for groups of women and applied to the individual women in each group.

specific age within childhood, but we refer to it as 'infant and child mortality', or sometimes 'child mortality' for brevity. We also do not know the age of the wife at the birth of any children not in the household, although in our regression models we control for both her age at census and that of her husband.

Our multivariable regressions include a suite of independent variables measured at the individual level, and also variables measured at two area levels. The rest of this section describes the variables and areas used, firstly at the individual level and then at the area level.

Table 1 provides summary statistics for the individual-level variables. We are particularly interested in the status or social class of couples and the effect that has on the survival of their children. We use the social class classification designed for analysis of the 1911 census and related vital statistics, which has five hierarchical levels and three separate occupational groups (Szreter 1996). The five levels range from professional and managerial occupations in social class 1 (high class) to unskilled manual occupations in social class 5 (low class), and the three separate occupations are textile workers, miners, and agricultural laborers. These three occupational groups were separated out by the Registrar General because of their atypical fertility and infant mortality experience. Many occupations within both the textile and mining industries were skilled, but infant mortality among these occupations was higher than among other skilled workers. In contrast agricultural laborers were considered unskilled, but their infant mortality was much lower than that for unskilled laborers working in other sectors. The mortality social class gradient in the five hierarchical levels is therefore much stronger when these groups are separated out than it would be if they were merged with their skill levels. Nevertheless, we use the scheme for comparability with previous research (Garrett et al. 2001; Preston and Haines 1991; Reid 1997). Social class is based on husband's occupation because the majority of women gave up work on marriage. We also identify those without a social class, which could indicate that the husband was unemployed, did not work because he could afford not to, or that his occupation was un-classifiable. This category is hard to interpret but numerically very small.

In order to tease out how social class affected mortality we would ideally use data on individual or household characteristics that could represent the mechanisms through which class affected survival. At a most proximate level these could include access to health services, water sources, and toilet facilities. Unfortunately, the census does not provide such information, nor does it provide indicators of education, income or wealth which would provide clues as to these mechanisms. We are limited to the information available – the number of servants, the size of house, birthplace, women's work status and position in the household – to represent differences in class.

Many middle-class families employed live-in servants and we use couples with none, one, two, and three or more live-in servants to distinguish levels of income among the upper- and middle-classes. The 1911 census recorded the number of rooms in each house (excluding kitchen and bathroom), and we treat this as a categorical variable, distinguishing those living in fewer than three rooms, those living in 3–5 rooms and those with at least six rooms in their household.[6] House size could reflect income, but may additionally represent facilities or relative crowding.

---

[6]Those reporting over 30 rooms are placed in the missing category together with other non-numeric answers as these are likely to be errors in recording or transcription.

Table 1. Summary statistics for individual-level variables

Variable	Category	Mean	Percentage	MI
Wife's age		32.00		
Husband's age		34.23		
Children ever born		2.68		
Mortality reference year		1907.40		
Social class	Social class 1*		9.89	0.671
	Social class 2		15.29	0.844
	Social class 3		24.24	0.949
	Social class 4		17.32	0.986
	Social class 5		16.15	1.200
	Textile workers		3.19	1.180
	Miners		9.26	1.270
	Agricultural laborers		3.62	0.790
	Missing social class		1.04	0.998
Wife's birth place	Great Britain*		94.91	1.000
	Ireland		1.00	1.060
	Eastern Europe		0.52	0.803
	Rest of Europe		0.45	0.844
	Other		0.51	0.781
	Missing		2.62	1.070
Husband's birth place	Great Britain*		94.87	1.000
	Ireland		1.18	1.140
	Eastern Europe		0.61	0.818
	Rest of Europe		0.71	0.919
	Other		0.50	0.884
	Missing		2.13	1.040
Wife's work status	Wife does not work*		91.99	0.969
	Wife works at home		1.95	1.170
	Wife works outside home		6.06	1.470
Servants	No live-in servants*		92.60	1.030
	1 servant		5.04	0.660
	2 servants		1.32	0.534
	3+ servants		1.04	0.449
Wife's household position	Wife of head*		98.04	0.997
	Not wife of head		1.96	1.180

(Continued)

Table 1. (*Continued*)

Variable	Category	Mean	Percentage	MI
Housing	<3 rooms		27.29	1.440
	3-5 rooms*		16.73	1.020
	6+ rooms		55.53	0.722
	Missing		0.45	1.000

Source: I-CeM dataset.

Notes: Couples married for less than 15 years, with valid information on fertility variables ($n = 2{,}295{,}993$). MI is the weighted mean for women in the category (weighted by number of children born ($n = 6{,}155{,}142$)).

Categories used as reference categories in regression analysis are identified by *.

We chose not to calculate a measure of persons per house because this has a negative relationship with child mortality, which we interpret as inverse causality: families with few surviving children tended to have less crowded houses at the time the crowding is measured (Garrett et al. 2001: 136–9). This serves as a reminder that this analysis combines cross-sectional independent variables with an outcome that took place up to 15 years before the census, and a family's circumstances may well have changed between the time a child was at risk of death and the census (Reid et al. 2016).

Wife's work is another variable that may have been particularly subject to change across the lives of her children. In this era, most women stopped working when they married or had children, and we interpret a working wife as a sign of poverty, but it is also possible that child mortality (through low effective parity) enabled a woman to return to work (Garrett and Reid 1994). The only women who can be identified as working are those who recorded a paid occupation at the time of the census, and instead of trying to distinguish different occupations, we have divided women into those with no recorded occupation (does not work), those who returned an occupation – other than housework or home duties – and said that they carried out that occupation at home, and those who worked outside the home.[7] We anticipate that those who were able to take paid work that they could do at home were better able to combine work with childcare.

We include indicators of husbands' and wives' countries of birth, distinguishing those born in Ireland, Eastern Europe, and the rest of Europe. Migrants are often selected for better health or social status, leading to a 'healthy migrant effect', and we expect most migrants to have lower mortality. We singled out those from Ireland and those from Eastern Europe as two interesting and potentially atypical groups. Migrants from Ireland tended to be low-skilled and poor and therefore might have had higher mortality. Many migrants from Eastern Europe in this period were ethnic Jews fleeing persecution and they were therefore less likely to be positively selected for health than other migrant groups. Nevertheless, we anticipate that they would have enjoyed a mortality advantage conferred by the high standards of hygiene that Jewish communities attached to food preparation (Derosas 2003;

---

[7]Unfortunately, the census does not distinguish between part-time and full-time paid work.

Marks [1994]) and their receptiveness to modern medical ideas (Riswick et al. [2022]). Other work has suggested that Jewish enclaves in European cities had low mortality because their relatively closed communities minimized the transfer of infectious diseases from outside the community (van Poppel et al. [2002]).[8]

Our final individual-level variable is the wife's household position. Here, we distinguish wives whose husband is not the head of the household, as such couples may have been disadvantaged by not being able to afford a household of their own.

Our data is multilevel: individual couples are nested within two levels of geographical area, which are used to examine variation in our multilevel models. We also assess the extent to which area-level variation can be explained by indicators, measured at the area levels, that are used as proxies for influences on health. We first describe British census geography and the areas we use, before explaining the area-level indicators.

