DEMOGRAPHIC RESEARCH

A peer-reviewed, open-access journal of population sciences

DEMOGRAPHIC RESEARCH

VOLUME 37, ARTICLE 58, PAGES 1861- 1890
PUBLISHED 12 DECEMBER 2017
http://www.demographic-research.org/Volumes/Vol37/58/
DOI: 10.4054/DemRes.2017.37.58

Research Article

The geography of early childhood mortality in England and Wales, 1881–1911

Hannaliis Jaadla

Alice Reid

Contents

Demographic Research: Volume 37, Article 58
Research Article

The geography of early childhood mortality in England and Wales, 1881–1911

Hannaliis Jaadla[1]

Alice Reid[2]

Abstract

BACKGROUND
Considerable regional variation existed in 19th-century infant mortality (IMR) in England and Wales.

OBJECTIVE
This study estimates early childhood mortality (ECMR) for over 2,000 registration subdistricts (RSDs) of England and Wales and analyses spatial and temporal variations in IMR and ECMR between 1881 and 1911.

METHODS
The combination of mortality statistics from the Registrar General and individual-level census data from the Integrated Census Microdata (I-CeM) project is used to estimate spatial models of the relationship between early childhood death rates and a range of district-specific contextual variables.

RESULTS
All regions of England and Wales experienced noticeable decline in early childhood mortality, but the spatial patterns were remarkably persistent, with high mortality in London and in the mining and textile centres. The earlier decline of childhood than infant mortality produced a widening gap between them, and in early phases this development was concentrated along the East-Midlands coastal area from Suffolk to North Yorkshire, and in mid-Wales. This gap continued to widen, and in 1911 IMR was at least twice as high as ECMR in most parts of England and Wales.

CONCLUSION
The changing spatial pattern of ECMR was influenced by a set of factors over and

[1] Cambridge Group for the History of Population and Social Structure, Department of Geography, University of Cambridge, UK and Estonian Institute for Population Studies, Tallinn University, Estonia. E-Mail: hj309@cam.ac.uk.
[2] University of Cambridge, UK.

above those that influenced IMR, and these were related more to the disease environment than to social and economic influences.

CONTRIBUTION

These new estimates of early childhood mortality, at a finer spatial scale than previously possible, highlight the vast spatial variation in mortality in England and Wales. It is likely that these regional differences also manifest in variation in other demographic outcomes.

1. Introduction

Recently there has been a resurgence of interest in spatial demography (Voss 2007; Gutmann et al. 2011; Matthews and Parker 2013; Howell, Porter, and Matthews 2016). In historical demography the analysis of spatial patterns and changes over time has largely focussed on fertility transition. This includes pioneering work by the Princeton European Fertility project (Coale and Watkins 1986) and more recent studies taking spatial econometric approaches to investigate the spread of fertility decline (Bocquet-Appel and Jakobi 1998; Haines and Hacker 2011; González-Bailón and Murphy 2013; Goldstein and Klüsener 2014; Klüsener, Dribe, and Scalone 2016). However, the spatial analysis of historical mortality has thus far mainly focussed on infant mortality and described subnational mortality patterns (Ramiro-Fariñas and Sanz-Gimeno 2000; Edvinsson, Brändström, and Rogers 2001; Thorvaldsen 2002; van den Boomen 2015) or focussed on specific urban environments (Thornton and Olson 2011; Reid, Garrett, and Szreter 2016; Connor 2017). A more limited amount of cross-country analysis has also concentrated on infant mortality (Edvinsson, Garðarsdóttir, and Thorvaldsen 2008; Klüsener et al. 2014).

Previous research on mortality changes in late 19th-century England and Wales has demonstrated that there was substantial geographic variation both in the levels of infant and early childhood mortality and in the timing and rates of their declines. These differences suggest that the geography of the relationship between infant and child mortality will also have changed over time. In a seminal paper on the relationship between infant and adult mortality, Woods (1993) mapped the ratio between infant mortality rates (IMR) and early childhood mortality rates (ECMR: the probability of death between exact ages 1 and 5) in England and Wales. He noted the ‘intriguing’ geographical relationship, whereby ECMR was higher than IMR both in major urban areas and in rural areas of the far West and North, and suggested that this was most likely attributable to the underregistration of infant deaths (Woods 1993: 206, 216). In a subsequent paper he and colleagues appear to have accepted the geographical and

temporal variation in the relationship between infant and child mortality as a real phenomenon and examined the relationship between infant and early childhood mortality using the English Life Tables (for England and Wales as a whole) and the decennial mortality rates for the 614 Registration Districts of England and Wales (Woods, Williams, and Galley 1997). They pointed out that in the first place the 'rule' encapsulated by the Coale–Demeny life tables (1983) that IMR would always exceed ECMR (only south levels 1 to 5 have ECMR above IMR) was unrealistic – as this occurred often for small geographical units such as urban areas but also for the whole country in a few particular years. Guillot et al. (2012) argue that high values of ECMR compared to IMR reflect high levels of mortality from infectious diseases as such conditions affect children rather than infants, and that this ratio therefore reflects disparities in local epidemiological contexts. Excess ECMRs in urban areas were also confirmed in 19th-century Spain by Ramiro-Fariñas and Sanz-Gimeno (2000).

Woods, Williams, and Galley (1997) noted that the relationship between IMR and ECMR was both temporally and geographically variable across the second half of the 19th century in England and Wales. This pattern was strongly related to population density and, like Guillot et al. (2012), they attributed it to the different epidemiological regimes in urban and rural areas. They argued that rapid urban growth without commensurate development of the urban sanitary infrastructure made towns and cities more vulnerable to outbreaks of diarrhoeal disease, particularly among infants and children in the summer months. They also noted that exposure to diseases sensitive to overcrowding, such as measles and scarlet fever, would be higher in urban areas, as such places would have a larger pool of susceptibles and greater ease of transmission (see also Woods, Watterson, and Woodward 1988, 1989; Williams and Mooney 1994). However, even their relatively well-fitting polynomial relationship between IMR and ECMR underpredicted ECMR in many urban areas with high population densities, and they did not revisit Woods' earlier observation of excess ECMR in rural areas of the North and West.

In line with most of the research on early age mortality, spatial analysis of mortality in England and Wales has concentrated on the first year of life. As stated above, in his initial analysis of geographical patterns in infant mortality in England and Wales, Woods argued that trends in mortality were strongly driven by urban areas, which contained the bulk of the population and where infant mortality stayed high until the end of the 19th century (Woods, Watterson, and Woodward 1989; Woods 1997). Others identified earlier declines in rural areas (Lee 1991; Williams and Galley 1995), and recent contributions using more sophisticated spatial analysis of registration districts in England and Wales confirmed both the importance of the urban story in driving the overall pattern of infant mortality decline over the second half of the 19th century and also the existence of a variety of rural trends (Gregory 2008; Atkinson et al.

2017a, 2017b). Rural areas in the South-East of the country experienced early mortality declines, but those in the North and West saw only minimal improvements.

Gregory's and Atkinson et al.'s analyses therefore highlight considerable regional variation in both levels and declines in infant mortality independent of the urban hierarchy. However, the choice of spatial scale can reduce variation, and the units used for England and Wales in all the above mentioned studies are counties or registration districts, and even the latter (the smaller of the two) often contained large populations spread over both urban and rural areas. Studies of selected smaller areas have shown that registration districts can contain considerable variation in the risk of death due to variations in population density, occupational and industrial make-up, environmental hazards, and local disease environments (Garrett and Reid 1994, 1995; Reid 1997; Garrett et al. 2001). Although infant mortality data for smaller units – registration subdistricts (RSDs) – do exist, these have rarely been used, except in smaller regional studies (Williams 1992; Mooney 1994a, 1994b; Sneddon 2006).

