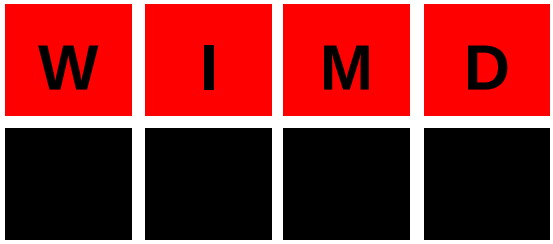


Welsh Index of Multiple Deprivation

Consultation on the proposed indicators for updating the Welsh Index of Multiple Deprivation

Welsh Assembly Government - Statistical Directorate

in association with

Contents

Introduction

The Welsh Index of Multiple Deprivation 2005 (WIMD) is the official measure of deprivation for small areas in Wales. It was developed for the Welsh Assembly Government by the Assembly's Statistical Directorate and the Local Government Data Unit ~ Wales. It replaced the index which was produced in 2000.

An updated index will be published in the summer of 2008, using more up-to-date data and improved indicators based on lessons learnt from the 2005 exercise.

The purpose of this consultation is to expose the current proposals for the domains and indicators for WIMD 2008 to critical review. We are consulting at this stage to allow time for consultation responses to be acted upon, where possible, prior to finalising the details for WIMD 2008.

This document details our current thinking on which domains will be included for 2008 as well as the indicators for each domain. Details of indicators considered but rejected for whatever reason (e.g. inappropriate, lack of data etc.) are also given including the reason for rejection.

The consultation is open to all individuals and organisations.

All consultation responses will be published on the Statistical Directorate web site along with a summary of the responses received in October 2007. The summary will include an initial reaction to the consultation responses, although for some issues more detailed investigation may be required. Normally, the name and address (or part of the address) of the author are published along with the response, as this gives credibility to the consultation exercise. If you do not wish to be identified as the author of your response, please state this expressly in writing to us.

The consultation exercise begins on 11 June 2007 and **7 September 2007 is the deadline for responses**.

WIMD Overview

The Welsh Index of Multiple Deprivation (WIMD) is the official measure of deprivation for small areas in Wales. WIMD 2005 was developed for the Welsh Assembly Government by the Assembly's Statistical Directorate and the Local Government Data Unit ~ Wales.

Multiple deprivation

Deprivation is a wider concept than poverty. Poverty means not having enough money (or other essentials) to get by. Deprivation refers to problems caused by a general lack of resources and opportunities (not just money).

Multiple deprivation, in the sense that it is used for the WIMD, cannot be measured directly—it is not some special sort of deprivation. It is a mixture of separate kinds of deprivation, each of which can be measured to some extent. But they cannot just be added together to make an overall index because they may interact and have more impact if found in certain combinations.

Index

An index is group of separate measurements which are combined into a single number. They are designed to show changes in a complicated variable like intelligence or industrial output. Values of the index can be compared with each other, sometimes over time and sometimes for different geographical areas.

For example, the Retail Price Index (RPI) is a well-known index based on the prices of goods bought by an average household. It includes everything from food and housing to entertainment. The RPI is used to measure changes in the cost of living.

Indexes are easy to work out and to understand. The problem is in deciding what to include and how much importance to give it—opinions could vary on both, and often do.

Construction of the WIMD

The WIMD 2005 is made up of seven separate domains (or kinds) of deprivation:

- income
- employment
- health
- education
- housing
- access to services
- environment

Each of them was based on a range of different indicators which meant that they were measured in different ways using different units. So before they could be combined the measurements had to be transformed to make them compatible. For example, if the height of something had been measured in metres and the weight in kilograms it would not make sense simply to add them together.

Income and employment were classed as the most important factors, and they were given the biggest weighting in the overall index. Although there is more to deprivation than poverty, not having enough money or a job is a big part of it.

England and Wales have been divided into Super Output Areas (SOA) each having roughly the same population. There are three levels: Lower Layer (the smallest), Middle Layer, and Upper Layer (the largest). There are 1,896 Lower Layer Super Output Areas in Wales each having about 1,500 people. Deprivation scores have been worked out for each of these areas: higher scores mean more deprivation. An area has a higher deprivation score than another one if the proportion of people living there who are classed as deprived is higher. An area itself is not deprived: it is the circumstances and lifestyles of the people living there that affect its

deprivation score. And it is important to remember that not everyone living in a deprived area is deprived—and that not all deprived people live in deprived areas.

What is the WIMD for?

It is important to be aware that the WIMD is not the only way to measure deprivation. WIMD has been developed for a particular purpose which is to measure concentrations of deprivation at a small area level. In trying to use the index there are two important questions to ask yourself:

- "am I interested in localised concentrations of deprivation or all deprived people?";
- " is deprivation actually concentrated in my area of interest?".

An analysis of the WIMD 2005 income and employment domains can illustrate these two issues.

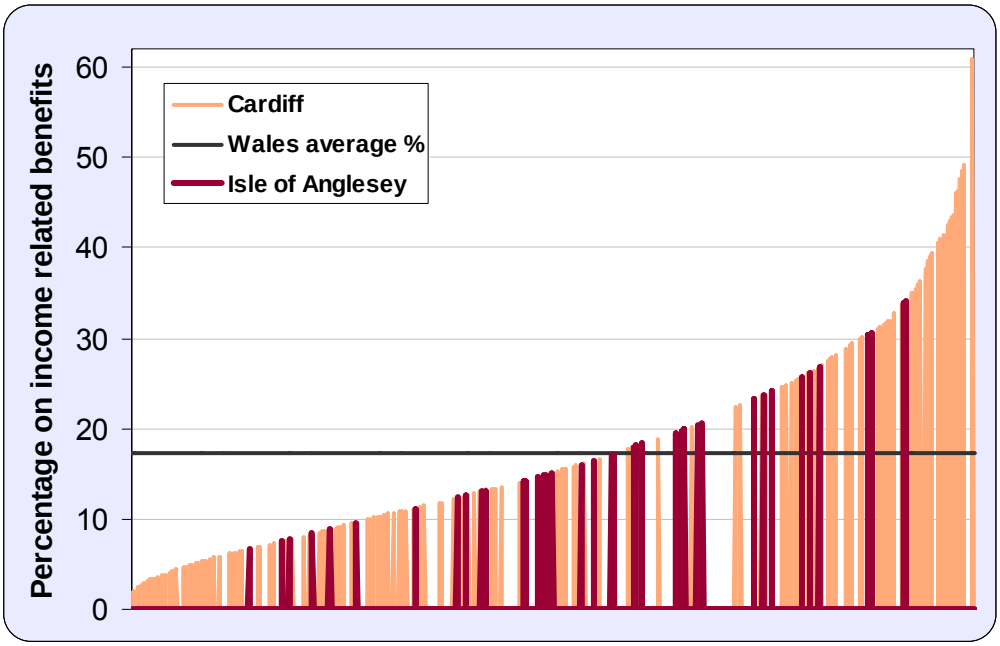
While there is a definite concentration of deprivation into a relatively small number of areas, by no means all deprived people live in an area of high deprivation. Almost a quarter of the income deprived live in areas within the most deprived 10% in Wales. This illustrates how well WIMD identifies clusters from within the overall population. Over half live in the most deprived 30%, but that still means that around a half of the income deprived in Wales are more sparsely spread over the remaining 70% of Wales. And even in the least deprived 10% of LSOAs there are still deprived people.

Table 1: Percentage of the deprived population living within different groupings of LSOAs, WIMD 2005

	Income domain	Employment domain
Most deprived 10% LSOAs	23%	19%
Most deprived 20% LSOAs	39%	35%
Most deprived 30% LSOAs	53%	48%
Least deprived 10% LSOAs	2%	4%

For WIMD 2005, the percentage on income related benefits in Wales was 17.4 per cent overall, and for Cardiff the figures was 17.9 per cent and 17.6 per cent in the Isle of Angelsey. So in terms of overall levels of income deprivation, both Cardiff and the Isle of Angelsey were comparable and also similar to the overall Wales average. However, as can be seen in Chart 1, the distribution of income deprivation within each local authority was quite different. While there was still a big difference between the most and least deprived LSOAs in the Isle of Anglesey, there was nothing like the range of variation seen in Cardiff. In particular, there were no LSOAs in Anglesey at the very extremes.

Chart 1: Percentage on income realted benefits by LSOA, ranked position in Wales, WIMD 2005



WIMD is suited to uses where the interest is in areas with high concentrations of deprivation. If all deprived individuals are the focus of attention then a different approach is needed, however, it is still going to be based on the same or similar indicators. In other words, WIMD is a way of combining deprivation measures for a particular purpose.

Uses of the index

The index can be used for:

- giving an overall deprivation score for each of the 1,896 Lower Layer SOAs in Wales;
- giving scores for the separate deprivation domains for each of the Lower Layer SOAs;
- comparing the deprivation scores for two or more of the Lower Layer SOAs;
- ranking the scores for all 1,896 SOAs or for a group of them (like those in a local authority), so that the SOAs can be put in order from the most deprived to the least;
- comparing two or more local authorities (or other groups of aggregated SOAs) by looking at the proportion of the SOAs in the local authority in the most deprived (say) ten per cent in all of Wales.

WIMD 2005 has actually been used for the following:

- The 100 most deprived electoral divisions in Wales as identified in the 2000 Index were included among the original 142 Communities First areas. Following the publication of WIMD 2005, a consultation exercise was launched on how the results could best be used and any possible expansion of the existing programme. 46 Lower Super Output Areas (LSOAs) in the top 10% of deprivation on the new index were identified as outside the existing Communities First footprint. The Minister invited the 46 new areas to apply for inclusion in the Communities First Programme in consultation with the local community by 30 March 2007. Submissions for all the prospective new areas have been received and

are currently being considered, or have requested additional time to prepare a more detailed proposal, for inclusion as Communities First areas. Additionally, 2 LSOAs ranked in the top 10% of deprivation but which contained only an existing pocket of deprivation have submitted proposals to extend their area of operation. To date 22 of the invited areas have now been formally included within the programme, these being the LSOAs in Cardiff (17 areas), Newport (1), Caerphilly (2) and Merthyr Tydfil (2).

