Community Needs Index 2023: Technical methodology paper

Oxford Consultants for Social Inclusion

Introduction and context

Analysis of the spatial distribution of 'Left-behind' areas is a growing concern for social policy with a range of studies starting to explore underlying factors contributing to areas being 'Left-behind' – including low levels of social mobility, low skills or declining industries and exclusion from external investment by the public and voluntary sector. In 2018 Local Trust commissioned Oxford Consultants for Social Inclusion (OCSI) to provide a quantitative definition of 'Left-behind' areas – with a Community Needs Index developed to capture the social infrastructure challenges experienced in deprived communities. The Index combined a series of indicators under three domains:

- Civic Assets: Capturing the presence of key community, civic, educational and cultural assets in and in close proximity to the area.
- Connectedness: Capturing connectivity to key services, digital infrastructure, isolation and strength of the local jobs market.
- Active and engaged community: Concerning the levels of third sector activity and volunteering and civic participation, social fabric and barriers to participation and engagement.

The findings of this research were published in 2019 in the paper Left behind? Understanding Communities on the edge. The research suggested that a lack of places to meet (whether community centres, pubs or village halls); the absence of an engaged and active community; and poor connectivity to the wider economy - physical and digital – make a significant difference to social and economic outcomes for deprived communities. Deprived areas which lack these assets have higher rates of unemployment, higher prevalence of ill health and higher levels of child poverty than other deprived areas. And they appear to be falling further behind them. This adds up to these areas being some of the most 'left behind'.

The Community Needs Index 2023 is the second release in a series of statistics produced to measure social infrastructure challenges at the small spatial scale. Following engagement with key stakeholders and consultation with wider users and a significant programme of work by the research team, the Community Needs Index 2023 retain broadly the same domain structures and thematic scope as the previous iteration; however, there has been a review of the underlying indicators, weighting methodology and units of geography used to construct the Index.
The overarching aim of this review is to provide a refined model, which makes use of the most up to date evidence and robust methodologies in order to provide further insights into the challenges and experiences of high needs communities and ensure areas are being identified using the most recent data which reflects the profound social changes experienced over the intervening years since the 2019 Index was released (particularly in the context of the global pandemic and cost of living crisis).

About this Technical Report

This report presents the conceptual framework of the Community Needs Index; the methodology for creating the domains and the overall Index; the component indicators and domains and the decisions taken to inform the methodological approach.

Methods

Overview of the methodology used to construct the Community Needs Index 2023

The construction of the Community Needs Index 2023 broadly consists of the seven following stages. As shown in Figure 1, these stages fulfil the purposes of defining the Indices, data inputs and data processing procedures, and producing the final Community Needs Index. Each stage is described in the following sections.

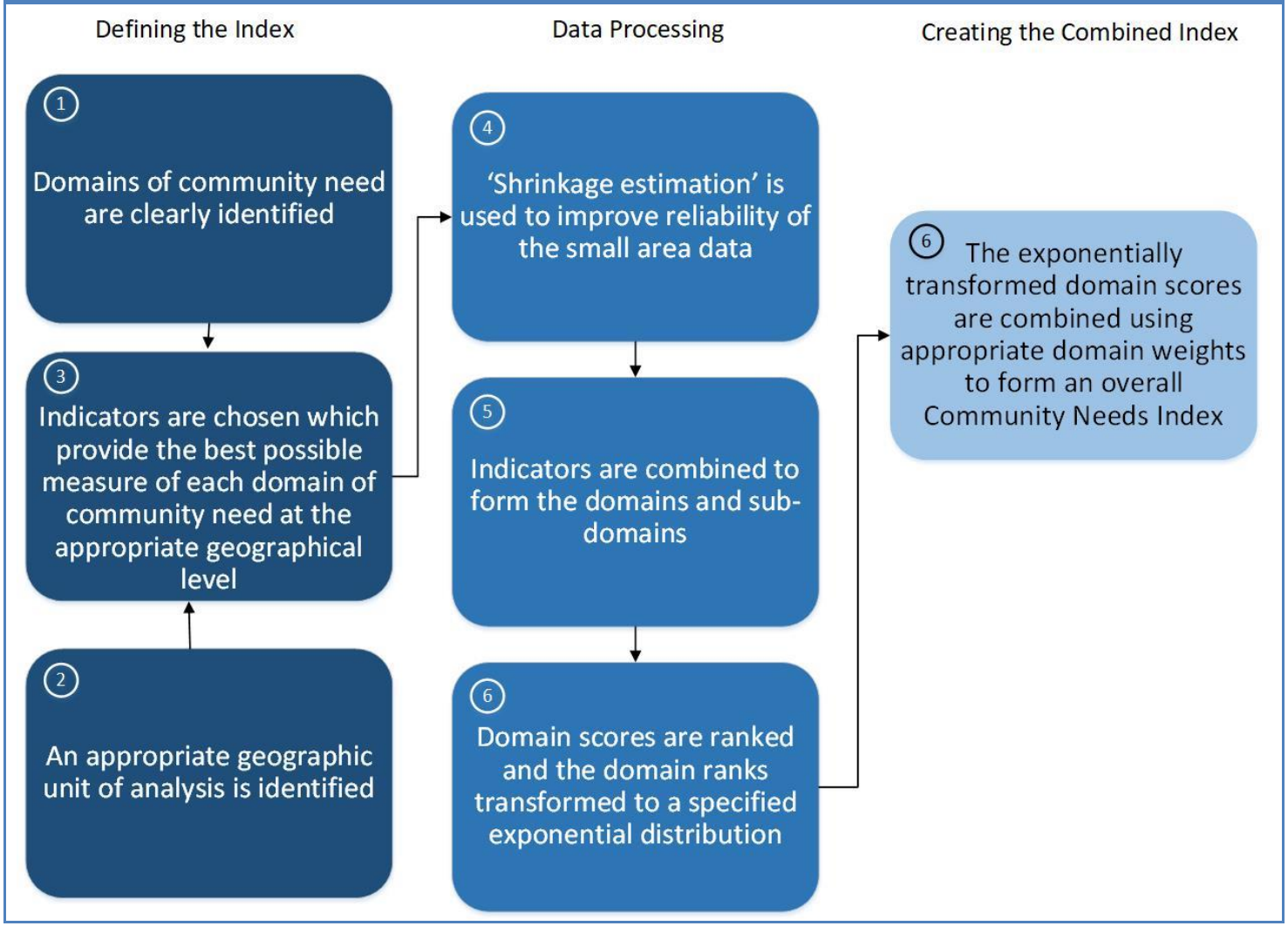
1. Dimensions (referred to as domains) of community need are identified.
2. The unit of geography is selected.
3. Indicators are chosen to provide the best possible measure of each domain of community need at the specified unit of geography.
4. 'Shrinkage estimation' is used to improve reliability of the small area data[1].
5. Indicators are combined to form the domains, generating separate domain scores. These can be regarded as indices in their own right – the domain indices[2].
6. Domain scores are ranked and the domain ranks are transformed to a specified exponential distribution[3].
7. The exponentially transformed domain scores are combined using appropriate domain weights to form an overall Community Needs Index at small area level. This stage completes the construction of the Community Needs Index 2023.

[1] See *Methods for data processing: Applying shrinkage to improve the robustness of indicators* and *Appendix C* for description of the shrinkage technique.

[2] In domains where there are sub-domains, this stage involves first combining the indicators into a sub-domain score. The sub-domain scores are then ranked and transformed to an exponential distribution before being combined into their respective domain scores.

[3] See *Methods for data processing: Step 9 Weighting domains* and *Appendix B* for description of the exponential transformation.

Figure 1. Overview of the methodology used to construct the Community Needs Index 2023



Defining the Index

Identifying dimensions of community need

In order to identify appropriate dimensions of community need, it is important to explore the conceptual framework which the index is intending to measure. The Community Needs Index is a quantitative measure of social infrastructure challenges at the small area level. The model of community need which underpins the Index is the same as that which underpinned its predecessor and is based on the idea of distinct dimensions of community need which can be recognised and measured separately.

Figure 2. Community Needs Index domains 2023

The social infrastructure challenges the index intends to measure concern issues including poor community and civic infrastructure, relative isolation and low levels of participation and engagement in the wider community. The focus on these aspects of community need arose from conversations Local Trust were having in Big Local Areas, where residents were consistently identifying lack of spaces to meet and poor connectivity as key priority issues, while an active and

engaged community was seen as a key necessary ingredient to help affect meaningful change and address challenges and develop the social fabric in deprived communities. Three dimensions of Community Need were identified to capture these challenges (see Figure 2 above).

The approach allows the separate measurement of different dimensions of community need – and these are combined to form the overall measure of community need.

Measuring different aspects of community need and combining these into an overall measure raises a number of questions. Perhaps the most important one is the extent to which area need in one dimension can be *cancelled* out by lack of need in another dimension. Thus, if an area is found to have high levels of civic assets but relatively low levels of participation, should the latter cancel out the former and if so to what extent? The Community Needs Index is essentially based on a weighted cumulative model and the methodology is designed to ensure that cancellation effects are minimised[4].

Another question concerns the extent to which the same people or households are represented in more than one of the dimensions of community need. The position taken in the Community Needs Index 2023 is that if an individual, family or area experiences more than one social infrastructure challenge this is 'worse' than experiencing only one dimension of community need. The aim is not to eliminate double counting *between* domains – indeed it is desirable and appropriate to measure situations where challenges occur on more than one dimension.

Selecting the geographic unit of analysis

This section explores the approach for selecting the geography to use as the building block in the construction of the 2023 Community Needs Index.

In the absence of data at the level of individuals or households, an Index that identifies areas with the greatest intensity of community need should be constructed from data on small geographic units of a standard population size. The Community Needs Index has therefore been developed as an area-based measure. An area can be characterised as deprived *relative to other areas* on a particular dimension of community need, on the basis that a higher proportion of people in the area are experiencing the type of need in question or if the area *as a whole* is lacking in assets and or infrastructure. In other words, both the experience of the people in an area and the areas' amenities and assets gives the area its community needs characteristics.

The selection of the unit of geography to use in an area-based analysis is important as it affects both the data we can draw from, and crucially, the focus areas for intervention and resource allocation which are identified as an outcome of the research.

