

## **How the LA Case Study areas were chosen**

### **Introduction**

The 2005/06 LA case studies are one of the key components of the Improving Migration and Population Statistics project (IMPS), the key objective of which is to identify sources and methods to improve ONS's mid-year population estimates (MYEs). In selecting LAs for the case studies, ONS considered that it was essential to use a method which was objective and transparent. The approach taken was to identify factors which may cause the MYEs to underestimate or overestimate the population as a whole, and/or particular population age/sex groups, LAs or types of LA. These factors were used to group Local Authorities into clusters with a greater or lesser risk to the MYEs. The aim of the case studies is to learn lessons that can be applied across all LAs, or groups of LAs, rather than to make adjustments only for those LAs being studied

### **Approach**

#### **Stage 1 - Identifying risk factors**

A framework was constructed to provide a comprehensive analysis of where there might be difficulties in estimating the population. ONS identified a long list of variables to represent these risk factors. However, several were rejected either because the data for the variables was not of a suitable quality or else the variables were highly correlated with others in the list. After eliminating all the unsuitable candidates, eight variables were used:

- Gross international migration 2002-03 (International Passenger Survey)
- Gross internal migration 2002-03 (Patient Registers)
- Armed forces 2002-03 (Home armed forces and foreign armed forces & their dependants)
- Non-white population (2001 Census)
- Full-time students aged 18-74 (2001 Census)
- Population density at mid-2003 (Regional Trends)
- Prisoners 2002-03 (Home Office)
- Men aged 20-39 (2003 mid-year estimate)

Apart from population density these were all expressed as a percentage of the population.

#### **Stage 2 - LAs grouped into clusters**

Cluster analysis was used to group LAs into clusters which were likely to have a similar risk to their MYEs. The advantage of this method is that it looks across the whole spectrum of risks to the MYEs, and also takes into account the specific characteristics of each area. Using the eight variables listed above, the cluster analysis allows us to group together LAs with similar combinations of characteristics which may affect the accuracy of their estimates, and to differentiate them from clusters with different combinations.

15 clusters were produced using these eight variables. The members of the clusters together with their characteristics are shown in the Annex.

### **Stage 3 - Clusters of LAs assessed using principal components**

We then used principal components analysis in order to ensure that the clusters that are the most difficult to estimate were selected in the case studies. This allowed us to investigate how the combination of variables in each cluster contributed to the overall measure of risk to the population estimates.

The first principal component (PC1), which accounts for just over half the variation among the eight variables, derived from a combination of migration, students, young men, non-whites and population density. The second principal component (PC2) was driven by the special populations - Armed Forces and prisoners.

The distribution of LAs by cluster and by their scores on the first two principal components can be seen in this [graph](#).

It can be seen from the graph that clusters 1, 4 and 14 have the highest values on the first principal component, followed by clusters 2, 6 and 11. Cluster 8 also has mainly positive scores on PC1. Clusters 12 and 15 have the highest scores on PC2.

### **Stage 4 - Cluster selection**

Of the seven clusters with high scores on PC1, cluster 14 consists of Oxford and Cambridge only, and there are some research topics which are dealing with student populations separately. One topic is currently investigating student registration with GPs, and another will be investigating population groups such as students which are known to be difficult to estimate. Therefore it was decided not to research cluster 14 as part of the case studies.

From the six remaining clusters with high PC1 values, ONS has resource to carry out research in four during 2005/06. Clusters 1 and 4 are entirely London-based and cluster 2 comprises mainly London Boroughs. Cluster 6 consists of two Outer London Boroughs and nine areas outside London, whilst clusters 8 and 11 consist entirely of non-London LAs. In order to maintain a balance between London and non-London areas we selected clusters 1, 2, 8 and 11 for study in 2005/06. This selection gives four characteristically distinct clusters.

### **Stage 5 - LA selection**

In selecting LAs from each of these clusters we chose the one which is closest to the centroid of its cluster. The centroid is where the average value of each of the variables is found within each cluster, and can be seen as the 'average' or 'most typical' of that cluster. The centroid itself does not represent a LA, but by studying the cluster member which is closest to the average, findings are most likely to be applicable to the other members of that cluster. The LAs closest to the centroid for the clusters we are initially studying are:

**Cluster 1** - Hammersmith and Fulham

**Cluster 2** - Barnet  
**Cluster 8** - Derby  
**Cluster 11** - Welwyn Hatfield

**Footnote - 2004 adjustments**

In 2004 reviews were carried out to check the Census figures of 32 Local Authorities and adjustments were made to the populations of 15 of these following an extensive analysis of Census results and processes. Derby was one of these 15 authorities, receiving an increase of 7,800. Hammersmith and Fulham, Barnet and Welwyn Hatfield were not among the 32 LAs reviewed.