When working with English and Welsh census data we are constrained by time, effort, and the availability of boundary data to use the geographical units that were used for census data collection. In this period, the English and Welsh census geography consisted of a number of nested units. At the highest level, the countries were divided into regions, with each region containing a variable number of administrative counties. Each county was divided into RDs: there were 634 of these across England and Wales in 1911. Each RD contained between 1 and 14, but typically three to six, RSDs. In 1911, the populations of RSDs ranged from 300 to over 150,000 inhabitants. The mean population per RSD was around 18,000 individuals, but the distribution was highly skewed, and the modal figure was just 3000–4000. There were 2009 RSDs in England and Wales in 1911, and each RSD was itself subdivided into up to 40 enumeration districts (EDs). There were around 35,000 EDs in total, with a mean population of 650, but this disguised a bi-modal distribution: rural EDs held a mean of 300 people, while urban EDs held 1400 inhabitants on average.[9]

RSDs are the smallest unit for which the Registrar General published infant mortality statistics and there have been several useful analyses of infant mortality at this scale (Williams [1992]; Mooney [1994]a, [1994]b; Sneddon [2006]). As already noted, RSDs could be very large and were therefore internally heterogeneous. Some cities were covered entirely or mainly by a single RSD. Even RSDs with fairly small populations could contain a number of diverse areas. Many towns were amalgamated with areas of surrounding countryside, while larger towns and smaller cities could be divided into two or more sections, each of which could be combined with a different part of the surrounding rural area. It is partly due to the internal heterogeneity of RSDs that we have chosen here to explore EDs as well as RSDs: the internal diversity of RSDs means that they are less likely than EDs to capture health effects that operate at relatively small scales. We examine both EDs

---

[8]It has also been suggested that Jewish minorities still in Russia had low-infant mortality because they were relatively well educated (Glavatskaya [2018]).

[9]Another geographic unit identified in the census is parish. We opted not to use parish in this analysis because although many parishes are small units, they were much more variable in size than EDs, with more very small (less than 50 people) and more very large (more than 6,000 people). Particularly populous parishes, usually urban areas which had grown rapidly over the 19th century, were coterminous with RSDs and are therefore likely to be less good at differentiating local neighborhoods.

Table 2. Summary statistics for area level variables

Variable	ED (*n* = 34,130)		RSD (*n* = 2,009)	
	mean	st dev	mean	st dev
Social class 1 (% of men 15–64 years)	8.71	9.68	7.34	5.16
Social class 5 (% of men 15–64 years)	13.73	10.59	13.15	6.19
Textile workers (% of men 15–64 years)	2.55	8.16	2.69	7.93
Miners (% of men 15–64 years)	5.69	16.00	6.71	15.45
Agricultural laborers (% of men 15–64 years)	10.72	15.63	13.23	13.29
Households with 1 servant (% of households)	3.92	6.21	5.98	3.21
Households with 2 servants (% of households)	1.28	2.68	1.63	1.52
Households with 3+ servants (% of households)	1.00	2.65	1.17	1.64
Households with <3 rooms (% of households)	9.28	12.94	8.30	9.49
Households with 6+ rooms (% of households)	30.07	20.02	31.95	12.30
Born in Eastern Europe (% of population)	0.17	1.99	0.08	1.17
Born in Ireland (% of population)	0.61	1.78	0.39	1.12
Population (number)	1,056	614	17,954	20,074

Notes: The average number of valid cases per ED is 67 women and 180 children, and 1143 women and 3064 children per RSD. *Source: I-CeM dataset.*

and RSDs because the literature presented above suggested that different influences on health could operate at different scales and capturing variation at these two levels allows us to test for the different sorts of influence even where we lack specific data. We therefore calculate the same area-level variables at both scales to allow us to judge at which scale they operate. Table 2 shows summary statistics for these variables at both ED and RSD levels.[10]

In the absence of detailed information about contextual influences on health such as sanitation, pollution, health services, and social cohesion, our area-level variables take the form of the percentages of: working men in particular social groups; people born in Eastern Europe; people born in Ireland; households with servants; in large houses; and in small houses. Once individual-level social class, housing, and birthplace are controlled for, these variables can represent the additional contextual effects of living in particular types of area, as discussed below. We would have liked to include population density, but although this is available for RSDs, EDs have not been geographically mapped and therefore their areas are not known. However, in the light of the facts that more densely populated RSDs tend also to have larger populations, and (as already noted) urban EDs also tend to have larger populations, the log of population was used as a proxy for population density. In regressions, these area-level variables were standardized so that the coefficients represent the increase in the MI associated with an increase of one standard deviation in the

---

[10]We combine EDs with fewer than 100 people with the preceding or following ED in enumeration order, ensuring that both were within the same RSD.

independent variable. In our final models, we also include a dummy variable for London.

In order to examine the effect of measuring variables at different spatial scales, we perform three sets of OLS regressions: two-level (household- and ED-level) regressions with random intercepts for EDs and area-level variables calculated at the ED level; two-level (household- and RSD-level) regressions with random intercepts for RSDs and area-level variables calculated at the RSD level; and three-level (household-, ED-, and RSD-level) regressions with random intercepts for both ED and RSD levels, and area-level variables calculated for both.[11]

Each set of regressions consists of 8 models:

- Model 0: Null model
- Model 1: Control variables (husbands' and wives' ages, parity, mortality reference date)
- Model 2: Control variables and husband's social class
- Model 3: Control variables and area-level occupational percentages
- Model 4: Control variables, husband's social class and area-level occupational percentages
- Model 5: Control variables, husband's social class, and other individual-level variables
- Model 6: As model 5, with area-level variables (except log of population)
- Model 7: As model 6, with log of population and dummy variable for London

These models are summarized in Tables A1, A2, and A3 in the online supplementary material. Table A4 provides the unadjusted coefficients from models that include each variable (or group of related variables) controlling only for wife's and husband's age, parity, and mortality reference date. As robustness checks, we ran versions of models 2 and 5 with fixed effects for EDs and RSDs respectively (Table A5), but these are not discussed in the text as they are very similar. We also ran versions of key models for women married for less than 10 years that are neither shown nor discussed but confirmed our results.

Descriptive results: class and place

We start by presenting some illustrative analyses. Figure 1 shows the MI calculated for the eight social classes aggregated across England and Wales, and for places classified into eight different types.[12] The upper panel indicates a clear mortality gradient within the five hierarchical social classes: the children of men belonging to Class 1 – professional and managerial occupations – had the lowest mortality, while the children of those in Class 5 – unskilled laborers – had the highest. The three singled-out occupational groups also showed distinctive mortality experiences.

---

[11]Previous analyses using the MI as an outcome variable compared ordinary least squares (OLS), tobit, and probit regressions, which all gave very similar results, leading authors to focus on OLS for ease of interpretation (Garrett et al. 2001: 468-70; Trussell and Preston 1982).

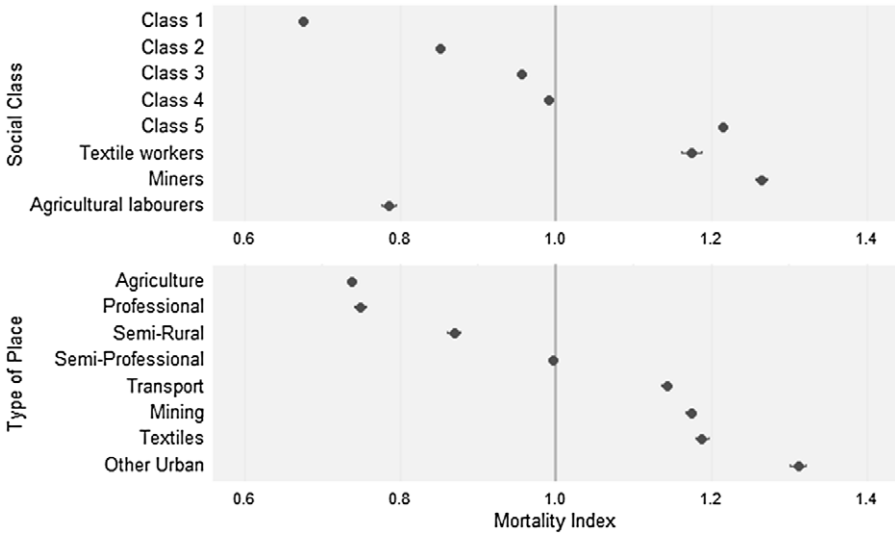[12]See www.PopulationsPast.org and Reid et al. (2020) for details of this categorization.