The following key principles have been considered when selecting the appropriate unit of geography for the Index:

- It should be possible to align the geography units to statistical geography boundaries in order to link key socio-economic indicators to the geography units.
- Geography units should be of sufficient size in order to ensure they are not smaller than the smallest standard statistical geographies (Output Areas), so that it is possible to obtain key socio-economic indicators to be used in the analysis.
- Geography units should be at a neighbourhood (sub-Local Authority) level in order to capture inequalities in social infrastructure provision and participation.
- Geography units should be relatively homogenous in population size so that it is possible to make direct comparisons between communities in terms of their relative needs and community and civic strength.
- Geography units should be meaningful and recognised as areas by the people residing in them.

[4] See Methods for data processing: Step 9 Weighting domains and Appendix B exponential transformation for details of how the Community Needs Index methodology minimises cancellation effects across the domains.

In the 2019 iteration of the Index, 2017 wards were selected as the unit of geography. A challenge with using wards is that ward boundaries are revised on an annual basis – leading to a lack of stability in the Index as old areas become redundant. As an illustration, 2,098 of the 7,445 wards used in the 2019 Community Needs Index have since been subject to boundary changes (28.2%). The change in ward boundaries impacts on both the Community Needs Index and on any frameworks that are combined with this index to identify areas of need.
Wards were initially selected as the unit of geography for a number of reasons:

- Ward boundaries are traditionally centred around established neighbourhoods, rather than cutting across existing neighbourhoods.
- Wards are administrative geographies used for electoral purposes and people are normally aware of the wards they are located in through engagement with the political process e.g. through voting in council elections or through dialogue with local councillors.
- Wards closely align to statistical boundaries and are sufficiently large and homogeneous in size that they can be used in comparative analysis.
- Until relatively recently, wards were the only geography at below neighbourhood level with names attached to them. This was useful for dissemination purposes, enabling us to provide a generally recognised name to each individual neighbourhood identified as 'left-behind'.

However, there are disadvantages with using wards.

- Ward boundaries change on an annual basis: This leads to constant changes in the list of 'left-behind' or 'at-risk' neighbourhoods as the research is updated. It also risks the names of the neighbourhoods becoming redundant in local communities as the ward names change.
- There are considerable variations in population size of wards in different parts of the country – the smallest ward in 2017 had a population of 971[5] while the largest ward had a population of 46,566. By using a unit of geography that varies considerably in size, there is a risk that in more metropolitan areas (where wards are generally larger) that the wards cover more than one neighbourhood. This increases the likelihood that a mixture of neighbourhoods with different contexts and needs are grouped together, masking pockets of need within the ward and increasing the likelihood that neighbourhoods with high need are overlooked.
- There are fewer datasets available at ward level: Because data is rarely published a ward level, additional steps are required to convert data to ward geographies – In the 2019 Community Needs Index, a best-fit Output Area to ward lookup table was used to achieve this.

There was general consensus in the consultation that we should move away from wards and that the Community Needs Index (CNI) should be constructed at the smallest practicable spatial scale and that the ideal geography should possess relatively even sized populations, with Lower-layer Super Output Areas (LSOAs) favoured by 58.3% of respondents.

LSOAs were identified as the preferred unit of measure for a number of reasons:

- They only change after every census, so they are more consistent over time. Even when changes are made following census updates, these changes are capped, with a minimum of 95% of boundaries remaining unchanged. They therefore represent a more stable geography than wards.
- LSOAs are designed to be fairly homogenous in size (averaging at 1,700 people).
- They also nest directly with smaller statistical geographies such as Output Areas without requiring a best-fit lookup.

[5] Excluding the micro wards in Isles of Scilly and City of London

- To facilitate links with the Indices of Deprivation which is produced at LSOA level.
- To support the identification of rural pockets of need, which tend to be revealed at smaller geographical scales. By contrast, single MSOAs and wards can encompass multiple villages with very different characteristics.
- Finally, LSOAs are a more commonly used geography and are increasingly used to disseminate key statistics releases[6]. Producing the Community Needs Index at LSOA level would therefore enable users to benchmark the performance of high need areas against a wider range of socio-economic measures. This should increase the utility and analytical value of the indicator and ultimately raise the profile of the work.

However, there are challenges with reverting from ward to LSOA geography in terms of backwards comparability. One consequence of changing the unit of geography is that it is difficult to compare results with the previous Community Needs Index.

Moreover, there are some indicators used in the Index where data is not available, robust or suitable at LSOA level.

In recognition of these challenges, we have adopted a hybrid solution. The Community Needs Index 2023 has been produced at 2021 LSOA level - however, where data for underlying indicators is not available or robust at 2021 LSOA level, we have used MSOA data to construct the indicator and all of the component LSOAs within the MSOA receive the same value. A similar approach has been taken where it makes conceptual sense to use MSOA geographies. For example, in the Civic Assets domain we have retained the buffer zone MSOA geography to represent the wider catchment around these assets, as in most cases, LSOAs are too granular to represent a catchment area around an asset. The process of converting data to LSOA geographies is described in more detail in the section: *Methods for data processing: Step 1 Convert all indicators to LSOA geography* (below).

Identifying indicators

For each of the three domains of community need, an assessment has been made about whether the indicators in the Community Needs Index 2019:

- are still appropriate measures of community need for that domain.
- can be updated.
- can be strengthened, for example due to better available data.

It was necessary to review the indicators in the Community Needs Index in light of the availability of a wider range of datasets in 2023 compared with 2018 (when the Index was first constructed). In approaching a review of potential indicators we have worked on the basis that the indicator set identified in the 2019 Community Needs Index is a baseline starting point and indicators should be retained and brought up to date where this is possible in order to aid backwards comparability. Indicators were only replaced where more robust or up to date indicators could be found from alternative sources. However, we also explored the inclusion of additional indicators where they added strength, thematic depth (capture a different facet of community need) or brought together more robust or up to date information on existing aspects of community need.

All new indicators have to meet the same criteria as for the Community Needs Index 2019. Indicators should:

- be 'domain specific' and appropriate for the purpose (as far as possible, being direct measures of that form of community need or social infrastructure challenge.
- measure major features of that domain (not conditions just experienced by a small number of people or areas).
- be up-to-date and (as far as possible) updateable[7].

[6] See for example, the daily COVID-19 caseload data https://coronavirus.data.gov.uk/

[7] Wherever possible, indicators are used that can be regularly updated. However not all indicators can be regularly updated, for example those based on Census 2021. Census data is used only when alternative data from administrative sources is not available.

- be statistically robust at the small area level.
- be available for the whole of England at a small area level in a consistent form.
- be available at sufficient granularity in order to make meaningful comparisons between areas.
- are non-disclosive and open.
- are sufficiently statistically robust to be included in a measure intended for use in resource allocation.

This section outlines the key indicators we have included in the 2023 Community Needs Index – highlighting key changes from the 2019 iteration of the model. **New indicators are presented in bold**.

The table below provides an overview of each of these indicators with metadata detailing:

- Source (included URL).
- Timepoints the data is available for.
- Geographical unit at which the data is published.
- Notes/Caveats associated with the indicator – including robustness issues to consider when incorporating the data.
- Details of change from the 2019 Community Needs Index.

Indicator	Details	Source	Date	Granularity	Notes/Caveats	Changes in 2023
Civic Assets						
CA1: Density of community space assets	This is conceptualised as the number of community and civic assets inside the local area or within 1km of the local area boundary, divided by the number of people living inside the local area or within 1km of the local area boundary. Rate is expressed per 100,000 population. The following assets are included: • Public / Village Hall / Other Community Facility • Youth Recreational / Social Club • Church Hall / Religious Meeting Place / Hall • Community Service Centre / Office • Place Of Worship	AddressBase https://www.ordnancesurvey.co.uk/business-government/products/addressbase	Feb 2023	Point Location	Details are not available on how accessible the assets are to the community.	Updated for 2023
CA2: Density of educational assets	This is conceptualised as the number of community and civic assets inside the local area or within 1km of the local area boundary, divided by the number of people living inside the local area or within 1km of the local area boundary. Rate is expressed per 100,000 population aged 20 and under. The following assets are included: • College • Further Education • Higher Education • Children's Nursery / Crèche • First School • Infant School • Junior School • Middle School • Primary School • Secondary School • University • Special Needs Establishment. • Other Educational Establishment	AddressBase https://www.ordnancesurvey.co.uk/business-government/products/addressbase	Feb 2023	Point Location	Details are not available on how accessible the assets are to the community.	Updated for 2023. Denominator changed to cover the population aged under 20 (in recognition that this age group are more likely to access these assets).
CA3a: Density of sport and leisure assets (address base)	This is conceptualised as the number of community and civic assets inside the local area or within 1km of the local area boundary, divided by the number of people living inside the local area or within 1km of the local area boundary. Rate is expressed per 100,000 population. The following assets are included: • Public House / Bar / Nightclub • Activity / Leisure / Sports Centre • Skateboarding Facility • Recreational / Social Club(Bingo) •Leisure Pier •Swimming facility •Public tennis court •Bandstand • Restaurant / Cafeteria	AddressBase https://www.ordnancesurvey.co.uk/business-government/products/addressbase	Feb 2023	Point Location	Details are not available on how accessible the assets are to the community. Some of the facilities identified will have a cost associated with access, which could potentially exclude those on lower incomes in the community.	Updated for 2023 and additional assets included: Leisure Pier, Swimming facility, Public tennis court, Bandstand and Restaurant / Cafeteria.
CA3b: Density of sport and leisure assets (Active places database)	**Active places database is compiled by Sport England and contains a record of all of the sports facilities in an area (grouped by type). The following facilities are included: Athletics Tracks, Health and Fitness Suite, Indoor Bowls, Indoor Tennis Centre, Grass Pitches, Sports Hall, Swimming Pool, Artificial Grass Pitch, Golf, Ice Rinks, Ski Slopes, Studio, Squash Courts, Tennis Courts, Cycling.** **An asset is included if it is located inside the local area or within 1km of the local area boundary and the figure is expressed as a rate per 100,000 people living inside the local area or within 1km of the local area boundary.**	**Active places database https://www.activeplacespower.com/OpenData/download**	**Feb 2023**	**Point Location**	**Details are not available on how accessible the assets are to the community. Some of the facilities identified will have a cost associated with access, which could potentially exclude those on lower incomes in the community.**	**New**

