

## Handling Migration & Commuting Flow Data (1):

# Using the Interaction Data from the SARs and the LS

Paul Norman <sup>1</sup> & Paul Boyle <sup>2</sup>

1. School of Geography, University of Leeds

2. School of Geography & Geosciences, University of St Andrews & Longitudinal Studies Centre – Scotland (LSCS)

### *Acknowledgements*

- SARs team at Centre for Census & Survey Research
- Permission of ONS to use the Longitudinal Study & help by staff at CeLSIUS, particularly Chris Marshall



## Handling Migration & Commuting Flow Data (1)

# Using the Interaction Data from the SARs and the LS

- Terminology & definitions
- Data & techniques
- Research examples
- Strengths & weaknesses

## Terminology & definitions

### The datasets

SARs = Samples of Anonymised Records

- 1991 & 2001 Censuses
- = cross-sectional data, a snapshot

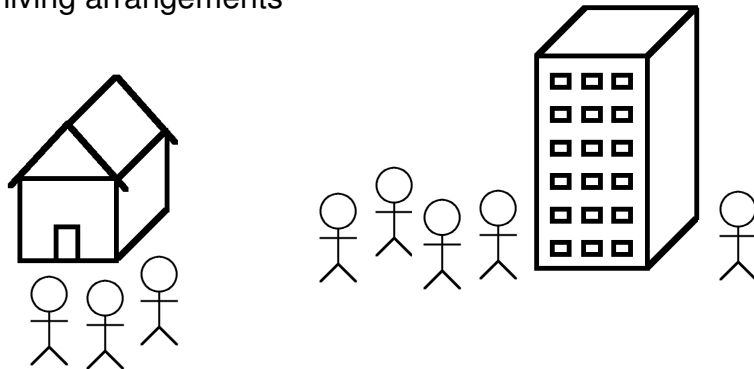
LS = Longitudinal Study

- England & Wales: 1971-2001 Censuses plus other data
  - Scotland (SLS)
  - Northern Ireland (NILS)
- = longitudinal data, tracks people over time

## Terminology & definitions

### Microdata, individual level data

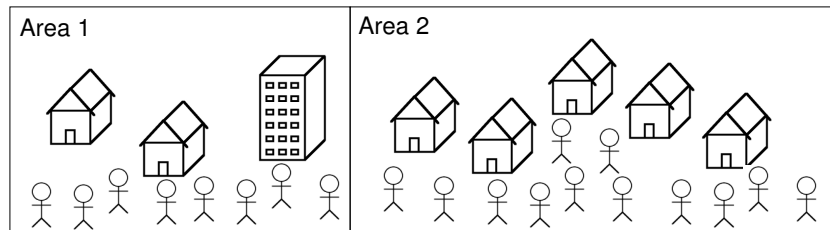
Individuals fill out census questionnaires & answer surveys & thus provide information about themselves & their living arrangements



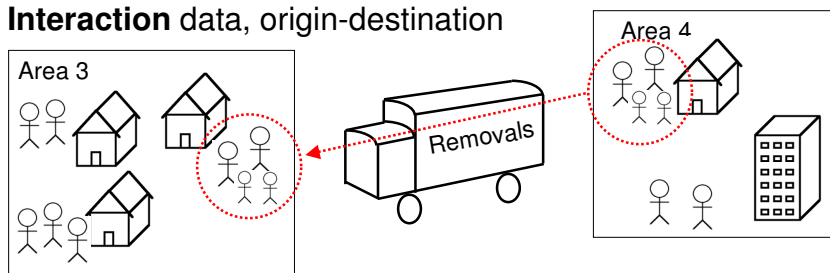
Norman (2003)

## Terminology & definitions

**Area** statistics, **aggregate** data, the sum of all individuals & houses in a geographically defined location



**Interaction** data, origin-destination



## Terminology & definitions

SARs & LS have rich individual level information with close equivalence to the Census form

Plus information:

- About others in the household
- Some 'space' (area type) & 'place' (named locations)

SARs & LS are 'samples'

- SARs: various files between 1% & 5% of population
- LS: c. 1% sample of England & Wales

High level of detail about individuals. Maintaining confidentiality high on agenda

- SARs: some files academics register to use, some apply
- LS: tight access arrangements

Some SARs files (download), other SARs & LS more 'remote'

## Some data & techniques

Microdata: each individual's ...