Figure 1. Mortality Index by social class (above) and type of place (below). *Source: I-CeM dataset. Notes:* horizontal lines show 95% confidence intervals.

The children of textile workers and miners had mortality levels just below and just above those of Class 5, respectively; perhaps higher than would be predicted from the skill level of those occupations. In contrast, the children of agricultural laborers had a very low risk of death; again not as predicted by the skill or status level of those workers.

The lower panel of Figure 1 shows that there was also a strong gradient between different types of place. Agricultural areas and suburbs with relatively high proportions of professional workers had low child mortality, while places specializing in mining or the textile industries, as well as other urban industrial areas, held far higher risks for infants and young children. Miners were, of course, disproportionately likely to live in mining areas, textile workers in textile areas, agricultural laborers in agricultural areas, and Class 1 in professional areas. Other urban and transport areas had high proportions of unskilled laborers. These results cannot identify whether the place pattern was produced by the concentration of people of particular social classes into different types of place (composition) or whether place exerted an independent effect on mortality (context).

Figure 2 starts to disentangle the mortality effects of class and place by presenting the MI for different class and place combinations. Exactly the same data points are shown in the left- and right-hand panels, but are arranged in a different order: the left-hand panel groups the social class results within each type of place while the right-hand panel allows easier comparison of each class across different types of place. The class gradient within each type of place is clear from the left-hand panel, but the fact that it does not entirely explain the differences in mortality by type of place suggests that geographical differences in mortality are likely to be the result of contextual as well as compositional effects. Within each social class, the differences in mortality according to residential location were at least as big as the differences
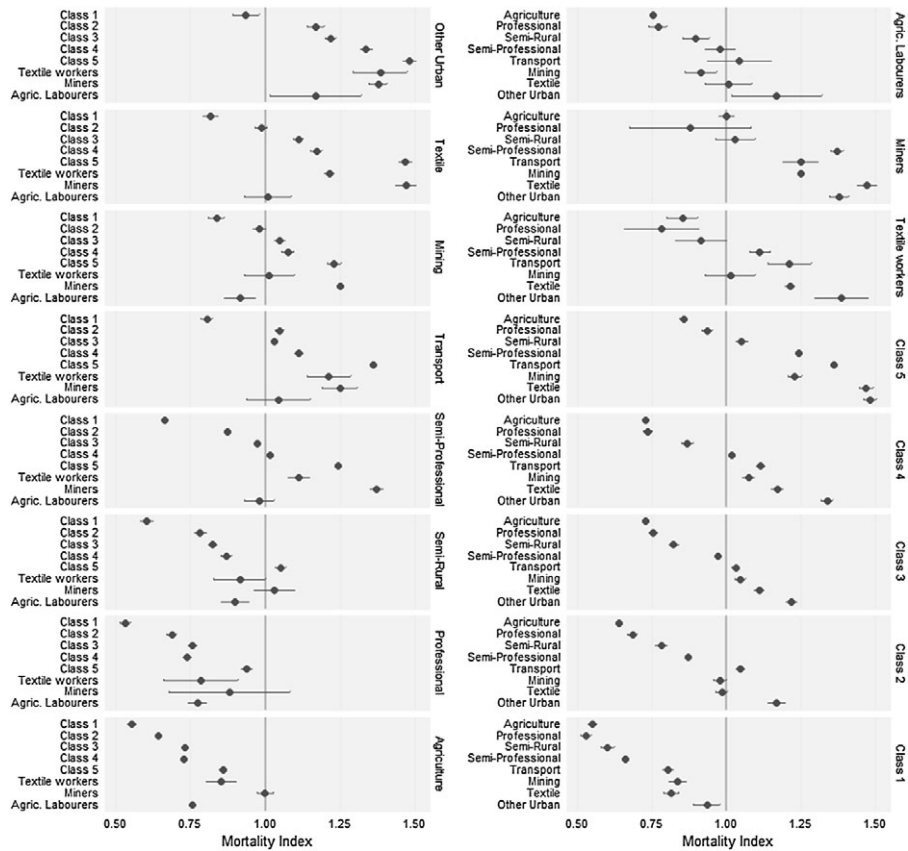
Figure 2. Mortality Index for social classes within (left) and across (right) types of place. *Source: I-CeM dataset*. *Notes:* horizontal lines show 95% confidence intervals.

between the classes within places. The multilevel regressions allow these contextual effects, and the scale at which they operate, to be explored in more detail.

Contextual results and the effect of scale

One way to examine the effect of scale is to consider the intraclass coefficients (ICCs), also known as variance partition coefficients, from the multilevel models (Castelli et al. 2013; Dribe et al. 2017; Monsalves et al. 2020). These show the proportions of unexplained variation that operate at each level and enable the detection and quantification of contextual effects.[13] Comparisons of the ICCs for the ED and RSD versions of each two-level model and for the three-level model can shed light on the scale at which those effects operate. ICCs for the different models are shown in Figure 3, where each pair of connected dots shows the ICC values for

---

[13]The formula for the ICC is $V_a/(V_a+V_i)$ where V_a is the variance between areas and V_i is the variance between individuals within areas (Merlo et al. 2018).
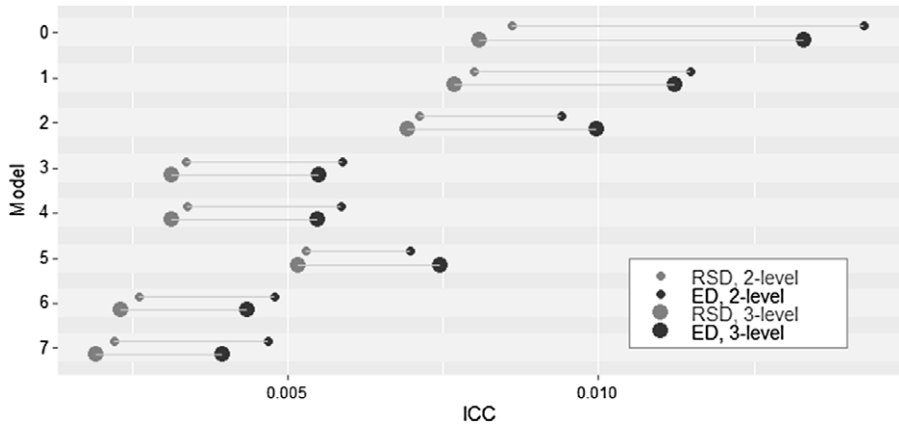
**Figure 3.** ICC values for different multilevel models. *Source: I-CeM dataset.*

RSD (light grey) and ED (dark grey) levels. The upper pair of dots for each model shows the values for the two-level models while the lower pair shows the values for the three-level models.

The ICC values shown in Figure 3 are low, indicating that the vast majority of the overall variation in mortality risk cannot be explained by area-level or contextual factors (at least not variation at the ED or RSD level). However, it is important to remember that models of mortality at an individual level are rarely able to explain much variation. Some variation will be attributable to known but unmeasured influences – here these will include sex, birth order, breast-feeding, season of birth, genetic factors, and the timing of epidemics in relation to the age of child. There is, however, also a strong element of chance affecting which children die young, and this will never be picked up in models, even with much more extensive data.[14] This means we do not expect contextual factors to explain high percentages of variation: we are more interested whether any contextual effects remain after controlling for individual-level effects, and the scale at which contextual effects operate.

