

## RESEARCH ARTICLE OPEN ACCESS

# An Inclusive Economy Classification of British Local Authority Districts and Assessment of Its Association With Life Expectancy and Lifespan Variation

Nik Lomax<sup>1</sup> | Hugh P. Rice<sup>1</sup>  | Andreas Hoehn<sup>2</sup>  | Ceri Hughes<sup>3</sup>  | Alison Heppenstall<sup>4</sup> | Corinna Elsenbroich<sup>2</sup> | Petra Meier<sup>2</sup>

<sup>1</sup>School of Geography, University of Leeds, Leeds, UK | <sup>2</sup>MRC/CSO Social & Public Health Sciences Unit, University of Glasgow, Glasgow, UK | <sup>3</sup>Manchester Institute of Education, University of Manchester, Manchester, UK | <sup>4</sup>School of Social and Political Sciences, University of Glasgow, Glasgow, UK

**Correspondence:** Nik Lomax ([n.m.lomax@leeds.ac.uk](mailto:n.m.lomax@leeds.ac.uk))

**Received:** 30 July 2024 | **Revised:** 22 September 2025 | **Accepted:** 8 November 2025

**Funding:** This study was supported by the UK Prevention Research Partnership (MR/S037578/2), which is funded by the British Heart Foundation, Cancer Research UK, Chief Scientist Office of the Scottish Government Health and Social Care Directorates, Engineering and Physical Sciences Research Council, Economic and Social Research Council, Health and Social Care Research and Development Division (Welsh Government), Medical Research Council, National Institute for Health Research, Natural Environment Research Council, Public Health Agency (Northern Ireland) and The Health Foundation and Wellcome.

**Keywords:** geodemographic classification | Great Britain | inclusive economy | life expectancy | lifespan variation | local government area

## ABSTRACT

Using primarily open-source data, we curated 13 indicators to consistently capture the concept of the inclusive economy across Local Authority Districts in Great Britain. We used unsupervised *k*-means clustering to identify clusters of similar Local Authority Districts with respect to their economic inclusivity. We identified four distinct clusters of Local Authority Districts, those which are: (1) more economically inclusive; (2) less economically inclusive; (3) closest to the national average; and (4) a mix of extremes and mostly represented within the London metropolitan area. For Local Authority Districts of all clusters, we then explored the relationship between life expectancy and lifespan variation. Here, results clearly indicated a key role of economic inclusivity for average population health alongside population health inequality: the least inclusive cluster had the lowest life expectancy and the highest lifespan variation, while the most inclusive cluster had the highest life expectancy and the lowest lifespan variation. This points towards a double disadvantage among the economically less inclusive cluster.

## 1 | Introduction

The role of economic inclusion as a key factor contributing to health inequalities is increasingly recognised (OECD 2019; Macintyre et al. 2024). This more recent understanding of inclusivity marks a move away from traditional research which primarily examined the relationship between deprivation and health. However, the dimensions of economic inclusivity are neither universally defined nor have been collectively agreed upon. A rapid review undertaken by Shipton et al. (2021) identified 22 approaches to operationalise the idea of an

inclusive economy (IE) in 56 articles and reports, distilling the evidence down into four attributes: (1) equitable design of the economy; (2) equitable distribution of societal benefits; (3) equitable access for all individuals to the resources needed to participate in the economy; and (4) that the economy operates within planetary boundaries. There also exists a competing economic narrative around inclusive growth which is focused on growth as a mechanism for enabling broader participation (Hill O'Connor et al. 2023). However, the definition of what constitutes inclusive growth is contested in the same way that it is for the IE. In both the IE and the inclusive growth literature,

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Population, Space and Place* published by John Wiley & Sons Ltd.

there are further challenges associated with quantifying the domains, as data availability and alignment with the concepts have proven to present challenges. For example, the Joseph Rowntree Foundation has produced an Inclusive Growth Monitor to measure the relationship between prosperity and poverty in England, which comprises 18 indicators across six domains (Beatty et al. 2016).

In the IE space, a collection of 13 indicators for Great Britain has been defined by Lupton et al. (2023). This collection draws on definitions set out in eight existing indicator sets (including those defined by the Joseph Rowntree Foundation) with refinement and input from policy partners at local and national government and community panels. The Lupton et al. (2023) work synthesised key findings from the wider indicator sets to focus on two key areas which represent how inclusive an economy is: (1) economic outcomes, where inclusive means that participation in economic activity is high/equitable and that the benefits of economic activity are widely shared; and (2) wider outcomes and enablers which represent both the end point (these would result from an economy being more inclusive) as well as enablers that help the economy become more inclusive (education, access to services, connectivity and wider structures/systems that enable inclusion). The Lupton et al. (2023) definition of the IE is pragmatic, outlining 13 domains that are meaningful to decision makers, possible to estimate and consistent over time at subnational levels.

Utilising the 13 indicators identified in Lupton et al. (2023), our objective was to provide a classification of Local Authority Districts ('lower tier', LADs) in Great Britain with respect to economic inclusivity. LADs are local government administrative units (in England and Wales comparable to Local Administrative Unit [LAU] level 1 areas in other European countries such as Germany's 'Kreise'). Our reference is the 2021 definition of these geographical zones. Operationalising the concepts set out in Lupton et al. (2023) required us to harmonise data for different nations of the United Kingdom to create a consistent yet comprehensive classification for LADs. This represents a novel undertaking which was possible for England, Scotland and Wales, but data consistency issues within Northern Ireland meant it had to be excluded from our work. Our motivation for the geographical dimension of the work stems from the understanding that spatial inequalities in both health (Pearce et al. 2010) and economic outcomes (Lupton et al. 2016) exist, and that people's economic circumstances impact their health outcomes (Shipton et al. 2021).

Typically, classifications of geographical zones are created to better understand how (dis)similar areas are to each other so that results can inform policy or explain differences in outcomes. Such classifications have been developed for a wide range of applications, such as to inform understanding of internal migration patterns (Dennett and Stillwell 2011), political sentiment (Clark et al. 2017) and for a range of commercial applications (A. D. Singleton and Spielman 2014). The potential for such classifications to improve service planning in the health sector is assessed by Abbas et al. (2009), who highlighted that classifications offer a tool for better understanding health inequalities. Abbas et al. (2009) argued that area classifications allow for assessment of outcomes against multiple attributes, rather than the more traditional disaggregation by univariate

data across different domains in an average score or rank, for example, quintiles or deciles of socioeconomic deprivation. Some specific applications where a classification has been used to understand differences in health outcomes include the prevalence of severe disease and mortality arising from COVID-19 infection (Clark et al. 2021), variations in spatial prevalence of diabetes by a lifestyle classification in the United States (Grubestic et al. 2014) and the association between geodemographic indicators and prevalence of Limiting Long Term Illness (Moon et al. 2019).

While it can be hypothesised, there is little evidence that more inclusive economies are associated with more favourable population health outcomes. To explore the association between inclusivity and average population health alongside within-population health inequalities, we looked at the relationship between our classification and LAD-level period life expectancy (LE) and lifespan variation (LV), as well as how these two measures are related to each cluster. Our focus on LE and LV offers a holistic perspective on population health by capturing both averages and inequality, as is increasingly recommended by Inter-Governmental Organisations such as the United Nations (Issifou and Pewitt 2022). While LE is an average measure of mortality, LV is a measure of mortality inequality. Generally, the association between LE and LV tends to follow a strong linear relationship (van Raalte et al. 2018). Cross-country comparisons have demonstrated that higher levels of LE tend to come at lower levels of LV (e.g., Németh 2017; Aburto et al. 2020). At the same time, studies exploring patterns within populations have indicated that, as levels of deprivation increase, LE tends to decrease while LV tends to increase (Woods et al. 2005; van Raalte et al. 2018). Here, the relationship between LE and LV becomes particularly revealing when comparing population subgroups. For example, Sasson (2016) assessed the relationship with educational attainment in the United States, and reported that LE was lower while LV was higher for noncollege-educated populations compared with college-educated populations. In support of this double-disadvantage theory, Permanyer et al. (2018) reported similar results for Spain, indicating a strong relationship between LE and LV across population subgroups.