- age, sex, economic activity, tenure, health, ethnicity, social class / NS-SEC, educational achievement, marital status, migration, commuting, etc

Data type: largely categorical

Microdata versatility & flexibility:

- Variables: re-code to application relevant categories
- Crosstabulations (not in standard area outputs)
- Statistical techniques not applicable with aggregate data

File formats:

- SPSS, Stata
- Learn syntax (.sps; .do) for audit trail, recycle variable & modelling code

## Some data & techniques

SARs example: Individual files for 1991 & 2001

Distance moved by migrants (year before the census)

1991

Variable Values

Value	Label
dismove -9.00 <sup>a</sup>	Not applicable
-8.00	Not stated
-1.00	From outside GB
.00	0-4 km
5.00	5-9 km
10.00	10-14 km
15.00	15-19 km
20.00	20-29 km
30.00	30-39 km
40.00	40-49 km
50.00	50-59 km
60.00	60-79 km
80.00	80-99 km
100.00	100-149 km
150.00	150-199 km
200.00	200 km and over

a. Missing value

2001

Variable Values

Value	Label
dismov0 -9 <sup>a</sup>	Not applicable (not a migrant or student living away)
1	0-2 km
2	3-4 km
3	5-6 km
4	7-9 km
5	10-14 km
6	15-19 km
7	20-29 km
8	30-49 km
9	50-99 km
10	100-149 km
11	150 - 199km
12	200 + km
13	From outside UK

a. Missing value

These are very detailed & not exactly equivalent

What is a short distance? What is a long way?!?

## Some data & techniques

SARs example: Individual files for 1991 & 2001

Distance moved to be compatible across censuses

1991

```
* Syntax for the 1991 individual SAR .
compute distcat = 0 .
```

```
if (distmove > -1 & distmove < 10 or distmove = -8) distcat = 1 .
if (distmove > 9 & distmove < 50) distcat = 2 .
if (distmove > 49 ) distcat = 3 .
if (distmove = -1) distcat = 4 .
```

```
variable label distcat "Distance Moved" .
```

```
add value labels distcat
```

```
0 "Non-migrant"
1 "Short distance migrant (0-9k)"
2 "Medium distance migrant (10-49k)"
3 "Long distance migrant (50+k)"
4 "Immigrant" .
```

```
execute .
```

2001

```
* Syntax for the 2001 licensed SAR .
compute distcat = 0 .
```

```
if (distmov0 > 0 & distmov0 < 5) distcat = 1 .
if (distmov0 > 4 & distmov0 < 9) distcat = 2 .
if (distmov0 > 8 & distmov0 < 13 ) distcat = 3 .
if (distmov0 = 13) distcat = 4 .
```

```
variable label distcat "Distance Moved" .
```