The ICC values for the null model (model 0) effectively identify the extent to which there is variation between different units, and this shows that there was nearly twice as much variation in mortality among EDs than among RSDs. This confirms our hypothesis that smaller areas (EDs) are better at picking up contextual or area-level effects on early age mortality in the early twentieth century. Nevertheless, the three-level model, which controls for variation at both ED and RSD levels, shows that there is independent variation at both these levels: some influences on mortality act at a broader geographic scale.

The ICCs are reduced by the addition of individual-level variables (indicated by the difference between model 0 and models 1, 2, and 5) suggesting that a small amount of the mortality effect of places is produced by the concentration of people with particular mortality risks in different areas, in other words by the composition

---

[14]As Daniel Scott Smith wrote, 'Quantitative analysis in history is more relevant and interesting when the differences between groups are small and the variance that is explained is low than when the opposite situation occurs. Large differences and a high R^2 quite often involve relationships that are obvious' (Smith 1984: 144).
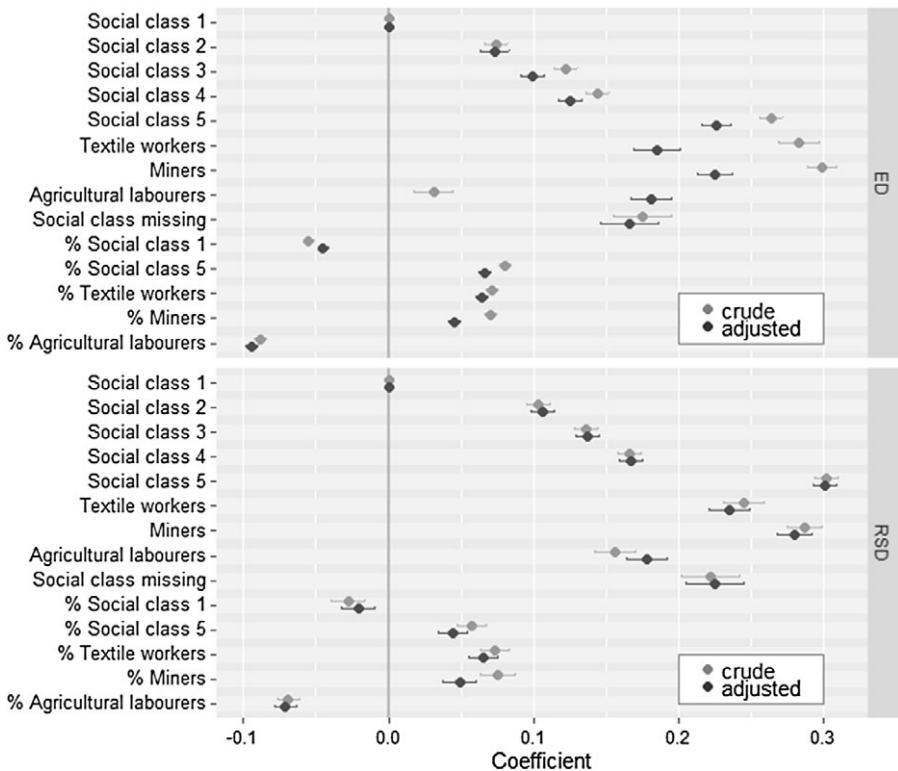
Figure 4. The effects of social class measured at individual level (Social class 1 to Social class missing) and area level (% Social class 1 to % Agricultural laborers), comparing area level variation when this is controlled, and effects are measured, at ED (upper panel) and RSD (lower panel) levels. Note: Crude coefficients include wife's and husband's ages, parity, mortality reference date and EITHER social class at individual level OR social class at area level. Adjusted coefficients include ages, parity, mortality reference date, and social class measured at both individual and area levels: i.e. these are the coefficients from multilevel models 2, 3, and 4 in Tables A1 and A2. Horizontal lines show 95% confidence intervals.

of those places. Understandably, it is the addition of independent variables capturing specific characteristics of geographic areas that has the largest effect on ICCs (captured by the difference between models 0, 1, 2, and 5 that only contain individual-level independent variables, and models 3, 4, 6, and 7), and we discuss what these might represent below.

The composition versus context issue can also be examined by comparing the effects of social class measured at the individual level and at area levels on early age mortality. This is illustrated in Figure 4 which shows the coefficients for individual-level social class and coefficients for the area-level percentages in classes 1, 5, and the three special classes. The crude models (light grey dots) show either the individual- or the area-level variables only (from models 2 and 3 in Tables A1 and A2), and the adjusted models (dark grey) include both (from model 5 in Tables A1 and A2).[15]

---

[15]These models also include controls for husband's and wife's age, parity, and the mortality reference date, but no other independent variables at either individual or area scales.

The crude results show a strong (although a little uneven) gradient in social classes 1–5, and also an effect of social class measured at the area level. If the area-level effects were simply the result of the fact that social classes with higher mortality risks tended to live in certain sorts of place, we would expect the magnitude of the area-level coefficients to reduce when individual social class was controlled. If, alternatively, some of the apparent effect of individual social class was due to the residential sorting of people into places with higher or lower risks – so that everyone in an area, whatever their class, faced the risks associated with that area – we would expect the magnitude of the coefficients for individual-level social class to reduce. This latter phenomenon is apparent when area effects are measured at the ED level: individual social class differences are reduced much more than those for area levels, confirming results using a small subset of this dataset (Reid 1997). In contrast, when the unit of measurement for contextual variables is RSDs, controlling for the percentages in social class groups has virtually no impact on the individual-level social class coefficients, suggesting that any contextual effects that operate at an RSD level are not closely linked to the occupational composition of the area. In other words, the scale at which area characteristics are measured is linked to the way that local environments affect health risks.

It is notable, although unsurprising, that the coefficients for the three special occupational groups (textile workers, miners, and agricultural laborers) change most on the addition of area-level indicators. This is largely because such workers tended to live in places with large occupational concentrations. It is possible that textile workers and miners incurred a mortality penalty due to high levels of air pollution, poor waste disposal, or insalubrious housing, although some mining, textile and railway companies built housing of a relatively high standard for their workers, so the pathways are not immediately obvious and would merit further investigation at a local level. In contrast, the coefficient for agricultural laborers increases in the adjusted model, indicating that much of their mortality advantage was not because of their own characteristics, but because they tended to live in benign rural environments where there were fewer local environmental hazards. In general, rural housing was more basic than that in urban areas, but this was probably compensated for by low population density which limited the transmission of both airborne and waterborne diseases. When the contextual effects of places where agricultural laborers lived are controlled (i.e. compared to others in such areas), the child mortality of agricultural laborers was more akin to that of other manual workers.