Previous analyses, beyond noting excess child mortality in urban (and sometimes rural) areas, have done little to explain how and why child mortality varies in relation to infant mortality. This paper will further describe these patterns as well as identify factors associated with differences in early childhood mortality and provide a more detailed geography of infant mortality for England and Wales than has previously been available. In our analyses we estimate early childhood death rates (deaths divided by mid-year population, or ${}_4M_1$) and for comparison also calculate the probability of dying between the exact ages of 1 and 5 (${}_4q_1$) used by Woods in his various works. This paper is part of a larger project analysing fertility decline in Victorian England and Wales.[3]

2. Data and methods

The main data sources for this study are the published mortality statistics from the Quarterly and Decennial Reports of the Registrar General, and the individual-level census data from an enhanced version of the Integrated Census Microdata (I-CeM) project (Higgs et al. 2013; Schürer and Higgs 2014).[4] This combination of mortality data is necessary because the Registrar General's statistics were not published for the same geographical units throughout the study period, nor did the details published

[3] This project uses the own children method to estimate age-specific fertility rates for subregistration districts and therefore requires not only the calculation of infant mortality for the same spatial units but also estimates of childhood mortality (www.geog.cam.ac.uk/research/projects/victorianfertilitydecline).

[4] Improvements to this enhanced version of I-CeM (I-CeM.2), produced by the members of the Fertility Atlas project and other members of the Cambridge Group for the History of Population and Social Structure, together with Kevin Schürer, mean that household variables are better specified and individuals are more accurately allocated to RSDs (Schürer et al. 2016).

remain the same.[5] Because we needed estimates of early childhood mortality for RSDs for our broader project, we first calculated spatial models of the relationships between infant mortality and early childhood death rates at the RD level and then used these to estimate early childhood mortality for subdistricts. As well as allowing us to estimate mortality measures for RSDs, these models afford a window onto the geographical relationship between IMR and early childhood mortality and indicate how this changed over time in the light of different geographical trajectories in mortality at different ages.

Our modelling strategy is to estimate separate ordinary least squares (OLS) regression models for 1881, 1891, 1901, and 1911. Formally, the model is specified as follows:

$$\gamma_i = X_i\beta_j + \varepsilon_i \,,$$

where γ_i, the outcome variable, is the log-transformed early childhood age-specific death rate ($_4M_1$, expressed per 1,000) in RD i provided in the Registrar General's Decennial Report for England and Wales.[6] The β_j refers to a vector of j regression parameters and the X to a matrix of independent variables. These include the infant mortality rate (IMR, expressed per 1,000 births) and the ratio of measles deaths to deaths under age 1.[7] Measles prevalence is used to capture differences in local early childhood disease environments, particularly those relating to diseases that are transmitted from person to person, such as measles. Other aspects of disease environments that affect food- and water-borne diseases, or those transmitted by animal or insect vectors, are captured using the log-transformed population density measured by people per acre.[8]

Independent variables related to the socioeconomic setting include the illegitimacy ratio, the social class structure, and measures of education and health employees. It is

[5] From 1869 onwards, deaths under the age of 1 were published for RSDs in the Quarterly Report, while deaths for single years up to the age of 5 were published annually (up until 1881) and decennially (until 1901) but only on a larger spatial scale – for the registration districts (RDs). The Decennial Report for the last decade of our study period (1901–1910) published mortality data for the RDs only for three age groups: under 1-, 1–2-, and 2–4-year-olds.

[6] In a few cases, data were missing from the Decennial Report and were replaced with an estimate aggregated up from the Quarterly Report: For 1881, Whitechapel (RD: 17) and Stratton (RD: 289) birth estimates were missing from the Decennial Report (1871–1880), so RD-level birth counts were estimated from the Quarterly Report (1876–1881). Toxteth Park (RD: 454) was missing entirely and the necessary information was drawn from the next Decennial Report (1881–1890). For 1901, Melton Mowbray (RD: 410) birth counts were missing from the Decennial Report (1891–1900), so the RD-level birth counts and infant mortality rate were estimated from the Quarterly Report (1896–1901).

[7] The ratio of measles deaths to deaths under age 1 was used to ensure comparability with the information available in the Quarterly Report.

[8] Acreages for the calculation of RDs were derived from the detailed GISs for each year specially created for this project by Joe Day (2016).

well established that illegitimately born children in this era had considerably higher mortality than those born to married parents, being around twice as likely to die (Reid 2002). This may have been for a variety of reasons, including poverty and household circumstances, feeding regimes, and stigmatisation (Reid 2005), and places with high levels of illegitimacy might be expected to have had higher levels of mortality overall. In order to capture the health risks that illegitimately born children might have been exposed to and their regional variation, we included in the models illegitimacy ratios averaged over the three years centred on each census year.[9]

The measures of occupational or social class structure provide the number of employed males and females in each social class per 100 inhabitants aged 15 and over. We used five classes: upper and middle class, unskilled workers, textile workers, miners, and agricultural labourers.[10] Our indicators of education and health employees measure the presence of education employees per 100 children aged 7–14 and health employees per 100 inhabitants,[11] and these are included as indicators of human capital. Educational attainment, or years of schooling, is a common measure of human capital, representing skills that might lead to better employment opportunities and specific knowledge and also greater access and receptivity to sources of information. Improving education may therefore lead to better childcare and domestic practices, and women's education in particular has been identified as one of the most important determinants of early age mortality in low- and middle-income countries (Hobcraft 1993). The British censuses of this era did not collect information on educational attainment or literacy, so we use the prevalence of employees in education as a proxy for the availability of education. Health employees may similarly signify access to healthcare; however, this variable must be interpreted with caution. Large health facilities generally had a much wider catchment than the RD or RSD, and deaths were registered to the district where they occurred rather than where the deceased lived. Therefore those dying in a hospital may not have lived in the same area, and hospitals are likely to have 'imported' deaths from other areas. It is thus possible that this variable may act as a proxy for mortality inflated by nonresidents (Mooney, Luckin, and Tanner 1999).

Overall, all these contextual variables are characteristics that can be calculated from the individual census returns, plus estimates of infant and measles mortality and illegitimate fertility, which are available from the Annual and Decennial Reports. The

[9] Illegitimacy ratios were calculated by Eilidh Garrett for every RSD in England and Wales from 1851 to 1911, using the numbers of births by legitimacy in the Registrar General's Annual Report. Reporting units changed radically in 1911, so the rates for 1911 are based only on the numbers of births in 1910.

[10] The social classes are defined by the classification introduced by the Registrar General in the 1911 'Fertility of Marriage' Census and are as follows: I upper and middle class, II intermediate class, III skilled workmen, IV intermediate class, V unskilled workmen, VI textile workers, VII miners, and VIII agricultural labourers.

[11] The category of education employees includes schoolmasters and teachers, school service, and others connected with education. The health employees are physicians, dentists, midwives, nurses, medical assistants, and others employed in medical service.

mortality data are drawn from the decade leading up to each census (see Table A-1 of the Appendix). Following initial explorations we also included dummy variables for three large cities – London, Liverpool, and Manchester – as we hypothesised that the extreme size of these places affected their disease environments or socioeconomic structure in ways that we could not capture using the variables at our disposal.

We include infant mortality as an independent variable because this produces the best models for predicting mortality at slightly older ages. Therefore our models cannot be interpreted as completely encapsulating the 'nonmortality' factors that influence childhood mortality because, to the extent that infant and child mortality are subject to the same influences, infant mortality is likely to capture the variation attributable to those shared influences. However, the models do allow us to examine the factors associated with the differences between infant mortality and mortality at older ages and factors correlated with such differences.

These inherently spatial models, where neighbouring RDs and RSDs are likely to share similar contextual characteristics, might be violating one of the basic assumptions of OLS regressions – that those observations are independent of one another. In order to measure these spatial dependencies we explored alternative spatial modelling techniques. Spatial econometric models make it possible to introduce into regression models spatial interactions among neighbouring observations in space. This approach is generally based on applying a spatial lag operator to the dependent variable (spatial lag model) or to the error term (spatial error model). Previous demographic studies have implemented both these spatial models to account for spatial autocorrelation.[12]

To investigate the extent of possible positive spatial autocorrelation and the bias it might introduce to our model estimates we applied the Global Moran's I[13] to the dependent variable and the model residuals. If the residuals showed evidence of statistically significant spatial autocorrelation, we fitted appropriate spatial autoregressive models (Anselin and Bera 1998; Duncan et al. 2012; Bivand, Pebesma, and Gomez-Rubio 2013). The Lagrange Multiplier test was applied to indicate the proper specification for the data and to choose between the spatial lag and the spatial

[12] Spatial lag models have been often used to study fertility change where theoretical motivation is grounded on spatial interaction, diffusion, and social influence (Montgomery and Casterline 1993; Goldstein and Klüsener 2014; Vitali and Billari 2017). Spatial error models, on the other hand, deal with spatial correlation in prediction 'errors' in terms of nuisance. The most common situation would be where data on important variables is missing and indicates a spatial structure (Anselin 2002). This problem is present in most spatial mortality studies and especially so in historical ones where the available data often struggles to capture the full variation within a local disease environment due to factors such as differences in public health and access to medical care.