Indicator	Details	Source	Date	Granularity	Notes/Caveats	Changes in 2023
CA4: Density of cultural assets	This is conceptualised as the number of community and civic assets inside or within 1km of the local area boundary divided by the number of people living inside or within 1km of the local area boundary. Rate is expressed per 100,000 population. The following assets are included: • Library • Reading Room • Museum/Gallery •Historical Site / Object •Historic Structure / Object •Monument Obelisk / Milestone / Standing Stone Statue •Castle / Historic Ruin • Permanent Art Display / Sculpture	AddressBase https://www.ordnancesurvey.co.uk/business-government/products/addressbase	Feb 2023	Point Location	Details are not available on how accessible the assets are to the community. Some of the museums will not be free to enter, which will exclude some sections of the community. Some of the libraries and reading rooms will not have open access.	Updated for 2023 and additional assets included: Historical Site / Object, Historic Structure / Object, Monument Obelisk / Milestone / Standing Stone Statue, Castle / Historic Ruin and Permanent Art Display / Sculpture
CA5a: Green assets (density)	This is conceptualised as the number of community and civic assets inside or within 1km it divided by the number of people living inside or within 1km of the local area boundary. Rate is expressed per 100,000 population. The following assets are included: • Public Park / Garden • Public Open Space / Nature Reserve • Playground • Play Area• Paddling Pool• Picnic / Barbeque Site• Allotment• Playing Field • Recreation Ground •Woodland • Lake / Reservoir • Forest / Pinetum	AddressBase https://www.ordnancesurvey.co.uk/business-government/products/addressbase	Feb 2023	Point Location	Details are not available on how accessible the assets are to the community. Some assets are not open- access to the whole community, e.g. allotments and some of the play areas/paddling pools. It is not possible to distinguish between these (though private parkland has been excluded). There is no information regarding the size or quality of the green space.	Updated for 2023. Additional assets included: Woodland, Lake / Reservoir and Forest / Pinetum.
CA5b: Green assets (Area of public green space)	**Area of public green space. This includes cemeteries, playing fields, public parks and gardens, religious grounds, plus Countryside Right of Way open access land. Based on Ordnance Survey Open Greenspace Map and Natural England CRoW Act 2000 - Open Access Mapping.**	**Friends of the Earth**	**Sep 2020**	**MSOA**	**Areas smaller than 2 hectares have been omitted.**	**New**
CA6: Retail assets	**Number of retail premises in the local area or within 1km of the local area boundary divided by the number of people living inside or within 1km of the local area boundary. The rate is expressed per 100,000 population. The following assets are included: Post Office, Market, Shop / Showroom and Garden Centre.**	**AddressBase https://www.ordnancesurvey.co.uk/business-government/products/addressbase**	**Feb 2023**	**Postcode**	**Does not take into account the size of the retail unit or how accessible it is to the local community. Excludes assets with negative community benefit.**	**New**
CA7: Community-owned assets	**Community owned assets divided by the number of people living inside or within 1km of the local area boundary. The rate is expressed per 100,000 population. Figures are compiled using data from Power to Change, the Community Land Trust Network, Co-operatives UK, Plunkett Foundation and Locality and Keep it in the Community.**	**Renaisi/ Plunkett Foundation/Locality**	**2022**	**Postcode**	**Some assets are geolocated based on the location of the organisation owning the assets rather than the assets themselves, and some postcodes containing multiple assets are listed as single assets in the database.**	**New**
Connectedness						
Physical connectivity (subdomain)						

Indicator	Details	Source	Date	Granularity	Notes/Caveats	Changes in 2023
CN1a: Travel time to key services by public transport/walk	Travel times in minutes to key services by public transport/walking and cycling. The following services are included: • Primary School • Employment centre (LSOA with more than 500 jobs) • Further Education Institution • GP • Hospital • Secondary School • Town Centre These statistics are derived from the analysis of spatial data on public transport timetables; road, cycle and footpath networks; population and key local services.	Department for Transport (DfT) https://www.gov.uk/government/collections/journey-time-statistics	2019	LSOA	Although the statistics are calculated to a high level of geographical detail, some assumptions and simplifications are necessary in the modelling (for example assigning the start point of journeys to a single point in each Output Area, road speeds, interchange times for public transport).	Updated for 2019 and supermarket removed due to large number of areas with equal value for this indicator.
CN1b: Access to Green/Blue Spaces	**Access to Green/Blue Spaces (both active and passive) is measured as mean road distance to these facilities (in km). Blue spaces are water features that can be positive amenities. Blue space indicator is based on the distance people need to travel to access their nearest water body such as a beach, a lake and a river. Blue space locations such as beaches were acquired from OpenStreetMap and the mainland water bodies (lakes, rivers) were retrieved from the European Settlement Map (ESM 2012) raster dataset at a 5 meters resolution.** **Active green spaces refer to recreational opportunities involving moderate to high intensity use requiring modification of natural landforms and the provision of service facilities, playing fields or equipment.** **Passive green space refers to recreational opportunities that occur in a natural setting requiring minimal development or facilities and providing areas for informal, self-directed activities for individuals or small groups.** **Open data from OS on Green spaces was used for preparing two variables related to the distance from the nearest green space (active) and the total green space areas available to each postcode in a range of a 900-meter buffer** **(passive) before creating LSOA level averages.**	**CDRC – Access to Health Assets and Hazards https://data.cdrc.ac.uk/dataset/ahah2**	**2022**	**LSOA**		**New**

Indicator	Details	Source	Date	Granularity	Notes/Caveats	Changes in 2023
CN2: Job access score	**This measure of connectivity developed by UK Onward includes the number of jobs accessible by car and public transport from every local area (LSOA) in the country across different time horizons. It incorporates TravelTime API, and the metric provides the reachable number of jobs and distance with 15 minutes, 30 minutes, 60 minutes and 90 minutes by both driving and public transport across Great Britain for each LSOA (in England and Wales) or Data Zone (in Scotland). The data incorporates a 'door-to-workplace' measure, including every journey stage from time spent walking to the car, driving, to parking and walking to an office - as well as average delays, timetabling and actual journey time on public transport. These measures have been combined into an overall Jobs access score, the weighted average job count, combining driving and public transport. A higher score indicates greater levels of job accessibility. For more information and a link to the research paper please see here: https://www.ukonward.com/reports/network-effects/**	**UK Onward (https://www.ukonward.com/reports/network-effects/)**	**2021**	**LSOA**		**New (replacement for Jobs Density measure)**
Wider connectivity (subdomain)						
CN3: Households with no car	The proportion of households who do not have a car or van. Figures are based on responses to the 2021 Census car ownership question, which asks for information on the number of cars or vans owned or available for use by one or more members of a household. It includes company cars and vans available for private use. This is included to supplement the accessibility of key services and labour market indicators in this domain, to take account of the additional challenges in accessing services for those without access to private transport.	Census 2021	2021	Output Area	The count of cars or vans in an area is based on details for private households only. Cars or vans used by residents of communal establishments are not counted.	Updated for 2021
CN4: Digital Exclusion Risk Index (DERI) score	**The Digital Exclusion Risk Index (DERI) is a dataset that explores the risk of digital exclusion at a localised geographic level. The score is based on three components to convey the breadth of the issue of digital exclusion; covering age, broadband access, and deprivation. It uses a number of indicators, which are normalised, weighted and summed, to create a DERI score: a number that identifies the risk of digital exclusion in an area. The score is calculated on a LSOA geography, where a higher score indicates a higher level of digital exclusion risk.**	**Greater Manchester Office of Data Analytics**	**2020**	**LSOA**		**New (replacement for broadband data)**
CN5: Loneliness (People living alone)	Shows the proportion of households that comprise one person living alone (as a proportion of all households). Figures are self-reported and taken from the household composition questions in the 2021 census.	Census 2021	2021	Output Area	This is included as a proxy measure of social isolation.	Updated for 2021

Indicator	Details	Source	Date	Granularity	Notes/Caveats	Changes in 2023
CN4b: Loneliness (Self-reported levels of loneliness)	**People who have self-reported that they 'feel lonely always or often' in the 2015/16 and 2016/17 Community Life Survey. Data is apportioned from national level to Output Area level based on Output Area Classification group.**	**Community Life Survey: DCMS/Output Area Classification 2011**	**2016 to 2021**	**Output Area**	**Data are constructed from a survey with a small sample size. Data has been apportioned down to Output Area level using Output Area Classification group membership – (which groups together Output Areas based on their shared socio-economic characteristics). Caution should be applied when interpreting these results at small-area level because of the small sample size of the survey. Two years of data were used to increase the size of the response rate.**	**New**
CN5: GP access score	**The GP access is a new measure introduced to capture connectivity in terms of the ability to access services based on service availability rather than location of services.** **The following measures are included:** **1) Patients registered at GP Surgery as ratio of all FTE GPs** **2) GP appointments 21 or more days after appointment is booked (as % of all appointments)** **3) GP appointments per 1,000 patients** **4) Same day GP appointments (as % of all appointments)** **NHS England provides data on the number of patients registered at the practice and the number Full-Time Equivalent (FTE) GPs. From NHS England it is also possible to identify the LSOA location of all patients registered at each GP. Using that information, it is possible to attribute the range of different GP access scores to each LSOA in England.**	**NHS England**	**2022**	**LSOA**		**New**
Active and engaged community						

Indicator	Details	Source	Date	Granularity	Notes/Caveats	Changes in 2023
AE1: Voter turnout at local elections	Valid voter turnout (%) at the most recent Local Council Elections.	Electoral Commission https://www.electoralcommission.org.uk/who-we-are-and-what-we-do/elections-and-referendums/past-elections-and-referendums/european-parliamentary-elections/report-may-2019-european-parliamentary-elections-and-local-elections	2016/2017/2018/2019/2021/2022	Ward	There is some local variation in the frequency and dates of local elections, with different parts of the country going to the polls at different times and at different intervals. Caution is therefore advised when drawing direct comparisons between local areas, as the socio-political context and weather conditions vary from year to year with associated impacts on turnout rates. Another factor affecting turnout is whether the local election is concurrent with other elections (for example, turnout is generally higher when general elections coincide with local ones. We have suggested steps to mitigate against this by adjusting estimates from previous years to the 2019 average turnout.	Updated for 2022.