Pathways from class and place to mortality: the effect of other variables

So far we have only considered the effect of social class measured at the individual and area levels, but more specific variables can indicate pathways through which class or context affect the risks of child mortality. The coefficients for these additional variables, taken from the final three-level multilevel model (model 7 in Table A3) are shown in Figure 5, together with unadjusted (crude) coefficients from three-level multilevel models (Table A4). Generally, the results are as expected: the individual-level categorical indicators tended to be associated with larger effects
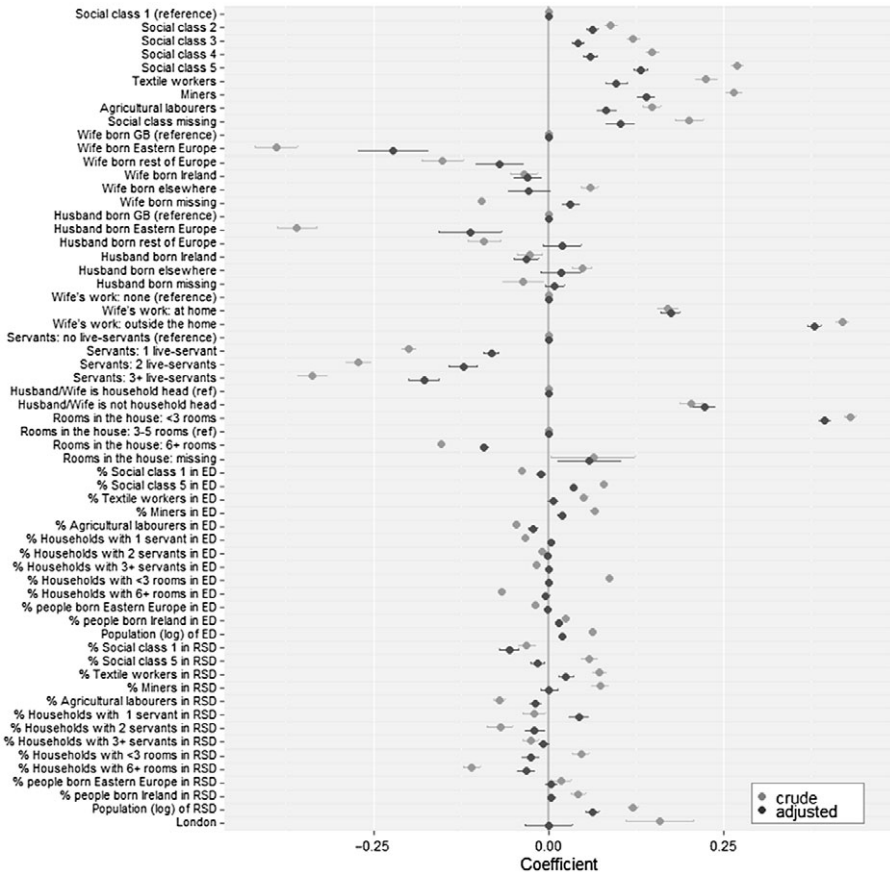
Figure 5. Influences on child mortality: unadjusted and final coefficients (I-CeM, 3-level multilevel models). *Notes:* Crude coefficients include wife's and husband's ages, and parity, and mortality reference date (from Table A4). Adjusted coefficients include all variables (model 7 of Table A3). Horizontal lines show 95% confidence intervals.

than the contextual variables, and individual-level social class is likely to have affected child mortality through many of these variables. For example, people of higher socioeconomic status lived in large houses and had live-in servants, and these statistically explain some of their mortality advantage. It is possible that these variables reflected aspects of housing and living circumstances that directly affected child health, but they could alternatively simply be capturing unmeasured aspects of resources, status, or knowledge better than social class based on husband's occupation. Working wives had higher child loss, with a particular disadvantage for those whose work took them outside the home.[16]

Results relating to the nativity of a child's parents are particularly interesting. As expected, most migrant categories (with the exception of those born outside Europe) tended to have better child survival than natives, and mother's nativity was generally

---

[16]As already noted, it is possible there is some reverse causality here.

more important than that of the father. The advantage for children with parents from Eastern Europe was particularly large. These migrants tended to live in close proximity to each other, and at the ED level a higher percentage of Eastern European migrants indicates lower mortality when other variables are not controlled. The fact this effect disappeared in the fully adjusted model indicates that this was a purely compositional effect: the mortality advantage of migrants from Eastern Europe was not strongly linked to their residential concentration or to neighborhood environments, but are more likely to have been the result of their behavioral and cultural practices, which did not spread to their locally-born neighbors. In contrast, there does appear to be a contextual effect associated with Irish-born couples: EDs with higher percentages of Irish-born tended to have higher mortality. Despite this, and the probable lower skill levels of the Irish-born, they also demonstrated a 'healthy migrant' effect for their children. Marks has attributed low infant mortality among both Irish and Jewish mothers in London to ethnically specific community organizations (Marks 1990).

This paper is particularly interested in the contextual influences on child mortality. It is notable that some of the area-level variables have a larger effect at the RSD level (percentage in Class 1, percentage in textiles, both variables measuring housing size, and population size), while others have a greater effect at the ED level (percentages in Class 5, miners, agricultural laborers, and those born in Ireland). It is not obvious how to interpret these differences, but it is possible that the RSD-level variables, once other variables are controlled, are measuring broad characteristics of the housing stock and amenities that may operate at a town rather than neighborhood level. It is notable here that at the RSD level high percentages of both very small and large houses are associated with lower mortality, and that high percentages of households with one servant are associated with higher mortality. More research is needed on the geography of housing provision and amenities to understand this.

A high percentage of textile workers in an RSD (but not an ED) is associated with higher child mortality, whereas the percentage of miners indicates high mortality in an ED (but not an RSD). This could reflect the size of the built-up areas housing these workers: mining was often concentrated in villages but whole towns were dominated by the textile industry. These variables, then, might be picking up facilities or hazards relating to whole textile or mining areas: avenues worth exploring might be industrial pollution or municipal water supply. In contrast, high percentages of men working as unskilled laborers (Class 5) are associated with high mortality in EDs, but (controlling for this and individual-level class) with lower mortality in RSDs. Within larger areas it seems that there were pockets of local disadvantage that could be related to uneven implementation of municipal facilities, which may have come last to the poorest communities within an area.

The log of population, used as a proxy for population density, had a larger effect when measured for RSDs than when measured for EDs. Although this might reflect genuine differences in the scale at which population density operated, it is also possible that population is a less good proxy for population density at ED level than at RSD level. London was not associated with higher mortality in the full model, demonstrating that the other variables did a good job of explaining urban disadvantage.

Limitations

Historical analysis is constrained by the form and content of data gathered many years in the past, and this introduces a number of limitations to our study. The information on early age mortality from the census is derived from the reports by married women of the numbers of children they had given birth to and had lost to mortality. We therefore have no information about the characteristics of the deceased children, such as their sex, birth order, or cause of death, and we cannot distinguish child- from family-level frailty although this has been shown to be important (Bengtsson and Dribe 2010). This also means we have no precise information on how old the children were when they died, and we are therefore unable to identify which influences act on mortality at different stages of infancy and childhood, although as explained in the methods section, our measure of mortality is strongly weighted towards infancy.

There is also an issue related to the fact that census data are cross-sectional, but the information on child survival is retrospective. This means that the children of women who had moved into an area shortly before the 1911 census may have lived for the bulk of their lives elsewhere, and therefore the area characteristics attributed to them may be inappropriate – an issue which might particularly impact migrants, urban dwellers and the lower classes (O'Campo 2003; Reid et al. 2016). This issue could be overcome by limiting the analysis to married couples who had been present in an area throughout their childbearing lives, but the information needed to determine that is not available. It would be shortsighted to exclude migrants, however, as both short- and long-distance movements affect the population composition of areas of origin and destination. Our paper considers only international migration, but future research could investigate ways in which the characteristics of areas acted as pull and push factors, driving migration which, in turn, further shaped these areas and the health outcomes of people living within them.

Further limitations arising from the census are that information on child survival is only given for married women, and that we have only used married women who could be linked programmatically to their husband in the same household. Illegitimate children are thus automatically excluded, as are those with at least one parent dead or living elsewhere, all of whom are likely to have been more vulnerable to mortality, and arguably more sensitive to their local circumstances. Nevertheless, there is currently no other systematic individual-level mortality data available for research on this period in England and Wales, so the data used here make a very important contribution to our knowledge of the influences on mortality in this era of British history.

Other limitations of our study relate to spatial issues. Firstly, without precise geocoding of individuals we were unable to define egocentric neighborhoods which might be best able to capture spatial effects (Xu et al. 2014). Secondly, any spatial autocorrelation between neighboring areas in our data might have the effect of over-estimating the significance levels of spatial effects (Manley et al. 2006; Xu et al. 2014). Although previous work on the data set used here established that spatial autocorrelation at the RSD level was not significant (Jaadla and Reid 2017), Xu et al. (2014) found that spatial autocorrelation mainly operates at a very small scale. We

were not able to investigate this as it was beyond the scope of our project to map the boundaries of all EDs in England and Wales.