[13] The formula for Moran's I index is defined as follows: $I = \left(\frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}}\right)\frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(y_i-\bar{y})}{\sum_{i=1}^{n}(y_i-\bar{y})^2}$, where n is the number of spatial units indexed by i and j (in this case RDs) and w_{ij} is a matrix of spatial weights. These weights are defined by the four nearest neighbours.

error models (Anselin 1988). The score tests were always highest for the spatial error model, which is specified as follows:

$$\gamma_i = X_i\beta_j + u_i\,, \qquad u_i = \rho W u_i + \varepsilon_i\,,$$

where the terms for X and β are the same as in the OLS model, but spatial autocorrelation is modelled by the error term u and the associated spatially lagged error term. In the latter, ρ is a spatial lag parameter, W is the spatial weight matrix, and ε_i denotes the vector of independently but not equally distributed error terms (Chi and Zhu 2008).[14]

Once the models were completed, we used them to predict RSD-level age-specific death rates, using values of the independent variables calculated for RSDs. The predicted death rates are estimated only for the five years prior to each census year.[15] For RSDs, infant mortality and the measles ratio were derived from the Quarterly Reports. The other variables were again calculated from the individual level census records (see Table A-1 of the Appendix).[16] Table 1 provides descriptive statistics of all variables at the RD- and RSD-level. As expected the variation in the measures, as captured by standard deviation (SD), is always larger for smaller spatial units.

[14] For all the data preparation and spatial data analysis we used R, and geovisualisation was conducted in ArcGIS version 10.

[15] To minimize the possibility of small numbers leading to random fluctuation in predicted levels of mortality, we averaged IMR and measles deaths over the five years preceding the census year.

[16] For four RSDs in 1901, data were missing from I-CeM: Wing (175.2), Chardstock (269.3), Topsham (271.4), and Haslingden (471.3). Population density for these places could be calculated from the published population figures, and we decided to use the occupational structure of those of the adjacent RSDs that had shared the most similar patterns with the missing RSDs in 1891. We used Edlesborough (175.4), Whitchurch Canonicorum (268.3), East Budleigh (271.1), and Rawtenstall (471.1) respectively.

Table 1: Descriptive statistics for registration districts (RDs) and registration subdistricts (RSDs) in England and Wales, 1881–1911

	1881		1891		1901		1911	
Variables	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Registration districts (RD)								
IMR	129.13	25.41	120.74	25.21	127.57	28.92	106.22	28.07
Measles	0.06	0.04	0.08	0.03	0.07	0.03	0.06	0.03
Illegitimacy ratio	60.72	23.83	54.06	21.93	50.24	19.18	52.60	29.25
Social classes per 100 inhabitants (15+)								
Upper and middle class	3.74	1.47	3.75	1.5	4.38	1.63	4.92	1.94
Unskilled workers	7.38	2.65	7.48	2.75	7.29	2.66	7.05	2.51
Textile workers	2.89	7.10	2.43	6.41	2.19	5.92	2.05	5.76
Miners	2.12	4.99	2.22	5.45	2.33	5.84	2.66	6.61
Agricultural labourers	9.28	7.03	8.13	6.56	6.40	5.39	6.67	5.63
Education employees per 100 children (age 7–14)	3.33	0.83	3.50	0.75	4.38	0.93	5.07	1.23
Health employees per 100 inhabitants	0.34	0.16	0.33	0.17	0.35	0.18	0.42	0.22
Population density (people per acre)	7.18	26.56	7.47	26.23	7.55	25.94	7.39	24.05
Registration subdistricts (RSD)								
IMR	127.02	28.85	123.91	31.85	128.49	35.90	98.14	28.46
Measles	0.06	0.05	0.08	0.06	0.06	0.05	0.07	0.06
Illegitimacy ratio	59.98	35.32	52.89	31.94	49.39	29.94	50.03	37.85
Social classes per 100 inhabitants (15+)								
Upper and middle class	3.57	1.83	3.61	1.89	4.24	2.11	4.84	2.51
Unskilled workers	7.02	3.46	7.21	3.57	7.01	3.34	6.77	3.20
Textile workers	3.92	9.28	3.37	8.55	3.07	7.92	2.85	7.62
Miners	2.17	5.66	2.33	6.22	2.46	6.59	2.91	7.49
Agricultural labourers	9.59	8.21	8.46	7.71	6.79	6.43	7.02	6.79
Education employees per 100 children (age 7–14)	1.89	1.82	1.80	1.83	2.05	2.11	2.67	2.61
Health employees per 100 inhabitants	0.33	0.24	0.31	0.23	0.33	0.25	0.40	0.26
Population density (people per acre)	12.29	39.88	11.63	36.75	10.94	33.73	9.98	27.85
Number of RDs	630		631		631		634	
Number of RSDs	2,175		2,110		2,064		2,009	

Source: Integrated Census Microdata V.2 (I-CeM.2), Registrar General's Annual, Decennial, and Quarterly Reports.

3. Results

3.1 Modelling of child mortality at RD level

Table 2 shows the results and model diagnostics of four OLS and four spatial models. An exploratory analysis of spatial patterns in the outcome variable and residuals using the Global Moran's *I* test indicated positive spatial autocorrelation in all OLS models. As expected, the early childhood death rate was highly spatially correlated with

Moran's *I* values around 0.68–0.70 for all four time points. The OLS model is better at explaining the spatial pattern of childhood mortality in 1891 and 1901 (Moran's *I* for residuals is 0.05 and 0.08) but is only capturing a limited extent of the spatial variation in 1881 (Moran's *I* 0.31). We find (although these maps are not shown) that the high negative residuals are notably spatially clustered in the East Midlands and Norfolk, indicating that the levels of childhood mortality were lower than predicted by the 1881 model.

Table 2: Model estimates for log-transformed early childhood age-specific death rates ($_4M_1$), registration districts in England and Wales, 1881–1911

	1881		1891		1901		1911	
Variables	**OLS**	**Spatial error**	**OLS**	**Spatial error**	**OLS**	**Spatial error**	**OLS**	**Spatial error**
(Intercept)	2.503 ***	2.292 ***	1.862 ***	1.868 ***	1.756 ***	1.752 ***	1.381 ***	1.448 ***
IMR	0.007 ***	0.007 ***	0.008 ***	0.008 ***	0.007 ***	0.007 ***	0.009 ***	0.008 ***
Measles	1.040 ***	0.668 ***	2.360 ***	2.348 ***	2.560 ***	2.465 ***	2.620 ***	2.542 ***
Illegitimacy ratio	–0.001 ***	–0.001 *	0.000	0.000	–0.001	–0.001	0.000	0.000
Social classes per 100 inhabitants (15+)								
Upper and middle class	0.015 *	0.011	–0.002	–0.002	0.002	0.002	0.010 *	0.006
Unskilled workers	–0.003	0.002	0.003 **	0.003 ***	0.000	0.000	0.009 **	0.010 **
Textile workers	0.002	0.002	0.002 ***	0.002 ***	0.004 **	0.003 **	0.004 **	0.004 **
Miners	0.006 ***	0.004 **	0.003 ***	0.003 ***	0.003 **	0.003 *	0.004 **	0.004 **
Agricultural labourers	–0.011 ***	–0.008 ***	–0.004 ***	–0.004 ***	–0.004 *	–0.004 *	–0.006 **	–0.005 *
Education employees per 100 children	–0.032 ***	–0.013	–0.006	–0.007	–0.008	–0.008	–0.011	–0.010
Health employees per 100 inhabitants	–0.294 ***	–0.187 ***	–0.001	0.003	–0.072 *	–0.054	–0.024	0.041
Log(population density)	0.053 ***	0.054 ***	0.038 ***	0.038 ***	0.037 ***	0.035 ***	0.034 ***	0.041 ***
London	0.064	0.048	0.092 ***	0.091 ***	0.129 ***	0.126 **	–0.021	0.014
Manchester	0.103	–0.010	0.020	0.019	0.050	0.047	0.029	0.018
Liverpool	0.249	0.066	0.030	0.029	0.333 *	0.311 *	0.206	0.112
Log-likelihood	369.755	439.747	1031.385	1033.005	412.627	418.220	343.453	357.927
Spatial dependence								
Moran's *I* dependent variable	0.682		0.696		0.678		0.681	
Moran's *I* residuals	0.310 ***	–0.034	0.046 *	–0.001	0.085 ***	–0.002	0.125 ***	–0.022

Note: *p < 0.05; **p < 0.01; ***p < 0.001.
Source: Integrated Census Microdata V.2 (I-CeM.2), Registrar General's Annual, Decennial, and Quarterly Reports.

