Supplementary Material
Modelling Socio-Economic Mortality at Neighbourhood Level

Jie Wen
Lloyds Banking Group

Andrew J.G. Cairns
The Maxwell Institute for Mathematical Sciences, and Department of Actuarial Mathematics and Statistics, School of Mathematical and Computer Sciences, Heriot-Watt University

Torsten Kleinow*
Research Centre for Longevity Risk
Faculty of Economics and Business, University of Amsterdam

S.1 Data - Details, Sources and Standardisation

S.1.1 Mortality Data

For the exposure data, E_{itx}, we used mid year population estimates by single LSOA and single year of age, which are available for the years 2001 to 2019 at:

- Office for National Statistics, 2020, *Lower layer Super Output Area population estimates (supporting information)*, [data collection], Accessed 11 September 2020. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/lowersuperoutputareamidyearpopulationestimates

The LSOA-specific number of deaths at single years of age during the years 2001 to 2018 are also available on the ONS website:

Deaths registered during 2001 to 2016:

- Office for National Statistics, 2017, *Deaths by lower super output area, age and sex: England and Wales, 2001 to 2016*, [data collection], Accessed 10 December 2017. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/adhocs/007807deathsbylowersuperoutputareaageandsexenglandandwales2001to2016

Deaths registered in 2017:

- Office for National Statistics, 2018, *Number of deaths registered in each Lower Super Output Area by sex and age, deaths registered in 2017*, [data collection], Accessed 20 November 2018. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/adhocs/009235numberofdeathsregisteredineachlowersuperoutputareabysexandagedeathsregisteredin2017

*corresponding author, t.kleinow@uva.nl

Deaths registered in 2018:

- Office for National Statistics, 2019, *Deaths by Lower level Super Output Area (LSOA), England and Wales, 2018 registrations*, [data collection], Accessed 02 December 2019. Available from: https://www.ons.gov.uk/peoplepopulationand community/birthsdeathsandmarriages/deaths/adhocs/10829deathsbylowerlevel superoutputarealsoaenglandandwales2018registrations

S.1.2 Individual Data sets for Predictive Variables

Old age income deprivation x_1: income deprivation affecting older people index (IDAOPI), see https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015, *File 3: supplementary indices - income deprivation affecting children index and income deprivation affecting older people index*

Employment deprivation x_2: subdomain "employment deprivation" of the IMD, see https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015, *File 2: domains of deprivation*

Education deprivation x_3: subdomain "Education, Skills and Training Deprivation" of the IMD, see https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015, *File 2: domains of deprivation*

Crime rate x_4: subdomain "crime" of the IMD, see https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015, *File 2: domains of deprivation*

Average number of bedrooms x_5: The average number of bedrooms in a household's accommodation. DOI: http://dx.doi.org/10.5257/census/aggregate-2011-2

Proportion of population born in the UK x_6: The proportion of usual residents that were born within the UK in an LSOA, calculated by dividing the number of UK-born residents by the total number of residents.
See https://www.nomisweb.co.uk/census/2011/ks204ew

Deprivation in housing/living environment x_7: This is the index for the"Wider Barriers Sub-domain" of the IMD domain "Barriers to Housing and Services", see https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015, *File 4: subdomains of deprivation*

Proportion in management position x_8: The proportion of usual residents that are in a managerial role (including both lower and higher management) at work, using the NS-SEC classification (split by age groups is also available). Unemployed and full-time students are included. DOI: http://dx.doi.org/10.5257/census/aggregate-2011-2

Proportion working more than 49h per week x_9: This is the proportion of full time workers (males and females) who worked more than 49 hours in the week before the 2011 census, out of all usual residents aged 16 to 74 in employment in an LSOA in the same week, see, https://statistics.ukdataservice.ac.uk/dataset/age-hours-worked-2011

Urban-Rural classification x_{10}: This is the level of urbanization following the 4-class rural/urban classification in the 2011 Census. We have added a fifth class for LSOAs in London. For LSOA-level data and guidance on the classification, see https://www.ons.gov.uk/methodology/geography/geographicalproducts /ruralurbanclassifications/2011ruralurbanclassification

Proportion of those aged 60+ in care homes with nursing x_{11}: This is the proportion of residents aged 60 and above that live in a care home with nursing services (0 for LSOAs without care homes). The data have been provided to the authors by the ONS and, to the best of our knowledge, are not publicly available.

