Consumer Data Research Centre

An ESRC Data Investment

Data Profile: CDRC Residential Mobility ("Population Churn") Index (2023)

Introduction

This data profile describes a dataset held by the CDRC which has been compiled from analysis on electoral registers and consumer registers originally supplied by CACI, DataTalk, and another provider. The data provides a yearly LAD and LSOA level estimate of the proportion of households that are different to those in 2023 going back to 1997.

The paper outlines the structure of the dataset and the variables that are available, highlights some potential uses of the data and provides some initial assessment of data quality.

Scale and Extent

Field	Value
Data Provider	CDRC
Analytical Units	LAD (2021 boundaries), LSOA (2011 boundaries)
Data Format	CSV
Temporal Extent	1997 – 2023
Geographical Extent	United Kingdom
Variables	Proportion of households that are different to those in 2023 by year
Observations	LAD, LSOA

Citation Information

The following statement should be included when citing the use of this dataset:

"The data for this research have been provided by the Consumer Data Research Centre, an ESRC Data Investment, under project ID CDRC [Project Number], ES/L011840/1; ES/L011891/1"

Data Classification and Access Summary

Both the LAD and LSOA estimates are classified as Open and are available for download.

Content

The data simply provides a ratio of the households that are different in each LAD or LSOA between the beginning of 2023 and the end of each year going back to 1997.

Residential mobility ("Population Churn") is estimated at the household level. Households' start and end dates are extracted from individual level data by combining individuals that at any point in time have shared time together in the same property or have a shared surname in the same property. First household member determines the 'start' date (household identified as moving in), last household member determines the 'end' date (household identified as moving out).

Novelty

This data enable research to explore annual variations in neighbourhood change at a small area geography. Crucially it also enables research 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 2011 Census of Population.

Quality, Representation and Bias

As the data is inputted from several different organisations, it is possible that some names and addresses are inconsistently formatted between datasets. This would have hampered data linkage when making the index. For instance, we estimate that 100,000 surnames are misspelled/recorded slightly differently in each register.

Data Profile: CDRC Residential Mobility ("Population Churn") Index (2023)

In addition, addresses (particularly flats) can be recorded differently. Addresses are recorded as address lines. This makes address matching to other data quite difficult as the number and composition of address lines varies by addresses, and between different versions of the data too.

Lastly, it is very difficult to determine the completeness of the data. It is possible that deceased adults are not removed immediately and often not until new individuals are registered at their address. This means that RMI values in the last few years (2018-23) should be interpreted with caution given the delay in the change of registered addresses for individuals. Generally, town centres are underrepresented though. The earlier consumer registers (2003-12) and later electoral registers (2018-20) tended to under represent the adult population relative to mid-year population estimates. However, efforts were made by the CDRC to fill in any gaps where possible. The data for Northern Ireland is also estimated to be less complete due to distinctive administrative procedures in the region. It is also possible that adults who reside in multiple addresses may have duplicate entries within the data.

Representation and Bias

Whilst the data providers have attempted to compile registers which are both as complete and accurate as possible, there are data biases that should be considered. First, the electoral register (up to 2023) is known to sufficiently under-represent the following groups: the younger age groups, the non-white British population and those in rented accommodation. Second, the counts in the original data also fluctuated relative to the population growth. CDRC research subsequently attempted to fill in the gaps and reweight changes based on additional data sources.

The counts of households in the original LCR data fluctuate according to data supplier in addition to actual population size changes. As such, meta data describing the annual distribution of household counts across LSOAs and LADs for the 2023 baseline year that have been used in calculating RMI are available below. RMI may be out of line with census counts and users should consult census statistics if they have concerns. This is particularly pertinent to the last few years (2018-23), especially 2019, in addition to areas such as Braintree.

LAD Household Distribution (2023)

Min	214
1st Quartile	16,423.25
Mean	29,149.78
Median	22,295.50
3rd Quartile	35,462
Max	173,394
Standard Deviation	21,206.49
Standard Error	1,096.65

LSOA Household Distribution (2023)

Min	17
1st Quartile	195
Mean	255.82
Median	254
3rd Quartile	309
Max	2109
Standard Deviation	91.90
Standard Error	0.45

Related Datasets

Linked Consumer Registers:

https://data.cdrc.ac.uk/dataset/linked-consumer-registers

Consumer Data Research Centre

An ESRC Data Investment

Data Profile: CDRC Residential Mobility ("Population Churn") Index (2023)

Data Overview

Source	Variable	Spatial granularity of comparator	Temporal granularity of comparator	Note(s)
LSOA / LAD	LSOA code / LAD code	LSOA	Annual	
chn1997	Proportion change between 2023 and 1997	LSOA	Annual	
chn1998	Proportion change between 2023 and 1998	LSOA	Annual	
…				
chn2023	Proportion change between 2023 and 2023	Same as above	Same as above	Defaults to 0.