Statistically significant spatial autocorrelation of OLS residuals and further diagnostic analysis identified that it was appropriate to use spatial error models as alternatives. The regression coefficients in the spatial error models are in most cases in the same direction as in the OLS models; only the magnitude and the significance levels

change. The residuals from the spatial error models do not exhibit spatial autocorrelation.

We find that the coefficients for infant mortality remain fairly constant across time in both the OLS and spatial error models. Unsurprisingly, an increase in infant mortality was associated with higher childhood mortality. A similar pattern emerges with population density, which was included in the models as a log-transformed independent variable. Population density was significantly and negatively associated with early childhood mortality in all four census years. The illegitimacy ratio, which was included to capture the health risks of illegitimately born children, surprisingly indicated lower child mortality for the years leading up to 1881. Areas of particularly high illegitimacy, such as Norfolk, Lincolnshire, and parts of Wales, were also places with high infant mortality but low levels of early childhood mortality, and the implication may be that the severe mortality penalty illegitimate children experienced during infancy did not persist into early childhood.

The measure of measles mortality was included as an indicator of the local disease environment. The coefficients for measles indicate that it is a stronger predictor of early childhood mortality in the years leading up to the later three censuses than in the earliest decade, and this fits well with the general pattern of measles mortality, which rose in the 1880s in England and Wales. The age-specific nature of measles mortality, where the risk of measles mortality is relatively low in infancy compared to the second year of life, helps to explain why the prevalence of measles has additional explanatory power to that of infant mortality. Measles mortality is usually correlated with population density (Woods and Shelton 1997), and the significance of both variables in the models indicates that different forms of disease transmission were involved in the higher rates of child mortality in urban areas. In the absence of a control for measles, it is likely that population density would have captured person-to-person transmission, and the fact that it remains strongly associated with child mortality when measles is controlled indicates that other aspects of the urban disease environment, such as sanitation and water supply, which affect food-, water-, and insect-borne diseases, also played a role in higher risks of death in children over age 1.

In terms of the occupational and social class structure, we find that mining and textile work were associated with higher mortality over the whole time period. Mining and textile areas were notorious for having high risks of infant death, but child mortality in these areas was at least as high and sometimes higher. As expected, the share of agricultural labourers was significantly associated with lower child mortality; rural areas in general experienced the lowest infant and early childhood mortality in the late 19th and early 20th centuries. By the years leading up to 1911, the models demonstrate that the share of inhabitants among upper and middle classes and unskilled workers were linked to higher mortality. Although it is perhaps surprising that places with more

upper and middle class people had higher child mortality in relation to infant, it is possible that this was again picking up the health risks in mostly urban environments where the majority of these classes lived.

The number of employees in the educational sector per 100 children aged 7–14 was a predictor of lower mortality with a fairly constant effect over time but was only statistically significant for 1881. The relative growth of educational employees was possibly reflecting the burgeoning middle class and growth of human capital in certain areas of England and Wales. Similarly, health employees were linked to lower childhood mortality, particularly in 1881. However, by 1911 the share of health employees in the spatial error model was associated with higher mortality although not statistically significant. The change in the sign may stem from the fact that by the early 20th century, deaths occurring in institutions (including hospitals) had increased considerably, but the Registrar General's Reports failed to redistribute these deaths back to the place of residence (Mooney, Luckin, and Tanner 1999). Therefore, it was more likely that high mortality areas coincided with large numbers of employees in hospitals. It is reasonable to expect this to affect child mortality more than infant, as most hospitals (even children's hospitals) at this time did not cater for infants.

Finally, the largest cities London, Manchester, and Liverpool had persistently high excess child mortality once other variables were controlled, although this was not always statistically significant. London is often heralded as having lower than expected infant mortality for its size and density (Williams and Mooney 1994); the same cannot be said about its child mortality.

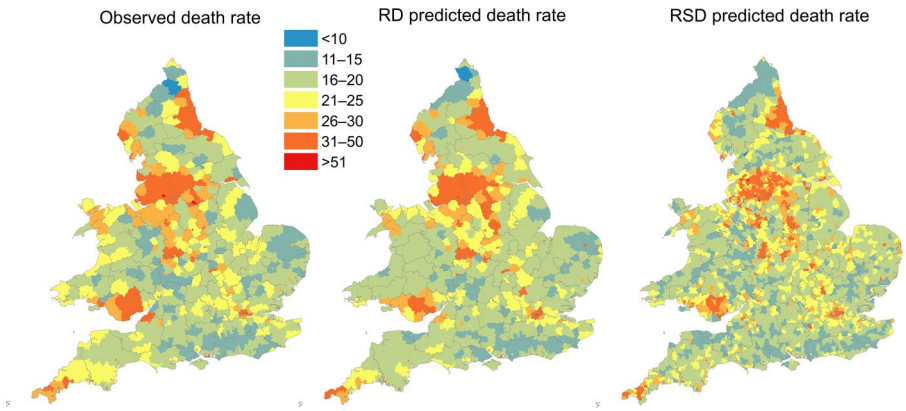
In our models we use IMR to capture a spectrum of influences on ECMR that are not captured by our other independent variables. This in-built endogeneity is not a problem for the predictive regression models, but it can cause problems if the models are interpreted as descriptive or (in particular) explanatory (Shmueli 2010). One of the major concerns is that the dependent variable may actually cause one or more of the independent variables, or in this case that early child mortality is likely to be jointly determined with infant mortality and measles mortality. Therefore we performed sensitivity analyses to analyse the robustness of our descriptive results by estimating another set of models using IMR and prevalence of measles mortality from the previous decade.[17] These yielded substantially similar estimates across all models with only minor differences in the statistical significance and magnitude of some variables (see Table A-2 of the Appendix), enabling us to conclude that observed correlations are not the consequence of short-term factors directly affecting mortality, such as epidemic cycles or particular weather conditions, but are more likely to be indicative of structural factors.

[17] Ideally we would have used an instrumental variable, but we were unable to identify any suitable variables for which data was available at the relevant scale.

3.2 Predicting child mortality at RSD level

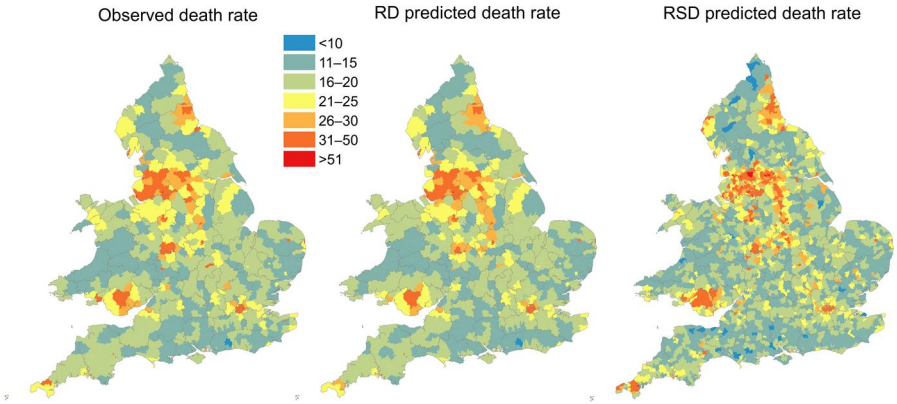
Figures 1 to 4 show observed early childhood death rates between exact ages 1 and 5 for RDs and predicted death rates for RDs and RSDs for each time period. Looking at the spatial visualization of our estimates, it is clear that our models predict each of the general patterns very well, although at the RD level they fail to predict some of the very low or high mortality areas. It is reassuring, however, that the finer detail in the RSDs, and therefore more extreme values of independent variables, leads to the recapture of some of the low and high mortality places. The RSD maps show that there is a considerable amount of local variation, probably due to population density and related contextual characteristics, which is lost through the use of RD figures, and we are confident, therefore, that our RSD predictions show plausible patterns of early age mortality. When looking at change over time it is clear that mortality in early childhood was declining fast, although mortality in the highly populated mining and textile areas and in London remained high, and this coloured the overall early childhood mortality picture. By 1911, there were very few RSDs where predicted ${}_4M_1$ was over 40, and most of those lay in Liverpool, Manchester, or London.