Proportion of those aged 60+ in care homes without nursing x_{12}: This is the proportion of residents aged 60 and above that live in a care home without nursing services (0 for LSOAs without care homes). The data have been provided to the authors by the ONS and, to the best of our knowledge, are not publicly available.

S.1.3 Standardised Data

In order to prevent the modelling outcome from being distorted by the different scales of different predictive variables, we standardise $x_1, \ldots, x_9$ in Table 1 in the main paper to be consistent with a standard normal distribution.

Denoting the original variable by A_i we define x_i as follows:

$$\begin{aligned} u_i &= \frac{rank(A_i)}{N+1} \quad \text{uniformized variables} \\ x_i &= \phi^{-1}(u_i) \quad \text{normalized to N(0,1)} \end{aligned} \tag{1}$$

where ϕ is the distribution function of the standard normal distribution and N is the number of LSOAs.

S.2 Correlation of Covariates by Urban-Rural Class

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{11}	x_{12}
x_1	1	0.8	0.8	0.58	-0.61	-0.22	0.72	-0.79	-0.63	0.04	0.05
x_2	0.8	1	0.79	0.56	-0.56	0.03	0.52	-0.8	-0.65	0.01	0.02
x_3	0.8	0.79	1	0.51	-0.56	0.02	0.57	-0.84	-0.66	0.04	0.03
x_4	0.58	0.56	0.51	1	-0.45	-0.25	0.54	-0.55	-0.39	0.03	0.05
x_5	-0.61	-0.56	-0.56	-0.45	1	0.01	-0.54	0.49	0.4	-0.01	-0.02
x_6	-0.22	0.03	0.02	-0.25	0.01	1	-0.49	0.01	0	-0.01	-0.04
x_7	0.72	0.52	0.57	0.54	-0.54	-0.49	1	-0.61	-0.43	0.02	0.03
x_8	-0.79	-0.8	-0.84	-0.55	0.49	0.01	-0.61	1	0.72	0.01	0.02
x_9	-0.63	-0.65	-0.66	-0.39	0.4	0	-0.43	0.72	1	0.03	0.05
x_10	0.04	0.01	0.04	0.03	-0.01	-0.01	0.02	0.01	0.03	1	0.06
x_11	0.05	0.02	0.03	0.05	-0.02	-0.04	0.03	0.02	0.05	0.06	1

Table 1: Correlations for urban-rural class 1.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{11}	x_{12}
x_1	1	0.81	0.77	0.63	-0.72	-0.1	0.68	-0.77	-0.53	0.02	0.09
x_2	0.81	1	0.76	0.63	-0.64	0.14	0.54	-0.81	-0.53	-0.02	0.06
x_3	0.77	0.76	1	0.52	-0.57	0.15	0.5	-0.84	-0.6	0.01	0.04
x_4	0.63	0.63	0.52	1	-0.52	-0.16	0.57	-0.57	-0.36	-0.02	0.07
x_5	-0.72	-0.64	-0.57	-0.52	1	0.06	-0.62	0.54	0.41	0	-0.07
x_6	-0.1	0.14	0.15	-0.16	0.06	1	-0.37	-0.14	-0.02	-0.04	-0.01
x_7	0.68	0.54	0.5	0.57	-0.62	-0.37	1	-0.54	-0.31	0.01	0.06
x_8	-0.77	-0.81	-0.84	-0.57	0.54	-0.14	-0.54	1	0.64	0.05	0.01
x_9	-0.53	-0.53	-0.6	-0.36	0.41	-0.02	-0.31	0.64	1	0.06	0.06
x_10	0.02	-0.02	0.01	-0.02	0	-0.04	0.01	0.05	0.06	1	0.07
x_11	0.09	0.06	0.04	0.07	-0.07	-0.01	0.06	0.01	0.06	0.07	1