Despite this overarching pattern, the exact relationship between LE and LV is often more complex. For example, different areas might share the same level of LE while levels of LV might differ significantly. In other words, similar levels of average population health can be achieved at different levels of population health inequality. This has been shown for Great Britain, with Scotland achieving similar levels of LE as England and Wales, but at a much later point in time and at much higher levels of LV (Seaman et al. 2016), a phenomenon caused primarily by relatively high levels of premature mortality in Scotland. A more detailed assessment of results by deprivation quintile in Scotland by Seaman, Riffe, Leyland, et al. (2019) revealed that area-level differences in LE widened between 1981 and 2011. According to Seaman, Riffe, Leyland, et al. (2019), this trend was driven by the least deprived areas experiencing a decrease in LV when the most deprived areas saw stagnation or a slight increase in LV. While the relationship between levels of deprivation and LE/LV has been established, there is increasing recognition of a relationship between economic inclusion and

health inequalities. In their post-COVID-19 pandemic recovery strategy, Public Health England (2021) set out the need for place-based action on inclusive (and sustainable) economies as a mechanism to reduce health inequalities, with a particular focus on action targeted towards the most disadvantaged areas. Further highlighting the importance of economic inclusivity, Hoehn et al. (2024) have reported that indicators of economic inclusion explained over half of the variation in quality-adjusted life expectancy (QALE) across LADs in Great Britain.

In this paper, we have addressed the following two research questions:

1. Can we establish a typology of LADs for all of Great Britain with respect to economic inclusivity?
2. How do LE and LV differ by level of economic inclusivity across LADs in Great Britain?

For this purpose, we created a classification of LADs within Great Britain that captured how inclusive the economy is when compared to other LADs. We then assessed the relationship between LE and LV within each of the resulting clusters of

LADs. By doing so, we took a dedicated, holistic, multi-domain view of the IE, with the aim of contributing to the growing debate around the relationship between economic inclusion and population health.

## 2 | Materials and Methods

### 2.1 | IE Indicators

The IE indicators used to create our classification are outlined in Table 1. Each of these variables has been constructed from open-source data, where this was possible, to allow for reproducibility and updatability as new data are released. Data are compiled for the period 2017 to 2021, and we computed the average value for that period as input to the classification. The rationale for each choice of indicator is described in Lupton et al. (2023). Each indicator can be attributed to one of the following categories: (1) those which represent economic outcomes; and (2) those which represent wider outcomes and enablers. We developed each proposed indicator through a multi-stage review process. In a first step, this process involved

**TABLE 1** | Overview of the selected 13 inclusive economy indicators.

| ID                              | Indicator                                    | Explanation                                                                                                                                                                     |
|---------------------------------|----------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Economic outcomes (A)           |                                              |                                                                                                                                                                                 |
| 1A                              | Participation in paid employment             | Percentage of working-age population (age 16–64) who are employed.                                                                                                              |
| 2A                              | Involuntary exclusion from the labour market | Percentage of working-age population (age 16–64) who are economically inactive due to ill health or disability.                                                                 |
| 3A                              | Wealth inequality                            | Ratio of maximum to minimum house sale prices within each Local Authority District.                                                                                             |
| 4A                              | Earnings inequality                          | Ratio of weekly earnings (residents in full-time work) between 80th and 20th percentiles.                                                                                       |
| 5A                              | Poverty                                      | Percentage of children living in low-income households (after adjustment for housing costs).                                                                                    |
| 6A                              | Decent pay                                   | Percentage of employee jobs that are not paid at or above the National Living Wage.                                                                                             |
| 7A                              | Job security                                 | Percentage of individuals in permanent employment out of all employed individuals aged 16–64.                                                                                   |
| Wider outcomes and enablers (B) |                                              |                                                                                                                                                                                 |
| 1B                              | Skills and qualifications                    | Percentage of adults aged 16–49 with a level 2 or higher National Vocational Qualification.                                                                                     |
| 2B                              | Digital connectivity/inclusion               | Percentage of individuals who are classified as e-withdrawn, passive and uncommitted internet users, or settled offline communities, based on the Internet User Classification. |
| 3B                              | Physical connectivity                        | Proportion of LSOAs/DZs within the Local Authority District that are among the 50% most accessible LSOAs/DZs for each devolved nation (56%/44% split in England).               |
| 4B                              | Housing affordability                        | Ratio of median house prices to median gross annual earnings (for residents) within each Local Authority District.                                                              |
| 5B                              | Cost of living                               | Percentage of individuals reporting that they are worried about getting the food they need.                                                                                     |
| 6B                              | Inclusion in decision-making                 | Percentage of individuals participating in local elections out of all eligible voters.                                                                                          |

a consultation with academic and policy experts who comprise the Systems Science in Public Health and Health Economics (SIPHER) consortium (Meier et al. 2019). In a second step, we sought feedback from an international IE advisory panel and community panels, which reflected areas that align with the SIPHER policy partners (Scotland, Greater Manchester Combined Authority and Sheffield City Council). The resulting indicator set derived from this process, and presented in this paper, was then aligned to be reflected through available data sets which capture each of the 13 dimensions specified. In a last step, all data sets to which the 13 indicators were aligned were reviewed by the SIPHER consortium. Further details on the calculation for each of the variables are provided in Supporting Information S1: Appendix 1, with technical documentation available at <https://doi.org/10.17605/OSF.IO/VNSUR>.

## 2.2 | Imputation of Missing Information

For all indicators where only one cross-sectional measurement was available (e.g., food insecurity), we assumed values to be constant within the period 2017–2021. Overall, we observed a level of missing information of 7.36% across the data set covering the 13 indicators of economic inclusion alongside supplementary demographic and health indicators for the period 2017–2021. Here, the level of missing information was highest among LADs with the smallest and most selected populations, such as the City of London and the Isles of Scilly. For context, LAD population sizes range between  $N = 2000$  and  $N = 1.14$  m, with an average of  $N = 178,500$ .

Once all data were gathered, harmonised and collated, we used the Amelia II package to reduce the amount of missing data to 0%. Implemented in R, Amelia II is a Bayesian multiple imputation algorithm, which imputes every missing data point while considering all existing as well as any additionally provided information for a particular data point (Honaker et al. 2011). In addition, Amelia II allows for automated detection of trends, lags and leads in time series data. One strength of Amelia II lies in its ability to utilise additional information to inform the multiple imputation process, leading to more reliable imputation results. We used additional information on the demographic composition and health outcomes within each LAD to inform the imputation of missing IE indicators. To represent the demographic composition of each LA, we used the sex ratio (number of all males divided by the number of all females) and the total dependency ratio (i.e.,  $N[\text{aged } 0-15] + N[\text{aged } 65+]/N[\text{aged } 16-64] * 100$ ). For health outcomes, we used LE and LV at birth for males, females and both sexes combined. All missing values were obtained by selecting the median of each imputed missing value based on 1000 runs of Amelia II. A detailed summary of the imputation is provided in Supporting Information S1: Appendix 2. Missingness was not evenly distributed across variables. For example, the level of missingness in the ‘inclusion in decision making’ variable was 52.56% as elections were not held in every LAD in every year. We performed sensitivity analysis on results presented in this paper by excluding this heavily imputed variable and re-running analysis. Description of this study is presented in Supporting Information S1: Appendix 2. In summary, the results presented later remained the same and the same

conclusions would be drawn on the relationship between IE and LE/LV.

## 2.3 | Calculating Health Outcome Measures

We investigated the association of LAD clusters with two health outcome measures: LE and LV at birth. By capturing LE and LV, representing outcomes of average mortality and mortality inequality that exists within each LAD, we provide a much more holistic perspective on population health than looking at one dimension in isolation.

LE provides a cross-sectional summary measure of age-specific mortality rates. It quantifies the average length of life in years, assuming that age-specific mortality rates, observed throughout a particular period of calendar time (e.g., 2019–2021) within a population, will remain constant throughout the entire life of a newborn man or woman of this population (Luy et al. 2020). The implicit age standardisation of the underlying life table method allowed for direct comparison over time and areas (Modig et al. 2020).

While LE is routinely published for subnational levels such as LADs, this is not the case for LV. As a comparison of both measures requires comprehensive and comparable life tables, we derived these for all LADs, using one consistent methodology. From these life tables, we then estimated LE and LV—the two health outcome measures explored in this study. Death counts and population-at-risk estimates were obtained from routine reporting of the respective statistical offices. This means that population-at-risk data for LADs in England, Wales and Scotland as well as death counts for LADs in England and Wales were queried through the Office for National Statistics’ NOMIS website. Death counts for LADs in Scotland were downloaded directly from the National Records of Scotland website.

LV is a measure of within-population mortality inequality and reflects the distribution of lifespans of the underlying life table population. A number of LV measures have previously been proposed, including absolute (e.g., e-dagger) and relative/dimensionless (e.g., coefficient of variation or standard deviation) measures (Van Raalte and Caswell 2013). We captured LV using the e-dagger method (Vaupel and Romo 2003). Intuitively expressed in years, e-dagger quantifies the average number of years an individual of the life table population dies prematurely relative to the average age at death of other individuals in this life table population who are exposed to that particular mortality schedule.