The fact that we wanted to investigate the whole of England and Wales at a small spatial level means that we were very limited in the independent variables available to us. While sources such as Medical Officer of Health reports can provide quite detailed information about a variety of local conditions (including water supply, sanitation, street cleaning and waste disposal, paving, and local nuisances) these were not produced in a standard format, and their survival is not uniform. More systematic data reflecting the installation of municipal services such as water supply and sewage systems is usually available only for large spatial units such as cities. In addition, the indicators available, such as the value of loans, often do not translate directly into facilities and are rarely available for units that match up with mortality statistics (Harris and Helgertz 2019). We were therefore constrained to use measures that we could calculate from the census itself. Despite the difficulties in working out which proximate determinants, or mechanisms of mortality causation, these represent, the multi-level structure has provided a useful overview and allowed the identification of some fruitful avenues for future research. In particular, further investigation to determine which contextual factors are important for infant and child mortality would benefit from detailed local studies as these would allow a wider variety of influences to be considered at fine spatial scales.

Conclusions

Our paper has undertaken a series of multi-level models of infant and child mortality in England and Wales in the early twentieth century. We have at least partially confirmed the findings of an earlier paper (Reid 1997) that argued that part of the infant mortality advantage for the higher social classes in this period could be attributed not to superior knowledge, attitudes, or the ability to purchase better health care or food, but to the opportunity to live in a more salubrious environment with fewer environmental hazards and better local amenities. In other words, places with high mortality did not have high mortality simply because they contained a lot of people whose individual characteristics put them at risk, but also because of various contextual aspects of those places. In common with other investigations into the effects of composition and context, we found that although there was a clear effect of geographic context on mortality, most variation in mortality risks operated at the individual level, but was not purely attributable to social class.

Our comparison of the contextual effects at different spatial levels indicated that these operated not only at the ED, or neighborhood, level but also at the larger RSD level, and that these were often mediated by different variables. Housing stock and population (density) affected mortality when measured across large, but not small areas, whereas relatively large proportions of unskilled laborers were associated with higher mortality only at the smaller spatial scale. More research is needed into how to interpret these indicators, but investigation of facilities such as water supply, sanitation, paving, street cleaning, and waste disposal and how these were implemented variously across districts within particular areas may be helpful.

The fact that the percentages of various social and occupation groups were still significant as contextual indicators, particularly at the smallest local scale, even after these were controlled at an individual level, indicates that the composition of a community might itself influence the context. This could operate through aspects such as local cohesion or community influence over local decisions and spending. These can impact the availability and standard of facilities and resources which in turn will influence rents and house prices which can be an important mechanism for sorting, or selecting, people into areas, making context and composition both conceptually and practically deeply entwined. In this period, as the wealthy moved out to healthy residential areas with higher rents, the poor were left in the areas with more environmental hazards and poorer facilities. Better-off people were able to lobby for the early introduction of new amenities, which therefore came later to the areas where the poor lived. Less good facilities kept rents comparatively low, making such areas more attractive to those with smaller budgets, thereby 'sorting' the poor into such areas.

The reality of the social class influences on early age mortality in the early twentieth century in England and Wales is far more complicated than is suggested by Preston and Haines's (1991) pronouncement that, in terms of the determinants of mortality differentials, social class represented in Britain what race represented in the USA. At a basic level, the use of the Registrar General's eight category classification, which separated out three large and anomalous occupational groups, artificially amplifies the gradient in the five hierarchical classes. This paper has also demonstrated that one of the main ways that class affected mortality was through the sorting of different people into areas that exposed them to different risks, but these areas were themselves molded by class.

References

Antonovsky, Aaron (1967) "Social class, life expectancy and overall mortality." Milbank Memorial Fund Quarterly 45 (2): 31–73.

Arcaya, Mariana C., Reginald D. Tucker-Seeley, Rockli Kim, Alina Schnake-Mahl, Marvin So, and S.V. Subramanian (2016) "Research on neighborhood effects on health in the United States: A systematic review of study characteristics." Social Science & Medicine (168): 16–29.

Atkinson, Paul, Brian Francis, Ian Gregory, and Catherine Porter (2017) "Spatial modelling of rural infant mortality and occupation in 19th century Britain." Demographic Research 36 (44): 1337–60.

Barker, David J. P., and Clive Osmond (1987) "Inequalities in health in Britain: specific explanations in three Lancashire towns." British Medical Journal (294): 749–52.

Bengtsson, Tommy, and Martin Dribe (2010) "Quantifying the family frailty effect in infant and child mortality by using median hazard ratio (MHR)." Historical Methods 43 (1): 15–27.

Bond Huie, Stephanie A. (2001) "The concept of neighbourhood in health and mortality research." Sociological Spectrum 21 (3): 341–58.

Boyle, Michael H., and J. Douglas Willms (1999) "Place effects for areas defined by administrative boundaries." American Journal of Epidemiology 149 (6): 577–85.

Cage, Robert A., and John Foster (2002) "Overcrowding and infant mortality: A tale of two cities." Scottish Journal of Political Economy 49 (2): 129–49.

Castelli, Adriana, Rowena Jacobs, Maria Goddard, and Peter C. Smith. (2013) "Health, policy and geography: Insights from a multi-level modelling approach." Social Science & Medicine (92): 61–73.

Clouston, Sean A. P., Marcie S. Rubin, Jo C. Phelan, and Bruce G. Link (2016) "A social history of disease: Contextualising the rise and fall of social inequalities in cause-specific mortality." Demography 53 (5): 1631–56.

Connor, Dylan S. (2017) "Poverty, religious differences, and child mortality in the early twentieth century: The case of Dublin." Annals of the American Association of Geographers 107 (3): 625–46.

da Silva Lopes, Katharina, Erika Ota, Prakash Shakya, Amarjargal Dagvadorj, Olukunmi Omobolanle Balogun, Juan Pablo Peña-Rosas, Luz Maria De-Regil, and Rintaro Mori (2017) "Effects of nutrition interventions during pregnancy on low birth weight: an overview of systematic reviews." BMJ Global Health 2 (3): e000389.

Derosas, Renzo (2003) "Watch out for the children! Differential infant mortality of Jews and Catholics in nineteenth-century Venice." Historical Methods 36 (3): 109–30.

Diez Roux, V. Ana, and Christina Mair (2010) "Neighborhoods and health". Annals of the New York Academy of Sciences 1186): 125–45.

Dribe, Martin, J. David Hacker, and Francesco Scalone (2020) "Immigration and child mortality: Lessons from the United States at the turn of the twentieth century." Social Science History 44 (1): 57–89.

Dribe, Martin, Sol Pía Juárez, and Francesco Scalone (2017) "Is it who you are or where you live? Community effects on net fertility at the onset of fertility decline: A multilevel analysis using Swedish micro-census data." Population, Space and Place 23 (2): e1987.

Duncan, Craig, Kelvyn Jones, and Graham Moon (1993) "Do places matter? A multi-level analysis of regional variations in health-related behaviour in Britain." Social Science & Medicine 37 (6): 725–33.

Edvinsson, Sören, Anders Brändström, and John Rogers (2001) "Regional variations in infant mortality in Sweden during the first half of the 19th century," in Lars-Göran Tedebrand and Peter Sköld (eds.) Nordic Demography in History and Present-Day Society. Umeå: Umeå University: 145–64.