Table 2: Correlations for urban-rural class 2.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{11}	x_{12}
x_1	1	0.8	0.76	0.37	-0.75	0.26	0.55	-0.75	-0.41	0.05	0.13
x_2	0.8	1	0.72	0.44	-0.69	0.41	0.46	-0.82	-0.51	0	0.07
x_3	0.76	0.72	1	0.37	-0.61	0.41	0.39	-0.83	-0.52	0.03	0.08
x_4	0.37	0.44	0.37	1	-0.31	0.04	0.36	-0.39	-0.17	0.03	0.05
x_5	-0.75	-0.69	-0.61	-0.31	1	-0.21	-0.55	0.68	0.44	0.01	-0.07
x_6	0.26	0.41	0.41	0.04	-0.21	1	-0.11	-0.45	-0.4	-0.05	-0.02
x_7	0.55	0.46	0.39	0.36	-0.55	-0.11	1	-0.46	-0.1	0.04	0.07
x_8	-0.75	-0.82	-0.83	-0.39	0.68	-0.45	-0.46	1	0.52	0.01	-0.05
x_9	-0.41	-0.51	-0.52	-0.17	0.44	-0.4	-0.1	0.52	1	0.02	0.03
x_10	0.05	0	0.03	0.03	0.01	-0.05	0.04	0.01	0.02	1	0.05
x_11	0.13	0.07	0.08	0.05	-0.07	-0.02	0.07	-0.05	0.03	0.05	1

Table 3: Correlations for urban-rural class 3.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{11}	x_{12}
x_1	1	0.66	0.68	0.13	-0.63	0.27	0.32	-0.63	-0.33	0.04	0.1
x_2	0.66	1	0.6	0.19	-0.62	0.39	0.21	-0.67	-0.37	0	0.08
x_3	0.68	0.6	1	0.14	-0.6	0.45	0.11	-0.69	-0.37	0.06	0.09
x_4	0.13	0.19	0.14	1	-0.09	-0.11	0.17	-0.05	-0.14	0.02	0.02
x_5	-0.63	-0.62	-0.6	-0.09	1	-0.33	-0.33	0.64	0.46	0.01	-0.01
x_6	0.27	0.39	0.45	-0.11	-0.33	1	-0.19	-0.44	-0.17	-0.04	0
x_7	0.32	0.21	0.11	0.17	-0.33	-0.19	1	-0.24	-0.11	0	0
x_8	-0.63	-0.67	-0.69	-0.05	0.64	-0.44	-0.24	1	0.25	0.04	-0.01
x_9	-0.33	-0.37	-0.37	-0.14	0.46	-0.17	-0.11	0.25	1	-0.01	-0.04
x_{10}	0.04	0	0.06	0.02	0.01	-0.04	0	0.04	-0.01	1	0.06
x_{11}	0.1	0.08	0.09	0.02	-0.01	0	0	-0.01	-0.04	0.06	1

Table 4: Correlations for urban-rural class 4.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{11}	x_{12}
x_1	1	0.76	0.6	0.53	-0.64	-0.52	0.8	-0.52	-0.35	-0.01	0
x_2	0.76	1	0.71	0.5	-0.45	-0.26	0.66	-0.73	-0.56	0	0.01
x_3	0.6	0.71	1	0.36	-0.26	-0.04	0.44	-0.75	-0.65	0.01	0.01
x_4	0.53	0.5	0.36	1	-0.31	-0.3	0.49	-0.34	-0.24	-0.04	0.01
x_5	-0.64	-0.45	-0.26	-0.31	1	0.19	-0.6	0.11	0	0	0
x_6	-0.52	-0.26	-0.04	-0.3	0.19	1	-0.64	0.31	0.23	0	0.03
x_7	0.8	0.66	0.44	0.49	-0.6	-0.64	1	-0.54	-0.35	0	0
x_8	-0.52	-0.73	-0.75	-0.34	0.11	0.31	-0.54	1	0.83	0.02	0.03
x_9	-0.35	-0.56	-0.65	-0.24	0	0.23	-0.35	0.83	1	-0.01	0.01
x_10	-0.01	0	0.01	-0.04	0	0	0	0.02	-0.01	1	0.03
x_11	0	0.01	0.01	0.01	0	0.03	0	0.03	0.01	0.03	1

Table 5: Correlations for urban-rural class 5.