We obtained all LE and LV estimates directly from previously estimated life tables. Instead of using raw mortality data for calculating these life tables, we derived life tables for men and women within each area based on modelled age- and sex-specific mortality rates for 1-year age groups. Estimated age- and sex-specific mortality rates for the age range 0–90 were modelled using a tool for projecting age-specific mortality rates using linear splines (TOPALS) models (de Beer 2011). Mortality rates above age 90 were estimated separately via a Kannisto model (Kannisto et al. 1994). For all estimations, we pooled data for 3 consecutive years and captured LADs in the exact same

boundary definitions (December 2021) as used for the 13 IE indicators. Our input data to estimate mortality rates reflected information for 5-year age groups with an upper age group 90+.

TOPALS models were used, as estimating LE and LV reliably for areas with small population sizes from raw mortality data only can be challenging. Key reasons are that mortality data are typically provided in broad age bands only, and because fluctuations in the number of observed deaths might cause raw age-specific mortality rates to be imprecise (Eayres and Williams 2004). Typically, this imprecision increases as the population size of the area decreases and only a small and fluctuating number of deaths (or no deaths at all) are observed among certain age groups. Our TOPALS models followed a general (non-Bayesian) approach for age-grouped data (Rau and Schmertmann 2020; Schmertmann 2023). For modelling purposes, we used age-specific mortality rates reflecting the general population of England and Wales as well as Scotland in 2010–2019 from the Human Mortality Database to represent the corresponding standard reference schedules by single years of age (Barbieri et al. 2015; HMD 2023). As TOPALS models are not well-suited for modelling mortality rates at higher ages (Rau and Schmertmann 2020), we modelled mortality rates above age 90 separately via a Kannisto model, based on modelled mortality rates for 1-year age groups obtained from TOPALS models. We then followed standard methodology to derive life tables and the outcome measures LE and LV (Chiang 1984). While the accompanying data set contains information for single calendar years, we calculated the average LE and LV for the 2017–2021 period for analysis presented in this paper, in line with our approach of handling all IE variables.

## 2.4 | Clustering Method

We used *k*-means classification (Hartigan and Wong 1979) to identify clusters of LADs which were similar in terms of the 13 indicators representing economic inclusivity. In brief, *k*-means is an unsupervised algorithm which groups together cases (here: LADs) based on their characteristics (here: the 13 IE variables), so that they are more similar to cases in the group than they are to cases in other groups. All clustering was undertaken in R using the NbClust package (Charrad et al. 2014).

Generally, *k*-means is better able to identify the underlying data structure if variables are normally distributed. To ensure that each variable provides a similar weight to the classification process, the data required a Z-standardisation. Data transformation was undertaken using the scale function in R, which calculates a Z-score by subtracting the mean of each indicator value for each case, then dividing by the standard deviation. As a result, we obtained a balanced data set where the mean value for each indicator is centred on zero.

Several methods exist to assess the optimal number of clusters within the data set. Figure 1 provides a visual representation of the 13 variables used in our clustering in two dimensions, achieved by undertaking principal component analysis (PCA) (Abdi and Williams 2010) on the input data set. As we are seeking to visually limit the overlap of observations, Figure 1 suggests a solution of between  $k = 3$  and  $k = 6$  to be optimal.

The clustergram method (Schonlau 2002) provides a visualisation of the weighted mean of the first component of a PCA for each cluster centre and is utilised as a diagnostic tool in similar clustering applications (e.g., see A. Singleton et al. 2020). We sought a number of clusters where mean values were dissimilar (i.e., the clusters don't overlap). As Figure 2 shows, there was a clear differentiation between cluster mean values for the two-, three- and four-cluster solutions. For five or more clusters, there was no clear definition as mean values started to overlap. Note that during development, we produced multiple clustergrams based on different runs of *k*-means using a random seed to assess cluster stability. The results were consistent over multiple runs.

For a further quantitative assessment of the optimal number of clusters, we used the NbClust package in R (Charrad et al. 2014), which offers 30 different indices with associated scores. Charrad et al. (2014) suggest that using a majority rule (i.e., the number of clusters that is suggested by the largest number of indices) is the best approach. Running NbClust on our data set, and taking the majority, the optimal number of clusters suggested is two or four, which both receive 9/30 indices recommendations. The next most recommended is a five-cluster solution (3/30). A summary of the indices extracted from NbClust is presented in Supporting Information S1: Appendix 3.

In summary, considering all available diagnostics, the optimal *k* value was between two and five clusters. However, deciding the number of clusters is an art as, ultimately, the solution needs to be meaningful and provide insight that is not available from the unclustered data. Therefore, results from all suggested options ( $k = 2$  to  $k = 5$ ) were further assessed against three criteria:

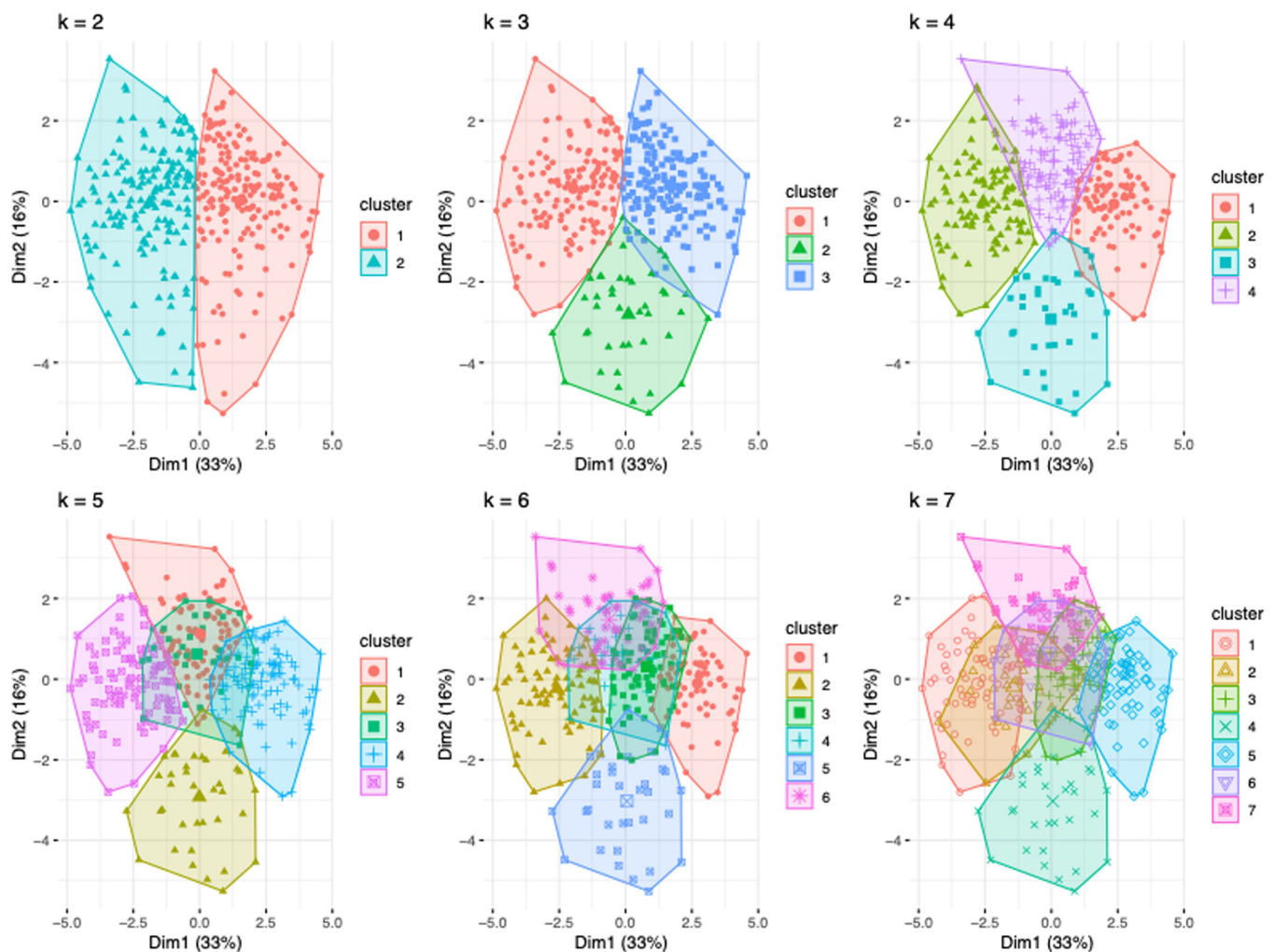
- Conceptual distinction: Whether the cluster is distinct (when compared to the other clusters) across all 13 dimensions of the IE indicators.
- Spatial separability: Whether there was sufficient and sensible spatial distribution of the clusters to provide insight.
- Association with health outcomes: Whether the classification enabled a suitable approach for differentiating groups of areas with distinct patterns with respect to LE and LV.

Against these qualitative criteria, the four-cluster solution performed best. Here, the four-cluster solution provided the greatest insight in terms of distinct cluster profiles, a well-defined and understandable spatial distribution and a meaningful differentiation with respect to LE and LV.

