Chapter 2
Estimating Life Expectancy in Small Areas, with an Application to Recent Changes in Life Expectancy in US Counties

Peter Congdon

Abstract Analysis of small area mortality contrasts via life tables, and estimation of functions such as life expectancies, raises methodological issues regarding a suitable model for the mortality data. Methodological assumptions may be relevant to assessing whether there are changes in spatial clustering or in spatial inequalities in life expectancy. Virtually all analyses of US small area mortality use conventional life table analysis, which takes no account of similarities between mortality rates for adjacent areas or ages, and is subject to potential instability of mortality rates involved in deriving life tables. The alternative strategy used here involves a statistical model that "borrows strength" by using random effects to represent correlations between adjacent ages and areas. The smoothed mortality rates from the model are used to derive male and female life expectancies in US counties for three periods: 1995–1998, 1999–2002 and 2003–2006. Changes in inequality measures (e.g. the concentration index) show an increase in income related inequality in county expectancies, while local spatial correlation indices show an enhancement of low expectancy clusters in the South Eastern USA.

Keywords Life expectancy · Spatial inequality · Clustering · Borrowing strength · Random effects · Bayesian

2.1 Background

Analysis of small area mortality variations raises questions about suitable techniques for estimating life table functions such as life expectancies. Conventional fixed effects methods for life tables are problematic for small areas (with under 10,000 population, and especially under 5,000 population), potentially resulting in implausible or even infinite life expectancy estimates, even with data pooled over a number of years.

P. Congdon (✉)
Department of Geography and Centre for Statistics,
Queen Mary University of London, Mile End Rd, London E1 4NS, UK
e-mail: p.congdon@qmul.ac.uk

J. Anson, M. Luy (eds.), *Mortality in an International Perspective,*
European Studies of Population 18, DOI 10.1007/978-3-319-03029-6_2,

Conventional life table analysis is subject to variance instability in estimated age-area specific mortality rates, leading on to wide confidence intervals for summary measures such as life expectancy (Anselin et al. 2006; Toson and Baker 2003). This issue is relevant to US counties, some of which have small populations: around 300 have populations under 5,000, and 700 (just under a quarter of all US counties) have populations under 10,000. Methodological choices are relevant to assessing growing spatial inequalities in life expectancy, and changes in spatial clustering of lower than average life expectancy.

Recent research points to widening geographic inequalities in US mortality (e.g. Murray et al. 2006; Ezzati et al. 2008) but, with the exception of Kulkarni et al. (2011), uses conventional life table analysis and various amalgamations of areas to alleviate the instability inherent in such analysis. The implicit statistical model for conventional life tables applied to areas involves a heavily parameterized fixed effect approach (i.e. NX parameters for N areas and X age groups), that ignores high correlations between mortality rates for adjacent ages, and spatial clustering in mortality risk (McLaughlin et al. 2007). As to amalgamation strategies, Ezzati et al. (2008), Murray et al. (2006) and Kulkarni et al. (2011) amalgamate the original 3141 US counties to 2068, 2072 and 2357 areas respectively. The area amalgamation approach may affect inferences regarding differentials in expectancies: for example, smaller counties in the US tend to be in rural areas, so amalgamation may affect inferences regarding the urban-rural gradient in expectancy.

The present analysis accordingly adopts a "structured effects" statistical model to estimate mortality rates, and hence life expectancies, in US counties between 1995–1998, 1999–2002 and 2003–2006. The original areas are retained, and a borrowing strength approach adopted for age and area mortality parameters, via random effects methods that recognize correlations between adjacent ages and areas (e.g. Jonker et al. 2012; Chambers et al. 2009). Bayesian estimation is used, with parameter estimates based on iterative Monte Carlo Markov Chain (MCMC) techniques, implemented via the WINBUGS program (Lunn et al. 2009). The effective number of parameters involved is considerably fewer than under conventional fixed effects approaches (Spiegelhalter et al. 2002; Zhu et al. 2006). The model includes spatial random effects that account for spatial clustering in high mortality (and low expectancy) (McLaughlin et al. 2007). A recent paper by Kulkarni et al. (2011) recognizes benefits from random effects borrowing of strength but adopts an area random effect which is not spatially configured.