- In the calculation of Standard Spending Assessments (SSAs) for the 22 unitary authorities. Just under 1 per cent of the total SSA is allocated on the basis of the overall index, but focussing on the most deprived areas. Around 25 per cent is allocated on the basis of other measures related to deprivation and 6 per cent on sparsity measures.
- The deprivation grant uses the most deprived areas, as defined by the overall index scores from WIMD 2000, to allocate this funding to the 22 unitary authorities. A review of the formula is currently being carried out and the use of the WIMD 2005 results, as opposed to other deprivation measures, forms part of that review.
- The education domain of WIMD is used to provide a deprivation uplift for funding allocations within the National Planning and Funding System . This uplift is designed to cover the additional cost of widening provision and delivering provision to those from educationally deprived backgrounds. The funding model applies the uplift to provision undertaken by learners residing in areas that are ranked in the top three deciles of the educational domain of the Welsh Index of Multiple Deprivation.
- The 'Flying Start' programme used the overall index values from WIMD 2005, ranked within each local authority, as a means to assess nominations for participating school catchment areas. Other information was also taken into account in assessing the nominations.
- The overall index, but with the health domain excluded, has been used to assess the progress of health gain targets across different levels of deprivation. This has been done by breaking down Wales level data into areas that fall with each of the quintiles of the WIMD distribution (that is, into five equal groups from the fifth most deprived down to the fifth least deprived).
- OFWAT have used the index for looking at the relationship between deprivation and levels of debt.
- The Lottery Fund have used the overall index to help allocate funds to deprived areas.
- More generally, WIMD has been used as supporting evidence in bids for European funding, targeting local service delivery and in analysis of the links between localised deprivation and various problems (such as health issues).

The index is not the only deprivation measure used for funding allocations. A number of Welsh Assembly Government grant allocations are based on formulae that use other measures of deprivation (ones not designed to identify concentrations) as well as measures of sparsity.

For income and employment it is possible to get the numbers of people classed as deprived living in an area (as well as a deprivation score). For the other five deprivation domains only the scores are available (because they are worked out using a range of indicators rather than just numbers of people). The data on which the index is based can be used to check for changes over time.

Limitations of the index

There are several ways in which it is tempting to use the index, but which do not really make sense or which lead to unreliable results.

Tiny differences in the deprivation scores do not mean anything. There's no point rummaging among the decimal places trying to show that one SOA is more deprived than another.

You cannot say how much more deprived one SOA is than another. If area A has a score of 40 and area B one of 20, it does not follow that A is twice as deprived as B. Or if A were ranked as the 100th most deprived area and B the 300th, it does not follow that A is three times more deprived than B. There can be a strong temptation jump to conclusions like these, but they are definitely wrong.

It is an index of deprivation not affluence. If one area is much lower down the ranked list than another then you can say that it is less deprived but you cannot say it is more affluent. The index is not based on the factors which mean that a place is affluent. Every area has people who are deprived and people who are affluent, but the index counts only those classed as deprived. It makes no difference whether the rest are nearly deprived but not quite, fairly well-off, or really rich.

The deprivation scores from the 2005 index cannot be compared with those from 2000. The areas on which the indexes are based are different (SOAs in 2005, electoral divisions in 2000), and the indexes were worked out in different ways. But even if these changes had not been made, comparisons would still not be valid. An area's score is affected by the scores of every other area; so it is impossible to tell whether a change in score is a real change in the deprivation level of an area, or whether it is a relative change due to the scores of other areas going up or down.

The deprivation scores cannot be compared with those from the deprivation indexes of other UK countries. To be able to compare an area in Wales with one in (say) England would mean having to design a joint England and Wales index—having two separate indexes would not work.

There are no official local authority deprivation scores. Local authority scores can be worked out—but there are several ways to do it, and there is no single obviously right way to do it.

Overview of development for 2008

To take forward the task of considering indicators for WIMD 2008, eight separate domain working groups were established:

Income

Employment

Education, Skills and Training

Health

Geographical Access to Services

Housing

Local Physical Environment

Community Safety

Members of these groups were drawn from a wide range of organisation with relevant expertise, and a list of members is given in Appendix A.

Each domain working group was tasked with examining the issues outstanding from WIMD 2005 plus any developments in the availability of data and to produce recommendations for possible indicators for WIMD 2008. Specifically, these sub-groups were expected to:

- Review the indicators and data sources used for WIMD2005;
- Discuss and agree proposed indicators for their domain, including the development of any new indicators as required;
- Ensure all indicators are tested and signed off as fit-for-purpose (including solutions to any anomalies identified and missing data);
- Oversee the production of technical guidance about the indicators chosen for inclusion in final reports including appropriate use of the individual indicators and the domain;

Alongside this process, the Welsh Assembly Government Statistical Directorate and the Local Government Data Unit ~ Wales hold six-monthly meetings with colleagues from the other 3 UK nations to share experiences. In this way developments in the other 3 nations, including relevant research projects, have been fed into the domain working group process in Wales.

Geographic Unit

It is proposed that the geographic areas used in the calculation of WIMD 2008 are the 1,896 Lower Layer Super Output Areas. These are the same areas used for WIMD 2005. It is a consistent statistical geography and is in line with the approach used by the other 3 UK nations for their indexes.

Super Output Areas

Following the 2001 Census of Population, the Office for National Statistics developed a new geographic hierarchy called Super Output Areas (SOAs). They were designed to improve the reporting of small area statistics in England and Wales. Their first statistical application was for the Index of Deprivation for England in 2004, which led to them being widely used within local government. They have been increasingly used for data on the Neighbourhood Statistics (NeSS) website. It is anticipated that they will eventually become a standard for the production of National Statistics and will be used more generally.

Electoral divisions (previously known as wards) were the basic geographical units used for the Welsh Index of Multiple Deprivation 2000, but there were disadvantages with this approach.

Electoral divisions vary greatly in size, from around 1,000 people to 20,000 (in Wales). This is not ideal for making comparisons throughout Wales, and it also means that data which can safely be released for larger electoral divisions may not be released for smaller electoral divisions due to disclosure rules (that is, the need to protect the confidentiality of individuals).

The boundaries of electoral divisions change. This creates problems when trying to compare data from different time periods. ONS decided to develop a range of areas that would be of consistent size and whose boundaries would not change. These would be built from groups of the Output Areas (OAs) used for the 2001 Census, and would be known as Super Output Areas (SOAs).

There are three layers of SOAs: Lower Layer, Middle Layer, and Upper Layer. This was because disclosure requirements mean that some sets of data could be released for much smaller areas than others. So to support a range of potential data requirements it was decided to create these three SOA layers. These are the constraints.

- A Lower Layer SOA must have a minimum population of 1,000. The mean size of all the Lower Layer SOAs must be close to 1,500. They are built from groups of Census OAs (usually between four and six).
- A Middle Layer SOA must have a minimum population of 5,000. The mean size of all the Middle Layer SOAs must be close to 7,200
- An Upper Layer SOA must have a minimum population of about 25,000: the formal definition has not been finalised.

There are 1,896 Lower Layer SOAs in Wales (34,378 in England and Wales). They were generated by a computer program which merged OAs taking into account population size, mutual proximity, and social homogeneity. The boundaries were released in February 2004.

The following table gives the number of Lower Layer SOAs in each local authority in Wales.

Table 2: Number of Lower Layer SOAs by local authority

	number of Lower Layer SOAs
Isle of Anglesey	44
Gwynedd	75
Conwy	71
Denbighshire	58
Flintshire	92
Wrexham	85
Powys	80
Ceredigion	47
Pembrokeshire	71
Carmarthenshire	112
Swansea	147
Neath Port Talbot	91
Bridgend	85
The Vale of Glamorgan	78
Cardiff	203
Rhondda, Cynon, Taff	152
Merthyr Tydfil	36
Caerphilly	110
Blaenau Gwent	47
Torfaen	60
Monmouthshire	58
Newport	94

The effect of changing from electoral divisions to LSOAs for WIMD 2005

It is not straightforward to give a brief summary of the effect of changing the geographic unit. It can only be attempted for the income or employment domains as they are the only domains made up of simple counts of individuals. While all LSOAs have roughly of the same population size, the numbers of people living in different electoral divisions vary a lot (from less than 1,000 to around 20,000 people). This means that:

- the number of people living in the 10 per cent most deprived areas can vary a lot depending on whether the deprived areas are large or small electoral divisions;
- the number of areas a local authority has within the top 10 per cent most deprived does not give a fair indication of the proportion of the population of the authority covered when electoral divisions are used;
- as there are many more LSOAs than electoral divisions, and this relationship varies across Wales, having more or fewer LSOAs in the top 10 per cent most deprived does not necessarily mean more of less of the authority's population has been included.

Reworking the WIMD 2005 income domain on 2001 Census wards (electoral divisions or combinations of these for Wales) provides a good illustration of the effects of the geography change. And Cardiff provides a good example of the difficulties highlighted above:

- Cardiff had almost 20 per cent of the most deprived LSOAs in Wales but only around 10 per cent of the most deprived electoral divisions in Wales;
- However, only 18 per cent of the LSOAs in Cardiff were in the 10 per cent most deprived while 28 per cent of electoral divisions in Cardiff would have been in the 10 per cent most deprived.

As even the number of people living in the most deprived areas depends so much on which electoral divisions come out as very deprived, the most sensible comparison is between the percentage of the population living in the most deprived areas. Table 3 shows the effect on the income domain. For most authorities the differences are quite small, but for authorities having electoral divisions with large populations the differences are more marked. The highlighted rows show local authorities that have a lower proportion of their population in the 10 per cent most deprived under an LSOA geography. However, it is quite apparent that there are no systematic pattern, either urban versus rural, cities versus other authorities or whether or not an authority has large or small electoral divisions.