## 3 | Results

### 3.1 | Cluster Results

Summary results from the four-cluster solution are presented in Figure 3. These radial plots demonstrate the mean cluster value for each of the 13 IE variables relative to the national (i.e., all-LAD, unweighted) average. Cluster mean values above zero demonstrate that values for the attribute are higher than average (and vice versa: below zero signifies lower than average). The sign of some indicators was reversed to aid interpretability so that values above zero indicate that the attribute for the cluster of LADs is more



**FIGURE 1** | A representation of the clustering of points in two dimensions following principal component analysis (PCA), where each data point is a Local Authority District (LAD).

inclusive than average, while values below zero are less inclusive than average. Figure 4 shows a standard projection and a hexagon cartogram of the distribution of clusters across LADs. The hexagon cartogram groups LADs into counties and other administrative areas, where the size of the LAD is relative to the size of its population, and makes it easier to see individual LADs and their position relative to other LADs (House of Commons Library 2022).

Considering the results shown in radial plots and maps, each cluster allowed for a meaningful interpretation with respect to economic inclusivity:

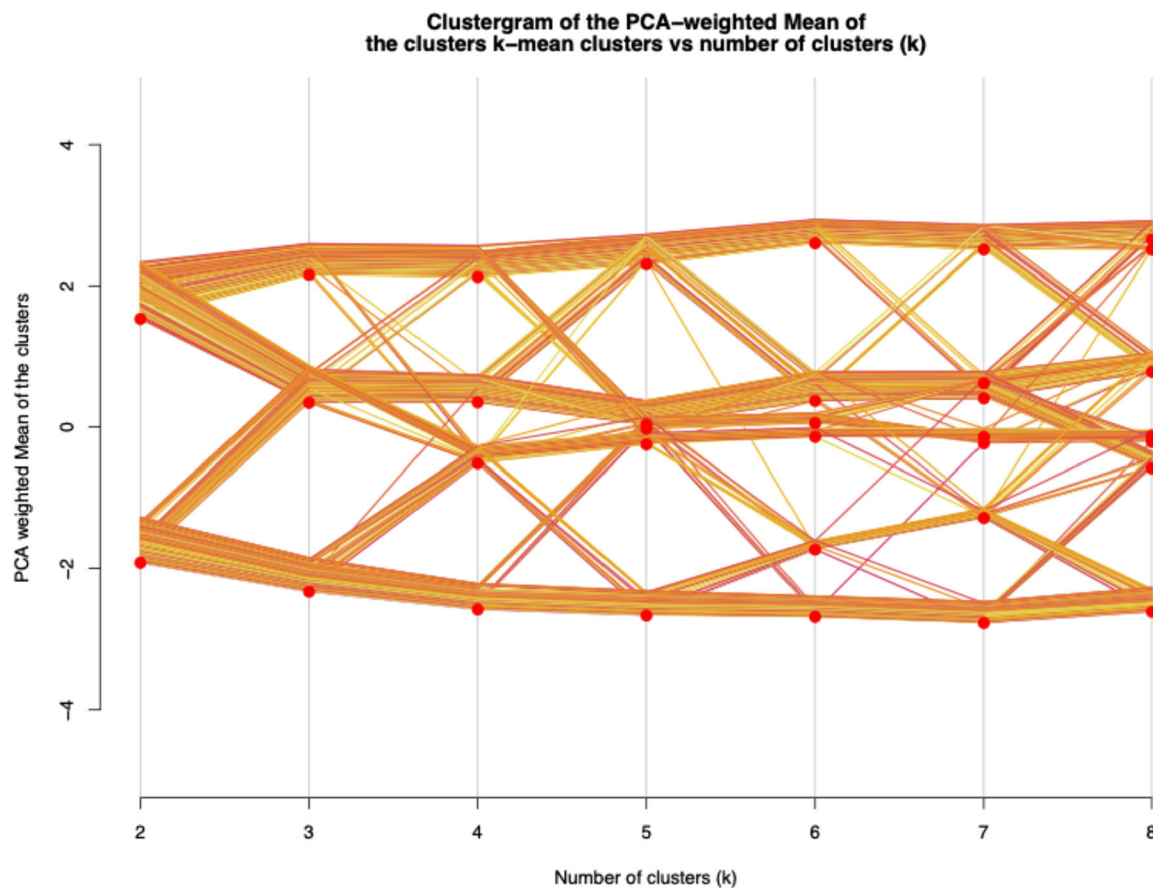
Cluster 1 was labelled ‘less economically inclusive’, as 10/13 indicators have a value below the national average. This cluster only performed well on physical connectivity (3B), earnings inequality (4A) and housing affordability (4B). The spatial distribution of LADs in this cluster shows strong representation in the north-east of England, a number of areas in northern England between Leeds and Manchester, the Scottish central belt connecting Edinburgh and Glasgow and in South Wales. In total, this cluster comprised 99 LADs.

Cluster 2 was labelled ‘more economically inclusive’, with values above the national average for all except three indicators.

Indicators below the average (representing less inclusive domains) are physical connectivity (3B), earnings inequality (4A) and housing affordability (4B). Many of the Cluster 2 LADs are in the south-east of England and there is also strong representation in Scotland. The cluster comprised a total of 104 LADs.

Cluster 3 represented the most ‘average’ cluster, with values across many of the indicators close to zero, pointing towards a strong similarity with the national average. We found that 6/13 indicators were above zero with only employment (Indicator 1A), housing Affordability (4B) and cost of living (5B) of note in terms of the magnitude of its positive scores. Of the 7/13 below zero, the most prominent features are poorer than average scores for skills and qualifications (1B), physical connectivity (3B) and decent pay (6A). The spatial distribution of Cluster 3 LADs is fairly evenly spread across Great Britain, with concentrations in the south-east of England, the Midlands, the north-west of England and southern Scotland. Cluster 3 comprises 122 LADs.

Cluster 4 was considered to reflect LADs which were a ‘mix of extremes, mostly London’—a highly polarised cluster—where values deviate substantially from the national average across



**FIGURE 2** | Clustergram of the cluster allocation for Local Authority Districts (LADs).

multiple variables and where there is a higher prevalence of less inclusive (7/13) than more inclusive (6/13) values. These LADs are on average more inclusive in the domains of physical connectivity (3B), wealth inequality (3A), digital inclusion (2B), involuntary exclusion from the labour market (2A), inclusion in decision making (6B) and skills and qualifications (1B). The cluster is less inclusive in terms of participation in paid employment (1A), job security (7A), decent pay (6A), cost of living (5B), poverty (5A), housing affordability (4B) and earnings inequality (4A). These LADs are predominantly located in London—30 of the 33 London Boroughs are within this cluster—along with Bristol, Brighton, Portsmouth, Southampton, Bournemouth, Cambridge, Exeter and Oxford which makes this our smallest cluster. However, while only capturing 38 LADs, these LADs comprise a significant part of the population of Great Britain.

### 3.2 | Assessment of LE and LV by Cluster

Figure 5 shows the distribution of LAD LE and LV values by cluster for males and females combined (results disaggregated by sex were very similar so not shown here). LADs within Cluster 1 ('less economically inclusive') have the lowest LE of all clusters (median = 79.83 years, IQR = 1.36) and the highest LV (median = 11.07 years, IQR = 0.46). This is contrasted with Cluster 2 ('most economically inclusive') in which LE is highest (median = 83.44 years, IQR = 1.09) while LV is lowest (median = 9.85 years, IQR = 0.57). Cluster 3 ('average') has LE