S.3 ASMR and ADSMR

S.3.1 ASMR using European Standard Population

For the age range relevant to our study, the European Standard Population (ESP) is given for five year age groups: 40 – 44, 45 – 49, etc. We therefore modify the general formula in the paper in equation (19)

$$\begin{aligned} \tilde{D}_{gta} &= D_{gta} + D_{gt,a+1} + ... + D_{gt,a+4} \quad \text{for } a = 40, 45, \ldots, 85 \\ \tilde{E}_{gta} &= E_{gta} + E_{gt,a+1} + ... + E_{gt,a+4} \quad \text{for } a = 40, 45, \ldots, 85 \\ \tilde{m}_{gta} &= \tilde{D}_{gta} / \tilde{E}_{gta} \end{aligned}$$

We then apply the formula in (19) to $\tilde{m}_{gta}$ with $\mathcal{X} \subseteq \{40, 45, 50, \ldots, 85\}$ and the standard exposures E^s_a referring to age group $[a, a+4]$.

S.3.2 ASMR by IMD and LIFE Deciles and Age Group

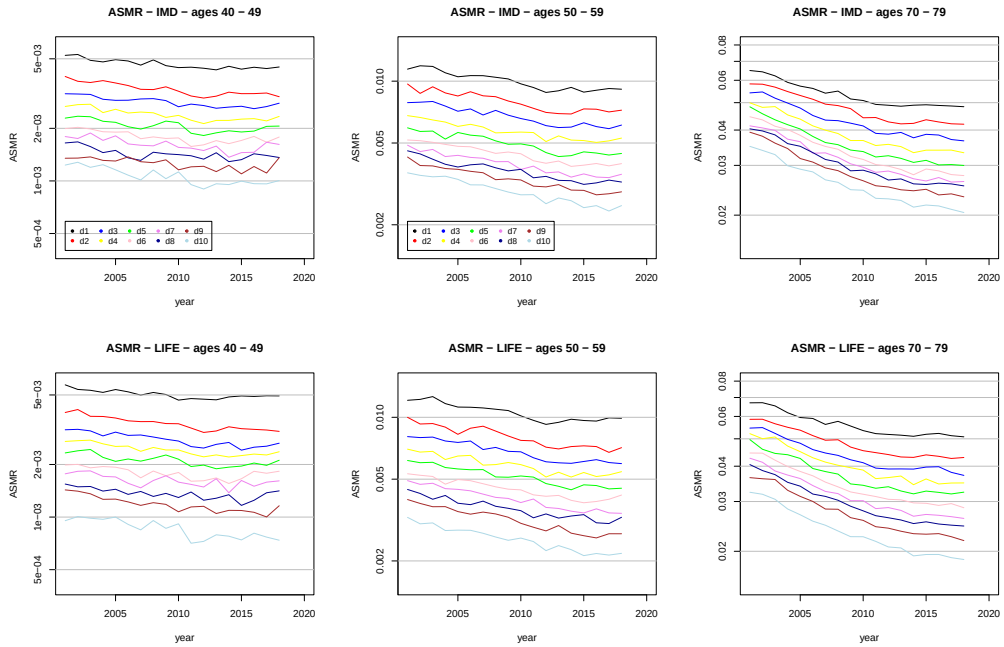

Figure 1: ASMR (log scale) by deprivation decile based on IMD scores (top) and LIFE scores (bottom). The LIFE index and the ASMRs have been calculated for age groups 40 – 49, 50 – 59 and 70 – 79.