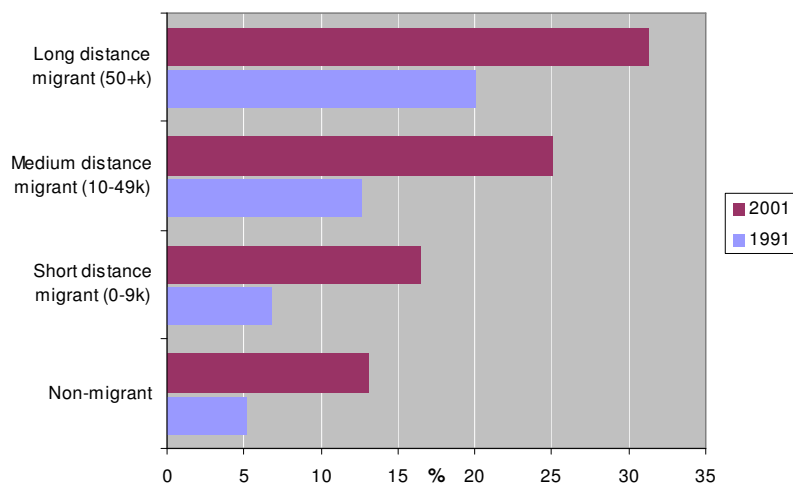
```
add value labels distcat
```

```
0 "Non-migrant"
1 "Short distance migrant (0-9k)"
2 "Medium distance migrant (10-49k)"
3 "Long distance migrant (50+k)"
4 "Immigrant" .
```

```
execute .
```

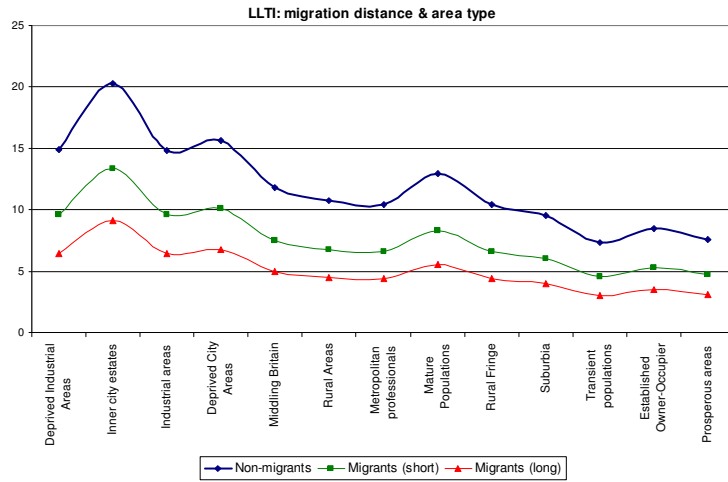
## Some data & techniques

Relationship between distance moved & education



## Research examples: SARs

Probability of migrating, in relation to health & area type

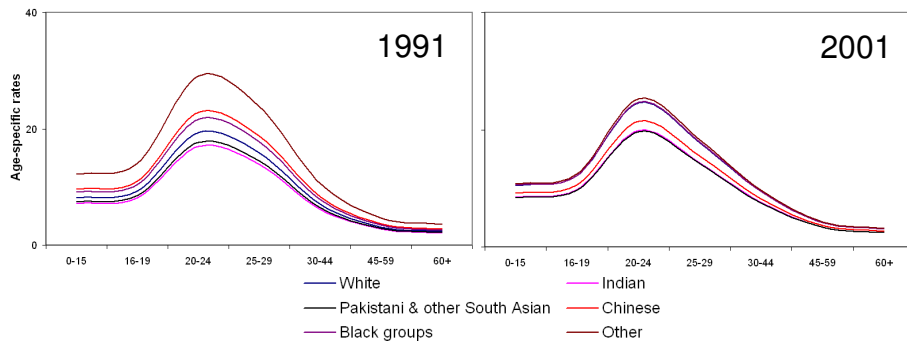


Boyle, Norman & Rees (2002)

## Research examples: SARs

Probability of migrating, relation to ethnic group, 1991 & 2001

- Controlling for Social Class, Tenure, Education, etc.