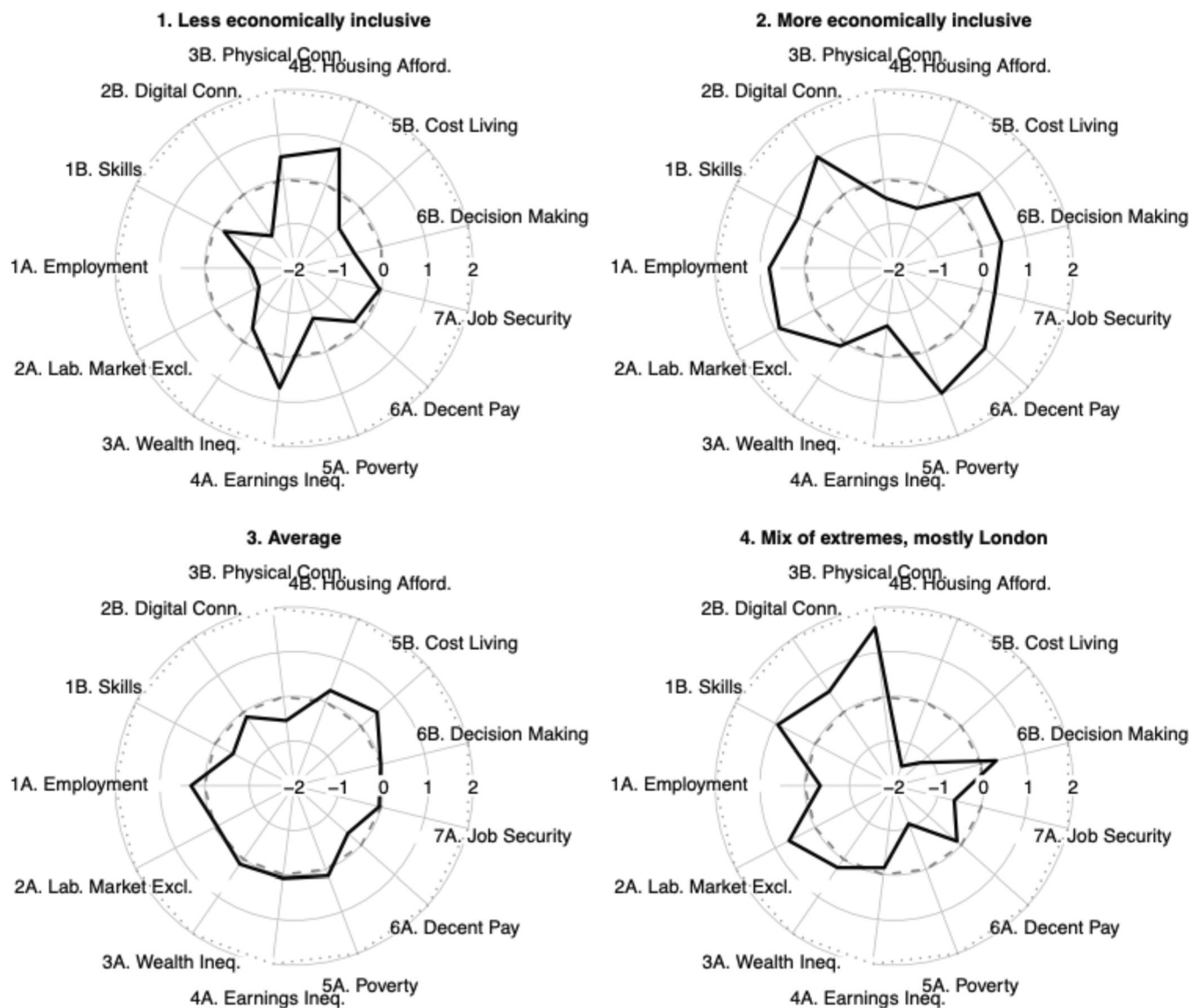
and LV values between the less and more economically inclusive clusters (LE: median = 81.82 years, IQR = 1.39; LV: median = 10.34 years, IQR = 0.73). Cluster 4 ('mix of extremes, mostly London') sees LE similar to the average group (median = 82.21, IQR = 1.30), but with higher levels of LV (median = 10.77 years, IQR = 0.57).

Because the relationship between LE and LV is revealing (Seaman et al. 2016), Figure 6 provides a visual representation of both outcomes as they relate to LADs within the four clusters. Each point corresponds to a single LAD, with the colour indicating cluster membership. To aid interpretability, we fitted a linear model for each cluster and showed the trend line with 95% confidence intervals (unweighted fit by population size of LADs).

Overall, the correlation between LE and LV across all LADs combined was strong and negative as expected ( $R = -0.87$ ,  $p < 0.01$ ). This held for the 'less economically inclusive' ( $R = -0.82$ ,  $p < 0.01$ ), 'more economically inclusive' ( $R = -0.87$ ,  $p < 0.01$ ) and 'average' ( $R = -0.85$ ,  $p < 0.01$ ) clusters. However, the relationship was weaker for the 'mix of extremes, mostly London' cluster ( $R = -0.47$ ,  $p < 0.01$ ).

## 4 | Discussion

In this paper, we established a typology of Local Authority Districts in Great Britain based on economic inclusivity and

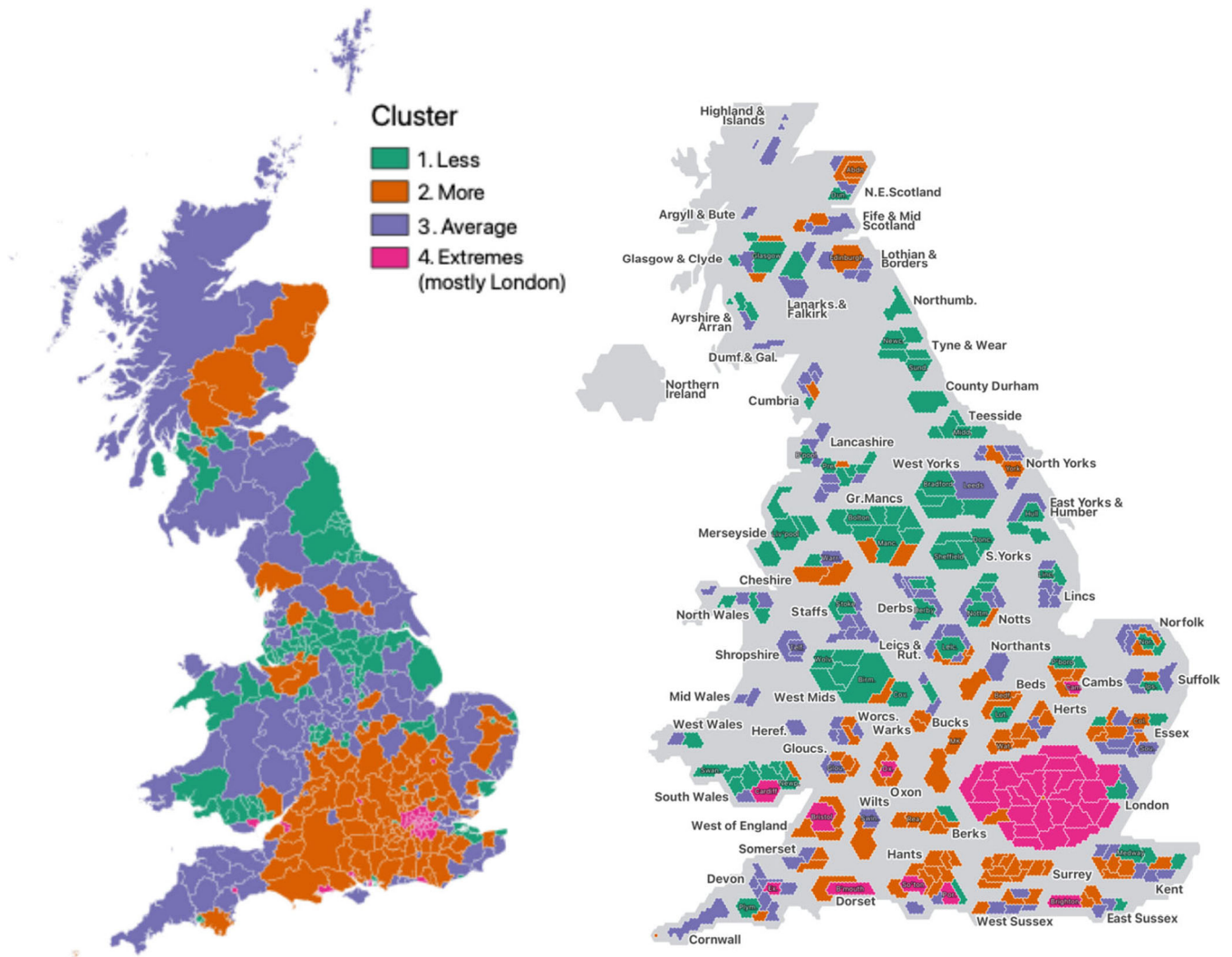


**FIGURE 3** | Radial plots of Z-scores for 13 inclusive economy indicators across four clusters. Note the sign of some indicators has been reversed so that values above zero represent more inclusive than average, while below zero represent less inclusive than average.