Indicator	Details	Source	Date	Granularity	Notes/Caveats	Changes in 2023
AE2: Civic participation (Self-reported measures of community and civic participation)	**The Community Life Survey contains key indicators of volunteering and civic participation.** **The recent iterations of the Community Life Survey are published with the associated Output Area Classification of each respondent in the survey. Using the Output Area Classification it is possible to apportion response rates to Output Area level, allocating response rates (%) to each Output Area based on their Output Area Classification group membership. Data is then aggregated from Output Area to provide estimated rates for key indicators for MSOAs. The following indicators are included:** • **People have not taken part in a consultation about local services or issues in their local area.** • **People have not taken part in community groups clubs or organisations e.g. children's education/schools, youth/children's activities, education for adults, Sport/exercise (taking part, coaching or going to watch), religion, politics, health, disability and social welfare, older people, safety, first aid, the environment, animals, justice and human rights, local community or neighbourhood groups, citizens groups, hobbies, recreation/arts/social clubs.** • **People have not taken part in any civic engagement.** • **People have not been engaged in formal or informal volunteering in the last month.** • **People definitely or tend to disagree that they can influence decisions in their local area.**	**Community Life Survey: DCMS/Output Area Classification 2011: ONS** **Licensed data – access via UK data archive https://www.data-archive.ac.uk/**	**2017 to 2021**	**Output Area**	**Data are constructed from a survey with a small sample size. Data has been apportioned down to Output Area level using the Output Area Classification group membership – (which groups together Output Areas based on their shared socio-economic characteristics). Caution should be applied when interpreting these results at small-area level because of the small sample size of the survey. Six years of data were used to increase the size of the response rate.**	**New**

Indicator	Details	Source	Date	Granularity	Notes/Caveats	Changes in 2023
AE3a Neighbourhood cohesion	**The Community Life Survey contains key indicators of Neighbourhood cohesion** **The 2015/16 and 2017/18 iterations of the Community Life Survey are published with the associated Output Area Classification of each respondent in the survey. Using the Output Area Classification it is possible to apportion response rates to Output Area level, allocating response rates (%) to each Output Area based on their Output Area Classification group membership. Data is then aggregated from Output Area to provide estimated rates for key indicators for MSOAs. The following indicators are included:** • **People do not feel that they belong very strongly to neighbourhood.** • **People disagree that they can borrow things or exchange favours with neighbours.** • **People never chat to their neighbours.** • **People feel fairly or very uncomfortable with asking a neighbour to keep a set of keys to their home for emergencies.** • **People feel fairly or very uncomfortable with asking a neighbour to collect a few shopping essentials if they were ill and at home on their own.** • **People disagree that people in this neighbourhood pull together to improve the neighbourhood.**	**Community Life Survey: DCMS/Output Area Classification 2011: ONS** **Licensed data – access via UK data archive https://www.data-archive.ac.uk/**	**2017 to 2021**	**Output Area**	**Data are constructed from a survey with a small sample size. Data has been apportioned down to Output Area level using the Output Area Classification group membership – (which groups together Output Areas based on their shared socio-economic characteristics). Caution should be applied when interpreting these results at small-area level because of the small sample size of the survey. Six years of data were used to increase the size of the response rate.**	**New, replaces Strength of local social relationships data.**

Indicator	Details	Source	Date	Granularity	Notes/Caveats	Changes in 2023
AE3b Estimates of social trust	**FocalData surveyed 42,696 adults across England. Respondents were asked how much they trust other people, using a standard question from the British Social Attitudes survey: "Generally speaking, would you say that people can be trusted or that you can't be too careful in dealing with people?" Respondents were asked to say whether they think that people can almost always or usually be trusted or that you usually or almost always can't be too careful in dealing with people.** **FocalData then ran a multilevel regression and post-stratification (MRP) to estimate average levels of trust in each neighbourhood, which we defined as a Middle-Layer Super Output Area (MSOA). The model included individual and area characteristics, including: overcrowded housing, ethnicity, health, qualification levels, housing tenure, social grade, age and gender.**	**UK Onward**	**2022**	**MSOA**	**These are not direct estimates. Data is modelled from sample survey data using a multilevel regression and post-stratification (MRP) to estimate average levels of trust in each neighbourhood. Variation in levels of social trust at MSOA level are influenced by the areas socio-demographic characteristics.**	**New**
AE3c CDRC Residential Mobility Index	**This index provides an estimate of the "churn" of the residential population in the UK - the proportion of households that have changed between the end of 2020 and the end of each of each year going back to 1997. The estimates were built from linking administrative and consumer data, including electoral registers, consumer registers and land registry house sale data. This data enables researchers to explore annual variations in neighbourhood change at a small area geography. Crucially it also enables researchers to focus on yearly data rather than relying on decadal census data to estimate change. It is even possible to observe trends that have occurred since the collection of the most recent census.**	**Consumer Data Research Centre https://data.cdrc.ac.uk/dataset/cdrc-residential-mobility-index**	**1999-2020**	**LSOA**	**This indicator is included as a measure of neighbourhood stability or cohesion**	**New**
AE3d Short term population turnover	**Proportion of people whose address at one year before the Census was different to that on Census Day. The information comes from responses to information on the usual address of a resident and the address one year ago.**	**Census 2021**	**2021**	**OA**	**This indicator is included as a measure of neighbourhood stability or cohesion. The previous residential address for children aged under one in households is determined by the status of their next of kin (defined as in order of preference, mother, father, sibling (with nearest age), other related person, Household Reference Person).**	**New**

Indicator	Details	Source	Date	Granularity	Notes/Caveats	Changes in 2023
AE5: Third sector organisations	**Non-overlapping count of** **1) Registered charities from Charity Base.** **2) Co-operative societies from Co-operatives UK.** **3) Charitable Incorporated Organisations, Community Interest Companies, PRI/LTD BY GUAR/NSC (Private, limited by guarantee, no share capital – excluding Property Management companies) and Registered Societies from Companies House .** **4) Co-operative societies, community benefit societies, and former industrial and provident societies from Financial Conduct Authority.** **Figure is expressed as a rate per 100,000 population.**	**Charities Commission https://charitybase.uk/chc . Co-operatives UK https://www.uk.coop/uk , Companies House http://download.companieshouse.gov.uk/en_output.html , from Financial Conduct Authority https://mutuals.fca.org.uk/**	**Feb 2023**	**Postcode**	**This is based on the location of organisations rather than on their area of operations (some will have a global focus). Larger charities ae excluded from this measure. This indicator is included in this theme to capture the level of third sector activity in the local area. Organisations with an exclusively national or international focus have been excluded, to ensure only organisations with a local focus are included. Some organisations appear on multiple registers – duplicate records have been stripped so only unique records remain. This will exclude smaller companies not registered and exclude co-operatives, community benefit societies, associations, trusts and partnerships (of varying types).**	**New**
AE6: National Lottery Community Fund	Combined total of grants made to local projects and organisations by the National Lottery Community Fund between 2004 and 2022 per 1,000 population (£). Figures are taken from data on grants made to projects and organisations in local areas in the UK by the Big Lottery Fund, from grants data published by Big Lottery in conjunction with the 360Giving initiative. Big Lottery used the 360Giving standard to produce a file of all the grants made in 2004-2022.	National Lottery (through 360 Giving) https://grantnav.threesixtygiving.org/	2004-2023	Ward level	Included in the active/engaged community theme to capture the level of third sector activity in the local area.	Updated for 2023

AE7: Grant funding per head from major grant funders	Combined grant funding from over 150 grant giving organisations whose data has been subject to the 360giving standard (per head of population). The following organisations are included: A B Charitable Trust, Access to Justice Foundation, Alan & Babette Sainsbury Charitable Fund, Alex Ferry Foundation, Andrew Lloyd Webber Foundation, ARCADIA, Architectural Heritage Fund, Barrow Cadbury Trust, BBC Children in Need, Brian Mercer Trust, Cabinet Office, CAF, Calouste Gulbenkian Foundation, UK Branch, CareTech foundation, Charles Hayward Foundation, CHK Foundation, Cloudesley, Comic Relief, Coop Foundation, Co-operative Group, Crisis UK, Crowdfunder, Culham St Gabriel's Trust, Daiwa Anglo-Japanese Foundation, Department for Business, Energy and Industrial Strategy, Department for Culture, Media and Sport, Department for Digital, Culture, Media & Sport, Department for Digital, Culture, Media and Sport, Department for Education, Department for Environment, Food and Rural Affairs, Department for International Development, Department for International Trade, Department for Levelling Up, Housing and Communities, Department for Transport, Department for Work and Pensions, Department of Health, Department of Health and Social Care, Disability Action, Esmée Fairbairn Foundation, Fenton Arts Trust, Foreign and Commonwealth Office, Foreign, Commonwealth & Development Office, Friends Provident Foundation, Garfield Weston Foundation, Gatsby Charitable Foundation, GMSP Foundation, Greenwood Place, Guy's and St Thomas' Charity, Hazelhurst Trust, HM Revenue & Customs, Home Office, Imperial Health Charity, Indigo Trust, J J Charitable Trust, Joffe Charitable Trust, John Ellerman Foundation, Joseph Levy Foundation, Joseph Rowntree Charitable Trust, Joseph Rowntree Foundation, Joseph Rowntree Housing Trust, Joseph Rowntree Reform Trust , JRSST-CT, Justice Together Initiative, KPMG Foundation, LandAid Charitable Trust, Lankelly Chase Foundation, LGBT Consortium, Lloyds Bank Foundation for England and Wales, Lloyd's Register Foundation, Macc, Mark Leonard Trust, Masonic Charitable Foundation, Maudsley Charity, Mercers' Charitable Foundation, Mercers' School Memorial Trust (Incorporating the Merrett Bequest), Millfield House Foundation, Ministry for Housing, Communities and Local Government, Ministry of Defence, Ministry of Housing, Communities & Local Government, Ministry of Justice, Mission 44, National Churches Trust, National Emergencies Trust, Nationwide Foundation, Nesta, Northern Rock Foundation, Nuffield Foundation, OVO Foundation, Paul Hamlyn Foundation, Pears Foundation, People's Health Trust, Power to Change Trust, Rank	360 Giving Grant Nav data https://grantnav.threesixtygiving.org/	2023	Postcode level[8]	Data are based on the location of grant recipients rather than the location of their beneficiaries. This indicator is included in this theme to capture the level of third-sector activity in the local area. Grants above £1m excluded to ensure capturing local initiatives rather than national activity. Measure expanded to include all grant funders which have a nationwide focus (e.g. not focused in one region of the country only[9]) where geographic information supplied.	Updated for 2023. Measure expanded to include all Grant Funders which have a nationwide focus (e.g. not focused in one region of the country only[10]) where geographic information supplied.