Figure 1: Observed and predicted early childhood age-specific death rates (${}_4M_1$) in England and Wales, 1881



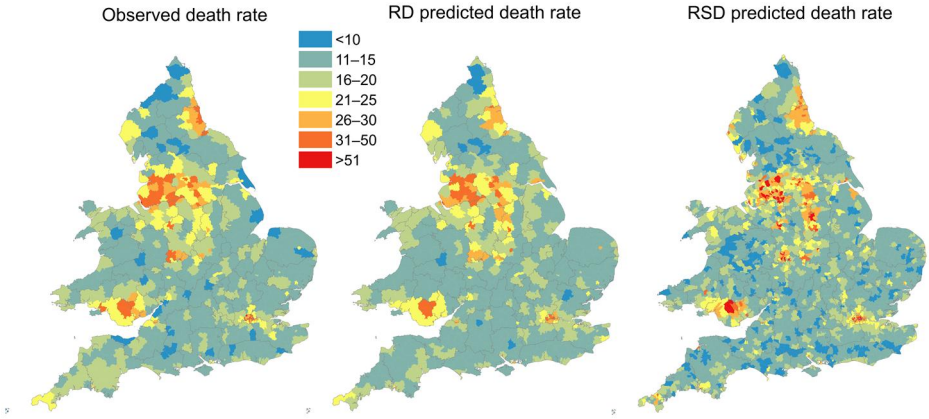
Source: Registrar General's Decennial and Quarterly Reports; own calculations from spatial error models.
Base maps: Day (2016).

Figure 2: Observed and predicted early childhood age-specific death rates ($_4M_1$) in England and Wales, 1891



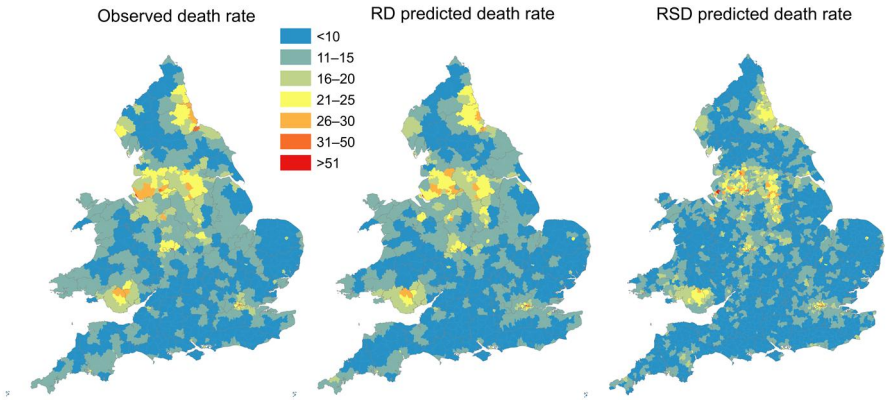
Source: Registrar General's Decennial and Quarterly Reports; own calculations from spatial error models.
Base maps: Day (2016).

Figure 3: Observed and predicted early childhood age-specific death rates ($_4M_1$) in England and Wales, 1901



Source: Registrar General's Decennial and Quarterly Reports; own calculations from spatial error models.
Base maps: Day (2016).

Figure 4: Observed and predicted early childhood age-specific death rates ($_4M_1$) in England and Wales, 1911



Source: Registrar General's Decennial and Quarterly Reports; own calculations from spatial error models.
Base maps: Day (2016).

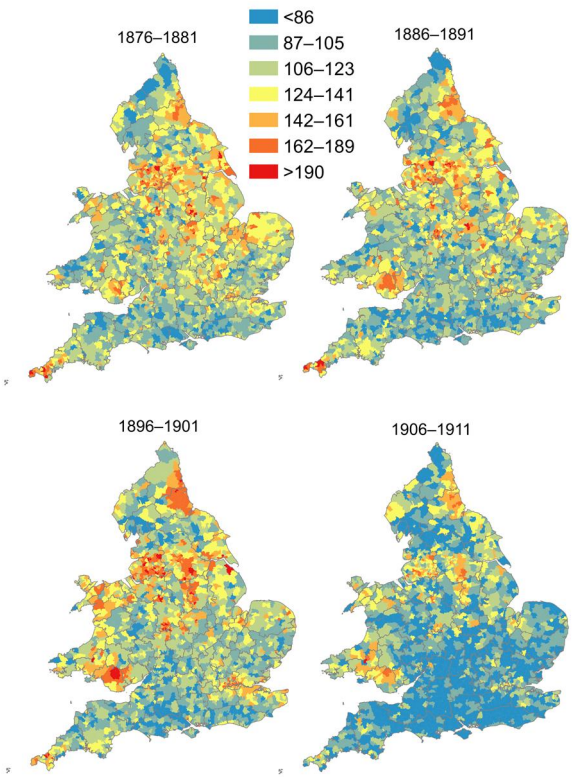
Despite these declines the geographic pattern of mortality risks was remarkably persistent, with high levels in London and the mining and textile centres. There are similarities here to the geography of infant mortality shown in Figure 5. However, comparison between Figures 1–4 and Figure 5 shows that at each time period IMR was far less concentrated than were early childhood age-specific death rates, with the result that there are a considerable number of areas with relatively high or moderate infant mortality but relatively low early childhood mortality. The years leading up to 1901, included in that map, cover some of the 'long hot summers' of the late 1890s, and the effect of the hot weather on raised levels of infant mortality in the urban industrial areas is clearly evident.

Figure 6 compares the ratio of the probability of dying during the first year of life (q_0 or IMR) to the probability of dying between the exact ages of 1 to 5 ($_4q_1$ or ECMR) in each of the four time periods.[18] Very few areas demonstrate excess ECMR in those time periods, and those are generally RSDs located in London, Liverpool, or Manchester. The bluer areas are those with particularly high $_4q_1$ in relation to IMR;

[18] The probability of dying between the exact ages of 1 and 5 ($_4q_1$) was estimated using the equation: $_4q_1 = (4*_4M_1)/(1+4*(1-f^*)*_4M_1)$, where f^* is a separation factor introduced by Chiang (1984). Deaths between exact ages 1 and 5 tend to be concentrated at the younger ages of that four-year age interval, and therefore it is not reasonable to assume linear survivorship through the interval. The f^* value for England and Wales was selected from 'West' Coale–Demeny model life tables – 0.34 for expectation of life at birth ranging from 50 to 55 (Coale, Demeny, and Vaughan 1983).

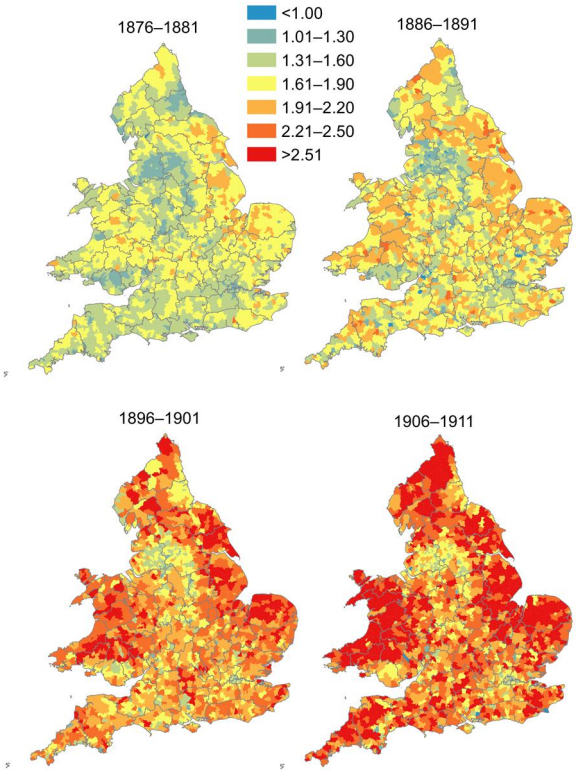
these places included many of the areas where early childhood mortality was particularly high – London, the textile districts, and the coalfields. In these places the risk of dying aged 1–4 years was similar to the risk of dying in infancy. However, over the years the earlier decline of childhood mortality produced a widening gap between IMR and ${}_4q_1$. From 1881 to 1891 and more strongly in 1901, this development was concentrated along the East-Midlands coastal area from Suffolk to North Yorkshire, and in mid-Wales, which are coloured red. Figure 5 shows that, in contrast to ECMR, IMR appears to have declined very little before 1901, although there are signs of reduction in some rural areas.