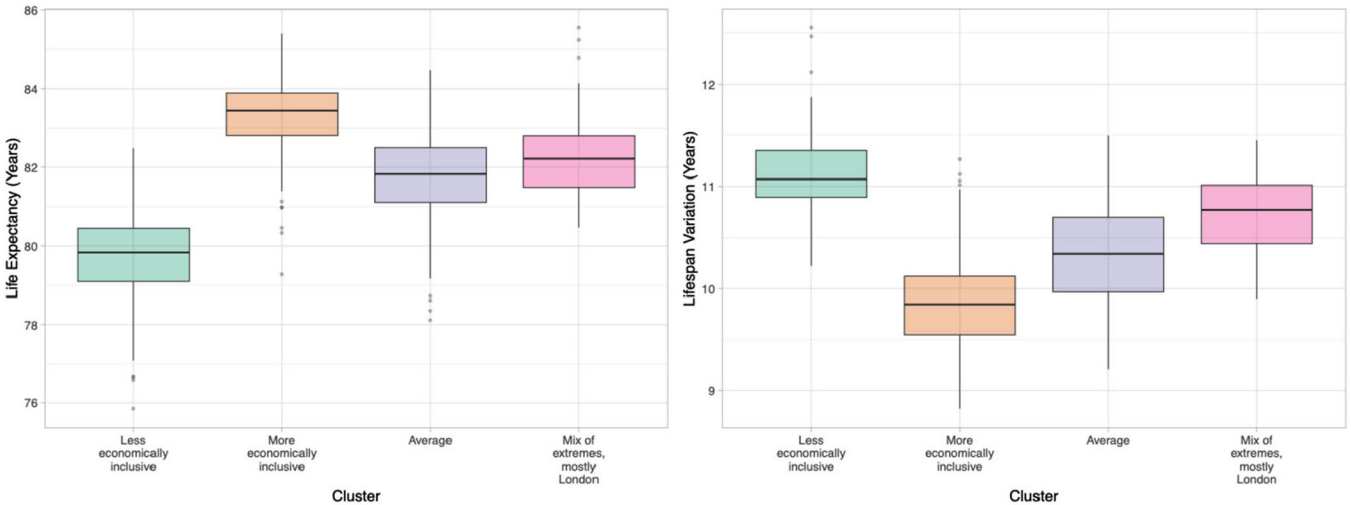
explored the relationship between identified clusters and average population health (captured as LE), as well as population health inequality (captured as LV). For this purpose, we first developed a classification of Local Authority Districts in Great Britain which captures 13 dimensions of economic inclusion set out in Lupton et al. (2023). The classification comprised four clusters of Local Authority Districts which are distinct in their profile across these 13 dimensions. Overall, we found a strong association between economic inclusion and both measures of population health. We found that Local Authority Districts within the 'less economically inclusive' cluster experienced the lowest LE and highest LV, while Local Authority Districts of the 'more economically inclusive' cluster had the highest LE and lowest LV. We found that the cluster defined as 'mix of extremes, mostly London' had a relatively high LE, but at higher levels of LV than we would expect. Overall, our results point very clearly towards a double disadvantage among the 'economically less inclusive' cluster. Here in particular, the high uncertainty around average lifespans presents a substantial challenge for

public planning, spanning from health or care needs to pensions and housing (van Raalte and Seaman 2019).

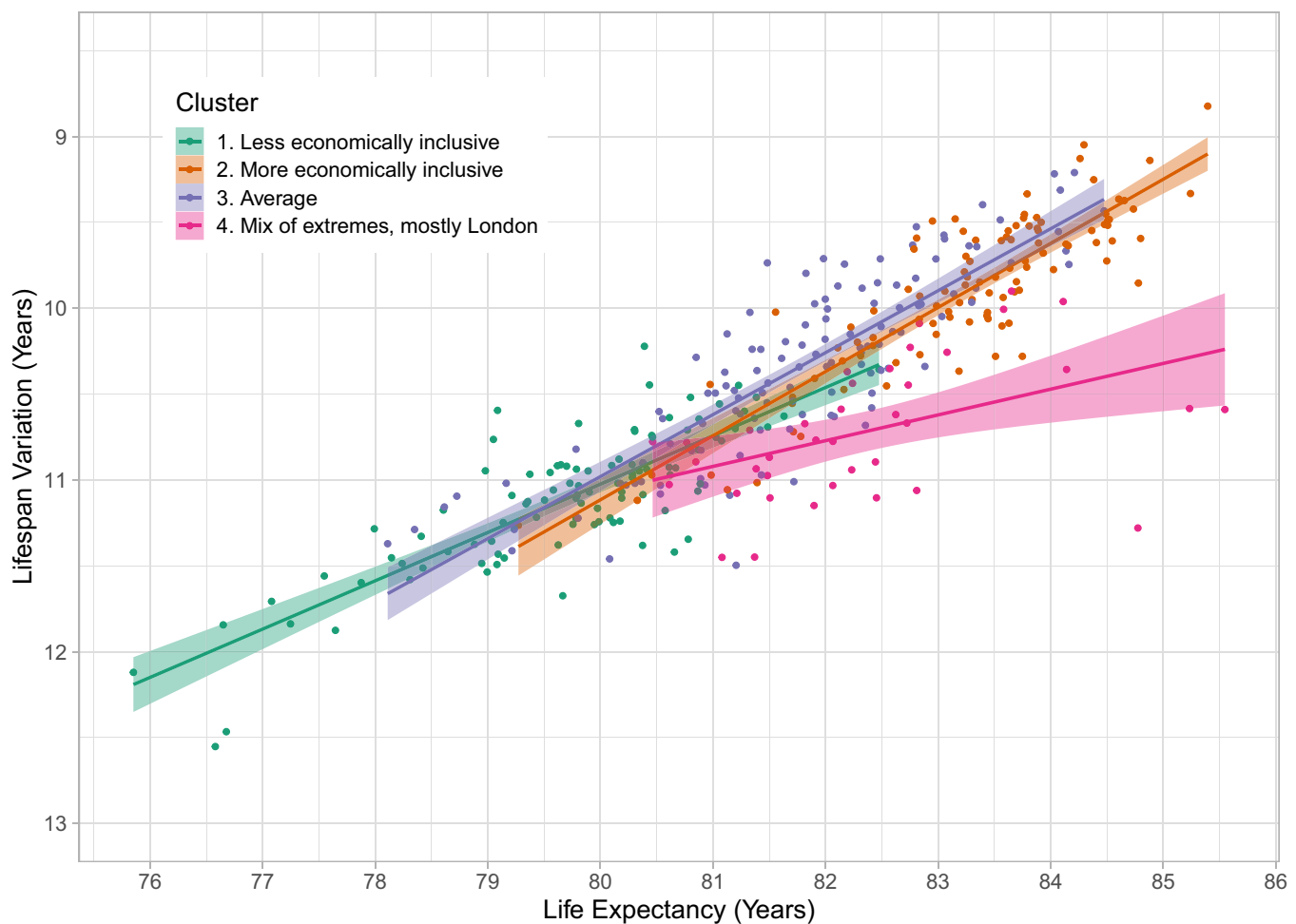
Consistent with previous literature on LE and LV, we found a strong relationship between the two measures at the studied subnational level: Local Authority Districts with high LE tend to have a lower LV. This held true for the 'more economically inclusive', 'less economically inclusive' and 'average' clusters. However, the 'mix of extremes, mostly London' cluster was an exception. Local Authority Districts in this cluster showed relatively high levels of LE, while levels of LV also remained high, indicating a high level of mortality inequality. While an exception, this finding is not surprising as it might be possible that the polarisation of economic inclusivity among this cluster, and its impact on the determinants of health, might translate to a polarisation of lifespans. Generally, high levels of LV could stem from both tails of the age-at-death distribution and reflect both: relatively high levels of premature mortality below the life table threshold age, as well as low levels of mortality among the oldest population subgroups (Gillespie et al. 2014). As both



**FIGURE 4** | Local Authority District (LAD) cluster allocations shown on a standard projection (left) and a hexagon cartogram (right).



**FIGURE 5** | Distribution of life expectancy (LE, left) and lifespan variation (LV, right) by Local Authority District (LAD) for the four inclusive economy clusters.



**FIGURE 6** | Association between life expectancy (LE) and lifespan variation (LV) for Local Authority Districts (LADs) in Great Britain, disaggregated by cluster.

phenomena can exist independently of each other, future research with a dedicated focus on decomposition is required to shed light on the role of age groups and causes of death contributing to this mix of population health outcomes.

Our analysis builds on work that assesses the association between LE/LV for specific groups (Sasson 2016), or by area-level deprivation (Seaman, Riffe, and Caswell 2019) by providing a perspective on these outcomes relative to multiple dimensions of economic inclusion. While shown in cross-country comparisons and for population subgroups within countries, the relationship between LE and LV has previously not been explored for small areas within countries, such as Local Authority Districts in Great Britain. The call from Public Health England (2021) for place-based action on inclusive (and sustainable) economies as a mechanism to reduce health inequalities demonstrates the first contribution of our work, which identifies the types of area (through classification) and in doing so individual areas within the classification that are most in need in terms of targeted intervention because of the interplay between their IE profile and population health outcomes. Investigation of the IE characteristics of individual Local Authority Districts that exhibit the lowest LE and highest LV should be the focus of further work.