Indicator	Details	Source	Date	Granularity	Notes/Caveats	Changes in 2023
	Foundation, Sam and Bella Sebba Charitable Foundation, Samworth Foundation, Sir Jules Thorn Charitable Trust, Spirit of 2012, St Paul's Schools Foundation, Staples Trust, Tedworth Charitable Trust, The AIM Foundation, The Archbishops' Council , The Badur Foundation, The Baring Foundation, The Bell Foundation, The Berkeley Foundation, The Bishop Radford Trust, The Blagrave Trust, The Childhood Trust, The Clothworkers Foundation, The David & Elaine Potter Foundation, The Dulverton Trust, The Dunhill Medical Trust, The EQ Foundation, The Fore, The Foyle Foundation, The Funding Network, The Greggs Foundation, The Grocers' Charity, The Henry Smith Charity , The Joseph Rank Trust, The Leathersellers' Company Charitable Fund, The Legal Education Foundation, The Michael And Betty Little Trust, The MSE Charity, The Pilgrim Trust, The Postlethwaite Music Foundation, The Rayne Foundation, The Road Safety Trust, The Seafarers' Charity, The Segelman Trust, The Triangle Trust 1949 Fund, The Trussell Trust, The Tudor Trust, The William Syson Foundation, Thirty Percy Foundation, Three Guineas Trust, True Colours Trust, Trusthouse Charitable Foundation, Tuixen Foundation, UK Community Foundations, Unbound Philanthropy, Virgin Money Foundation, Vision Foundation, Wates Family Enterprise Trust, Wates Foundation, William Grant Foundation, Wolfson Foundation, Woodward Charitable Trust, Youth Music, ZING,					
AE8a: SME lending by banks	Total value of lending to SME businesses from key financial lenders (Barclays, CYBG, Lloyds Banking Group, HSBC, Nationwide Building Society, Royal Bank of Scotland and Santander UK in Great Britain).	UK Finance	Novemb er 2021	Postcode sector	The measure covers four quarters of lending data at postcode sector level. The data is modelled from postcode sector to Output Areas using a weighted lookup built from the numbers of shared postcodes between a postcode sector and Output Area in combination with the working age population per Output Area. Data is then aggregated to local area level to get total value of SME lending at local area level.	Updated for November 2021.

[8] Comic relief and Children in Need only supply references of Local Authority recipients. This data will be allocated to local area by apportioning

[9] Grant givers with a specific area focus e.g. Community Foundations have been excluded to mitigate against reflecting the extent to which local grant givers have submitted data to GrantNav e.g. Not all Community Foundations have submitted data to Grant Nav and we want to guard against introducing systematic bias into the data by including data for some regions and excluding others.

[10] Grant givers with a specific area focus e.g. Community Foundations have been excluded to mitigate against reflecting the extent to which local grant givers have submitted data to GrantNav e.g. Not all Community Foundations have submitted data to Grant Nav and we want to guard against introducing systematic bias into the data by including data for some regions and excluding others.

Indicator	Details	Source	Date	Granularity	Notes/Caveats	Changes in 2023
AE8b: Small businesses: Local Business Units with 0-4 employees	**Small businesses: VAT registered local businesses with 0-4 employees per 10,000 population**	**Inter Departmental Business Register (IDBR)**	**2022**	**MSOA**		**New**

Methods for data processing

This section outlines the proposed step by step approach to developing the 2023 Community Needs Index.

Step 1 Convert all indicators to LSOA geography

The 2023 Community Needs Index has been produced at Lower layer Super Output Area (LSOA) geography (using the updated 2021 version of the LSOA boundaries which were developed as part of the Census 2021 outputs). However, the majority of the indicators in the Index are not published at 2021 LSOA level. It is therefore necessary to convert these indicators to 2021 LSOA geographies.

The table below outlines our approach to converting indicators to 2021 LSOA level:

Geography	Indicators	Approach to conversion
Postcode/Point Location	Grant funding per head from major grant funders Third sector organisations	Use GIS software to overlay point data against 2021 LSOA polygon boundaries and using a point-in-polygon method to calculate the number of points that fall within each LSOA.
Polygon	Green assets: b) Area of public green space	Overlay Output Area boundaries and Address Base points against polygon boundaries. If the majority of residential addresses within an Output Area fall within the polygon, the Output Area has been identified as part of the polygon. A best-fit approach is taken with no splitting across multiple polygons or apportioning in/out of a polygon and the entire OA is included in a single LSOA. The Output Area 2021 to LSOA 2021 Lookup table developed by the ONS has been used to aggregate data from Output Area to LSOA level.
2011 Output Area	Self-reported measures of community and civic participation Neighbourhood cohesion Self-reported levels of loneliness	Overlay the 2011 and 2021 Output Area boundaries with individual residential postcode data from the ONS Postcode directory. Apply a point-in-polygon method to calculate the number of postcodes that fall within each 2011 and 2021 Output Area boundaries. Use this to weight the extent of overlap between 2011 and 2021 Output Areas. Apply this weighting to generate 2021 Output Area estimates. Aggregate from 2021 Output Area to 2021 LSOA using the Output Area to LSOA level lookup table from the ONS Census 2021 geography products.
2021 Output Area	Households with no car People living alone Short term population turnover	Use the Output Area to LSOA level lookup table from the ONS Census 2021 geography products to aggregate to LSOA.

2011 LSOA	Travel time to key services by public transport/walk Access to green and blue spaces Job access score Digital Exclusion Risk Index (DERI) Number of GPs per 1,000 patients CDRC Residential Mobility Index	Use the 2011 Output Area to 2011 LSOA Look-up table to apportion data to 2011 Output Area. Overlay the 2011 and 2021 Output Area boundaries with individual residential postcode data from the ONS Postcode directory. Apply a point-in-polygon method to calculate the number of postcodes that fall within each 2011 and 2021 Output Area boundaries. Use this to weight the extent of overlap between 2011 and 2021 Output Areas. Apply this weighting to generate 2021 Output Area estimates. Aggregate from 2021 Output Area to 2021 LSOA using the Output Area to LSOA level lookup table from the ONS Census 2021 geography products.
MSOA	Leisure and cultural participation Small businesses: VAT registered local businesses with 0-4 employees per 10,000 population Estimates of social trust	Use the 2011 2011 Output Area to 2011 MSOA Look-up table to apportion data to 2011 Output Area. Overlay the 2011 and 2021 Output Area boundaries with individual residential postcode data from the ONS Postcode directory. Apply a point-in-polygon method to calculate the number of postcodes that fall within each 2011 and 2021 Output Area boundaries. Use this to weight the extent of overlap between 2011 and 2021 Output Areas. Apply this weighting to generate 2021 Output Area estimates. Aggregate from 2021 Output Area to 2021 LSOA using the Output Area to LSOA level lookup table from the ONS Census 2021 geography products.
Electoral ward	Voter turnout at local elections Big Lottery funding per head	Apportion data from relevant ward to Output Area (using ONS open geography portal Output Area to ward lookup tables). Overlay the 2011 and 2021 Output Area boundaries with individual residential postcode data from the ONS Postcode directory. Apply a point-in-polygon method to calculate the number of postcodes that fall within each 2011 and 2021 Output Area boundaries. Use this to weight the extent of overlap between 2011 and 2021 Output Areas. Apply this weighting to generate 2021 Output Area estimates. Aggregate from 2021 Output Area to 2021 LSOA using the Output Area to LSOA level lookup table from the ONS Census 2021 geography products.

In addition, there are a small number of indicators where data has been produced at MSOA level, with a 1km buffer zone applied around the MSOA boundary. This applies to all of the indicators in the Civic Assets domain, with the 'buffer zone MSOA' geography representing the wider catchment area around each of these assets, as in most cases, LSOAs are too granular to represent a catchment area around an asset. The table below lists these indicators and highlights the approach taken to create data at this geographical level:

Indicators	Approach to conversion
Density of community space assets Density of educational assets Density of sport and leisure assets Density of cultural assets Density of retail assets Density of community owned asserts Density of green assets	Use PostGres SQL to produce a 1km buffer zone around each of the 2021 MSOA boundaries in England. Use GIS software to overlay point data against 2021 MSOA buffer zone polygon boundaries and use a point-in-polygon method to calculate the number of points that fall within each buffer zone MSOA.

Step 2 Quality assurance of the data

The next step was to comprehensively check the distributions of all the indicators at LSOA level to ensure that all indicators have passed the relevant fitness tests and are “fit for purpose”. These tests include excluding indicators with high numbers of zeros or equal upper limits (for example where a large number of areas have values of 100%) which would distort the Index. All of the indicators selected above passed these quality assurance tests.

Step 3 Applying shrinkage to improve the robustness of indicators

Where a rate or other measure of community need for a small area is based on small numbers, the resulting estimate may be unreliable, with an unacceptably high standard error. The technique of shrinkage estimation is used to ‘borrow strength’ from larger areas to increase the reliability of small area data; the impact of shrinkage will tend to move an LSOA’s score towards that of their parent higher-level area. Shrinkage moderates the levels of unreliability in the dataset and reduces the impact of chance fluctuations from year to year. Such scores occur most commonly where numbers are small at LSOA level and the event is thus relatively rare. This may be the case for the indicator as a whole or only for particular LSOAs. In shrinkage estimation the score for a small area is estimated as a weighted combination of that small area’s score and the mean value for a larger area from which the smaller areas within the larger area borrow strength. The most up to date set of Local Authority Districts as the larger area (this was the larger area used in the Indices of Deprivation shrinkage calculations). LSOAs within a single Local Authority District share issues relating to local governance. To a certain extent, they may also share issues relating to labour market sub-climates.

Further details about the shrinkage technique are given in Appendix C.

Step 4 Ensuring that all indicators are “pointing in the same direction”

In order to combine the indicators into domains, it is necessary for each of the indicators to be orientated in the same direction. However, for some of the indicators included in the Community Needs Index, a *high* value indicates *low* levels of need – for example an area with high levels of grant funding would be measured as having low levels of need. By contrast, for other indicators, a high score denotes high levels of need – for example areas with high travel times to key services. It is necessary therefore to ‘reverse the polarity’ for some scores to ensure that a high value is negative for all indicators – so they can be consistently combined.

Step 5 Producing composite indicators

A small subset of the indicators have been amalgamated to provide composite indicators before combining with the other indicators to create domain scores. The purpose of creating composite indicators is to produce more robust data that captures multiple facets of what the indicator is intending to measure.