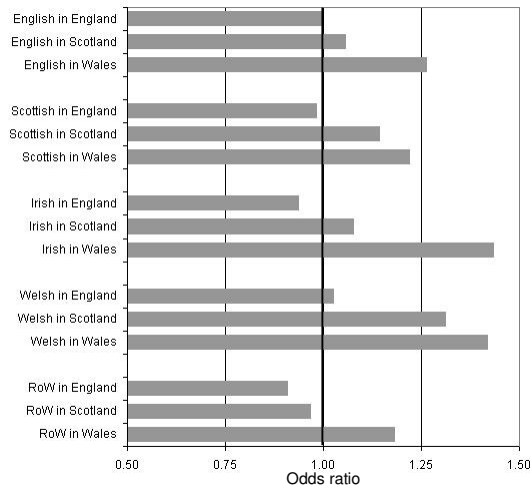


Norman, Stillwell & Hussein (2007)  
See also Finney & Simpson (2008)

## Research examples: SARs

Geographical detail in SARs rather coarse & about 'origins' relatively poor. Some 'region' information, but limited utility

Longer term: Country of birth (origin) & residence (destination)



Norman, Boyle & Brown (2006)

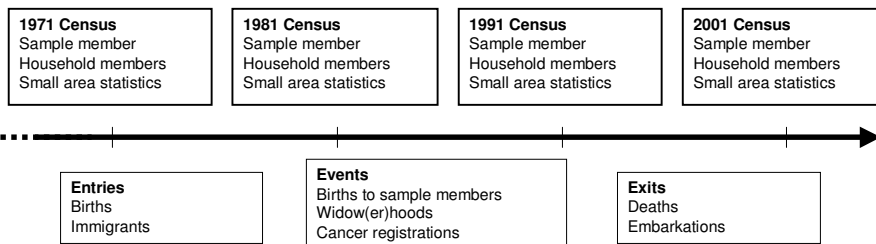
## Research examples: LS

The SARs very informative on migrant characteristics, fairly useful about destinations, largely uninformative on origins

The LS tracks people over time ...

### Record linkage study c.1% population England & Wales

- Census microdata about individuals 1971, 1981, 1991, 2001
- Information equivalent to a person's census form
- Aggregate-level data available about each person's locations



## Research examples: LS

Whilst 'places' of origin and destination are recorded on LS database, only regional geography likely to be undisclosive  
Interaction between different types of 'space' can be informative:

- People moving between urban-suburban-rural areas
- Tracking people through differently deprived areas

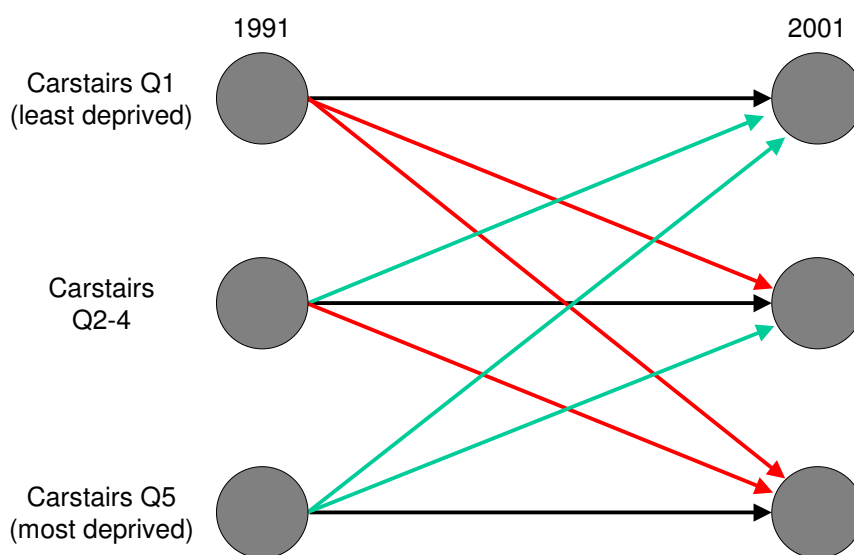
Census area data, SMS & SARs use usual address one year ago to inform on migrant transitions & status

LS has this information, but:

- 10 year transitions, people & their circumstances
- 1971-1981; 1981-1991; 1991-2001