## **Annex - Details of the 15 clusters**

All 376 lower tier local authorities in England and Wales were analysed into 15 clusters, with the exception of the Isles of Scilly.

The characteristics of the clusters with the highest risk to the accuracy of MYEs are as follows:

### ***Cluster 1***

Central London Boroughs with the highest population density, rates of migration and young males, and well above national average proportions of students and non-white population:

Camden	Islington	Wandsworth
City of London	Kensington and Chelsea	Westminster
Hammersmith and Fulham	Lambeth	

### ***Cluster 2***

Outer London Boroughs together with some areas outside London. These LAs also exhibit above average population density and migration and rates of young males and students, but not as high as Cluster 1. The proportions of their population who are non-white are typically higher than in Cluster 1.

Barking and Dagenham	Harrow	Merton
Barnet	Hillingdon	Redbridge
Croydon	Hounslow	Slough UA
Enfield	Leicester UA	Waltham Forest
Greenwich	Luton UA	Watford

### ***Cluster 4***

Other Inner London Boroughs with the highest rates of non-white populations and above average population density and proportions of migrants, students and young males:

Brent	Haringey	Southwark
Ealing	Lewisham	Tower Hamlets
Hackney	Newham	

### ***Cluster 6***

University towns and cities with high proportions of students, young males and migrants, and somewhat above average rates of non-white populations:

Brighton and Hove UA	Manchester	Reading UA
Bristol UA	Norwich	Richmond upon Thames
Exeter	Nottingham UA	Southampton UA
Kingston upon Thames	Portsmouth UA	

### ***Cluster 8***

Large cities in the North and Midlands with above average proportions of non-white populations, students and young males but low rates of internal migration:

Birmingham	Derby UA	Preston
Blackburn with Darwen UA	Kirklees	Sandwell
Bradford	Leeds	Sheffield
Coventry	Liverpool	Wolverhampton

### ***Cluster 11***

Mainly university towns with high proportions of students but slightly below average population density and non-white populations:

Bath and North East Somerset UA	Cheltenham	Oadby and Wigston
Bournemouth UA	Durham	Runnymede
Canterbury	Guildford	Warwick
Cardiff	Lancaster	Welwyn Hatfield
Ceredigion	Lincoln	York UA
Charnwood	Newcastle upon Tyne	

### ***Cluster 14***

Oxford and Cambridge, which have the highest rates of students and internal migration and large proportions of young males. Research projects have already been set up to study student populations.

The remaining clusters, where the risk of incorrectly estimating the population is lower, can generally be characterised as follows:

### ***Cluster 3***

Mainly urban areas spread across England with above average population density and young males but lower than average values of the other variables:

Basildon	Havering	South Gloucestershire UA
Basingstoke and Deane	Hyndburn	Southend-on-Sea UA
Bedford	Ipswich	Stevenage
Bexley	Kettering	Stoke-on-Trent UA
Bromley	Kingston upon Hull UA	Sutton
Broxbourne	Medway UA	Swindon UA
Bury	Middlesbrough UA	Tamworth
Cannock Chase	Milton Keynes UA	Telford and Wrekin UA
Chelmsford	Northampton	Thurrock UA
Crawley	Pendle	Trafford
Gloucester	Peterborough UA	Worcester
Gravesham	Plymouth UA	Worthing
Harlow	Redditch	
Hastings	Salford	

### ***Cluster 5***

Mainly Southern parts of England with above average internal migration and Armed Forces:

Blackpool UA	Gosport	South Derbyshire
Bracknell Forest UA	Hart	South Oxfordshire
Broxtowe	Hertsmere	Spelthorne
Cherwell	Kennet	Surrey Heath
Chester	Mid Bedfordshire	Test Valley
Colchester	North Hertfordshire	Three Rivers
Dartford	North Kesteven	Tunbridge Wells
East Cambridgeshire	North Wiltshire	Vale of White Horse
East Hertfordshire	Reigate and Banstead	West Berkshire UA
East Northamptonshire	Rushcliffe	West Oxfordshire
Eastleigh	Rushmoor	Winchester
Elmbridge	St. Albans	Windsor and Maidenhead UA
Epping Forest	Salisbury	Woking
Epsom and Ewell	South Bedfordshire	Wokingham UA
Fareham	South Cambridgeshire	Wycombe