The following indicators are grouped together:

Original indicators	Combined indicator
• Density of sport and leisure assets part 1 (from AddressBase) • Density of sport and leisure assets part 2 (from Active places database)	Density of sport and leisure assets
• Density of green assets • Area of public green space	Green assets

• Travel time to employment centre (LSOA with more than 500 jobs) • Travel time to Further Education institution • Travel time to GP • Travel time to hospital • Travel time to Primary School • Travel time to Secondary School • Travel time to town centre • Access to blue spaces • Access to green spaces (active) • Access to green spaces (passive)	Access to services
• People living alone • Self-reported levels of loneliness	Loneliness
• Neighbourhood cohesion • Estimates of social trust • CDRC Residential Mobility Index • Short term population turnover	Self-reported measures neighbourhood strength
• SME lending by banks • Local business units with 0-4 employees	Small businesses
• Patients registered at GP Surgery as ratio of all FTE GPs • GP appointments 21 or more days after appointment is booked (as % of all appointments) • GP appointments per 1,000 patients • Same day GP appointments (as % of all appointments)	GP access score

Before combining each of the individual indicators to produce an overall composite indicator, the indicators have first had shrinkage applied (to reduce any standard errors associated with small numbers). The indicators are then standardised (by ranking and transforming to a normal distribution) – as each of the composite indicators are on a different scale (step 6 below describes the standardisation process in more detail). Indicators are then weighted before combining to produce the composite indicators. Where composite indicators were constructed from two underlying indicators, each of these indicators were given equal weighting. Where composite indicators were constructed from three or more underlying indicators, *Maximum Likelihood Factor Analysis* was used to weight each component indicator (see step 7 for more details regarding *Maximum Likelihood Factor Analysis*).

Step 6 Standardisation

When combining measures, it is important to ensure that indicator scores are comparable and that the weighting of domains is not distorted by the variation in distribution across different indicators . The indicators in the Community Needs Index are based on different metrics and each indicator in the Index needs to be standardised to ensure that they have a common distribution, so that indicators can be combined, without a single indicator dominating due to having a wide distribution. Indicators in the Community Needs Index 2023 have been standardised by ranking each of the indicators and then transforming to a normal distribution.

Step 7 Weighting

Because the Community Needs Index is a compositional measure, decisions have to be made as to the weight given to the various indicators and domains of the Index. There are a number possible approaches to weighting the indicators in a domain.

Option 1 is to provide equal weightings to each of the indicators in a domain. This was the approach taken in the 2019 Community Needs Index.

Option 2 is to apply different weights depending on theoretical judgements regarding the suitability of indicators in the model. Examples of this approach include applying higher weightings to indicators which are constructed from more robust administrative data sources and lower weightings to data from modelled data sources. Alternatively higher weightings can be applied to indicators which more closely match the issue that is being captured – this can be ascertained through a Discrete Choice Experiment (DCE) – a survey of key stakeholders and people from impacted communities identifying the relative importance of indicators.

Option 3 is to introduce a statistical technique called Maximum Likelihood Factor Analysis to determine the weights of the indicators within each domain (Appendix D provides a more detailed explanation of the process). Factor analysis works most effectively where there is a single overwhelming factor which explains the performance on a set of indicators within a domain[11] and where indicators within a domain exert an influence on one another. The outcome of applying factor analysis is that not all indicators in the domain will have equal weights, with the weights affected by the extent to which each of the indicators within a domain measure the underlying aspect that the domain is trying to capture. A key advantage of using factor analysis, is that it takes into account 'double-counting' within domains. However, if there is no underlying factor common among the indicators in a domain, factor analysis is less effective. One way to get around this is to split domains into subdomains which share a common factor.

Having reviewed the indicators in each domain, it was determined that all of the indicators in the Civic Asset domains have close associations (they are all measuring aspects of the same issue – the density of assets of community benefit or community value in a local area) and that factor analysis could be safely applied. However, the *Connectedness* and *Active and engaged community* domains measure conceptually distinct subsets of indicators and needed to be split into subdomains in order for factor analysis to be applied.

The *Connectedness* domain explores connectivity both in terms of access to services and wider measures of connectivity such as access to transport, digital connectivity and isolation - which do not necessarily have strong associations with the more physical concepts of connectivity.
The indicators in this domain were therefore grouped into two subdomains:

Subdomain	Indicators
Physical connectivity	Access to services Job access score
Wider connectivity	Households with no car

[11] For example, in the 2019 Indices of Deprivation, factor analysis was used to weight indicators in the health domain because there was an underlying factor (general health) that impacted on the range of measures from prevalence of long term illness, hospital admissions to premature mortality.

	Digital Exclusion Risk Index (DERI) Loneliness GP access score

The Active and engaged community domain consists of measures concerning self-reported participation and engagement, alongside measures of the strength of the community sector. Again, it made conceptual sense to group these into separate subdomains as follows:

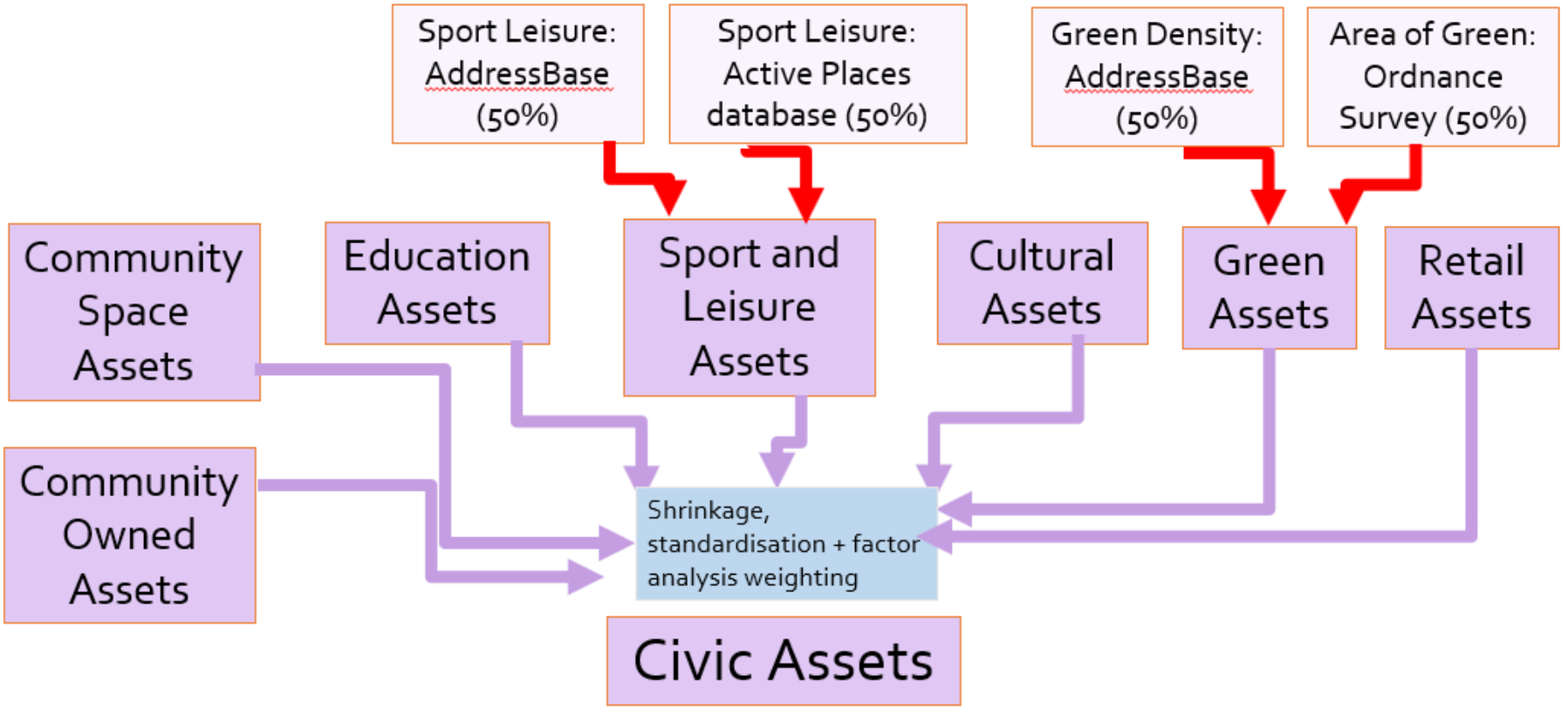
Subdomain	Indicators
Civic participation	Voter turnout at local elections Self-reported measures of community and civic participation Self-reported measures neighbourhood strength
Civic activity	Third sector organisations per head Big Lottery funding per head Grant funding per head from major grant funders Small businesses

The weighted and standardised indicators are combined to form subdomain scores (in the case of indicators in the *Active/engaged community* and *Connectedness* domains) and domain scores (in the case of the Civic assets domain – which would not contain any subdomains).

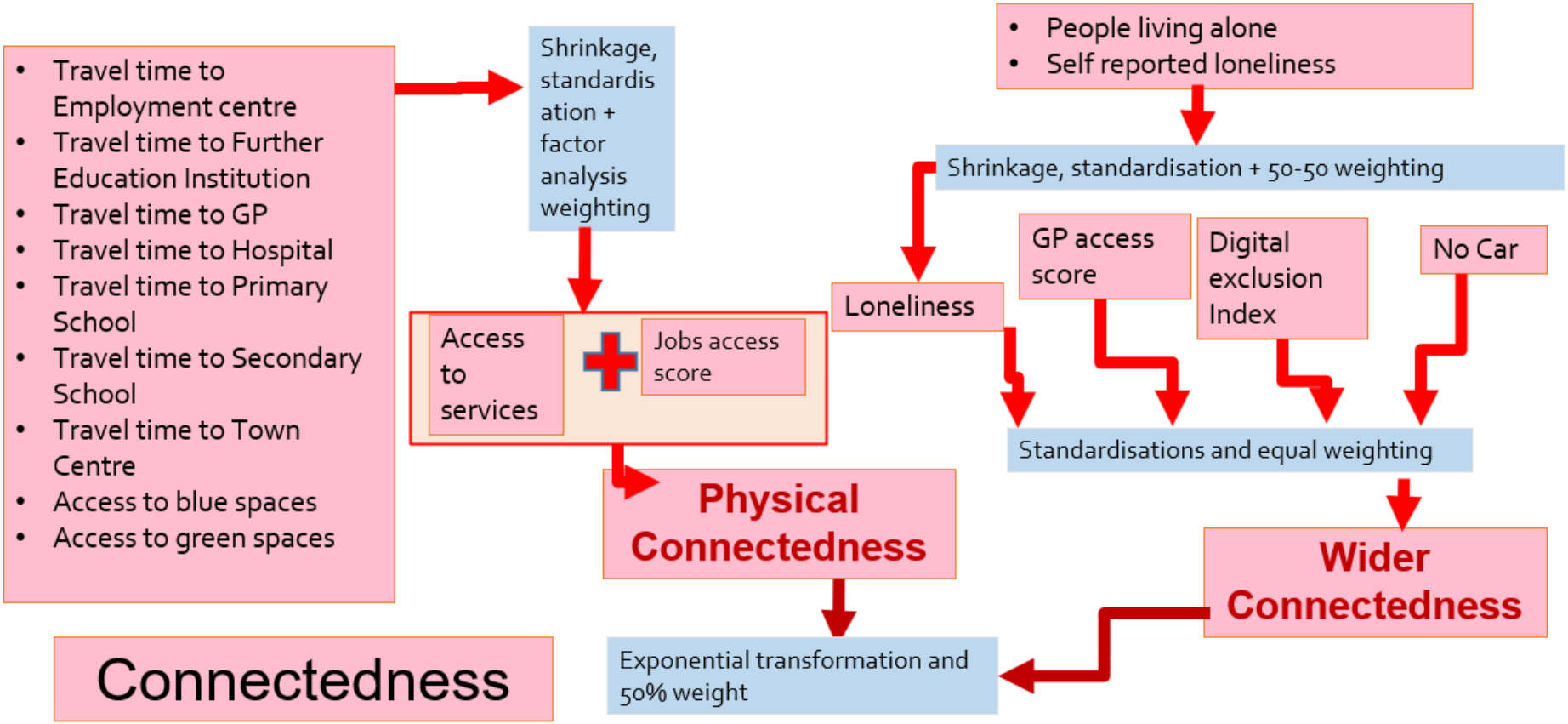
The combination process involves summing each of the weighted indicator scores (the standardised indicator scores * weight) together for all of the indicators within a domain/subdomain. The subdomains are then standardised (using the exponential transformation method outlined in step 8 below) and added together to form domain scores.