Figure 5: Registration subdistrict level variation in infant mortality, England and Wales, 1876–1911



Source: Registrar General's Quarterly Report.
Base maps: Day (2016).

Figure 6: Registration subdistrict level variation in the ratio of the infant mortality rate (q_0) to early childhood mortality rate ($_4q_1$), England and Wales, 1876–1911



Source: Registrar General's Decennial and Quarterly Reports; own calculations from spatial error models.
Base maps: Day (2016).

The lack of change in IMR, combined with the rapidly declining ECMR, produces stark change in the ratio of IMR to $_4q_1$. By 1901 most of the country had relatively high levels of infant mortality in relation to early childhood mortality. Even in London and the urban industrial heartlands where child mortality continued to be relatively high, the balance had also shifted towards infant mortality. Despite the onset of decline in IMR in

the years leading up to 1911, the gap continued to widen until IMR was twice as high as or more than ECMR in most parts of England and Wales.[19]

4. Discussion

This paper extends the existing literature on regional variation of early childhood mortality in England and Wales to the level of registration subdistricts. We have also explored the relevance of different demographic, socioeconomic, and contextual characteristics for explaining early childhood mortality variation at the registration district level. Our models confirm previous findings that levels of childhood mortality were strongly related to infant mortality (Woods, Williams, and Galley 1997). Over and above the correlations with infant mortality, the measure of measles mortality, which we interpret as capturing overcrowding and local disease environments although it might also reflect household overcrowding and fertility (Reves 1985), was positively associated with early childhood mortality. Despite this, population density was still independently significant, probably capturing the fact that sanitation and water supply struggled to keep up with population growth in Victorian and Edwardian cities. In line with previous research, the analysis of local occupational and social class structure confirmed that mining and textile districts experienced poor childhood survival, and the districts determined by large share of agricultural workers had relatively low mortality in early childhood (Gregory 2008; Atkinson et al. 2017a, 2017b).

The findings also demonstrate that our models are generally better at explaining the spatial pattern of childhood mortality for the years leading up to 1891, 1901, and to some extent 1911. However, the model is less effective for 1881. This could be related to different causes but might be reflecting our study's general limitations in the ability to measure all district-level differences in the development of public health and sanitation, local infectious disease environment, access to basic healthcare, or women's work and childcare. For example, the labour intensity of women's agricultural work in eastern England has been identified as a possible reason for high levels of infant mortality in that region (Sneddon 2006; Hinde and Fairhurst 2015; Atkinson 2017b). Women's agricultural work may have curtailed breastfeeding and so increased infant mortality but had a less detrimental effect on older children. Another argument is that for the earliest period the absence of a measure representing the prevalence of scarlet fever in different areas might cause additional spatial clustering of measurement errors. Mortality from scarlet fever did not decline markedly until after the 1880s, but

[19] It is important to note that although IMR suffered a resurgence in 1911 due to a particularly hot summer this is not captured in our data, as we use IMR data only up to (and including) the first quarter of 1911.

combined relatively low mortality rates in infancy with significantly higher mortality rates in each of the subsequent four years of age (Woods and Shelton 1997: 76).[20]

The predicted early childhood death rates for RSDs demonstrated considerable regional variation, which is concealed when using RD level estimates. Comparisons of our results with the patterns observed at RD level indicate that our procedures do not produce significant random fluctuations due to small numbers.

It should be remembered that our models do not describe the influences on child mortality per se but excess child mortality over and above the factors that are also correlated with variation in infant mortality. However, this allows us to confirm that the relationship between IMR and ECMR had a very distinct spatial pattern and regional variation across England and Wales, which changed considerably over the course of the late 19th and early 20th centuries. In the years leading up to 1881, excess ECMR was still present in some highly populated districts in Liverpool, Manchester, and London, where the local epidemiological context was mostly driven by high levels of mortality from infectious diseases (Woods and Shelton 1997; Woods, Williams, and Galley 1997). Over time the changes in urban environment and in local sanitary infrastructure contributed significantly to the limitation of infectious disease transmission and improved survival in these large cities. Indicators of social capital – education and health employees – increased over time but do not appear to have been as strongly predictive of excess child mortality as the local disease environment. Even where the levels of infant and child mortality remained relatively high compared to rural areas, the gap between IMR and ECMR widened. Similar patterns of change were present in the urban industrial heartlands of England and Wales. This rapidly changing relationship between infant and child mortality, defining the process of mortality decline all over Europe, cautions against assuming a fixed relationship across space or time (Ramiro-Fariñas and Sanz-Gimeno 2000).

This article contributes to the growing literature using spatial approaches to study historical demographic changes. The analysis of spatial patterns of mortality has become increasingly achievable thanks to constant improvement in the quantity and quality of available data sources and the development of new geographical information system (GIS) databases. Our findings highlight the vast spatial variation in mortality in England and Wales when analysing mortality data at a finer spatial scale than previously possible and suggest that it is likely that these regional differences also manifest in variation in other demographic outcomes such as fertility, nuptiality, illegitimacy, and migration that occurred over the late 19th century.

[20] Scarlet fever was also more common in urban areas (Woods 2000: 258).

5. Acknowledgements

This work was supported by the Economic and Social Research Council (ES/L015463/1). The authors are grateful to Joe Day, Eilidh Garrett, and Kevin Schürer for their help with data and comments on the earlier version of the paper. The authors would also like to thank two anonymous reviewers whose useful comments and suggestions helped to improve this paper.

References

Anselin, L. (1988). Lagrange multiplier test diagnostics for spatial dependence and spatial heterogeneity. *Geographical Analysis* 20(1): 1–17. doi:10.1111/j.1538-4632.1988.tb00159.x.

Anselin, L. (2002). Under the hood issues in the specification and interpretation of spatial regression models. *Agricultural Economics* 27(3): 247–267. doi:10.1111/j.1574-0862.2002.tb00120.x.

Anselin, L. and Bera, A.K. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. *Statistics Textbooks and Monographs* 155: 237–290.

Atkinson, P.D., Francis, B.J., Gregory, I.N., and Porter, C. (2017a). Patterns of infant mortality in rural England and Wales, 1850–1910. *Economic History Review* 70(4): 1268–1290. doi:10.1111/ehr.12488.

Atkinson, P.D., Francis, B.J., Gregory, I.N., and Porter, C. (2017b). Spatial modelling of rural infant mortality and occupation in 19th-century Britain. *Demographic Research* 36(44): 1337–1360. doi:10.4054/DemRes.2017.36.44.

Bivand, R.S., Pebesma, E., and Gómez-Rubio, V. (2013). *Applied spatial data analysis with R*. New York: Springer. doi:10.1007/978-1-4614-7618-4.

Bocquet-Appel, J.P. and Jakobi, L. (1998). Evidence for a spatial diffusion of contraception at the onset of the fertility transition in Victorian Britain. *Population* 10(1): 181–204.

Chi, G. and Zhu, J. (2008). Spatial regression models for demographic analysis. *Population Research and Policy Review* 27(1): 17–42. doi:10.1007/s11113-007-9051-8.

Chiang, C.L. (1984). *The life table and its applications*. Malabar: Robert E. Krieger.

Coale, A., Demeny, P., and Vaughan, B. (1983). *Regional model life tables and stable populations*. New York: Academic Press.

Coale, A.J. and Watkins, S.C. (1986). *The decline of fertility in Europe*. Princeton: Princeton University Press.