### *Cluster 7*

Urban areas in the North Midlands and Wales with below average values on all the risk factors:

Allerdale	Ellesmere Port and Neston	St. Helens
Barnsley	Flintshire	Sefton
Barrow-in-Furness	Gateshead	South Tyneside
Blaenau Gwent	Halton UA	Stockport
Blyth Valley	Hartlepool UA	Stockton-on-Tees UA
Bolton	Isle of Anglesey	Sunderland
Bridgnorth	Knowsley	Swansea
Caerphilly	Merthyr Tydfil	Tameside
Calderdale	Neath Port Talbot	Torfaen
Carlisle	Newport	Wakefield
Carmarthenshire	North East Lincolnshire UA	Walsall
Chesterfield	North Lincolnshire UA	Wansbeck
Copeland	North Tyneside	Warrington UA
Corby	Nuneaton and Bedworth	Wigan
Darlington UA	Oldham	Wirral
Derwentside	Redcar and Cleveland UA	Wrexham
Doncaster	Rhondda, Cynon, Taff	Wyre Forest
Dudley	Rochdale	
Easington	Rotherham	

### ***Cluster 9***

More rural areas with low population density and well below average values on the other risk factors:

Alnwick	Herefordshire UA	Rugby
Amber Valley	High Peak	Ryedale
Arun	Hinckley and Bosworth	Scarborough
Ashfield	Kerrier	Sedgefield
Ashford	King's Lynn and West Norfolk	Sedgemoor
Berwick-upon-Tweed	Macclesfield	Selby
Bolsover	Maldon	Shepway
Braintree	Mansfield	Shrewsbury and Atcham
Broadland	Melton	Solihull
Burnley	Mendip	South Holland
Caradon	Mid Suffolk	South Kesteven
Castle Point	New Forest	South Lakeland
Chester-le-Street	Newcastle-under-Lyme	South Ribble
Congleton	North Devon	South Somerset
Conwy	North East Derbyshire	Staffordshire Moorlands
Crewe and Nantwich	North Norfolk	Stroud
Denbighshire	North Somerset UA	Suffolk Coastal
Dover	North Warwickshire	Taunton Deane
Eden	North West Leicestershire	Tendring
Erewash	Oswestry	Thanet
Forest of Dean	Pembrokeshire	Tynedale
Gedling	Penwith	Vale of Glamorgan
Great Yarmouth	Powys	Vale Royal
Gwynedd	Restormel	Waveney
Hambleton	Ribble Valley	Wear Valley
Harrogate	Rochford	West Lancashire
Havant	Rossendale	West Wiltshire

### ***Cluster 10***

Sparsely populated areas with average rates of internal migration and above average proportions of prisoners:

Aylesbury Vale	East Riding of Yorkshire UA	Newark and Sherwood
Bassetlaw	East Staffordshire	North Dorset
Blaby	Fenland	North Shropshire
Boston	Fylde	St. Edmundsbury
Breckland	Harborough	South Staffordshire
Bridgend	Huntingdonshire	Stafford
Bromsgrove	Lichfield	Teignbridge
Dacorum	Maidstone	Wellingborough
Derbyshire Dales	Monmouthshire	Wychavon

### ***Cluster 12***

Areas with the highest proportions of prisoners average rates of migration and Armed Forces but below average on the other risk factors:

Castle Morpeth	Isle of Wight UA	Teesdale
Chorley	Rutland UA	West Devon
Daventry	Swale	Weymouth and Portland

### ***Cluster 13***

Mainly retirement areas with the lowest proportions of young males but above average internal migration:

Adur	Lewes	Stratford-on-Avon
Babergh	Malvern Hills	Tandridge
Brentwood	Mid Devon	Tewkesbury
Carrick	Mid Sussex	Tonbridge and Malling
Chichester	Mole Valley	Torbay UA
Chiltern	North Cornwall	Torridge
Christchurch	Poole UA	Uttlesford
Cotswold	Purbeck	Waverley
Craven	Rother	Wealden
East Devon	Sevenoaks	West Dorset
East Dorset	South Bucks	West Lindsey
East Hampshire	South Hams	West Somerset
East Lindsey	South Norfolk	Wyre
Eastbourne	South Northamptonshire	
Horsham	South Shropshire	

### ***Cluster 15***

Richmondshire and Forest Heath, which have the highest rates of Armed Forces and are above average for young males, but have few students.

This [map](#) shows the distribution of the clusters across England and Wales.