Edvinsson, Sören, Ólöf Garðarsdóttir, and Gunnar Thorvaldsen (2008) "Infant mortality in the Nordic countries, 1780–1930." Continuity and Change 23 (3): 457–85.

Farr, William (1859) "On the construction of life-tables, illustrated by a new life-table of the healthy districts of England." Philosophical Transactions of the Royal Society 149: 837–78.

Fernihough, Alan, and Mark E. McGovern (2014) "Do fertility transitions influence infant mortality declines? Evidence from Early Modern Germany." Journal of Population Economics 27 (4): 1145–63.

Fildes, Valerie (1998) "Infant feeding practices and infant mortality in England, 1900–1919." Continuity and Change 13 (2): 251–80.

Flowerdew, Robin, David J. Manley, and Clive E. Sabel (2008) "Neighbourhood effects on health: Does it matter where you draw the boundaries?" Social Science & Medicine 66 (6): 1241–55.

Garrett, Eilidh, Alice Reid, Kevin Schürer, and Simon Szreter (2001) Changing Family Size in England and Wales: Place, Class and Demography, 1891–1911. Cambridge: Cambridge University Press.

Garrett, Eilidh M., and Alice M. Reid (1994) "Satanic mills, pleasant lands: Spatial variation in women's work, fertility and infant mortality as viewed from the 1911 census." Historical Research 67: 157–77.

Glavatskaya, Elena, Julia Borovik, and Gunnar Thorvaldsen (2018) "Urban infant mortality and religion at the end of the nineteenth and in the early twentieth century: the case of Ekaterinburg, Russia." The History of the Family 23 (1): 135–53.

Gregory, Ian N. (2008) "Different places, different stories: Infant mortality decline in England and Wales, 1851–1911." Annals of the Association of American Geographers 98 (4): 773–94.

Haines, Michael R. (1995) "Socio-economic differentials in infant and child mortality during mortality decline: England and Wales, 1890–1911." Population Studies 49 (2): 297–315.

Hanlon, W. Walker (2022) "London Fog: A century of pollution and mortality, 1866–1965." The Review of Economics and Statistics 1-49. https://doi.org/10.1162/rest_a_01190

Harris, Bernard, and Jonas Helgertz (2019) "Urban sanitation and the decline of mortality." The History of the Family 24 (2): 207–26.

Haynes, Robin, Konstantinos Daras, Richard Reading, and Andrew Jones (2007) "Modifiable neighbourhood units, zone design and residents' perceptions." Health & Place 13 (4): 812–25.

Heady, J. A., C. F. Stevens, C. Daly, and J. N. Morris (1955) "Social and biological factors in infant mortality. IV: The independent effects of social class, region, the mother's age and her parity." The Lancet March (5): 499–503.

HMSO (1917) Census of England and Wales, Vol XIII, Fertility of Marriage (Part I), 1911. London: HMSO.
—— (1923) Census of England and Wales, Vol XIII, Fertility of Marriage (Part II), 1911. London: HMSO.
Jaadla, Hannaliis, and Allan Puur (2016) "The impact of water supply and sanitation on infant mortality: Individual-level evidence from Tartu, Estonia, 1897–1900." Population Studies 70 (2): 163–79.
Jaadla, Hannaliis, and Alice Reid (2017) "The geography of early childhood mortality in England and Wales, 1881–1911." Demographic Research, 37 (1): 1861–90.
Klüsener, Sebastian, Isabelle Devos, Peter Ekamper, Ian N. Gregory, Siegfried Gruber, Jordi Martí-Henneberg, Frans Van Poppel, Luís Espinha da Silveira, and Arne Solli (2014) "Spatial inequalities in infant survival at an early stage of the longevity revolution: A pan-European view across 5000 regions and localities in 1910." Demographic Research (30): 1849–64.
Lee, Clive H. (1991) "Regional inequalities in infant mortality in Britain, 1861–1971: Patterns and hypotheses." Population Studies 45 (1): 55–65.
Link, Bruce G., and Jo C. Phelan (1995) "Social conditions as fundamental causes of disease." Journal of Health and Social Behavior 35 (Extra Issue): 80–94.
Macintyre, Sally, Anne Ellaway, and Steven Cummins (2002) "Place effects on health: how can we conceptualise, operationalise and measure them?" Social Science & Medicine 55 (1): 125–39.
Macintyre, Sally, Sheila MacIver, and Anne Sooman (1993) "Area, class and health: should we be focusing on places or people?" Journal of Social Policy (22): 213–34.
Manley, David, Robin Flowerdew, and David Steel (2006) "Scales, levels and processes: studying spatial patterns of British census variables." Computers, Environment and Urban Systems 30 (2): 143–60.
Marks, Lara (1990) "Irish and Jewish women's experience of childbirth and infant care in East London, 1870–1939." PhD diss., University of Oxford.
—— (1994) Model mothers: Jewish mothers and maternity provision in East London, 1870–1939. Oxford: Clarendon Press.
Meijer, Mathias, Jeannette Röhl, Kim Bloomfield, and Ulrike Grittner (2012) "Do neighborhoods affect individual mortality? A systematic review and meta-analysis of multilevel studies." Social Science & Medicine 74 (8): 1204–12.
Merlo, Juan, Philippe Wagner, Peter C. Austin, S. V. Subramanian, and George Leckie (2018) "General and specific contextual effects in multilevel regression analyses and their paradoxical relationship: A conceptual tutorial." SSM - Population Health (5): 33–37.
Monsalves, Maria Jose, Ananta Shrikant Bangdiwala, Alex Thabane, Shrikant Ishver Bangdiwala (2020) "LEVEL (logical explanations & visualizations of estimates in linear mixed models): recommendations for reporting multilevel data and analyses." BMC Medical Research Methodology 20 (1): 3.
Mooney, Graham (1994a) "Did London pass the sanitary test? Seasonal infant mortality in London, 1870–1914." Journal of Historical Geography 20 (2): 158–74.
—— (1994b) "Stillbirths and the measurement of infant mortality rates c. 1890–1930." Local Population Studies (53): 42–52.
Morgan, Nigel (2002) "Infant mortality, flies and horses in later-nineteenth-century towns: a case study of Preston." Continuity and Change 17 (1): 97–132.
Morris, J. N., and J. A. Heady (1955) "Social and biological factors in infant mortality—V. Mortality in relation to the father's occupation, 1911–1950." The Lancet March (12): 554–60.
Mosley, W. Henry, and Lincoln C. Chen (1984) "An analytical framework for the study of child survival in developing countries." Population and Development Review 10 (Supplement: Child Survival: Strategies for Research): 25–45.
O'Campo, Patricia (2003) "Invited commentary: Advancing theory and methods for multilevel models of residential neighborhoods and health." American Journal of Epidemiology 157 (1): 9–13.
Oliver, Lisa N., and Michael V. Hayes (2007) "Does choice of spatial unit matter for estimating small-area disparities in health and place effects in the Vancouver Census Metropolitan Area?" Canadian Journal of Public Health 98 (Suppl 1): S27–34.
Openshaw, Stan (1983) The Modifiable Areal Unit Problem. Concepts and Techniques in Modern Geography No. 38. Norwich: Geobooks.
Owen, Gwilym, Richard Harris, and Kelvyn Jones (2016) "Under examination: Multilevel models, geography and health research." Progress in Human Geography 40 (3): 394–412.
Pamuk, Elsie R. (1985) "Social Class Inequality in Mortality from 1921 to 1972 in England and Wales." Population Studies 39 (1): 17–31.

Pickett, Kate E., and Michelle Pearl (2001) "Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review." Journal of Epidemiology and Community Health (**55**): 111–22.

Preston, Samuel, and Michael Haines (1991) Fatal Years: Child Mortality in Late Nineteenth Century America. Princeton: Princeton University Press.