The flow chart below shows the overall structure of these indicators and subdomains:

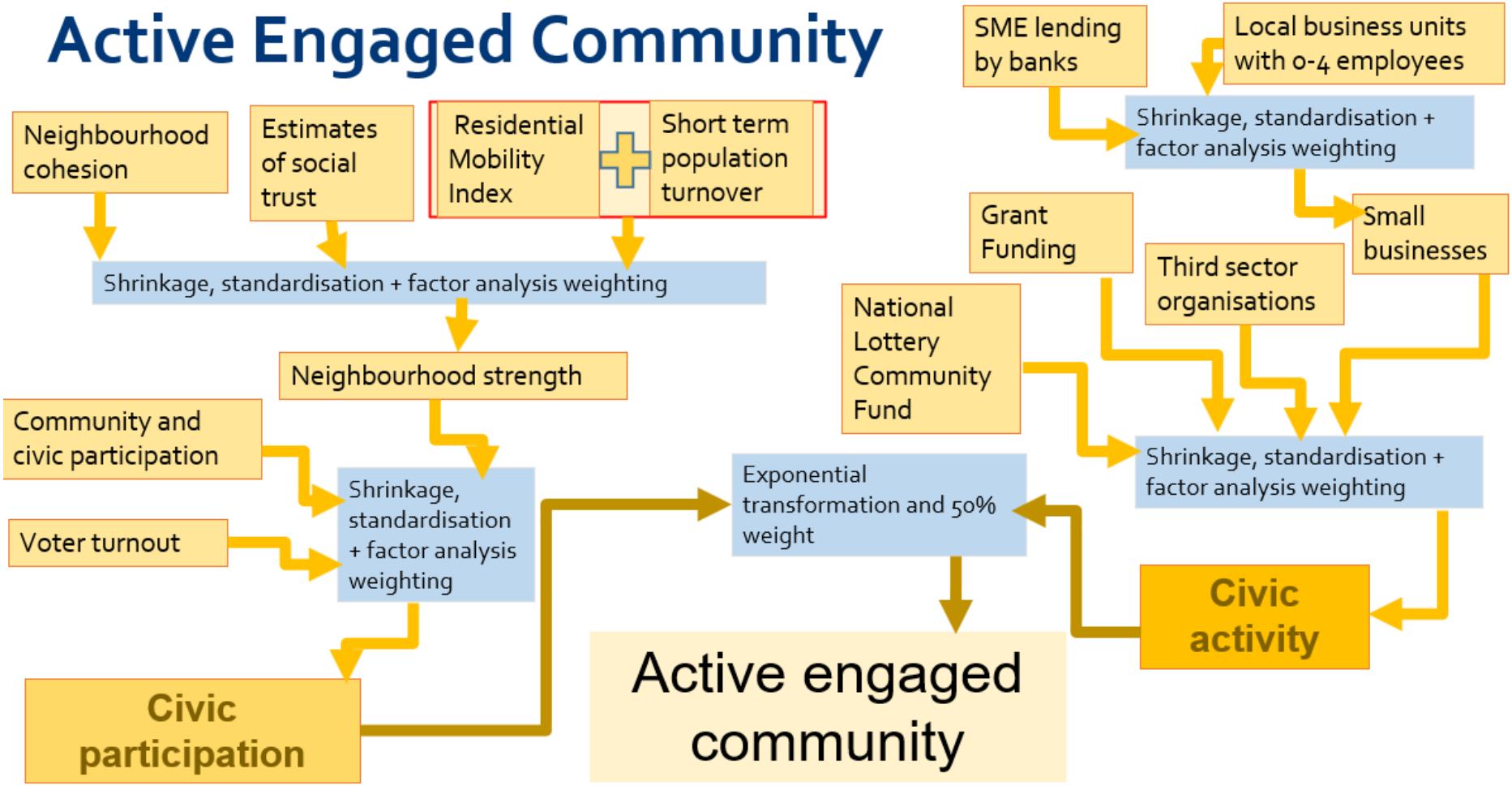
Civic Assets



Connectedness



Active Engaged Community



Step 8 Standardising domains

The three domain scores are then combined to produce the overall Community Needs Index.

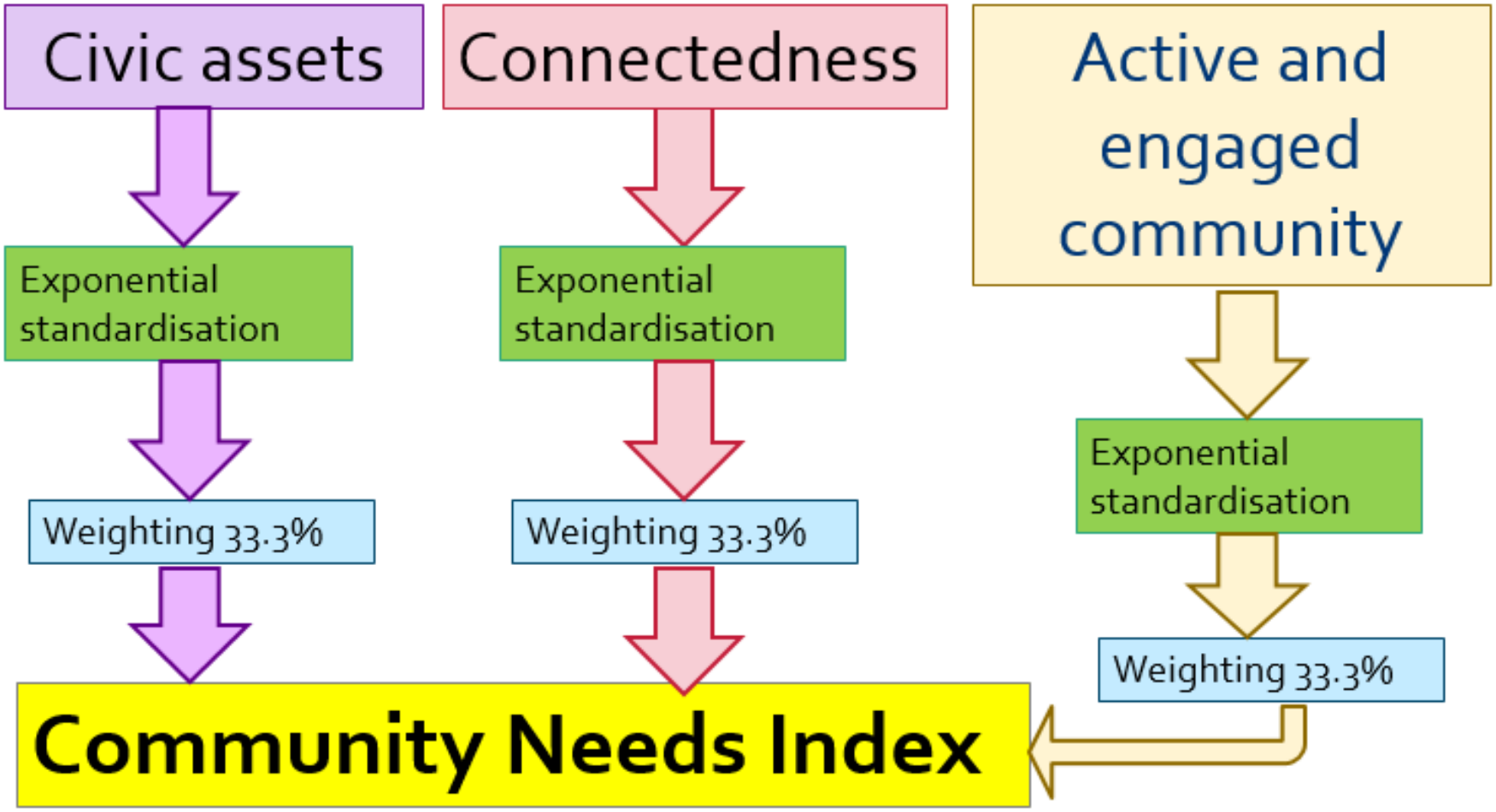
However, each of the domains will be on a different scale to one another, with two of the three domains produced from combined subdomain scores, while the *Civic assets* domain is produced from combined weighted indicators.

It was therefore necessary to standardise the domain scores before combining. As with the 2019 Community Needs Index, the method of standardisation that was adopted in the 2023 Index was to transform the domains to a specified **exponential distribution** using an *exponential transformation* function (see Appendix B for details). The exponentially transformed subdomain/domain scores were then be combined to form an overall 'community need' measure at LSOA level.

The *exponential transformation* method of standardisation differs from the normal distribution method as it gives more emphasis on the top end of the distribution (the areas with the highest scores) and so facilitates identification of the areas with the highest levels of need. This was the method of standardisation applied in the Indices of Deprivation in order to control cancellation effects (e.g. high levels of deprivation in one domain are not completely cancelled out by low levels of deprivation in a different domain) and ensures that areas that perform particularly badly on one aspect of community need are moved closer to the high end of the community need spectrum even when they show positive outcomes on other indicators.