Once stabilized estimates of life expectancy are obtained from the statistical model, the extent of spatial inequality and clustering is considered, both cross-sectionally and through time. Before this analysis of trends, a sensitivity analysis compares the structured effects model against the conventional life table approach, and other "non-conventional" smoothing methods for life tables, albeit within a Bayesian framework. Specifically a spatial version of the Heligman-Pollard model (Heligman and Pollard 1980), involving county specific parameters is considered, and a spatial adaptation of the relational model (Brass 1974).

2.2 Methods: Structured Effects Mortality Model

Studies applying conventional life tables to analyse area mortality use moment estimates of central death rates to derive life table schedules such as death probabilities q_{cx}, (for areas $c = 1, \ldots, N$; ages $x = 1, \ldots, X$), and average years of life remaining at exact age x, E_{cx}. Although the underlying statistical model is typically not made explicit, it has important limiting features. Assuming an underlying binomial sampling model for deaths y_{cx} in relation to populations at risk, P_{cx}, namely

$$y_{cx} \sim Bin(P_{cx}, m_{cx}),$$

then conventional life table involves as many fixed effect parameters (central mortality rates) m_{cx} as there are data points, namely NX. Moment maximum likelihood estimates of these parameters under the conventional life table method are simply, $\hat{m}_{cx} = y_{cx}/P_{cx}$. For the US application with $N = 3139$ counties (in 2003–2006) and $X = 13$, there are 40807 parameters (i.e. mortality rate estimates) under the fixed effects model, providing a large model dimension.

This drawback is combined with potentially unstable estimates of rates under the conventional life table estimation approach, with rates for smaller areas showing most variability (Anselin et al. 2006; Riggan et al. 1991), leading on to wide confidence intervals for life expectancies in each area. A related feature is that no account is taken of similarities in mortality between areas or ages, so each parameter is estimated separately without reference to those for other ages or areas. The conventional life table estimation procedure has particular drawbacks including potential overestimation of life expectancy for small populations, and problems (infinite left expectancies) produced by zero death rates in the final age interval in a particular area, which can only be remedied by ad hoc devices (e.g. replacing the zero death rate by the national death rate for the final age interval) (Eayres and Williams 2004).

By contrast, the pooling strength approach views effects for particular ages or areas as drawn from a larger population of random effects, following a particular overarching density. Additionally these effects may be *structured* to reflect real world correlations between mortality at adjacent ages and in neighbouring areas. Thus in the model for age-county mortality rates in Eq. (2.1) below, the county parameters γ_c represent smoothly varying effects of spatially correlated unobserved risk factors. The underlying spatially smooth process represented by these effects straddles arbitrary county boundaries which are unrelated to the underlying continuous spatial process (Best 1999; Fotheringham et al. 2002).

Spatial dependence in the γ_c follows a conditional autoregressive form (Besag et al. 1991),

$$\gamma_c|\gamma_{[-c]} \sim N\left(\frac{1}{\sum_d w_{cd}} \sum_{d \neq c} w_{cd}\gamma_d, \frac{\delta_\gamma}{\sum_d w_{cd}}\right),$$

where w_{cd} represents spatial interaction between counties c and d, and δ_γ is a variance parameter. Suppose $w_{cd} = 1$ if counties c and d are contiguous, and $w_{cd} = 0$

otherwise, let m_c be the number of counties contiguous to county c, and let A_c denote the neighbourhood of county c (namely the set of counties adjacent to it). Then this spatial scheme (implemented as the car.normal density in WINBUGS) becomes

$$\gamma_c|\gamma_{[-c]} \sim N\left(\frac{1}{m_c}\sum_{d\in A_c}\gamma_d, \frac{\delta_\gamma}{m_c}\right).$$

Other spatial schemes may be used, such as neighbourhoods including second order as well as first order neighbours (see Appendix 1); however, the above scheme is much the most commonly used for human health applications. For age effects, life table death rates typically display very high correlation between rates for adjacent ages, and a scheme representing such dependence is provided by a local level or random walk model (Durbin and Koopman 2001)

$$\alpha_x \sim N(\alpha_{x-1}, \delta_\alpha), \qquad x > 1$$

with the initial age effect α_1 modelled separately.