Table 3: Local authority share of the population living in an area within the 10 per cent most income deprived areas.[1]

	LSOA geography	Electoral division geography
Isle of Anglesey	1.1	1.3
Gwynedd	1.7	1.2
Conwy	2.0	0.5
Denbighshire	3.2	2.3
Flintshire	1.5	1.6
Wrexham	3.1	2.1
Powys	0.0	0.8
Ceredigion	0.4	0.6
Pembrokeshire	1.6	2.1
Carmarthenshire	2.7	3.7
Swansea	14.7	10.7
Neath Port Talbot	6.8	7.9
Bridgend	3.1	3.2
The Vale of Glamorgan	3.1	3.5
Cardiff	19.5	22.6
Rhondda Cynon Taf	11.7	15.2
Merthyr Tydfil	3.6	2.5
Caerphilly	6.0	6.5
Blaenau Gwent	4.5	3.6
Torfaen	1.2	1.6
Monmouthshire	0.0	0.0
Newport	8.3	6.3

source: WIMD 2005 income domain and the same data reworked on 2001 Census wards

[1] Calculated as the population within each authority living in an area within the 10% most deprived in Wales, divided by the total population in Wales living in such areas.

Domains and domain indicators

Appendix B shows the indicators and domains used for WIMD 2005. The following sections provide details of the work done since the publication of WIMD 2005 to examine possible improvements and new indicators for the existing domains as well as progress on developing a new community safety domain. Each section provides a summary of the current proposals for the domain as well as a list of issues that we particularly wish to receive comments on.

Income

The purpose of this domain is to capture the extent of deprivation relating to income at a small area level across Wales, and focuses on the proportion of people living in households with income below a defined threshold or claiming benefits relating to low incomes.

Proposed indicators

The indicators used in the WIMD2005 were reviewed, together with those indicators rejected for WIMD 2005. Potential new indicators were also discussed. The options for the proposed indicators are listed below.

Preferred Option

1. **LSOA Income estimates** - a project currently being undertaken by Essex University is looking at modelling income estimates at the LSOA level, using data from the 2001 Census and the annual Family Resource Survey. The results will give the income estimates for each LSOA, the proportion of the households below a threshold income level, e.g. 60% of median household income (DWP standard measure). This is considered the preferred option as it looks directly at income rather than proxies for income e.g. income related benefits. Results for Wales are due during the consultation period and if robust will be proposed as the sole indicator for the income domain for WIMD 2008.

Alternative Option

2. **The second option are based on those indicators used for WIMD 2005, namely:**

- **Adults and Children in Income Support Households** - Income Support is intended to help people on low incomes who do not have to be available for employment. It can normally be claimed by people in certain circumstances (mainly lone parents and the sick and disabled) who are:
  - ages 18-60;
  - working less than 16 hours a week (and / or with a partner working less than 24 hours a week);
  - not required to be available for full-time employment;
  - and in receipt of insufficient income to meet prescribed needs.
- **Adults and Children in Pension Credit Households** - is intended to help people on low incomes who do not have to be available for employment. It can normally be claimed by people in certain circumstances (mainly the sick and disabled) who are:
  - ages 60 and over;
  - working less than 16 hours a week (and / or with a partner working less than 24 hours a week);
  - not required to be available for full-time employment;
  - and in receipt of insufficient income to meet prescribed needs.
- **Adults and Children in Income-Based Job Seekers Allowance households** - this indicator captures all persons living in a household dependent on Income-Based Job Seekers Allowance (JSA). JSA can be claimed by people aged 16 and over who are

available for and actively seeking employment, including those in remunerative work for less than 16 hours a week on average, and by people on a government training scheme.

- **Adults and Children in Tax Credit (Child Tax Credit and Working Tax Credit) households below a low income threshold** - All persons in a household dependent on the following Tax Credits:

  1) Child Tax Credit - an income-based benefit for low-income families who have responsibility for a child(ren) under 16, or under-20 if in full-time non-advanced education or approved training.

  2) Working Tax Credit - an income-based benefit for working adults available to households with adults in one of 4 categories:

     - Work 16 hours or more a week and meet one of the following criteria:

       a) have responsibility for a child;

       b) have a disability that puts you at a disadvantage in getting a job;

       c) qualify for a 50-plus element;

     OR

     - Work 30 hours or more a week and are 25 or over.

- **National Asylum Support Service (NASS) supported asylum seekers in Wales in receipt of subsistence only and accommodation support** - represents the presence of asylum seekers in Wales and contains people who are not eligible for the other benefits in the domain, yet are income deprived.

Other indicators discussed

Housing Benefits - this was a suggested indicator for WIMD 2008, resulting from feedback from WIMD 2005. However this indicator would only apply to rented accommodation. Also, people already claiming other benefits included in WIMD would have to be excluded from this count, resulting in either lack of availability of data or very small numbers. For these reasons this indicator was not proposed.

Outstanding issues

The only outstanding data issues surround the income estimates and tax credits. The results of the income estimates project are due out by the end of June 2008 and these will be analysed for their robustness for inclusion in WIMD 2008. Should the income estimates prove reliable and robust these will be the indicator used for WIMD 2008. With the new tax credits to those used for WIMD 2005, the new data will also need to be analysed for its robustness before being included in WIMD 2008 (if needed).

Issues for response

There are specific questions that will help shape the development of the WIMD2008. These questions are listed below. They are not exhaustive, nor should limit any response that is relevant to this domain.

- Is the definition for the domain appropriate?
- Are the indicators chosen appropriate?
- Are the definitions given appropriate for the indicators?
- Have any essential indicators been missed, providing the data set is available on a consistent national, reliable basis?

Employment

The purpose of this domain is to capture the extent of deprivation relating to employment at a small area level across Wales, and focuses on the proportion of working age people claiming out-of-work benefits.

Proposed indicators

The indicators used in the WIMD2005 were reviewed, together with those indicators rejected for WIMD 2005. Potential new indicators were also discussed. The proposed indicators are listed below.

Claimants of Unemployment-related benefits - all persons of working age (18-59 for women and 18-64 for men) who are claiming Job Seekers Allowance, available to those who are actively seeking work, but not in work. It was considered including those aged 16 and above, however unrounded LSOA data for the 16-17 age group is particularly hard to get hold of, also those that are aged 16 and 17 are most unlikely to be able to claim such benefits when living at home with parents.

Claimants of Incapacity Benefit/Severe Disablement Allowance - these two benefits were separately counted for WIMD 2005. However, due to the ever decreasing numbers of those receiving Severe Disablement Allowance (since 2001 there have been no new claimants), it was decided to combine the two groups. Incapacity benefit claimants are those of working age who are unable to work due to illness or disability and who meet certain contribution conditions. All 3 rate categories of claimants are included (short-term lower, short-term higher and long term). Severe Disablement Allowance is available to those people who are incapable of work and do not satisfy the contribution conditions of Incapacity Benefit.

Participants on New Deal for Young People and Intensive Activity Period (for New Deal 25+) not included in unemployment-related benefit counts - young people (18-24) who have been claiming JSA for at least 6 months and persons 25+ who have been claiming JSA for at least 2 years and have therefore moved into the New Deal programme.

Participants on New Deal for Lone Parents - a new indicator, included to incorporate a large and statistically important group (lone parents), not included in other indicators for this domain.

Participants on New Deal for those aged 50+ - a new indicator, again included to incorporate a major group of working age people who are seeking work.

Other indicators discussed

There were no other indicators discussed as none were suggested either as part of WIMD 2005 feedback or new for WIMD 2008.

Outstanding issues

Data is to be collected and tested for the 2 new indicators to ensure the data is robust and provides reliable results at the LSOA level. Should the data prove not to be robust for each indicator that individual indicator will not be used in the index.

Issues for response

There are specific questions that will help shape the development of the WIMD2008. These questions are listed below. They are not exhaustive, nor should limit any response that is relevant to this domain.

- Is the definition for the domain appropriate?
- Are the indicators chosen appropriate?
- Are the definitions given appropriate for the indicators?
- Have any essential indicators been missed, providing the data set is available on a consistent national, reliable basis?

Education

The purpose of this domain is to capture the extent of deprivation relating to education, training and skills. It is designed to reflect the 'stock' and 'flow' or educational disadvantage within an area, by capturing low attainment among children and young people and the lack of qualifications and skills in adults.

Proposed indicators

The indicators used in the WIMD2005 were reviewed, together with those indicators rejected for WIMD 2005. Potential new indicators were also discussed. The proposed indicators are those used for WIMD 2005 with two additions. These are listed below.

Key Stage 2, Average points scores - 3 years worth of teacher assessment scores for pupils being taught in the National Curriculum Year group 6 (in 2005, 2006 and 2007) to be used thus reducing the impact of small numbers at the LSOA level. Pupils attaining level 'N' are to be excluded from the calculations because a grading of N denotes no information available so no accurate point score can be applied to such pupils. Pupils attaining level 'N' were included for WIMD 2005, with a score of 15, following the same methodology of the Department for Education and Skills (DfES) in England. Independent schools to be excluded for WIMD 2008, because it is not statutory for them to submit KS2 information to the Welsh Assembly Government and also as they do not complete PLASC there is no information available at postcode level for individual pupils. However this should be re-visited for possible inclusion for future updates of WIMD.

Key Stage 3 Average points scores - 3 years worth of teacher assessment scores of pupils being taught in the National Curriculum Year group 9 (in 2005, 2006 and 2007) thus reducing the impact of small numbers at the LSOA level. Pupils attaining level 'N' to be excluded from the calculations. As for the Key Stage 2 indicator, Independent schools to be excluded for WIMD 2008, because it is not statutory for them to submit KS2 information to the Welsh Assembly Government and also as they do not complete PLASC there is no information available at postcode level for individual pupils. However this should be re-visited for possible inclusion for future updates of WIMD.