Ramiro-Fariñas, Diego and Alberto Sanz-Gimeno (2000) "Structural changes in childhood mortality in Spain, 1860–1990." International Journal of Population Geography 6 (1): 61–82.

Reid, Alice (1997) "Locality or class? Spatial and social differentials in infant and child mortality in England and Wales, 1895–1911," in Carlo A. Corsini and Pier Paulo Viazzo (eds.) The Decline of Infant and Child Mortality: The European Experience, 1750–1990. The Hague: Martinus Nijhoff: 129–54.

—— (2017) "Infant feeding and child health and survival in Derbyshire in the early twentieth century." Women's Studies International Forum 60: 111–19.

—— (2021) "Why a long-term perspective is beneficial for demographers." Population Studies 75 (Sup.1): 157–77.

Reid, Alice, Eilidh Garrett, and Simon Szreter (2016) "Residential mobility and child mortality in early twentieth century Belfast," in Diego Ramiro-Fariñas, and Michel Oris (eds.) New Approaches to Death in Cities during the Health Transition. Dordrecht: Springer: 55–76.

Reid, Alice, Hannaliis Jaadla, Eilidh Garrett, and Kevin Schürer (2020) "Adapting the Own Children Method to allow comparison of fertility between populations with different marriage regimes." Population Studies 74 (2): 197–218.

Riswick, Tim, Sanne Muurling, and Katalin Buzasi (2022) "Exploring the mortality advantage of Jewish neighbourhoods in mid-19th century Amsterdam." Demographic Research 46 (25): 723–36.

Schürer, Kevin, and Edward Higgs (2014) Integrated Census Microdata (I-CeM), 1851–1911 [Data collection SN-7481-1]. Colchester: UK Data Service.

Smith, Daniel Scott (1984) "A mean and random past: The implications of variance for history." Historical Methods 17 (3): 141–48.

Smyth, Fiona (2008) "Medical geography: understanding health inequalities." Progress in Human Geography 32 (1): 119–27.

Sneddon, Sam (2006) "A double penalty? Infant mortality in the Lincolnshire Fens, 1870–1900," in Eilidh Garrett, Chris Galley, Nicola Shelton, and Robert Woods (eds.) Infant Mortality: A Continuing Social Problem. Aldershot: Ashgate: 79–97.

Stafford, Mai, Oliver Duke-Williams, and Nicola Shelton (2008) "Small area inequalities in health: Are we underestimating them?" Social Science & Medicine 67 (6): 891–99.

Szreter, Simon (1996) Fertility, Class and Gender in Britain, 1860–1940, Cambridge: Cambridge University Press.

Szreter, Simon, and Michael Woolcock (2004) "Health by association? Social capital, social theory, and the political economy of public health." International Journal of Epidemiology 33 (4): 650–67.

Tarkiainen, Lasse, Pekka Martikainen, Mikko Laaksonen, and Alastair H. Leyland (2010) "Comparing the effects of neighbourhood characteristics on all-cause mortality using two hierarchical areal units in the capital region of Helsinki." Health & Place 16 (2): 409–12.

Thornton, Patricia and Sherry Olson (2011) "Mortality in late nineteenth-century Montreal: Geographic pathways of contagion." Population Studies 65 (2): 157–81.

Thorvaldsen, Gunnar (2002) "Rural infant mortality in nineteenth century Norway." Hygiea Internationalis 3 (1): 75–113.

Trussell, James, and Samuel Preston (1982) "Estimating the covariates of childhood mortality from retrospective reports of mothers." Health Policy and Education (**3**): 1–36.

United Nations (1983) Manual X: Indirect Techniques of Demographic Estimation. New York: UN Department of Economic and Social Affairs.

Van den Boomen, Nynke, and Peter Ekamper (2015) "Denied their 'natural nourishment': Religion, causes of death and infant mortality in the Netherlands, 1875–1899." History of the Family 20 (3): 391–419.

van Poppel, Frans, Jona Schellekens, and Aart C. Liefbroer (2002) "Religious differentials in infant and child mortality in Holland, 1855–1912." Population Studies 56 (3): 277–89.

Wang, Huanhuan, Peter Brimblecombe, and Keith Ngan (2021) "A numerical study of local traffic volume and air quality within urban street canyons." Science of the Total Environment (**791**): 148138.

Watterson, Patricia A. (1988) "Infant mortality by father's occupation from the 1911 Census of England and Wales." Demography 25 (2): 289–306.

Williams, Naomi (1992) "Death in its season: Class, environment and the mortality of infants in nineteenth-century Sheffield." Social History of Medicine **5** (1): 71–94.

Williams, Naomi, and Chris Galley (1995) "Urban–rural differentials in infant mortality in Victorian England." Population Studies **49** (3): 401–20.

Williams, Naomi, and Graham Mooney (1994) "Infant mortality in an 'age of great cities': London and the English provincial cities compared, c. 1840–1910." Continuity and Change **9** (2): 185–212.

Woods, Robert, and Naomi Williams (1995) "Must the gap widen before it can be narrowed? Long-term trends in social class mortality differentials." Continuity and Change **10** (1): 105–37.

Woods, Robert I., Patricia A. Watterson, and John H. Woodward (1988) "The causes of rapid infant mortality decline in England and Wales, 1861–1921 Part I", Population Studies **42** (3): 343–66.

—— (1989) "The causes of rapid infant mortality decline in England and Wales, 1861–1921. Part II", Population Studies **43** (1): 113–32.

Xu, Hongwei, John R. Logan, and Susan E. Short (2014) "Integrating space with place in health research: A multilevel spatial investigation using child mortality in 1880 Newark, New Jersey." Demography **51** (3): 811–34.

Kevin Schürer is a Principal Research Associate at the Cambridge Group for the History of Population and Social Structure, University of Cambridge. He is a Fellow of the Academy of Social Sciences and before returning to Cambridge worked as Director of the UK Data Archive (University of Essex) and as Pro-Vice-Chancellor for Research (University of Leicester). He has published widely on a range of topics within historical demography, migration and household structure and is currently preparing a digital version of the 1921 census returns of England and Wales for future research.

Alice Reid is Professor of Demography at the Department of Geography, University of Cambridge, UK, and Co-Director of the Cambridge Group for the History of Population and Social Structure. She specialises in the demography of the British Isles in the 19th and early 20th centuries and has recently published on maternal mortality (Annales de Demographie Historique, Social Science and Medicine), infant and child mortality (Historical Life Course Studies, Demographic Research), fertility (Population Studies, Demography) and the value of a long-term perspective in demography (Population Studies).

Hannaliis Jaadla is a senior research associate at the Cambridge Group for the History of Population and Social Structure, University of Cambridge and a research fellow at the Estonian Institute for Population Studies, Tallinn University. Her research broadly addresses the dynamics and mechanisms of demographic change in historical context and has recently published on social inequalities in health and mortality (The Economic History Review and Population Studies) and on fertility transition (Population Studies, Demography, Spatial Demography).

Eilidh Garrett is currently a Senior Researcher at the Scottish Centre for Administrative Data Research (SCADR), University of Edinburgh. Her research interests lie in the 19th and 20th century historical demography of the British Isles. She has extensive experience of working with civil registers, large-scale, individual-level census data and published statistics. Her recent publications include co-authored articles in Populations Studies, International Journal for Population Data Science, and Historical Life Course Studies.

Sarah Rafferty is a Senior Research Officer at the UK Office for National Statistics, currently working on health and pandemics. Additionally, Sarah recently defended her PhD thesis on infant and child mortality in London (1896–1910) at the Cambridge Group for the History of Population and Social Structure, University of Cambridge. She has published on historical demography and infectious diseases in Local Population Studies, the Journal of Historical Geography and Social Science and Medicine.