However, heterogeneity among individuals within age-area units implies that age and area effects alone may not account for extra-binomial variation (Appendix 2). So the model also includes unstructured random effects u_{cx} specific to area-age combinations, with the full model being

$$logit(m_{cx}) = \lambda + \alpha_x + \gamma_c + u_{cx} \tag{2.1}$$

This is subsequently denoted as the "structured effects model", which is later compared with spatially adapted versions of the Heligman-Pollard and relational models. It is likely that heterogeneity is greater in some age bands than others, so it is assumed that $u_{cx} \sim N(0, \phi_x)$. In model (2.1), gamma priors are adopted for $1/\delta_\gamma$ and $1/\delta_a$, and for the age-specific precisions $1/\phi_x$, with index and shape parameters equal to 1. A $N(0, 100)$ prior is used for λ.

Let age group x have interval length n_x, and denote the average fraction of the interval survived as a_x. Then life table death probabilities by area and age may be estimated (Bell and Miller 2005) from the smoothed m_{cx} as

$$q_{cx} = n_x m_{cx}/(1 + n_x(1 - a_x)\, m_{cx}).$$

From these one may estimate the survivorship and years-lived functions l_{cx} and L_{cx}, and in turn the average number of years of life E_{cx} remaining at exact age x. The case study below considers expectancies at birth, E_{c0}, and expectancies at age 65, E_{c65}, but other functions can be obtained such as survivor probabilities over particular age spans, such as the county specific probability of surviving from age 45 to 75, ${}_{30}p_{c,45} = 1 - {}_{30}q_{c,45} = l_{c,75}/l_{c,45}$.

As part of the MCMC sampling output, one may obtain full posterior densities for the life table functions in each county. This is not possible using conventional life table methods, which provide variance estimates using large sample approximations. One may also assess interval hypotheses on life table functions, for example the probability that the life expectancy in a particular county, E_{c0}, is below the US wide

average, $\overline{E}_0$. This involves using indicator functions, such as $I(E_{c0} < \overline{E}_0)$, and totalling MCMC iterations where the condition in the indicator function is satisfied. Additional advantages are the greater precision (narrower confidence intervals) of this method compared to conventional methods, especially in areas with relatively small populations. Furthermore the borrowing strength method provides sensible estimates for small counties with small observed death totals. Examples are King County in Texas with 5 male deaths and a male population-years total of 655 in 2003–2006, and Arthur County (Nebraska) with 14 male deaths and a population-years total of 757. Conventional methods give implausible life expectancies under 40 for these counties. Bayesian methods avoid the *ad hoc* adjustments needed for conventional life tables when there are zero deaths in the last age interval, such as using national death rates instead.

It may be noted that in addition to small US counties, there are also many large counties in the data used, with large within-area samples of deaths and large populations at risk. The idea behind borrowing strength methods is to use information on mortality provided by all counties to provide stabilized estimates for small counties where within-area samples are relatively small. The large observation sample across areas, both in terms of over 3000 counties over 4 year periods, and around 10 million deaths (for males and females combined) in each 4 year period, will both assist in providing stabilized estimates for small areas and ensure that the data will dominate any prior assumptions. For example, borrowing of strength applied to the common age structure effects α_x will combine information about the mortality age gradient over all counties, so providing precise estimates. Similarly, the borrowing of strength random effects approach to estimate county-age interactions u_{cx} penalizes extreme parameter values that can occur under fixed effects methods.

2.3 Sensitivity Analysis: Comparison with Other Models for Smoothing Mortality and Conventional Life Tables

Other approaches have been used to smooth irregular mortality data. One widely used approach is to use parametric equations to smooth mortality rates, as in the eight parameter Heligman-Pollard model. Expressed in terms of odds of death rates, this model is

$$\frac{m_x}{1-m_x} = A^{(x+B)^C} + Dexp\left[-E\left\{\log\frac{x}{F}\right\}^2\right] + GH^x = R_{1x} + R_{2x} + R_{3x} = R_x.$$

The three components represent respectively child mortality, young adult mortality and mature-age mortality. The application here generalizes this parametric model to explain age-county mortality rates m_{cx}, and allows spatial variation in particular parameters. To this end, one may represent the level of child mortality as $A = \exp(\omega)$, and the level of young adult mortality as $\log(R_{2x}) = \eta_1 + \eta_2(\log x - \log F)^2$. Furthermore one may write the third component as

$$logit(R_{3x}) = \beta_1 + \beta_2 x,$$

since

$$R_{3x} = \frac{e^{\beta_1} e^{\beta_2 x}}{1 + e^{\beta_1} e^{\beta_2 x}} = \frac{GH^x}{1 + GH^x} (with\ G = e^{\alpha}, H = e^{\beta}).$$