Previous work by Hoehn et al. (2024) has assessed an alternative outcome: QALE, in relation to the IE indicators defined by Lupton et al. (2023). While Hoehn et al. (2024) found that indicators explained more than half of the variance in QALE at the Local Authority District level, the relationship was relatively complex as directions of associations were not always in line with the initial hypothesis for several indicators. Hoehn et al. (2024) suggest that this might be due to residual confounding in the relationship between indicators. This highlights a second key contribution of our work: by creating a multidimensional classification of the IE, we were able to present a more holistic view of what constitutes the economic inclusivity of an area. However, none of the four clusters were universally inclusive or uninclusive across all 13 dimensions. Instead, the profiles present a mixed picture relative to other clusters. This highlights the difficulties of defining clear markers of an IE (Shipton et al. 2021), identifying the variables that capture the desired definition and calculating consistent metrics across areas. An example is the value of three indicators in the 'more economically inclusive' and 'less economically inclusive' clusters that run counter to expectation: physical connectivity (3B), earnings inequality (4A) and housing affordability (4B). These perform poorly (relative to the average) in the 'more economically inclusive' cluster, and well in the 'less economically inclusive' cluster. In addition to the residual confounding previously

reported by Hoehn et al. (2024), some within-indicator mismatch of geographies might have contributed to this unexpected result. For example, housing affordability captures the place of residence house prices against workplace earnings, and could therefore be conditional on patterns across two different areas. In this particular case, high ratios of house prices against workplace earnings could be indicative of affluent suburban areas with few job opportunities in the area itself.

A third contribution of our work is the creation of a data set that represents 13 dimensions of the IE which allows for direct comparison of subnational areas in the three nations in Great Britain (Scotland, England, Wales) which would not otherwise be possible. In producing these data, underlying theoretical concepts as defined in Lupton et al. (2023) were reviewed and the selection and calculation of indicators involved coproduction between researchers and policy and practice partners. The process for creating the data has been extensively documented, including the harmonisation of indicators, and the code and final data set have been made publicly available at <https://doi.org/10.17605/OSF.IO/VNSUR>. As such, we have curated a data set that can be used for further research and that contributes to the ongoing debate around how best to define and measure the IE (Shipton et al. 2021).

The relationship between LE, LV and deprivation has been well documented (Woods et al. 2005; van Raalte et al. 2018). In a sensitivity analysis, we therefore performed a post hoc analysis to establish that results from our cluster solution were not simply replicating deprivation relationships observed across areas. Looking at the crossover of our four clusters with the 2019 English Indices of Deprivation (McLennan et al. 2019) by stratifying the Local Authority District level summaries into quartiles, we found that areas identified as 'less economically inclusive' in our clustering were more heavily represented in the two most deprived quartiles, while 'more economically inclusive' were more heavily represented in the two least deprived quartiles. While we would expect some relationship, given that some of the indices of deprivation domains are represented in our clustering (income, employment, education, barriers to housing), there was no one-to-one mapping, suggesting that the classification presented in this paper captured more dimensions than just deprivation. Our 'average' cluster was spread across all deprivation quartiles, while the 'mix of extremes, mostly London' group of Local Authority Districts were spread across quartiles 1 (most deprived) to 3. These results suggest that the relationship between IE and deprivation is complex, and that the two sets of indicators are indeed capturing different concepts.

While our results are specific to the geographical context studied, our findings have wider implications. Our study provides strong evidence for an association between economic inclusivity and average lifespan, as well as inequality captured by LV within and between clusters. Our results demonstrate this relationship for a spatial resolution of administrative units at which deliberate resource allocation decisions can be made to affect access to services and jobs. Future research translating our approach to other geographical and welfare state contexts provides a valuable extension. This echoes previous calls for further work on the assessment of interventions and policies targeting economic inclusivity and health equity (Macintyre et al. 2024). Our work aligns with these calls by identifying how various dimensions of

an IE impact average health and health inequality and could be extended with a specific focus on particular policy areas. Our study is not without limitations. We aimed to use the best available data to calculate indicators reflecting the IE definition. However, pragmatic decisions had to be made to ensure consistent coverage as there is no perfect set of data that represents all covered domains. Some indicators or raw data contributing to the estimation of indicators and health outcome measures (e.g., mid-year population estimates used to estimate LE) might be subject to retrospective changes in the future (e.g., census readjustment). In addition, some indicators might have been available in a more granular format (with respect to age or geography) via bespoke requests or safeguarded data sets. Indicators were constructed from a mix of underlying data sources. This included survey data which may be impacted by sampling strategies and systematic nonresponse (e.g., we make use of the UK Labour Force Survey). In their raw form, some data sets were not consistent between the nations of Great Britain (e.g., the underlying data used to calculate the transport accessibility metric were different in England, Wales and Scotland). This required some harmonisation to produce a comparable indicator. We also had to impute some missing data, where the missingness for some variables was high. Sensitivity analysis presented in the paper suggests that imputation did not impact conclusions drawn from the work, but more complete data would be preferable if they were available. Data availability constraints meant that we had to exclude Northern Ireland from our analysis, which is the reason for the focus on Great Britain rather than the United Kingdom.

While a more spatially detailed analysis than the Local Authority District level may be desirable, data constraints are prohibitive. Data from non-synthetic sources required to calculate most of the 13 IE indicators is not consistently available below the Local Authority District level. Similar data constraints existed when choosing and calculating the health outcomes of interest. In the methodology, we describe issues when calculating LE and LV for the smallest Local Authority Districts, and as such, the methods could not be applied to smaller geographical zones within Great Britain, such as Wards or Output Areas. The degree of separation identified using *k*-means clustering and the subsequent variation in outcomes provides valuable insight at the Local Authority District scale, while balancing data availability and reproducibility.

Finally, the geographical zones that constitute Local Authority Districts in Great Britain are subject to periodic revision. We set our reference time point as the 2021 boundaries. However, a recent 2023 revision saw the aggregation of some spatial zones to create new boundaries for which there are not currently data available across the 13 IE indicators. Further work could reproduce our study according to these new boundaries or indeed by following the methods set out in this paper to reproduce results for other countries, spatial zones or other area-level and sociodemographic characteristics of interest where suitable data are available.

## Acknowledgements

This study was supported by the UK Prevention Research Partnership (MR/S037578/2), which is funded by the British Heart Foundation,

Cancer Research UK, Chief Scientist Office of the Scottish Government Health and Social Care Directorates, Engineering and Physical Sciences Research Council, Economic and Social Research Council, Health and Social Care Research and Development Division (Welsh Government), Medical Research Council, National Institute for Health Research, Natural Environment Research Council, Public Health Agency (Northern Ireland) and The Health Foundation and Wellcome. The authors thank Mark Nyack at the Electoral Commission for providing local election results that were not available online, Alistair Grant at the Scottish Boundary Commissions' Secretariat for information about Scottish ward and local authority definitions, Mitchel Langford and Gary Higgs of the University of South Wales for their Welsh transport metric (WABI) data (Indicator 3B) and Alex Singleton and Alexandros Alexiou at the University of Liverpool for giving permission to make an aggregated version of the Internet User Classification (Indicator 2B) available in the accompanying data set.

## Conflicts of Interest

The authors declare no conflicts of interest.