Key Stage 4 Average points scores - 3 years worth of attainment of pupils aged 15 as at 31st August (in 2003-04, 2004-05 and 2005-06). All approved GCSE and GNVQ results are included for 2003-04 and 2004-05, with wider points scores (which include all externally approved pre-16 qualifications) used for 2005-06.

Proportion of people aged 16 – 17 not in Further Education - a new indicator for WIMD 2008, formerly part of a combined people aged 16-19 in further or higher education in WIMD 2005. Only 16-17 years olds are included for Further Education, as this has the more robust data available. Only 2006/07 data to be used.

Proportion of people not entering Higher Education aged 18 – 19 - a new indicator for WIMD 2008. Only 2005-06/07 data to be used as this is all that is available. For further updates of WIMD, 3 years worth of data should be considered.

Primary School all absence rate - a new indicator from WIMD 2005. It is agreed that this is an important indicator and should include all absences, as for Secondary Schools. This indicator will provide a balance with the Secondary School Absence rates previously included. Only 2 years worth of data (2005-06 and 2006-07) will be used for WIMD 2008 due to data reliability and coverage issues for earlier years, however future WIMD updates should use 3 years worth of data.

Secondary School all absence rate - as for WIMD 2005 all absences (authorised and unauthorised) will be included in calculations. 3 years worth of data will be used (2004-05, 2005-06 and 2006-07). Independent schools still not included due to unavailability of data for all schools.

Number of adults aged 25 – 59/64 with no qualifications - all adults aged 25 to retirement age (59 for women and 64 for men) who reported at the time of the 2001 Census as holding no qualifications or qualifications which do not reach the standard to be categorised as Level 1 qualifications.

Other indicators discussed

School exclusions - although this is a good indicator of educational deprivation the data is not robust enough, with the available only at LEA level data and therefore unable to be disaggregated to LSOA level.

Proportion of those under 21 not entering higher education - This indicator is used in the English Index of Multiple Deprivation, however for Wales data is unable to be modelled reliably from census 2001 data and subsequent mid year estimates. Part of this indicator is capture in the indicator 'Proportion of people not entering Higher Education aged 18-19' which can be robustly calculated.

Literacy and Numeracy - Considered a good indicator, but no data available on this at the LSOA level.

Post 16 basic skills - similarly a good indicator but with no data available at the LSOA level.

Pupil mobility - considered a relevant factor but current data not robust enough for it to be included for WIMD 2008.

Lifelong learning - covering adults from 25 to pensionable age that are not currently undertaking any education or skills training. However, data sources are limited and this indicator could also highlight areas as deprived which in reality contain a large number of people educated to graduate or even postgraduate level, if they finished their training before the age of 25 or have not undertaken training in the year in which the data was taken. The indicator capturing no or low qualifications was deemed a better determinant of educational deprivation for people aged 25 to pensionable age.

Other issues

For WIMD 2005 shrinkage was used. Subsequently in a review of the Scottish Index of Deprivation, it was recommended that shrinkage should no longer be used. Tests on Welsh data have shown that not using shrinkage has no significant impact on the groupings of the LSOAs e.g. the vast majority top 10% of LSOAs stay within the top 10% within the education domain. The use of multiple years of data will also help reduce the impact of small numbers.

Issues for response

There are specific questions that will help shape the development of the WIMD2008. These questions are listed below. They are not exhaustive, nor should limit any response that is relevant to this domain.

- Is the definition for the domain appropriate?
- Are the indicators chosen appropriate?
- Are the definitions given appropriate for the indicators?
- Have any essential indicators been missed, providing the data set is available on a consistent national, reliable basis?

Health

The purpose of this domain is to capture the degree to which people are deprived of good health.

Proposed indicators

The indicators used in the WIMD2005 were reviewed, together with those indicators rejected for WIMD 2005. Potential new indicators were also discussed. The proposed indicators are the same as those used for WIMD 2005 with one addition and are listed below:

Limiting long-term illness (LLTI) - while this is a subjective measure, it is still a valid indicator of morbidity in the community. The way people feel about their health is important and affects their quality of life.

Standardised all-cause death rate - a very robust indicator, with poor health manifesting itself both through a poorer quality of life but also in lower life expectancy that can be captured through age and sex standardised death rates. Due to small numbers at the LSOA level 10 years worth of data will be used.

Standardised cancer incidence rate - robust data and easily available rates cancer incidence have a strong link with deprivation. Due to small numbers at the LSOA level 10 years worth of data will be used.

Singleton low birth weights - the numbers are low, however there is 10 years worth of data available. Evidence suggests that low birth rate is linked to the mother's lifestyle and health. Low birth rate can also cause problems for a baby in later life increasing the risk of chronic diseases. It is also one of the child poverty objectives. There were data quality issues for WIMD 2005, however the data is now thought to be more robust and issues have been resolved. The data is to be run as it was for WIMD 2005, to check the data, and low birth weight rates will be included if the data is robust enough.

Other indicators discussed

Diabetes - no suitable data available.

Mental Health - no suitable data available.

Oral health - data is not available at the lower area level.

Immunisations - influenced by media campaigns and can be opted out of. Data is also not available at the lower area level.

Teenage conceptions - Data only available at the ward level.

Coronary Heart Disease - looking at mortality from this was discussed, however this was thought to be influenced by availability, quality and effectiveness of treatment and therefore would measure a service and that is not part of this domain.

Cirrhosis of the liver - data not available at the lower area level.

Emergency admissions - it was suggested that this could be an indicator of deprivation with the suggestion that less deprived being more likely to be proactive about their health and visit their doctor more regularly. However elective care elective care depends on waiting times and this is a measure of health care rather than health.

Breastfeeding - no reliable data available at the lower area level.

Childhood obesity - an ideal indicator for this domain, however a suitable data collection system not due to be implemented until 2008 and should therefore be considered for future publications of WIMD.

Sexual health - data is only available on those that seek treatment, thus not providing an accurate enough picture at the lower area level.

Substance misuse - data only available on those that seek treatment and is also influenced by the availability of the treatment, thus not providing an accurate picture of prevalence at the lower area level.

Injury mortality - evidence suggest that injury rates are higher in more deprived areas, however due to the small numbers involved this indicator cannot be included.

Outstanding issues

The method of standardisation is still to be decided. The standardisation technique used for WIMD 2005 was direct standardisation. However, following a review of the index of deprivation in Scotland, indirect standardisation was used for their most recent publication in 2006. Comments on the method of standardisation to use will be part of the consultation.

Not using the shrinkage technique for the Education domain had little impact on the results, but the issue of small numbers is that much more important for the Health domain and removing shrinkage has a large impact on the ranking of LSOAs. Current proposals are to retain shrinkage for the Health domain.

Issues for response

There are specific questions that will help shape the development of the WIMD2008. These questions are listed below. They are not exhaustive, nor should limit any response that is relevant to this domain.

- Is the definition for the domain appropriate?
- Are the indicators chosen appropriate?
- Are the definitions given appropriate for the indicators?
- Have any essential indicators been missed, providing the data set is available on a consistent national, reliable basis?

Access to Services

The access to services domain is to illustrate the deprivation as a result of a household's inability to access a range of services, considered necessary for day-to-day living. The access will be by bus and walking, to reflect the members of the household will not all be able to use private motorised transport (children, people with disability and when the vehicles available to a household are already in use).

The way that this domain measures deprivation is different from the others in Index. This domain shows areas where people may have difficulty accessing services. The number of people who actually experience difficulty is not known. This difference may be because the people have access to their own transport or because specific individuals do not require a particular service.

Proposed indicators

The indicators used in the WIMD2005 were reviewed to ensure they were still appropriate, available and of sufficient quality to use. Indicators rejected for WIMD 2005, and potential new indicators were also discussed. The proposed indicators are listed below.

Food shops – the shop should be capable of providing basic provisions e.g. bread, milk. These items are required frequently and need to be within a short travel time. For WIMD 2005 this was 10 minutes travel time.

GP Surgeries – this is to cater for local primary health care needs. Full, branch and part time surgeries would be included as in WIMD2005. For WIMD 2005 a travel time of 15 minutes was used.

Primary Schools – excluding nursery, special and independent. Primary schooling is for 4 to 11 year old children and in view of the age is considered a local facility with 15 minute travel time used in WIMD 2005.

Post office – many people in a community require weekly or more frequent access to a post office. This provides communications, advice, assistance, bill payments, collecting cash and other services. The travel time would reflect both frequency of use and range of services and people the post office covers. There are now mobile offices and the stopping points of these will be mapped wherever possible. For WIMD 2005 a 15 minute travel time was used.

Public library – In 2005 this was for static libraries only and will continue as this for 2008. Libraries offer services in addition to book lending and are another part of the communications network with the public. A travel time would be chosen to reflect the relatively local nature of this service, for WIMD 2005 this was 15 minutes.

Leisure centre – these would be included to provide opportunities to cater for the health and well-being of the community. The travel time would be set to recognise this facility could not be local, but should be accessible within a reasonable period. For WIMD 2005 a travel time of 20 minutes was used.

NHS Dentist – This indicator works on the presence of a NHS surgery, not whether the practice has capacity for more people to register. The travel time would be chosen to recognise the service cannot be totally local, but within a reasonable travel time. This travel time used for WIMD 2005 was 20 minutes.

Secondary Schools - excluding special and independent. Secondary schools are not just used by school age pupils and provide other resources for the wider community throughout the week. The travel time would be chosen to recognise the resource cannot be totally local, but within a reasonable travel time. Thus a 30 minutes travel time was used for WIMD 2005.

Transport nodes – A new indicator was proposed to show access to long distance transport services. The proximity of transport nodes to each household could be computed for coaches and rail routes. The travel time proposed is 30 minutes, to reflect the strategic, rather than local nature of long distance services.