Then spatial variation, following the same conditional autoregressive scheme as discussed above, involves the model

$$\frac{m_{cx}}{1 - m_{cx}} = A_c^{(x+B)^C} + D_c\, exp\left[-E\left\{\log \frac{x}{F}\right\}^2\right] + G_c H_c^x$$
$$= R_{1cx} + R_{2cx} + R_{3cx} = R_{cx},$$

and four sets of spatial effects, $\{s_{1c}, \ldots, s_{4c}\}$, which are centred to have average zero:

$A_c = exp\,(\omega + s_{1c})$ (varying child mortality level);
$D_c = exp\,(\eta_1 + s_{2c})$ (varying young adult mortality level);
$G_c = exp\,(\beta_1 + s_{3c})$ (varying mature adult mortality level);
$H_c = exp\,(\beta_2 + s_{4c})$ (varying mature adult mortality slope).

Then

$$logit(m_{cx}) = \log(R_{cx}) + u_{cx}, \tag{2.2}$$

where the $u_{cx}$are unstructured county-age random effects, as in the structured effects model of (2.1).

Another widely used approach to smoothing mortality data involves the use of standard age schedules in relational models (e.g. Himes et al. 1994), and relational models have also been applied to smoothing fertility and migration rates (e.g. De Beer 2011). Here logits of county-age model death rates for period t are related to US-wide logit death rates in period $t-1$, so avoiding double use of the same data. Thus, for a county mortality model in 2003–2006, the standard (m_{sx})is provided by US-wide death rates in 1999–2002. Additionally intercepts and slopes $\{\alpha_c, \beta_c\}$ in the relational model are taken to vary by county, and be spatially structured (according to the above conditional autoregressive scheme). Then one has

$$y_{cx} \sim Bin(P_{cx}, m_{cx}),$$
$$logit(m_{cx}) = \alpha_c + \beta_c\, logit(m_{sx}) + u_{cx}, \tag{2.3}$$

where the u_{cx} are county-age random effects, as discussed above, to account for extra-binomial variation.

The structured effects model (2.1) is compared to these alternative borrowing strength specifications (2.2) and (2.3), and to conventional life table methods, using male deaths and populations $\{y_{cx}, P_{cx}\}$ for counties $c = 1, \ldots, N$, and ages $x = 1, \ldots, X$ in 2003–2006. There are $N = 3139$ counties, excluding Clifton Forge (Virginia), and amalgamating two very small counties, namely Kalawao (Hawaii) with Maui, and Loving (Texas) with Winkler (Texas). There are $X = 13$ age groups

as provided at the CDC Wonder site (http://wonder.cdc.gov/), namely, under 1, 1–4, 5–9, 10–14, 15–19, 20–24, 25–34, 35–44, 45–54, 55–64, 65–74, 75–84 and 85 +.

The conventional life table method involves unrelated fixed effects and is operationalized by assuming Beta(1,1) (i.e. uniform) priors on the death rates m_{cx}. Additionally for the conventional life table approach, if there are zero deaths in the final age interval, the national death rate for over 85's is substituted as the estimate for m_{cX}, in order to avoid infinite life or implausibly large life expectancies.

A particular focus is on the extent of differences between the four methods in estimates of life expectancies at birth and of probabilities of surviving from age 45 to 75. Aspects of sensitivity considered are the similarity/dissimilarity (inter-method correlation) across the 3139 counties; the precision of the estimates, with more precise estimates preferred; and model fit, assessed using the log pseudo marginal likelihood, abbreviated as LPML (Christensen et al. 2011; Kim et al. 2012). Precision is based on the total of posterior variances of the parameter estimates over the counties, which has been used by some analysts as a measure of model complexity (Gelfand and Ghosh 1998, p. 5).

An important additional aspect is the extent of spatial correlation (global and local) in county life table functions. For example, let $z_c = E_{c0} - \overline{E}_0$ be deviations in posterior mean life expectancies at birth in county c from the US wide average. Then a global measure of spatial correlation is provided by the Moran index,

$$I = \frac{N}{S_0} \sum_c \sum_d w_{cd} z_c z_d \Big/ \sum_c z_c^2,$$

where $S_0 = \sum_c \sum_d w_{cd}$. The localised version of this index (Anselin 1995) is

$$I_c = \frac{z_c}{m_2} \sum_d w_{cd} z_d,$$

where $m_2 = \sum_c z_c^2 / N$. For considering spatial patterning, an additional perspective using conventional methods is provided by the standard mortality ratio (SMR): the ratio of deaths in a county to expected deaths in the county population if US-wide age-specific death rates prevailed.