Connor, D.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): 625–646. doi:10.1080/24694452.2016.1261682.

Duncan, D.T., Aldstadt, J., Whalen, J., White, K., Castro, M.C., and Williams, D.R. (2012). Space, race, and poverty: Spatial inequalities in walkable neighborhood amenities? *Demographic Research* 26(17): 409–448. doi:10.4054/DemRes.2012.26.17.

Edvinsson, S., Brändström, A., and Rogers, J. (2001). Regional variations in infant mortality in Sweden during the first half of the 19th century. In: Tedebrand, L.-G. and Sköld, P. (eds.). *Nordic demography in history and present-day society*. Umeå: Umeå University: 145–164.

Edvinsson, S., Garðarsdóttir, Ó., and Thorvaldsen, G. (2008). Infant mortality in the Nordic countries, 1780–1930. *Continuity and Change* 23(3): 457–485. doi:10.1017/S0268416008006917.

Garrett, E. and Reid, A. (1994). Satanic mills, pleasant lands: Spatial variation in women's work, fertility and infant mortality as viewed from the 1911 census. *Historical Research* 67(163): 157–177. doi:10.1111/j.1468-2281.1994.tb01822.x.

Garrett, E. and Reid, A. (1995). Thinking of England and taking care: Family building strategies and infant mortality in England and Wales, 1891–1911. *International Journal of Population Geography* 1(1): 69–102. doi:10.1002/ijpg.6060010106.

Garrett, E., Reid, A., Schürer, K., and Szreter, S. (2001). *Changing family size in England and Wales: Place, class and demography, 1891–1911*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511495816.

Goldstein, J.R. and Klüsener, S. (2014). Spatial analysis of the causes of fertility decline in Prussia. *Population and Development Review* 40(3): 497–525. doi:10.1111/j.1728-4457.2014.00695.x.

González-Bailón, S. and Murphy, T.E. (2013). The effects of social interactions on fertility decline in nineteenth-century France: An agent-based simulation experiment. *Population Studies* 67(2): 135–155. doi:10.1080/00324728.2013.774435.

Gregory, I. (2008). Different places, different stories: Infant mortality decline in England and Wales, 1851–1911. *Annals of the Association of American Geographers* 98(4): 773–794. doi:10.1080/00045600802224406.

Guillot, M., Gerland, P., Pelletier, F., and Saabneh, A. (2012). Child mortality estimation: A global overview of infant and child mortality age patterns in light of new empirical data. *PLoS Med* 9(8): 1–15. doi:10.1371/journal.pmed.1001299.

Gutmann, M.P., Deane, G.D., Merchant, E.R., and Sylvester, K. (2011). *Navigating time and space in population studies*. Dordrecht: Springer.

Haines, M.R. and Hacker, J.D. (2011). Spatial aspects of the American fertility transition in the nineteenth century. In: Gutmann, M.P., Deane, G.D., Merchant, E.R., and Sylvester, K. (eds.). *Navigating time and space in population studies*. Dordrecht: Springer: 37–63. doi:10.1007/978-94-007-0068-0_3.

Higgs, E., Jones, C., Schürer, K., and Wilkinson, A. (2013). *The Integrated Census Microdata (I-CeM) guide*. Colchester: UK Data Archive.

Hinde, A. and Fairhurst, V. (2015). Why was infant mortality so high in eastern England in the mid nineteenth century? *Local Population Studies* 94(1): 48–66.

Hobcraft, J. (1993). Women's education, child welfare and child survival: A review of the evidence. *Health Transition Review* 3(2): 159–175.

Howell, F.M., Porter, J.R., and Matthews, S.A. (2016). *Recapturing space: New middle-range theory in spatial demography*. Cham: Springer. doi:10.1007/978-3-319-22810-5.

Klüsener, S., Devos, I., Ekamper, P., Gregory, I., Gruber, S., Martí-Henneberg, J., van Poppel, F., Silveira, L., and Solli, A. (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(68): 1849–1864. doi:10.4054/DemRes.2014.30.68.

Klüsener, S., Dribe, M., and Scalone, F. (2016). Spatial and social distance in the fertility transition: Sweden 1880–1900. Rostock: Max Planck Institute for Demographic Research (MPIDR working paper; WP-2016-009).

Lee, C.H. (1991). Regional patterns in infant mortality in Britain, 1861–1971: Patterns and hypotheses. *Population Studies* 45(1): 55–65. doi:10.1080/0032472031000145086.

Matthews, S. and Parker, D.M. (2013). Progress in spatial demography. *Demographic Research* 28(10): 271–312. doi:10.4054/DemRes.2013.28.10.

Montgomery, M.R. and Casterline, J.B. (1993). The diffusion of fertility control in Taiwan: Evidence from pooled cross-section time-series models. *Population Studies* 47(3): 457–479. doi:10.1080/0032472031000147246.

Mooney, G. (1994a). Did London pass the sanitary test? Seasonal infant mortality in London, 1870–1914. *Journal of Historical Geography* 20(2): 158–174. doi:10.1006/jhge.1994.1013.

Mooney, G. (1994b). Stillbirths and the measurement of infant mortality rates c. 1890–1930. *Local Population Studies* 53: 42–52.

Mooney, G., Luckin, B., and Tanner, A. (1999). Patient pathways: Solving the problem of institutional mortality in London during the later nineteenth century. *Social History of Medicine* 12(2): 227–269. doi:10.1093/shm/12.2.227.

Ramiro-Fariñas, D. and Sanz-Gimeno, A. (2000). Structural changes in childhood mortality in Spain, 1860–1990. *International Journal of Population Geography* 6(1): 61–82. doi:10.1002/(SICI)1099-1220(200001/02)6:1<61::AID-IJPG168>3.0.CO;2-C.

Reid, A. (1997). Locality or class? Spatial and social differentials in infant and child mortality in England and Wales, 1895–1911. In: Corsini, C.A. and Viazzo, P.P. (eds.). *The decline of infant and child mortality: The European experience, 1750–1990*. The Hague: Martinus Nijhoff: 129–154.

Reid, A. (2002). Infant feeding and post neonatal mortality in Derbyshire, England, in the early twentieth century. *Population Studies* 56(2): 151–166. doi:10.1080/00324720215926.

Reid, A. (2005). The influences on the health and mortality of illegitimate children in Derbyshire, 1917–1922. In: Levene, A., Nutt, T., and Williams, S. (eds.). *Illegitimacy in Britain, 1700–1920*. Basingstoke: Palgrave Macmillan: 168–189.

Reid, A., Garrett, E., and Szreter, S. (2016). Residential mobility and child mortality in early twentieth century Belfast. In: Ramiro-Fariñas, D. and Oris, M. (eds.). *New approaches to death in cities during the health transition*. Dordrecht: Springer: 55–76. doi:10.1007/978-3-319-43002-7_4.

Reves, R. (1985). Declining fertility in England and Wales as a major cause of the twentieth century decline in mortality: The role of changing family size and age structure in infectious disease mortality in infancy. *American Journal of Epidemiology* 122(1): 112–126. doi:10.1093/oxfordjournals.aje.a114070.

Shmueli, G. (2010). To explain or to predict? *Statistical Science* 25(3): 289–310. doi:10.1214/10-STS330.

Sneddon, S. (2006). A double penalty? Infant mortality in the Lincolnshire Fens, 1870–1900. In: Garrett, E., Galley, C., Shelton, N., and Woods, R. (eds.). *Infant mortality: A continuing social problem*. Aldershot: Ashgate: 79–97.

Thornton, P. and Olson, S. (2011). Mortality in late nineteenth-century Montreal: Geographic pathways of contagion. *Population Studies* 65(2): 157–181. doi:10.1080/00324728.2011.571385.

Thorvaldsen, G. (2002). Rural infant mortality in nineteenth century Norway. *Hygiea Internationalis* 3(1): 75–113. doi:10.3384/hygiea.1403-8668.023175.

van den Boomen, N. (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. doi:10.1080/1081602X.2015.1022199.

Vitali, A. and Billari, F.C. (2017). Changing determinants of low fertility and diffusion: A spatial analysis for Italy. *Population, Space and Place* 23(2): e1998. doi:10.1002/psp.1998.

Voss, P.R. (2007). Demography as a spatial social science. *Population Research and Policy Review* 26(5–6): 457–476. doi:10.1007/s11113-007-9047-4.