Other indicators discussed

Mobile libraries - were proposed to be included as part of the libraries indicator. Although some data is available the exact location of stops is changing regularly. For this reason it was considered unsuitable for inclusion for WIMD 2008.

The mobile phone network - was proposed for inclusion. There are issues with the supply of such data from the operators, who will not release mapped data of their services, they only permit inquiries at individual points. This is the operator's interpretation of a ruling that coverage data should be made available. As most of Wales is covered by the networks anyway, it was considered that the indicator would not offer a distribution of values suitable for an index of this sort.

Public open space - was considered as a health and social well-being service to the public. There is a considerable range of types of open space, not all of which are suitable for public access. When in more rural areas it is arguable whether land which is in private ownership, but where access is not prohibited could be considered alongside registered public open space or common land. The problems with both definitions and reliable data sets to identify accessible land have led to a rejection of this indicator.

Cash point machines - are important for the supply of cash for day-to-day living. The locations of machines can be obtained from a commercial data set for the whole of Wales. The issue of cash is not only from cash machines, but more and more from cash back in retail premises. It is not possible to differentiate between machines that charge a percentage or flat fee for the service and the free ones. This led to an inconsistent indicator which was not felt to be usable.

Access to forested land - was considered as a health and social well-being service to the public. The forested land is not always suitable for public access from a health and safety viewpoint, so was not considered suitable for this domain. This may also be a function of affluence in terms of affording the travel cost to the forest area.

Child care facilities - offer opportunities to undertake employment or education. The facilities are often dependent on the ability to pay, rather than location. As only publicly accessible services (like NHS dentists rather than private ones) are proposed for inclusion, it was felt this should go into another domain or be excluded.

Citizens advice bureau - was considered as a valuable service available to all, but the availability of telephone services means it is not strictly a local service.

Telephone boxes - are under a requirement for cover over Wales, thus would not reflect deprivation, but a uniform distribution.

The legal services - were thought not to be a local service, nor essential to a "normal day-to-day" life. The data on legal practices is available from Neighbourhood Statistics, but was not considered for inclusion.

Other issues

Travel times

- **Reasonable time -** For WIMD 2005 the travel time thresholds chosen were those times that provided a good distribution of LSOAs when looking at the proportion of households within an LSOA that are above the time threshold. These times were based on estimates of what was thought was a reasonable time to be able to access any particular service, depending on those using the service, the importance of the service to everyday life and how often an individual would use that service. The times used for WIMD 2005 will form the basis for starting the calculations for WIMD 2008, with times possibly moving up or down according to the distributions the analysis gives.
- **Measuring travel times for lower super output areas -** The WIMD 2005 used the proportion of households in a LSOA which were able to travel to a service within a set time. WIMD 2008 proposes to use this again but in conjunction with the average of the time for every household in a LSOA to travel to the services identified in the domain. The change is to try to provide a better distribution of times and not to ignore those journeys which were marginally above the journey cut off time (e.g. 21 minutes with a cut off of 20 minutes).
- **Computer model for calculating travel times -** The production of travel times for complex geographic areas is extremely computation intensive. The WIMD 2005 access to services domain was produced from the Accession GIS computer package (www.accessiongis.org.uk) commissioned by Department for Transport in England (DfT). Since that time the DfT has worked with partners to produce a new computer system which works on sparse matrix techniques. The new methodology provides a less computational intensive technique which can reduce the need for approximations to reduce the computational burden. The issue will be whether this technique produces results in a way comparable to that used in 2005.

Buffer areas

In order to catch spillage over the border with England, a buffer area of 30 minutes travel time has been considered appropriate. This should pick up trips by residents of Wales to services located just over the border into England.

Flexible public transport

The existence of other than timetabled public transport services was noted. Dial-a-ride, taxi and other flexible, yet public transport, schemes exist, but a method of including such services has not been established. It may be possible to model the delay to service arrival as the time from base location to a particular location, but this in itself is a modelling exercise and assumes the vehicle is available and not used for journeys shared with other passengers (thus extending the route / time). A taxi type service may not always start from the same base point, so again would be impossible to model. There are also issues of obtaining Wales wide reliable data sets.

Day of week for travel

To model the journey to services by timetabled bus and walk, it is necessary to choose a day of the week and time of day for the trip. This is so bus timetables can be picked up to model the journey. The day of the week taken is usually a Tuesday, Wednesday or Thursday, being less influenced by the distorting effect of a weekend. However, considerable discussion took place about on which day the leisure centre trip should be considered to take place. It was recognised that many trips would be desired in the evening, after work or education, however this would produce a distorted picture, as public transport tends not to be available them. This would produce a uniform lack of access, not a distribution of difficulty which is what is needed for the

index. After much deliberation, it was decided to use a Saturday morning service, this being a day when people could potentially use daytime services for the trip.

Issues for response

There are specific questions that will help shape the development of the WIMD 2008. These questions are listed below. They are not exhaustive, nor should limit any response that is relevant to this domain.

- Is the definition for the domain appropriate?
- Are the services chosen appropriate?
- Are the definitions given appropriate for the indicators?
- Have any essential services been missed, providing the data set is available on a consistent national, reliable basis?
- Are the travel times chosen appropriate?
- Will a buffer of 30 minutes travel into England be sufficient for the analysis proposed?

Housing

The purpose of this domain is to capture deprivation through a lack of access to adequate housing.

Proposed indicators

The indicators used in the WIMD 2005 were reviewed, together with those indicators rejected for WIMD 2005. Potential new indicators were also discussed. The proposed indicators are listed below.

Fuel Poverty – households in fuel poverty are defined as those that spend more than 10% of their income on maintaining satisfactory heating regimes. Therefore it is expected that deprived areas would have a higher proportion of households in fuel poverty.

Fuel poverty estimates from the Living in Wales survey are being modelled down to the LSOA level via a project commissioned by the Welsh Assembly Government's Housing Directorate. It is proposed that this indicator would provide proxy information on energy efficiency and quality of houses.

Cost of housing per square metre - proportion of properties within a LSOA whose cost per square metre falls below a specified threshold. It is proposed that more deprived areas would have a higher proportion of larger, cheaper houses than less deprived areas. This indicator would therefore give an indication of the availability of cheaper housing, with more square metres. Data for this indicator will be obtained from the Valuation Office Agency (VOA).

Other indicators discussed

Repair costs were suggested as a possible indicator, modelling data from both the Living in Wales survey and from the Valuation Office Agency (VOA). Analysis has shown there is a good correlation between repair costs and the age and tenure of a building. Further investigation found that up to date tenure information at LSOA level was not available and therefore the repair costs could not be successfully modelled to the small area.

SAP energy ratings were thought to be a good indicator of deprivation, however there is currently not data available at the LSOA level and no plans to model information down to this level, therefore this indicator cannot be considered.

General Physical Condition was considered from the Living in Wales survey, however due to the sample size used of the survey information is only available at the all Wales level.

Availability of central heating was no longer considered a robust indicator of deprivation, with the data not having been available since the 2001 census. It was also considered that due to the relatively small numbers that this affects there would not be a good distribution of the LSOAs across Wales for the purposes of the index.

Similarly **Availability of various basic amenities** was also rejected to the timeliness and availability of data at the LSOA level as well as the low incidence.

Overcrowding suitable data to model overcrowding at the LSOA level is not available and therefore this indicator can not be used at present. However once local authority level figures are available from the Living in Wales survey the figures could possibly by modelled to LSOA level.

There is also no suitable data available for **Homelessness** or **Social Housing** as data is only collected at the local authority level.

Housing benefit was discussed and it was agreed that this is more appropriate for the Income domain.

Vacant dwellings were considered as an indicator however there were a few issues of concern. It was proposed that a dwelling might be vacant due to it being a second home or an investment property, rather than an indication of deprivation. Also data on vacant dwellings is only currently collected for social housing at the local authority level.

Affordability was also considered although there are data difficulties for this and it was also considered there could be considerable overlap with the Income domain.

Issues with indicators

The domain group agreed that if data availability were not an issue the following areas would best indicate the level of housing deprivation within an area:

- Overcrowding
- House prices in relation to income
- Availability of social housing
- Physical condition
- Voids/Lack of demand

Several alternative indicators were discussed, including those above, with the general lack of data being the main cause for not being able to include them as indicators. The domain group would recommend a focus be set on increasing the collection and coverage of data concerning the theme of housing at the smaller area level across Wales.

Issues for response

There are specific questions that will help shape the development of the WIMD2008. These questions are listed below. They are not exhaustive, nor should limit any response that is relevant to this domain.

- Is the definition for the domain appropriate?
- Are the indicators chosen appropriate?
- Are the definitions given appropriate for the indicators?
- Have any essential indicators been missed, providing the data set is available on a consistent national, reliable basis?

Physical Environment

The purpose of this domain is to measure environmental factors that may impact on quality of life in an area. Negative impacts on quality of life are judged to be a form of deprivation. Environmental deprivation is generally not correlated with social or economic deprivation in Wales. (ref: Walker et al 2003). This domain does not capture aspects of deprivation such as health inequalities (this is an objective of the Health Domain). Data covering Wales is not available on actual impacts on quality of life; so factors were chosen that were judged to indicate an increased potential for reduced quality of life and had data available.

Proposed indicators

The indicators used in the WIMD2005 were reviewed, together with those indicators rejected for WIMD 2005. Potential new indicators were also discussed. The proposed indicators are listed below.

Flood Risk – This indicator was used in WIMD 2005. It is proposed that for WIMD 2008 different levels of risk will be taken into account, as is done with insurance companies with 3 levels of risk; significant, moderate and low risk. The previous problem that existed was the lack of flood defence information in the data and that the Index would need to be sensitive to these interventions. In consequences, The Environment Agency now has risk maps including most flood defences and has developed a way of mapping flood risk based on an analysis of the proportion of households in each flood risk zone. The risk is based on frequency rather than level of damage caused by any flooding.