## Research examples: LS

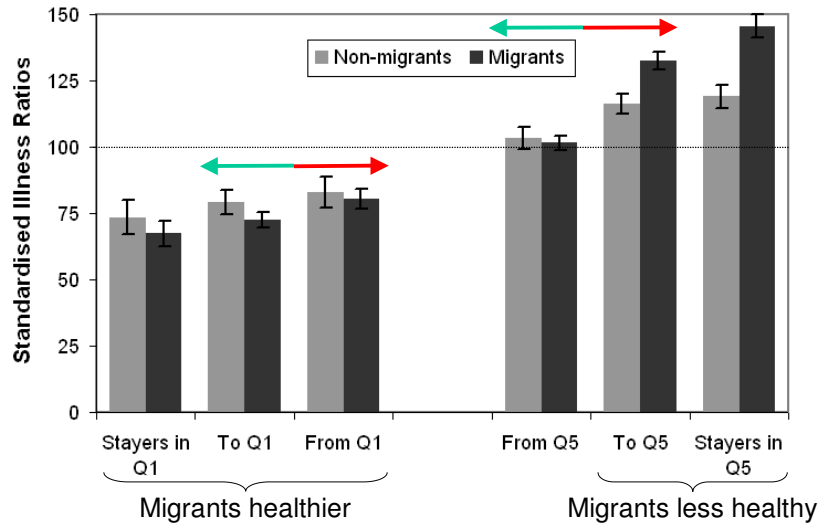
Transitions between Carstairs quintiles





## Research examples: LS

Are migrants healthier than non-migrants?



Norman, Boyle & Rees (2005); Boyle, Norman & Rees (2004)

## Strengths & weaknesses

Individual level microdata very versatile & flexible

- Derive application-relevant variables & user-defined study population
- Crosstabulations not in standard outputs (quicker & cheaper than commissioned tables)
- Useful statistical techniques applicable

But ...

- SARs & LS are samples (though *very large* samples)
- Geographical information for migration somewhat limited
- LS and most detailed SARs challenging to access

Though ...

- Downloadable SARs for 1991 & 2001 great utility

## Resources

### Samples of Anonymised Records

- <http://www.ccsr.ac.uk/sars/>
- User support by SARs team at CCSR

### ONS Longitudinal Study for England & Wales

- <http://celsius.census.ac.uk/>
- User support by CeLSIUS at LSHTM

### Scottish Longitudinal Study

- <http://www.lscs.ac.uk/>
- User support by Longitudinal Study for Scotland team

### Northern Ireland Longitudinal Study

- <http://www.nisra.gov.uk/nils/default.asp.htm>
- User support by NISRA

## Recommended reading

Dale A, Fieldhouse E and Holdsworth C (2000) *Analysing Census Microdata*. Arnold: London  
Hattersley L. and Creeser R (1995). *Longitudinal Study 1971-1991: History, Organisation and Quality of Data*. OPCS Series DS 15, HMSO: London

## References

Boyle P, Norman P & Rees P (2002) Does migration exaggerate the relationship between deprivation and limiting long-term illness? A Scottish analysis. *Social Science & Medicine* 55: 21-31

Boyle P, Norman P & Rees P (2004) Changing places: do changes in the relative deprivation of areas influence limiting long-term illness and mortality among non-migrant people living in non-deprived households? *Social Science & Medicine* 58: 2459-2471

Finney N & Simpson L (2008) Internal migration and ethnic groups: evidence for Britain from the 2001 Census. *Population, Space & Place* 14: 63-83

Norman P (2003) What are individual-level microdata and aggregate-level area census data? FAQ 11 Individual versus Aggregate. Online:  
<http://www.chcc.ac.uk/overview/faq11/frame.html>

Norman P, Boyle P & Brown M (2006) Which influences the self-reporting of health, country of birth or country of residence? British analysis using individual-level data. Presentation at the International Population Geography Conference, Liverpool, June

Norman P, Boyle P & Rees P (2005) Selective migration, health and deprivation: a longitudinal analysis. *Social Science & Medicine* 60(12): 2755-2771