Inferences are based on the second halves of two chain runs of 10,000 iterations from dispersed initial values, with convergence achieved before iteration 5,000 using Brooks-Gelman criteria (Brooks and Gelman 1998). Table 2.1 shows that the best fitting approach is the relational model, but that more precise estimates of life expectancies and of 45–75 survival probabilities are obtained under the structural effects model. In fact, there is a very high correlation (over 0.99) between life expectancies and survival probabilities under these two methods (see panel B in Table 2.1), such that in practical terms they are effectively interchangeable. A slightly worse fit is obtained for the spatial Heligman-Pollard model, with correlations of 0.94–0.95 between its life expectancy estimates and those of the relational and structural effects models.

Table 2.1 Sensitivity analysis according to method, male mortality, 2003–2006

(A) Precision, Fit and Distributional Aspects

	Structural Effects	Spatial-Relational	Spatial Heligman-Pollard	Conventional (Fixed Effects)
Precision of Estimates (total posterior variances over all counties)				
Life Expectancies, age 0	1649	1965	3139	9044
Probability of surviving from 45 to 75	1.066	1.121	1.553	2.721
Measures of Fit				
LPML	-119482	-118207	-121510	-140793
Distributional features: Expectancies at Birth				
Mean	74.28	74.19	74.44	72.34
1st percentile	68.31	67.51	68.63	57.87
5th percentile	69.99	69.43	70.31	65.95
95th percentile	78.20	78.32	78.46	77.76
99th percentile	79.69	79.76	81.09	79.22
Distributional features: survival probabilty $_{30}p_{45}$				
Mean	0.625	0.623	0.624	0.618
5th percentile	0.519	0.517	0.515	0.493
95th percentile	0.717	0.717	0.723	0.731

(B) Correlations between Life Table Functions

Life Expectancies, age 0	Structural Effects	Spatial-Relational	Spatial Heligman-Pollard
Spatial-Relational	0.994		
Spatial Heligman-Pollard	0.953	0.942	
Conventional (Fixed Effects)	0.541	0.558	0.433
Survival probabilty $_{30}p_{45}$	Structural Effects	Spatial-Relational	Spatial Heligman-Pollard
Spatial-Relational	0.997		
Spatial Heligman-Pollard	0.978	0.977	
Conventional (Fixed Effects)	0.936	0.939	0.943

(C) Spatial Sensitivity

Global Moran Spatial Correlation	E_0, Structural Effects	E_0, Spatial-Relational	E_0, Spatial Heligman-Pollard	E_0, Conventional (Fixed Effects)	SMR
	0.66	0.65	0.54	0.41	0.51

Correlation between Local Moran Statistics	E_0, Structural Effects	E_0, Spatial-Relational	E_0, Spatial Heligman-Pollard	E_0, Conventional (Fixed Effects)
E_0, Spatial-Relational	0.985			
E_0, Spatial Heligman-Pollard	0.886	0.853		
E_0, Conventional (Fixed Effects)	0.188	0.206	0.172	
SMR	0.935	0.931	0.885	0.168

Fig. 2.1 Estimated life expectancies, males 2003–2006, structural effects method

The conventional fixed effects approach has a markedly worse fit, and correlations only around 0.5 between its life expectancy estimates and those obtained under the three borrowing strength models, though its survival probability estimates are closer to those under other methods. It may be noted that in some counties with small death totals and populations at risk, the conventional method provides estimated life expectancies under 50 (in 16 counties), the lowest expectancies being for King County (Texas) and Arthur County (Nebraska).

Panel C of Table 2.1 considers global and local Moran spatial correlations for life expectancy estimates. It shows that higher (but similar) global Moran coefficients of around 0.65 are obtained under the structural effects and spatial-relational methods. The Moran coefficient for the SMR is slightly lower at around 0.51, though still indicates strong spatial clustering. There is also a high concordance in local Moran coefficients in terms of correlation between the local Moran coefficients over the 3139 counties. The correlation between the SMR-based local Moran coefficients and the local coefficients obtained using the structural effects model is around 0.94. Life expectancies estimated by the conventional life table method show a relatively low concordance with other methods in terms of localised spatial patterning.