Williams, N. (1992). Death in its season: Class, environment and the mortality of infants in nineteenth-century Sheffield. *Social History of Medicine* 5(1): 71–94. doi:10.1093/shm/5.1.71.

Williams, N. and Galley, C. (1995). Urban–rural differentials in infant mortality in Victorian England. *Population Studies* 49(3): 401–420. doi:10.1080/0032472031000148746.

Williams, N. and Mooney, G. (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. doi:10.1017/S0268416000002265.

Woods, R. (1993). On the historical relationship between infant and adult mortality. *Population Studies* 47(2): 195–219. doi:10.1080/0032472031000146976.

Woods, R. (1997). Infant mortality in Britain: A survey of current knowledge on historical trends and variations. In: Bideau, A., Desjardins, B., and Brignoli, H. (eds.). *Infant and child mortality in the past*. Oxford: Clarendon: 74–88.

Woods, R. (2000). *The demography of Victorian England and Wales*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511496127.

Woods, R. and Shelton, N. (1997). *An atlas of Victorian mortality*. Liverpool: Liverpool University Press.

Woods, R., Watterson, P., and Woodward, J. (1988). The causes of rapid infant mortality decline in England and Wales, 1861–1921: Part 1. *Population Studies* 42(3): 343–366. doi:10.1080/0032472031000143516.

Woods, R., Watterson, P., and Woodward, J. (1989). The causes of rapid infant mortality decline in England and Wales, 1861–1921: Part 2. *Population Studies* 43(1): 113–132. doi:10.1080/0032472031000143876.

Woods, R., Williams, N., and Galley, C. (1997). Differential mortality patterns among infants and other young children: The experience of England and Wales in the nineteenth century. In: Corsini, C.A. and Viazzo, P.P. (eds.). *The decline of infant and child mortality: The European experience, 1750–1990*. The Hague: Martinus Nijhoff: 57–72.

Appendix

Data sources

Day, J.D. (2016). Registration sub-district boundaries for England and Wales 1851–1911. This dataset was created by the 'Atlas of Victorian Fertility Decline' project (PI: A.M. Reid) with funding from the ESRC (ES/L015463/1) [GIS dataset].

The Day dataset has been created using Satchell, A.E.M., Kitson, P.M.K., Newton, G.H., Shaw-Taylor, L., and Wrigley, E.A. (2016), *1851 England and Wales census parishes, townships and places*. The Satchell et al. dataset is an enhanced version of Burton, N., Westwood, J., and Carter, P. (2004), *GIS of the ancient parishes of England and Wales, 1500–1850*, Colchester: UK Data Archive, SN 4828, which is a GIS version of Kain, R.J.P., and Oliver, R.R. (2001), *Historic parishes of England and Wales: An electronic map of boundaries before 1850 with a gazetteer and metadata*, Colchester: UK Data Archive, SN 4348.

Schürer, K. and Higgs, E. (2014). *Integrated Census Microdata (I-CeM), 1851–1911* [data collection]. Colchester: UK Data Archive, SN 7481. doi:10.5255/UKDA-SN-7481-1.

Schürer, K., Higgs, E., Reid, A.M., and Garrett, E.M. (2016). *Integrated Census Microdata V.2 (I-CeM.2), 1851–1911* [data collection].[21]

[21] The Integrated Census Microdata V.2 (I-CeM.2) dataset is an enhanced version of the I-CeM dataset (Schürer and Higgs 2014). The original I-CeM project produced a standardised, integrated dataset of most of the censuses of Great Britain for the period 1851 to 1911 for use by academic researchers. Version 2 contains the same data, with further enhancements and improvements. Please note that values of some coded variables may differ to those in the original version. The authors would like to thank the following individuals for their input into the product: Joe Day, Hannaliis Jaadla, Christine Jones, and Amanda Wilkinson.

Table A-1: The data sources for variables used in the regression models for registration districts (RDs) and registration subdistricts (RSDs) in England and Wales, 1881–1911

Variables	Data source
Registration districts (RD)	
Early childhood death rate (dependent variable)	RG Decennial Reports
IMR	RG Decennial Reports
Measles (ratio of measles deaths to deaths under age 1)	RG Decennial Reports
Illegitimacy ratio	RG Annual Reports
Social classes per 100 inhabitants (15+)	
Upper and middle class	I-CeM: individual-level census data
Unskilled workers	I-CeM: individual-level census data
Textile workers	I-CeM: individual-level census data
Miners	I-CeM: individual-level census data
Agricultural labourers	I-CeM: individual-level census data
Education employees per 100 children (age 7–14)	I-CeM: individual-level census data
Health employees per 100 inhabitants	I-CeM: individual-level census data
Population density (people per acre)	Census published reports
Registration subdistricts (RSD)	
IMR	RG Quarterly Reports
Measles (ratio of measles deaths to deaths under age 1)	RG Quarterly Reports
Illegitimacy ratio	RG Annual Reports
Social classes per 100 inhabitants (15+)	
Upper and middle class	I-CeM: individual-level census data
Unskilled workers	I-CeM: individual-level census data
Textile workers	I-CeM: individual-level census data
Miners	I-CeM: individual-level census data
Agricultural labourers	I-CeM: individual-level census data
Education employees per 100 children (age 7–14)	I-CeM: individual-level census data
Health employees per 100 inhabitants	I-CeM: individual-level census data
Population density (people per acre)	Census published reports

Table A-2: Sensitivity analysis: model estimates for log-transformed early childhood age-specific death rates ($_4M_1$), registration districts in England and Wales, 1881–1911

	1881		1891		1901		1911	
Variables	**OLS**	**Spatial error**	**OLS**	**Spatial error**	**OLS**	**Spatial error**	**OLS**	**Spatial error**
(Intercept)	2.916 ***	2.787 ***	2.424 ***	2.44 ***	2.024 ***	2.023 ***	1.237 ***	1.330 ***
IMR (previous decade)	0.004 ***	0.003 ***	0.005 ***	0.005 ***	0.007 ***	0.007 ***	0.009 ***	0.008 ***
Measles (previous decade)	1.594 ***	0.771 **	0.360 ***	0.271 ***	0.963 ***	0.9 ***	1.129 ***	1.049 ***
Illegitimacy ratio	–0.001 *	0.000	–0.001 **	–0.001 *	–0.001	–0.001 *	0.001 *	0.000
Social classes per 100 inhabitants (15+)								
Upper and middle class	0.003	0.009	0.001	0.001	0.004	0.005	0.011	0.008
Unskilled workers	–0.006 *	0.003	0.006 *	0.006 **	0.004	0.005	0.013 **	0.013 **
Textile workers	0.002 *	0.005 *	0.003 ***	0.003 ***	0.004 **	0.004 **	0.005 ***	0.006 **
Miners	0.008 ***	0.008 ***	0.005 ***	0.005 ***	0.009 ***	0.008 ***	0.007 ***	0.007 ***
Agricultural labourers	–0.014 ***	–0.010 ***	–0.009 ***	–0.009 ***	–0.006 **	–0.006 **	–0.008 ***	–0.006 *
Education employees per 100 children	–0.040 ***	–0.022 *	–0.028 ***	–0.028 ***	–0.016 *	–0.015 *	–0.025 *	–0.025 **
Health employees per 100 inhabitants	–0.300 ***	–0.199 **	–0.058 *	–0.053	–0.054	–0.045	0.055	0.168 ***
Log(population density)	0.086 ***	0.096 ***	0.066 ***	0.068 ***	0.064 ***	0.063 ***	0.014	0.023 *
London	0.048	–0.004	0.127 ***	0.123 ***	0.095 *	0.092 *	0.079	0.103
Manchester	0.062	0.007	0.108	0.085	0.144	0.112	0.025	0.015
Liverpool	0.302	0.163	0.299 **	0.277 **	0.167	0.155	0.297	0.156
Log-likelihood	311.179	360.445	584.915	588.898	314.171	317.207	248.070	272.551
Spatial dependence								
Moran's *I* dependent variable	0.682		0.696		0.678		0.681	
Moran's *I* residuals	0.237 ***	–0.027	0.076 **	–0.007	0.063 **	–0.002	0.166 ***	–0.020

Note: *p < 0.05; **p < 0.01; ***p < 0.001.