## References

- Abbas, J., A. Ojo, and S. Orange. 2009. "Geodemographics: A Tool for Health Intelligence?" *Public Health* 123, no. 1: 35–39.
- Abdi, H., and L. J. Williams. 2010. "Principal Component Analysis." *WIREs Computational Statistics* 2, no. 4: 433–459.
- Aburto, J. M., F. Villavicencio, U. Basellini, S. Kjærgaard, and J. W. Vaupel. 2020. "Dynamics of Life Expectancy and Life Span Equality." *Proceedings of the National Academy of Sciences of the United States of America* 117, no. 10: 5250–5259.
- Barbieri, M., J. R. Wilmoth, V. M. Shkolnikov, et al. 2015. "Data Resource Profile: The Human Mortality Database (HMD)." *International Journal of Epidemiology* 44, no. 5: 1549–1556.
- Beatty, C., R. Crisp, and T. Gore. 2016. *An Inclusive Growth Monitor for Measuring the Relationship Between Poverty and Growth*. Joseph Rowntree Foundation (JRF).
- Charrad, M., N. Ghazzali, V. Boiteau, and A. Niknafs. 2014. "Nbclust: An R Package for Determining the Relevant Number of Clusters in a Data Set." *Journal of Statistical Software* 61, no. 6: 1–36.
- Chiang, C. L. 1984. *The Life Table and Its Applications*. Krieger Publishing Company.
- Clark, S., N. Lomax, M. Morris, F. Pontin, and M. Birkin. 2021. "Clustering Accelerometer Activity Patterns From the UK Biobank Cohort." *Sensors* 21, no. 24: 8220.
- Clark, S., N. Lomax, and M. A. Morris. 2017. "Classification of Westminster Parliamentary Constituencies Using e-Petition Data." *EPJ Data Science* 6, no. 1: 16.
- De Beer, J. 2011. "A New Relational Method for Smoothing and Projecting Age-Specific Fertility Rates: Topals." *Demographic research* 24: 409–454.
- Dennett, A., and J. Stillwell. 2011. "A New Area Classification for Understanding Internal Migration in Britain." *Population Trends* 145: 146–171.
- Eayres, D., and E. S. Williams. 2004. "Evaluation of Methodologies for Small Area Life Expectancy Estimation." *Journal of Epidemiology and Community Health* 58, no. 3: 243–249.
- Gillespie, D. O. S., M. V. Trotter, and S. D. Tuljapurkar. 2014. "Divergence in Age Patterns of Mortality Change Drives International Divergence in Lifespan Inequality." *Demography* 51, no. 3: 1003–1017.
- Grubestic, T. H., J. A. Miller, and A. T. Murray. 2014. "Geospatial and Geodemographic Insights for Diabetes in the United States." *Applied Geography* 55: 117–126.
- Hartigan, J. A., and M. A. Wong. 1979. "Algorithm as 136: A k-Means Clustering Algorithm." *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 28, no. 1: 100–108.
- Hill O'Connor, C., K. Smith, C. Hughes, P. Meier, and R. Purshouse. 2023. "Operationalizing Inclusive Growth: Can Malleable Ideas Survive Metricized Governance?" *Public Administration* 102, no. 1: 114–130.
- HMD. 2023. "Human Mortality Database." June 1. [www.mortality.org](http://www.mortality.org).
- Hoehn, A., N. Lomax, H. Rice, et al. 2024. "Estimating Quality-Adjusted Life Expectancy (QALE) for Local Authorities in Great Britain and Its Association With Indicators of the Inclusive Economy: A Cross-Sectional Study." *BMJ Open* 14, no. 3: e076704.
- Honaker, J., G. King, and M. Blackwell. 2011. "Amelia II: A Program for Missing Data." *Journal of Statistical Software* 45: 1–47.
- House of Commons Library. 2022. "Geographical Templates for Non-Contiguous Cartograms of the UK." <https://github.com/houseofcommonslibrary/uk-hex-cartograms-noncontiguous>.
- Issifou, I., and J. Pewitt. 2022. "On the Importance of Monitoring Inequality in Life Expectancy."
- Kannisto, V., J. Lauritsen, A. R. Thatcher, and J. W. Vaupel. 1994. "Reductions in Mortality at Advanced Ages: Several Decades of Evidence From 27 Countries." *Population and Development Review* 20: 793–810.
- Lupton, R., C. Hughes, and N. Lomax. 2023. "SIPHER Inclusive Economy Indicator Set: Technical Paper." [https://www.gla.ac.uk/media/Media\\_970680\\_smx.pdf](https://www.gla.ac.uk/media/Media_970680_smx.pdf).
- Lupton, R., P. Obolenskaya, and A. Fitzgerald. 2016. "Spatial Inequalities." In *Social Policy in a Cold Climate*, 291–316. Policy Press.
- Luy, M., P. Di Giulio, V. Di Lego, P. Lazarevič, and M. Sauerberg. 2020. "Life Expectancy: Frequently Used, but Hardly Understood." *Gerontology* 66, no. 1: 95–104.
- Macintyre, A. K., D. Shipton, S. Sarica, G. Scobie, N. Craig, and G. McCartney. 2024. "Assessing the Effects of Population-Level Political, Economic and Social Exposures, Interventions and Policies on Inclusive Economy Outcomes for Health Equity in High-Income Countries: A Systematic Review of Reviews." *Systematic Reviews* 13, no. 1: 58.
- McLennan, D., S. Noble, M. Noble, E. Plunkett, G. Wright, and N. Gutacker. 2019. "The English Indices of Deprivation 2019: Technical Report."
- Meier, P., R. Purshouse, and M. Bain. 2019. "The SIPHER Consortium: Introducing the New UK Hub for Systems Science in Public Health and Health Economic Research." *Wellcome Open Research* 4, no. 174: 174.
- Modig, K., R. Rau, and A. Ahlbom. 2020. "Life Expectancy: What Does It Measure?" *BMJ Open* 10, no. 7: e035932.
- Moon, G., L. Twigg, K. Jones, G. Aitken, and J. Taylor. 2019. "The Utility of Geodemographic Indicators in Small Area Estimates of Limiting Long-Term Illness." *Social Science & Medicine* (1982) 227: 47–55.
- Németh, L. 2017. "Life Expectancy Versus Lifespan Inequality: A Smudge or a Clear Relationship?" *PLoS One* 12, no. 9: e0185702.
- OECD. 2019. *Health for Everyone?: Social Inequalities in Health and Health Systems*. OECD Publishing.
- Pearce, J. R., E. A. Richardson, R. J. Mitchell, and N. K. Shortt. 2010. "Environmental Justice and Health: The Implications of the Socio-Spatial Distribution of Multiple Environmental Deprivation for Health Inequalities in the United Kingdom." *Transactions of the Institute of British Geographers* 35, no. 4: 522–539.
- Permanyer, I., J. Spijker, A. Blanes, and E. Renteria. 2018. "Longevity and Lifespan Variation by Educational Attainment in Spain: 1960–2015." *Demography* 55, no. 6: 2045–2070.
- Public Health England. 2021. "Inclusive and Sustainable Economies: Leaving No-One Behind. Supporting Place-Based Action to Reduce Health Inequalities and Build Back Better. Technical Report, PHE."

- Rau, R., and C. P. Schmertmann. 2020. "District-Level Life Expectancy in Germany." *Deutsches Ärzteblatt International* 117, no. 29–30: 493.
- Sasson, I. 2016. "Trends in Life Expectancy and Lifespan Variation by Educational Attainment: United States, 1990–2010." *Demography* 53, no. 2: 269–293.
- Schmertmann, C. 2023. "Topals Fitting With Age-Grouped Data." May 10. <https://github.com/schmert/>.
- Schonlau, M. 2002. "The Clustergram: A Graph for Visualizing Hierarchical and Non-Hierarchical Cluster Analyses." *Stata Journal: Promoting Communications on Statistics and Stata* 2, no. 4: 391–402.
- Seaman, R., A. H. Leyland, and F. Popham. 2016. "Increasing Inequality in Age of Death at Shared Levels of Life Expectancy: A Comparative Study of Scotland and England and Wales." *SSM-Population Health* 2: 724–731.
- Seaman, R., T. Riffe, and H. Caswell. 2019. "Changing Contribution of Area-Level Deprivation to Total Variance in Age at Death: A Population-Based Decomposition Analysis." *BMJ Open* 9, no. 3: e024952.
- Seaman, R., T. Riffe, A. H. Leyland, F. Popham, and A. van Raalte. 2019. "The Increasing Lifespan Variation Gradient by Area-Level Deprivation: A Decomposition Analysis of Scotland 1981–2011." *Social Science & Medicine (1982)* 230: 147–157.
- Shipton, D., S. Sarica, N. Craig, et al. 2021. "Knowing the Goal: An Inclusive Economy That Can Address the Public Health Challenges of Our Time." *Journal of Epidemiology and Community Health* 75, no. 11: 1129–1132.
- Singleton, A., A. Alexiou, and R. Savani. 2020. "Mapping the Geodemographics of Digital Inequality in Great Britain: An Integration of Machine Learning Into Small Area Estimation." *Computers, Environment and Urban Systems* 82: 101486.
- Singleton, A. D., and S. E. Spielman. 2014. "The Past, Present, and Future of Geodemographic Research in the United States and United Kingdom." *Professional Geographer* 66, no. 4: 558–567.
- Van Raalte, A. A., and H. Caswell. 2013. "Perturbation Analysis of Indices of Lifespan Variability." *Demography* 50, no. 5: 1615–1640.
- Van Raalte, A. A., I. Sasson, and P. Martikainen. 2018. "The Case for Monitoring Life-Span Inequality." *Science* 362, no. 6418: 1002–1004.
- Van Raalte, A. A., and R. J. Seaman. 2019. "The Social Determinants of Life Expectancy and Inequality in Lifespan." In *Encyclopedia of Biomedical Gerontology*, 239–246. Academic Press.
- Vaupel, J. W., and V. C. Romo. 2003. "Decomposing Change in Life Expectancy: A Bouquet of Formulas in Honor of Nathan Keyfitz's 90th Birthday." *Demography* 40: 201–216.
- Woods, L. M., B. Rachet, M. Riga, N. Stone, A. Shah, and M. P. Coleman. 2005. "Geographical Variation in Life Expectancy at Birth in England and Wales Is Largely Explained by Deprivation." *Journal of Epidemiology and Community Health* 59, no. 2: 115–120.

## Supporting Information

Additional supporting information can be found online in the Supporting Information section.  
 Appendices\_REVISED.