Air Quality - Air quality has been chosen for inclusion as an indicator for WIMD 2008, as it was for WIMD 2005. It is believed to be a good proxy measure of the quality of the surrounding environment. Poor air quality suggests proximity to certain activities such as traffic, domestic combustion and industrial sites – activities that could have a negative impact on quality of life, the local environment and health.

The mapped data for each pollutant, together with Air Quality Management Area data, will be normalised using air quality standards (See Appendix C) and exponentially transformed to create the indicator in one combined map. This helps overcome some of the issues encountered in WIMD 2005, such as when using an average measure of all pollutants, information on single pollutants at high concentrations can be lost in the deprivation score.

Air Emissions - The reasons for inclusion of air emissions as an indicator are similar to those for air quality. Air emissions act as a proxy measure for undesirable activities in a neighbourhood, in the same way that air quality data does. Air Emission datasets were for pollutants for which mapped concentrations data was not available and includes pollutants that may not be present in air quality monitoring data or considered in the designation of Local Air Quality Management Areas, but nevertheless are relevant in terms of human health and food safety. Air emissions data therefore provides a good set of complimentary data covering pollutants not included in the Air Quality indicator. Mapping of air emissions data will be carried out in the same way as for air quality data. (See Appendix C for normalising constants).

Proximity to waste disposal and industrial sites – The Proximity to waste disposal and industrial sites indicator was used in WIMD 2005. Some sites have a greater potential to impact on quality of life than others – a direct quantification of this is not available. In a change from WIMD 2005, WIMD 2008 will classify sites in order to define different buffer zones according to hazard assessments. This will be done using a more scientific and robust process, which takes into account process type and size and operator performance. Where available the Environment

Agency 'Environment Protection Operator and Pollution Risk Appraisal' (EPOPRA*) scores will be used as a basis for categorising waste disposal and industrial sites. It is proposed that Integrated Pollution Prevention and Control Directive (IPPC) sites, A1and A2 processes sites are to be included, along with Control of Major Accident Hazards (COMAH) sites, nuclear sites, waste management sites and Sewage treatment works. Many of these types of sites do not have an EPOPRA score. Where there is not an EPOPRA score, a proxy score will be allocated by the Environment Agency relating to size and processes and buffer zones calculated accordingly. The Proximity to Waste Disposal and Industrial sites indicator was agreed as an indicator, which should relate sites with a EPOPRA or proxy score to where people live. If more people live in close proximity to an industrial or waste site (which would be defined by a buffer) then the greater the level of deprivation. The banding proposal detailed below allocates a greater weighting and greater number of buffers (greater distance effect) to sites in a higher band. Those in band '0' are judged to have no significant effect (weighting = 0); those in band '5' are judged to have the greatest effect (weighting varies from 5 to 0 with distance); If a household is within two buffers (overlapping) then the proximity score will be counted twice accordingly.

1. Assign each site into one of 6 bands, depending on the OPRA score where available.
2. Use the bands to identify the size and weighting of the buffers to apply around each site using the weighting matrix:

Distance from site (km)	Band created from EPOPRA score (Industrial and waste management sites)					
	5	4	3	2	1	0
<=0.5	5	4	3	2	1	0
>0.5 to 1	4	3	2	1	0	0
>1 to 1.5	3	2	1	0	0	0
>1.5 to 2	2	1	0	0	0	0
>2 to 3	1	0	0	0	0	0
>3	0	0	0	0	0	0

Sites without EPOPRA scores: Nuclear sites - Band 4, A2 sites - Band 2 (affect up to 1 km), STWs - Base on population. Largest population in Band 2

STW Population Equivalent	Band for profile
>100,000	2
>10,000 - 100,000	1
<= 10,000	0

- EP OPRA is a risk-screening tool that the Environment Agency use to regulate operators under the Integrated Pollution Prevention and Control (IPPC) regulations and the Waste Management Licensing (WML) regulations. It has been used for IPPC since 2003 and for WML since April 2005. IPC sites also have EPOPRA scores, which are derived in a simpler way. EP OPRA is used to help target regulatory effort at those activities that present the greatest risk to the environment.

Ground movement risk – This is a new indicator for WIMD 2008 relating to data from a British Geological Survey (BGS) dataset called G–BASE looking at contaminated land. BGS have developed a dataset called Geo Sure, which assesses the risk of ground movements in the UK. Data for the whole of Wales is available and two potential indicators for this domain have been identified. They are 'shrink swell' and 'slope instability'. The data is currently being assessed for is robustness and suitability for use in WIMD 2008. Should the data prove suitable both the indicators will be included. The methodology used could be the same as flood risk, establishing the proportion of households in an LSOA with a certain risk.

Other indicators discussed

Intensive farming – Where these sites are regulated under PPC, they are included in the proximity to industrial sites indicator. Other sites have no suitable way to identify and score them and so it was decided not to include them in the index.

Quarries - there is no data and a different regulatory systems across Wales and therefore has not been included.

Fly Tipping – data is still only available at Local Authority level and will therefore not be usable for WIMD 2008.

Water Quality – It is felt that once the Water Framework Directive is implemented this maybe suitable for consideration for future updates of WIMD.

Contaminated Land – the Geochemical Baseline Survey of the Environment (G-BASE) database contains data on contaminated land. However, the data only exist with precision for Cardiff and Swansea. This data however could be useful for the future updates of WMD as more of Wales is covered in detail. The European Framework Directive also collates an inventory of contaminated land, but this will not go through until 2008, which is too late for WIMD 2008. Data on derelict land was also considered, with Local Authority only getting information about contaminated land if they receive a planning application for change of use. This indicator will be therefore reassessed for the next index.

Transport – this indicator would capture the pollution and other effects of traffic that could deprive people of a good local environment in which to live and enjoy. Therefore, the proximity to motorways and A1 roads using noise, pollutant and road safety as deprivation factors was investigated. Data availability is generally the main constraint on producing flow maps for Wales. Suitable datasets have not so far been identified to allow an indicator to be produced.

Noise - this indicator would capture all sources of noise which would deprive people of a good environment in which to live. Noise mapping techniques are being developed and may be available for future revisions of the WIMD.

Green Spaces – including natural shading in street, however data is not available. The data would also be anecdotal rather than have all Wales coverage. This could be included in future revisions as a positive affect i.e. an indicator of less deprived people and a mechanism for counteracting negative effects.

Part B processes - these were considered as part of the proximity to industrial sites indicator. The data is not currently available. Consideration would have to be given to a weighting factor to apply a buffer zone around each site.

Acid and nutrient deposition – the potential for including acid and nutrient deposition with other aspects of air pollution is being investigated. If included, this would be a measure of potential pollution impacts on local areas of natural and semi-natural habitat and consequent effects on quality of life.

Outstanding issues

Ground movement risk data is still being investigated and only if this proves to be a suitable and viable dataset will the indicators be included in WIMD 2008.

COMAH sites - this data set needs to be made available. Consideration will need to be given to the quality of the geographical data before it can be included in the index.

Issues for response

There are specific questions that will help shape the development of the WIMD2008. These questions are listed below. They are not exhaustive, nor should limit any response that is relevant to this domain.

- Is the definition for the domain appropriate?
- Are the indicators chosen appropriate?
- Are the definitions given appropriate for the indicators?
- Have any essential indicators been missed, providing the data set is available on a consistent national, reliable basis?

Community Safety

The community safety domain is to reflect where people are deprived of a safe community in which to live. Safety includes levels of household and personal crime and quality of experiences in public places compatible with access to ordinary work, leisure and social relationships.

The indicators used in the WIMD2005 were reviewed to ensure they were still appropriate, available and of sufficient quality to use. In addition further indicators were considered to test whether they too should be included in the domain. The indicators are listed below.

Proposed indicators

Police force recorded crime - During the preparation of the WIMD2005 police recorded crime data was collected, but data issues prevented the inclusion of this data at that time. It is still felt appropriate to include police recorded crimes for four groups of offences[2]; burglary, violent crime, theft and criminal damage. The basic definitions are those used for the English Index of Deprivation. There were additional violent crime codes[3] identified that were considered to be relevant but not included. It was decided to investigate whether their inclusion would improve the definitions.

Youth Offenders - Offenders are often drawn from the most deprived section of society. Offenders at liberty is a measure to indicate issues relating to young people that indicate their need and the social problems they cause. This indicator complements the recorded crime data by including the offenders at liberty in an area. Youth offender data was successfully collected in 2004/5 and proved reliable as a national data set for Wales. It is proposed to include the "criminogenic need assessment", which is collected through the ASSET tool, in this data set. The criminogenic need will provide valuable additional explanation for the unlawful behaviour, like alcohol or drug dependency.

Adult offenders - As with youth offenders, adult offenders are often drawn from the most deprived section of society. Offender information will complement crime records and is collected by the probation service. Criminogenic need for adults is collected in the OASys tool, but is not to date a mature technique. It is not considered possible to use criminogenic need for adult offenders in the 2008 index.

Fire safety - Incidents requiring call out of fire and rescue services are related to deprivation[4] and more likely within disadvantaged groups. Primary fires[5] include "all fires in buildings, vehicles and outdoor structures or any fire involving casualties, rescues, or fires attended by five or more appliances". Secondary fires are "the majority of outdoor fires including grassland and refuse fires unless they involve casualties or rescues, property loss or five or more appliances attend". It is proposed to use primary fires which relate better to property and people, with the addition of " derelict vehicle" fires that tend to occur in deprived areas. The secondary " derelict

---

[2] For details of offence codes see Appendix D
Burglary defined by National Crime Recording Standard (NCRS) numbers : 28-31
Violent crime: NCRS 1,2,4.1-2, 5, 8A, 8C-E, 34A-B, 37.1, 105A-B
Theft: NCRS 37.2, 39, 45, 48, 126
Criminal damage: NCRS 56, 58A-H, 59
NCRS definitions can be found at: http://www.homeoffice.gov.uk/rds/countrules.html

[3] Additional violent crime codes for test: NCRS 3, 4.7, 11, 12, 81, 90

[4] Reference in the introduction, page 5 at: http://www.scotland.gov.uk/Resource/Doc/36496/0024964.pdf

[5] Definitions are found in the FDR1 (Fire Damage Report) Guidance document introduction section or page 115 of Fire Statistics UK 2005 at:
http://www.communities.gov.uk/pub/25/FireStatisticsUnitedKingdom2005_id1509025.pdf

vehicle" fires will only be included if available at the LSOA level. Data may be further analysed by motive (Deliberate, Accidental) where available.