Figure 2.1 shows estimated life expectancies under the structural effects method, and Fig. 2.2 shows local clustering coefficients. The latter have high values in regions with concentrations of adjacent high expectancies (e.g. in northern parts of the mid-West), and also in regions with concentrations of adjacent low expectancies (e.g. in Mississippi and Louisiana). Figure 2.3 shows local clustering coefficients based on standard mortality ratios, which show a similar pattern to Fig. 2.2. Table 2.2 includes results from the structural effects model applied to both males and females in 2003–2006, and shows the concentrations of low expectancies at birth in states of the mid-South and South-East USA.

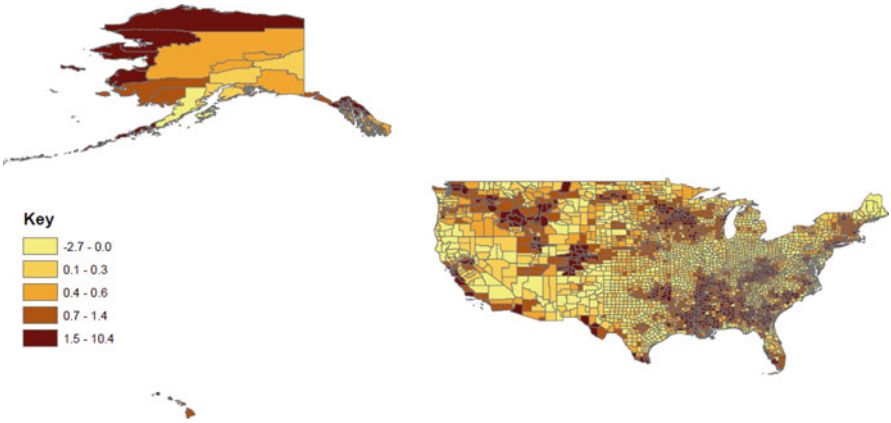

Fig. 2.2 Local Moran statistics, male expectancies 2003–2006, structural effects method

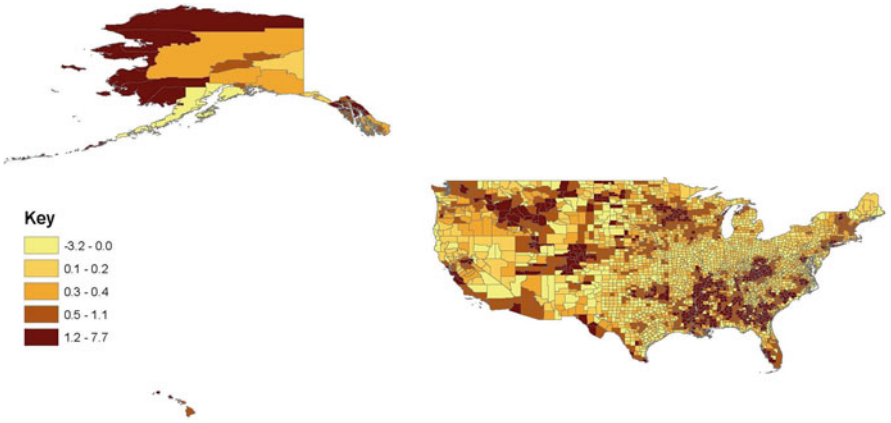

Fig. 2.3 Local Moran statistics, male SMRs 2003–2006

2.4 Assessing Trends in Inequality and Spatial Clustering

The previous analysis has shown the utility of the structured effects model and this model is now applied to assessing trends in spatial inequality. Using estimates of county life expectancy based on smoothed estimates of mortality rates, m_{cx}, one may assess changes in life expectancy gradients. We consider life expectancy at birth and at age 65. Three periods are considered: 1995–1998, 1999–2002, and 2003–2006, with data from the CDC Wonder site (http://wonder.cdc.gov/) – see Appendix 3 for details of areas.

One approach to assessing inequality is in terms of measures of area socioeconomic status, such as county average household income. Then inequality measures of social-group disparity are relevant (Harper and Lynch 2005), such as the slope

http://www.springer.com/978-3-319-03028-9

Mortality in an International Perspective
Anson, J.; Luy, M. (Eds.)
2014, XI, 359 p. 85 illus., 3 illus. in color., Hardcover
ISBN: 978-3-319-03028-9