Step 9 Weighting domains

The final stage for producing the Community Needs Index is to assign weights to the three domains that have been created – to apply to the domain scores before importing. It is important to note that all potential combinations of domains involve weights. If, after standardisation, the domains are simply added together, this gives each domain an equal weight. It is intended that the weights should be explicit and based on clear criteria. Part of this commitment to transparent weights involves the standardisation of the domain Indices as outlined above. This ensures that the domains can be combined without 'hidden' weights. Having standardised the domains, it is then necessary to choose explicit weights. Equal weights have been applied to each of the three domains to mirror the approach taken in the 2019 Community Needs Index where equal weights were selected to reflect the equal importance afforded to each of the dimensions of Community Need. Once each domain was weighted, the domains were combined to produce the overall Community Needs Index. The combination process involves summing each of the weighted standardised domain scores (the exponentially transformed domain scores * weight) together to produce an overall Community Needs Index score (see flow chart below).


Civic assets
Connectedness
Active and engaged community
Exponential standardisation
Exponential standardisation
Exponential standardisation
Weighting 33.3%
Weighting 33.3%
Weighting 33.3%
Community Needs Index

Appendix A Combining the Community Needs Index with other needs frameworks

This section explores how the Community Needs Index can be combined with other frameworks in order to identify key communities at risk – where social infrastructure challenges co-inside with other key socio-economic challenges.

In 2019, the Community Needs Index was combined with the Index of Multiple Deprivation (IMD) 2019 to identify 'left-behind' areas. The decision to use the IMD in combination with the Community Needs Index arose from out theoretical conception of 'left-behind' areas as areas which experienced a "dual disadvantage":

a) High levels multiple deprivation (measured by the IMD)

b) Poor connectivity, low levels of civic assets and community participation and engagement (measured by the Community Needs Index).

Areas were identified as 'left behind'; if they ranked among the most deprived **10%** of Wards in England on *both* the 2019 Community Needs Index and the 2019 IMD. It is important to note that the development of the Community Needs Index was conceptually shaped by the interactions with the Index of Multiple Deprivation. The Index was created in order to capture the social infrastructure challenges in areas that also experienced multiple deprivation challenges (with a recognition that social infrastructure challenges are more acute for those experiencing low income, worklessness or poor health) and indicators were only considered for the Community Needs Index where they were not already included in the IMD.

As part of the 2023 update, we have not combined the Community Needs Index with the IMD or any other framework in order to identify at 'risk groups'.

However, we have produced the Community Needs Index 2023 at LSOA level in order to facilitate the combination of the measure with the Index of Multiple Deprivation or other key frameworks.

There are a number of potential options for combining the 2023 iteration of the Community Needs Index with the IMD. The following are presented as examples only:

- Raising the threshold at which areas are identified as at risk e.g. a two-tier categorisation of 'left behind', with those in the top 20% on both IMD and the Community Needs Index being regarded as 'moderately left behind' in addition to the 10% identified as 'severely left behind' (which would therefore be a subset of the moderately left behind group).
- Identifying areas as 'left behind' at MSOA level if they contain at least one LSOA ranked in the most deprived 10% on the IMD 2019 or CNI 2023.
- Identify a set number of areas with the highest combined scores on the Community Needs Index and IMD.
- Defining cut-points based on cumulative population rather than number of MSOAs.
- Identify Local Authorities as high risk if they have a high proportion of LSOAs ranked in the top 10% on the CNI 2023 and a high proportion of LSOAs ranked in the top 10% on the IMD 2019.

Appendix B: Exponential transformation

In order to combine the domains into an overall measure of need, the domain scores first need to be standardised. Any standardisation and transformation should meet the following criteria:

- Standard distribution. It must ensure that each domain has a common distribution, so that domains can be combined, without one domain dominating due to a much larger distribution.
- Identify areas of need. It must facilitate the easy identification of the areas with highest levels of need.
- Scale independent. It must not be scale dependent (in other words confuse population size with level of need).

One possible standardization approach involves each of the domain scores being ranked, and then the ranks are transformed to an exponential distribution. The exponential distribution has a number of properties that satisfy the criteria above.

Standard distribution

The exponential distribution transforms each domain so that they each have a common distribution, the same range and identical maximum / minimum values. The process starts by ranking the scores in each domain to standardise the domain scores (from 1 for the lowest need to 6,791 for the most highest need), before applying the exponential transformation procedure to create a standardised domain score ranging from 0 (lowest need) to 100 (highest need).

Cancellation

The exponential transformation procedure gives control over the extent to which lack of need in one domain cancels or compensates for high need in another domain. It allows precise regulation, although not elimination, of these cancellation effects. The scaling constant (23) used produces roughly 10 per cent cancellation. This means that in the extreme case, an LSOA which was ranked most deprived on one domain but least deprived on another would overall be ranked at the 90th percentile in terms of levels of need. This compares to the 50th percentile if the untransformed ranks or a normal distribution had been used instead.

Identify deprived areas

The exponential transformation effectively spreads out that part of the distribution in which there is most interest - that is the ‘tail’ which contains the areas with the highest levels of need in each domain. The scaling constant ensures that the most deprived 10 per cent of areas cover 50 per cent of the distribution of scores (in other words, scores between 50 and 100 after exponential transformation).

Scale independent

The transformation is not affected by the size of the LSOA’s population.

The exponential transformation calculation

The transformation used is as follows:

For any LSOA, denote its rank on the domain R, scaled to the range [0,1]. R=1/N for the least deprived and R=N/N (in other words R=1) for the most deprived, where N=the number of LSOAs in England.

The transformed domain score X is given by:

$$X = -23\ln\left(1 - R\left(1 - \exp^{-100/23}\right)\right)$$

where 'ln' denotes natural logarithm and 'exp' the exponential or antilog transformation

Appendix C: Shrinkage

Improving the reliability of small area data values using shrinkage estimation

The shrinkage technique is designed to deal with the problems associated with small numbers in an LSOA. In some areas – particularly where the at-risk population is small – data may be 'unreliable', that is more likely to be affected by sampling and other sources of error.

The technique of shrinkage estimation (in other words empirical Bayesian estimation) is used to 'borrow strength' from larger areas to avoid creating unreliable small area data. Shrinkage estimation involves moving LSOA scores towards another more robust score, often relating to a higher geographical level. All LSOA scores will move somewhat through shrinkage, but those with large standard errors (in other words the most 'unreliable' scores) will tend to move the most. The LSOA score may be moved towards a 'higher need' or 'lower need' score through shrinkage estimation. Without shrinkage, some LSOAs would have scores which do not reliably describe the community need in the area due to chance fluctuations from year to year.

It could be argued that shrinkage estimation is inappropriate for administrative data which are, in effect, a census. This is not correct. The problem exists not only where data are derived from samples but also where scans of administrative data effectively mean that an entire census of a particular group is being considered. This is because such censuses can be regarded as samples from 'super-populations', which one could consider to be samples in time. All the data from administrative sources and the 2011 Census are treated as samples from a super-population in this way, and the shrinkage technique was applied to indicators which use this data. The exceptions are the indicators supplied at Local Authority District level.

Selecting the larger areas from which unreliable small area data can borrow strength

The principle for selecting the larger area should be that the LSOAs within them share characteristics. In the current shrinkage methodology, Local Authority Districts are used. The LSOAs within a single district share issues relating to local governance and possibly to economic sub-climates. To a certain extent, they may also share issues relating to labour market sub-climates.

The shrinkage calculation

The actual mechanism of the shrinkage procedure is to estimate deprivation in a particular LSOA using a weighted combination of (a) data from the LSOA, and (b) data from another more robust score (in the case of the Indices, this is the Local Authority District score). The weight attempts to increase the efficiency of the estimation, while not increasing its bias. All LSOA scores are adjusted to some degree through the shrinkage process, but the magnitude of the adjustment will be greatest for areas with the least reliable scores. The amount of movement depends on both the size of the standard error and the amount of heterogeneity amongst the LSOAs in a Local Authority District.

The 'shrunk' estimate of a LSOA level proportion (or ratio) is a weighted average of the two 'raw' proportions for the LSOA and for the corresponding District. The weights used are determined by the relative magnitudes of within-LSOA and between-LSOA variability.

If the rate for a particular indicator in LSOA j is r_j events out of a population of n_j, the empirical logit for each LSOA is:

$$m_j = \log\left[\frac{(r_j + 0.5)}{(n_j - r_j + 0.5)}\right]$$

whose estimated standard error s_j is the square root of:

$$s_j^2 = \frac{(n_j + 1)(n_j + 2)}{n_j(r_j + 1)(n_j - r_j + 1)}$$

The corresponding counts r out of n for the district in which LSOA j lies gives the district-level logit:

$$M = \log\left[\frac{(r + 0.5)}{(n - r + 0.5)}\right]$$

The 'shrunk' LSOA level logit is then the weighted average:

$$m_j^* = w_j m_j + (1 - w_j)M$$

where w_j is the weight given to the 'raw' LSOA-j data and $(1-w_j)$ the weight given to the overall rate for the district. The formula used to determine w_j is:

$$w_j = \frac{1/s_j^2}{1/s_j^2 + 1/t^2}$$

where t^2 is the inter-LSOA variance for the k LSOAs in the district, calculated as:

$$t^2 = \frac{1}{k-1}\sum_{j=1}^{k}(m_j - M)^2$$

Thus large LSOAs, where precision $1/s_2^j$ is relatively large, have weight w_j close to 1 and so shrinkage has little effect. The shrinkage effect is greatest for small LSOAs in relatively homogeneous districts.

The final step is to back-transform the shrunk logit m_j^* using the 'anti-logit', to obtain the shrunk LSOA level proportion for each LSOA:

$$z_j = \frac{\exp(m_j^*)}{1 + \exp(m_j^*)}$$

Appendix D: Factor Analysis methodology

Factor analysis is used as a method for combining indicators, by finding appropriate weights for combining indicators into a single score based on the inter-correlations between all the indicators.

Factor analysis is only used in domains where 'latent variables' are hypothesised to exist and where the indicator variables are 'effect indicators', i.e. indicators that are influenced by the latent variable.

There are many candidates in terms of types of factor analysis. Two of the main contenders are maximum likelihood factor analysis (as used in the current and previous versions of the Indices of Deprivation) and Principal Components Analysis. The distinction between maximum likelihood factor analysis and Principal Components Analysis is a technical one. In brief, the assumptions underpinning Principal Components Analysis are that the indicators going into the analysis are perfectly reliable and measured without error. Maximum likelihood factor analysis requires no such assumption.

The process of combining indicators using factor analysis comprises three stages:

- All indicators are converted to the standard normal distribution.
- The standardised scores were factor analysed (using the Maximum Likelihood method), deriving a set of weights.
- The indicators were then combined using these weights.