Other indicators discussed

Domestic violence - considered as an indicator, but suffers from under reporting, small numbers and inconsistency. For those reasons it was not felt possible to use it for this index.

Anti-social behaviour - the reporting and handling of anti-social behaviour is likely to differ between local authority areas and within small geographic area. The perception of residents of less deprived areas may be more sensitive to behaviour problems considered routine in other more deprived areas. The records of stage 2 and above in the process towards Anti-social Behaviour Orders (ASBOs) will be collected for the index. Although the inclusion of this indicator was desired, following investigation it was found that the recording of the incidents was too inconsistent to be used.

Road traffic crashes - considered for this domain, but there are issues of poor reporting and data quality. In addition to those problems, it was considered that pedestrian and cycling incidents should be the focus of any indicator, but the number of records of such crashes are very small or zero for many lower super output areas, so would not be a reliable measure.

The fear of crime - may contribute to "spirals of decline in high-crime areas" (Hough 1995) but attitude measures have not been used in any other domain in the index, so it was felt inappropriate to use that type of indicator in the community safety domain. Additional difficulties would arise from the modelling required to take the data from local authority level to LSOAs. The variability of the perception of safety is affected by publicity campaigns, for example in newspapers, so is not reliable. Perception of safety is also a function of the local environment. In some areas where the threat or incidence is low the perception is of lack of safety, more so than areas with higher incidence where local perception is of greater safety.

Outstanding issues

There are issues about small counts in LSOAs which produce unreliable data, subject to significant "noise effects". This may be overcome by using more than one year's data to increase the counts in each cell or statistical techniques (shrinkage) to increase the reliability.

Issues for response

There are specific questions that will help shape the development of the WIMD2008. These questions are listed below. They are not exhaustive, nor should limit any response that is relevant to this domain.

- Is the definition for the domain appropriate?
- Are the indicators chosen appropriate?
- Are the definitions given appropriate for the indicators?
- Have any essential indicators been missed, providing the data set is available on a consistent national, reliable basis?

Methodology

The proposal for WIMD 2008 is to continue with the same basic methodology as used for the WIMDs 2000 and 2005, with a small number of enhancements as described below.

Overview

The methodology adopted for the WIMDs 2000 and 2005 focuses on the creation of a series of robust domain sub-indexes; that is, for income, health, education, and other factors separately. These domain sub-indexes can be validated, either because they directly measure the factor itself (for example, in the case of means-tested benefit, reliance on unemployment benefit), or by making comparisons with other research studies.

This approach results in a more complex index based on more data sources than earlier indexes of deprivation. It also requires domain-specific indicators; that is, indicators which are cogent measures of that domain's deprivations and are not just vaguely and imprecisely related to the domain. It also requires procedures for combining indicators and the related data within any domain according to well-defined algorithms.

The WIMD 2005 was made up of seven separate domains (or kinds) of deprivation:

- income
- employment
- health
- education
- housing
- access to services
- environment

Each of them was based on a range of different indicators which meant that they were measured in different ways using different units. So before they could be combined the measurements had to be transformed to make them compatible (see Appendix E). For example, if the height of something had been measured in metres and the weight in kilograms it would not make sense simply to add them together.

To overcome any problems of unreliable data for some areas (where numbers are particularly small), a statistical technique called shrinkage was used, detail can be found in Appendix F. This technique was only used for the Education and Health domains for WIMD 2005.

To combine indicators within a domain when there were more than 2, and they could not simply be added (as they could for the Income and Employment domains), a statistical technique called factor analysis was used. Details of this approach are given in Appendix G.

Proposed improvements for WIMD 2008

- the inclusion of a new domain, Community Safety (for details see page 34);
- the removal of shrinkage from the Education domain (for details see page 20).

Weighting of domains

Virtually every reputable study on deprivation shows that the financial component in any overall scale of deprivation is of great significance. This was taken into account in the weighting of the components which make up the overall index for the Welsh Index of Multiple Deprivation (WIMD) 2005. This was also true for the WIMD 2000, and for the similar indexes of the other UK countries.

Weights for WIMD 2005

Income and employment were classed as the most important factors, and they were given the biggest weighting in the overall index. Although there is more to deprivation than poverty, not having enough money or a job is a big part of it. The weights used for WIMD 2005 were based on those for WIMD 2000, with the weight for the Housing domain split between Housing and the new Physical Environment domain.

Income deprivation	25%
Employment deprivation	25%
Health deprivation and disability	15%
Education, skills and training deprivation	15%
Geographical access to services	10%
Housing deprivation	5%
Physical Environment	5%

Proposed weights for WIMD 2008

While the body of evidence used to inform the choice of domain weights for WIMD 2000 and WIMD 2005 has not changed substantially, the introduction of a new domain (community safety) means that there has to be some adjustment to the weighting of domains for WIMD 2008. The proposal is to take a similar approach to that used for the Scottish index, which is:

- Decide on a weight for the new domain;
- Scale other domain weights so that the total of all domain weights remains as 100.

However, to avoid unnecessary and spurious decimals, as well to retain the importance of domains that currently have fairly low weights, the proposal is that only the 4 larger domains are adjusted.

This approach produces the following domain weights:

Income deprivation	23.5%
Employment deprivation	23.5%
Health deprivation and disability	14%
Education, skills and training deprivation	14%
Geographical access to services	10%
Housing deprivation	5%
Physical Environment	5%
Community safety	5%

Views are welcome on both the weight that should be applied to the new community safety domain as well as the overall weighting of domains. However, suggestions should be backed up with supporting evidence, either quantitative or qualitative.

Possible sub-indexes

There have been numerous calls for sub-indexes to look at certain population groups; children and older people in particular. Indeed, the 2000 index had a child index published alongside it.

Difficulties with data for WIMD 2005 meant that the data relating to children for the income domain was not as up-to-date as the main indicator data – due to the effects of new tax credits and the change of responsibility between government departments. Because of this it was decided not to publish a child index at that time. Subsequently, research carried out on behalf of the Scottish Executive showed that using the approach used for the 2004 SIMD to create a child index was not particularly robust because it was based on a small number of indicators and smaller population denominators. In addition, it was no more efficient than the SIMD at identifying areas with concentrations of children receiving income benefits; the combined area child measure was better at identifying areas with lower educational attainment.

Research undertaken for the Department of Communities and Local Government into a completely separate Index of Child Well Being, including its own set of domains, provides a potential way forward. Once this is published in late 2007, it will be examined to assess the possibility of producing a similar index for Wales. The timing of such an index (whether alongside WIMD 2008 or shortly after) would depend on the practicalities of obtaining equivalent data. An announcement will be made by early 2008. Any views on potential sub-indexes and relevant indicators will be considered as part of any final decisions on the approach for WIMD 2008.

How can I respond?

The consultation will run from 11 June 2007 **until 7 September 2007**.

Responses are welcome either in writing or by email. We would be grateful if responses could be sent by email as this will make the publishing of responses much easier. However, we will give equal consideration to responses whether by e-mail or in writing.

The address for responses is:

Social Inclusion and Equality Statistics
Welsh Assembly Government
Room 2-002
Cathays Park
CARDIFF
CF10 3NQ

or by email:

stats.inclusion@wales.gsi.gov.uk

During the consultation events will be run to help explain the details set out in this consultation event. Two events have been arranged in advance and these are:

4th July 2007 at Dylan Thomas Centre, Somerset Place, Swansea, SA1 1RR

10th July 2007 at the Conwy Business Centre, Llandudno Junction, Conwy, LL31 9XX

And if you wish to attend one of these events please email or send your details using the contact details above (that is, the same as given for consultation responses).

We are happy to receive comments on any aspect of the proposals set out in this document. When giving your views on particular indicators it would be helpful to group together comments by domain.