S.3.3 ASMR by Region and Age Group

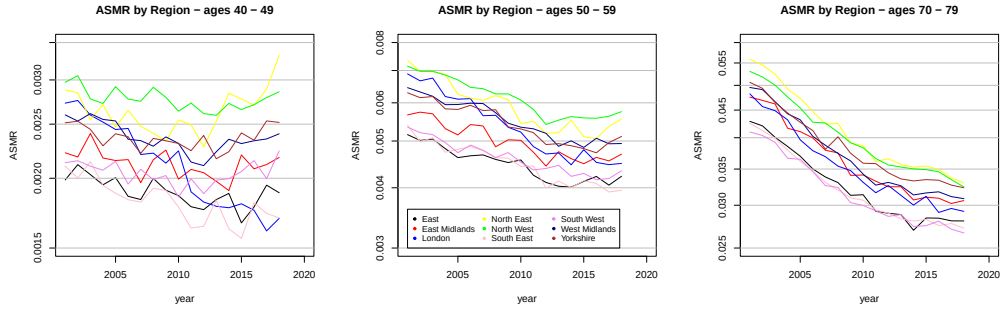

Figure 2: ASMRs by region for mortality data for ages 40 – 49, 50 – 59 and 70 – 79 (log scale).

S.3.4 ADSMR by Region and Age Group

In addition to the ADSMRs published in the main paper, we provide the ADSMRs for three further age groups. An interesting feature is that the ADSMRs for the youngest age group 40 – 49 are very flat except for those in London indicating very low mortality improvement rates outside London. We also find that the results based on IMD deciles and LIFE deciles are more similar for younger populations showing that the LIFE index is better able to explain mortality differences with socio-economic factors for older ages. This might be due to the lower number of deaths at young ages leading to more fluctuations in the observed relative risk at those ages.

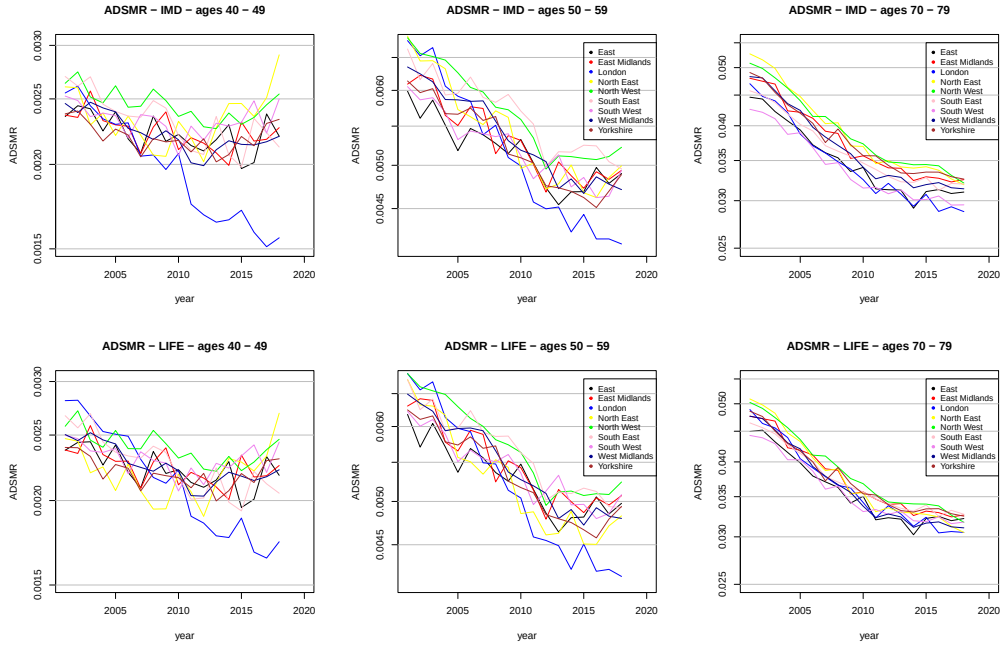

Figure 3: ADSMRs on the basis of IMD deciles (top) and LIFE deciles (bottom) where mortality data for different age groups has been used (log scale).