Domain Group membership

Income and Employment

Emma Arnell-Smith (Chair)	Statistical Directorate - Welsh Assembly Government
Nigel Brough	Department for Work and Pensions
Alan McIntyre	Her Majesty's Revenue and Customs
Stephen King	Swansea County Council
Julian Revell	Statistical Directorate - Welsh Assembly Government
Peter Sloane	Welsh Economy Labour Market Evaluation and Research Centre (WELMERC)
Edward Sheriff	Economic Advice Division - Welsh Assembly Government

Education

Emma Arnell-Smith (Chair)	Statistical Directorate - Welsh Assembly Government
Glyn Jones	Statistical Directorate - Welsh Assembly Government
Stephen Hughes	Statistical Directorate - Welsh Assembly Government
Sioned Moffett	Statistical Directorate - Welsh Assembly Government
Stephanie Howarth	Statistical Directorate - Welsh Assembly Government
Mike Jones	Swansea County Council
Helen James	Local Government Data Unit ~ Wales
Scott Clifford	Planning and Learning Division (DELLS) - Welsh Assembly Government
Emma Williams	Performance and Improvement Division (DELLS) - Welsh Assembly Government
Alan Lowndes	Performance and Improvement Division (DELLS) - Welsh Assembly Government
Robin Jones	Qualifications and Curriculum Group (DELLS) - Welsh Assembly Government
Jackie McDonald	Business Development Division (DELLS) - Welsh Assembly Government
Elizabeth Heal	Higher Education Funding Council Wales (HEFCW)

Health

Emma Arnell-Smith (Chair)	Statistical Directorate - Welsh Assembly Government
Cath Roberts	Statistical Directorate - Welsh Assembly Government
Nathan Lester	National Public Health Service
Nick Holmes	Local Government Data Unit ~ Wales
Gwyneth Thomas	Statistical Directorate - Welsh Assembly Government
Gareth John	Health Solutions Wales
Chris Tudor-Smith	Public Health Improvement Division - Welsh Assembly Government
Eva Elliot	Cardiff University
Sue Denman	Research and Development, Health & Social Care - Welsh Assembly Government

Access to Services

Chris Williams (Chair)	Statistical Directorate - Welsh Assembly Government
Vincent James	Cartographics - Welsh Assembly Government
Stuart Neil	Statistical Directorate - Welsh Assembly Government
Nick Holmes	Local Government Data Unit ~ Wales
Nicolas Theuray	Local Government Data Unit ~ Wales
Sean White	Rural Observatory
Natalie Grohmann	Sustainable Futures Division - Welsh Assembly Government
Diana Greaves	Powys County Council

Physical Environment

Chris Williams (Chair)	Statistical Directorate - Welsh Assembly Government
Havard Prosser	Technical Services Division - Welsh Assembly Government
John Houlgreave	Environment Agency
Julie Boswell	Environment Agency
Nicola Theuray	Local Government Data Unit ~ Wales
Robert Hartshorn	Cardiff County Council

Housing

Chris Williams (Chair)	Statistical Directorate - Welsh Assembly Government
Rhiannon Caunt	Statistical Directorate - Welsh Assembly Government
Ralph Bourke	Caerphilly County Borough Council
David James	Torfaen County Borough Council
Nick Holmes	Local Government Data Unit ~ Wales
Steve Wilcox	York University
Amanda Oliver	Welsh Federation of Housing Associations

Community Safety

Chris Williams (Chair)	Statistical Directorate - Welsh Assembly Government
Nick Holmes	Local Government Data Unit ~ Wales
Nicolas Theuray	Local Government Data Unit ~ Wales
Alison Dixon	Merthyr Youth Offending Team
Ian Clark	Swansea BCU - South Wales Police
Kerry Citric	Community Safety Division - Welsh Assembly Government
Madog Williams	Statistical Directorate - Welsh Assembly Government
Rachel Morgan	Welsh Local Government Association
Jocelyn Kynch	Independent Consultant
Sian Vowles	South Wales Fire & Rescue Service
Dusty Kennedy	Youth Justice Board

Indicators used for WIMD 2005, by domain

Income

The indicators used were:

- Income Support claimants (and their children and partners)
- income-based Jobseeker's Allowance
- Working Families' Tax Credit
- Disabled Person's Tax Credit
- National Asylum Support Service (NASS) supported asylum seekers in Wales in receipt of subsistence only and accommodation support

They were combined using straightforward addition.

Employment

The indicators used were:

- claimants of unemployment related benefits.
- claimants of Incapacity Benefit
- Severe Disablement Allowance (for women under 60 and men under 65)
- participants on New Deal for Young People and Intensity Activity Period (for New Deal 25+)

They were combined using straightforward addition.

Education

The indicators used were:

- Key Stage 2, average point scores
- Key Stage 3, average point scores
- Key Stage 4, average point scores
- proportion of adults with low or no qualifications
- proportion of 17 and 18 year olds not entering further or higher education
- secondary school absence rates

They were combined using weights derived by factor analysis.

Health

The indicators used were:

- limiting long-term illness
- deaths
- cancer incidence

They were combined using weights derived by factor analysis.

Housing

The indicators used were:

- lack of central heating
- overcrowding (excluding all student households)

They were combined directly assuming equal weighting.

Physical environment

The indicators used were:

- air quality
- air emissions
- living within 1 km of a waste disposal site
- proportion of people living within 1 km of an Environment Agency regulated industrial source
- proportion of people living in an area with a significant risk of flooding

They were combined in three sets: air quality and emissions; waste disposal site and industrial source; and flooding.

Access to services

The indicators used were:

- access to food shop (10 minutes)
- access to GP surgery (15 minutes)
- access to primary school (15 minutes)
- access to post office (15 minutes)
- access to public library (15 minutes)
- access to leisure centre (20 minutes)
- access to NHS dentist (20 minutes)
- access to secondary school (30 minutes)

They were combined using weights derived by factor analysis.

Air quality substances and standards used to normalise (source: Air quality strategy objective)

Pollutant	Normaliser (= standard)	Unit
1,3 Butadiene (Bu)	2.25	ug / m3
Benzene (Bz)	5	ug / m3
NO2	40	ug / m3
SO2	20	ug / m3
PM10	20	ug / m3
CO_8hr	10	mg / m3
SO2_999	266	ug / m3
O3_8hr	10	days

Air emission substances and standards used to normalise

Pollutant	Normaliser	Unit
Ammonia	8	ug / m3
Arsenic	6	ug / m3
B[a]P	0.25	ug / m3
Cadmium	5	ug / m3
Chromium	0.5	ug / m3
Dioxins	0.3	ug / m3
Lead	250	ug / m3
Mercury	1000	ug / m3
Nickel	20	ug / m3
NMVOC	5	ug / m3
NOx	30	ug / m3
Vanadium	200	ug / m3

Categories of crime defined in - Home Office Counting Rules for Recorded Crime April 2007

Violent crime (Violence against the person & Robbery)

1	Murder
2	Attempted Murder
3	*Threat or Conspiracy to Murder**
4 / 1	Manslaughter
4 / 2	Infanticide
4/3	*Child Destruction*
4/4,	*6,8 Causing Death by Dangerous or Careless Driving*
4/7	*Causing or Allowing Death of Child or vulnerable Person**
4/9	*Causing Death by Driving: Unlicensed, Disqualified or Uninsured Drivers*
5	Wounding or Other Act Endangering Life
6	Endangering a Railway Passenger
7	Endangering Life at Sea
8A	Other Wounding
8B	*Possession of Weapons*
8C	Harassment
8D	Racially or Religiously Aggravated Other Wounding
8E	Racially or Religiously Aggravated Harassment
11	*Cruelty to and Neglect of Children**
12	*Abandoning Child Under Two Years **
13	*Child Abduction*
14	*Procuring Illegal Abortion*
15	*Concealment of Birth*
37 / 1	Causing Death by Aggravated Vehicle Taking
104	*Assault on a Constable*
105A	Common Assault
105B	Racially or Religiously Aggravated Common Assault
34A	Robbery of Business Property
34B	Robbery of Personal Property
81	*Firearms Act 1968 and other Firearms Act**
90	*Other Knives Offences**

Burglary

28	Burglary in a Dwelling

29 Aggravated Burglary in a Dwelling

30 Burglary in a Building other than a Dwelling

31 Aggravated Burglary in a Building other than a Dwelling

Theft (& handling stolen goods)

37/2 Aggravated Vehicle Taking

38 Proceeds of Crime

39 Theft from the Person

40 Theft in a Dwelling other than from an Automatic Machine or Meter

41 Theft by an Employee

42 Theft or Unlawful Taking of Mail

43 Abstracting Electricity

44 Theft or Unauthorised Taking of a Pedal Cycle

45 Theft from a Vehicle

46 Theft from a Shop

47 Theft from an Automatic Machine or Meter

48 Theft or Unauthorised Taking of a Motor Vehicle

49 Other Theft

54 Handling Stolen Goods

126 Vehicle Interference and Tampering

Criminal damage

56 Arson

58A Criminal Damage to a Dwelling

58B Criminal Damage to a Building other than a Dwelling

58C Criminal Damage to a Vehicle

58D Other Criminal Damage

58E Racially or Religiously Aggravated Criminal Damage to a Dwelling

58F Racially or Religiously Aggravated Criminal Damage to a Building other than a Dwelling

58G Racially or Religiously Aggravated Criminal Damage to a Vehicle

58H Racially or Religiously Aggravated Other Criminal Damage

59 Threat or Possession with Intent to Commit Criminal Damage

The codes in *italic print* were not included in the previous test for WIMD2005 nor the English Index of Deprivation. The codes in *italics* with a star (*) appended are being tested to see whether their inclusion would improve the overall definition of the indicators.

These codes and full explanations of the definitions can be found at:
http://www.homeoffice.gov.uk/rds/countrules.html

Exponential Transformation of the Domain Indices

The precise transformation involved is as follows. For any LSOA, denote its rank on the domain, scaled to the range [0,1], by R (with R=1/1896 for the least deprived, R=1896/1896=1 for the most deprived).

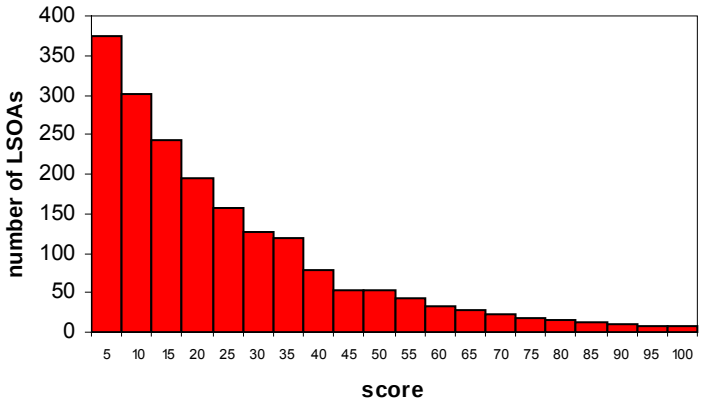
The transformed domain, (X) equals:

-23*log{1 - R*[1 - exp(-100/23)]}

where log denotes natural logarithm and exp the exponential or antilog transformation, and * denotes multiplication. This formula may at first sight seem complicated, but it is very straightforwardly calculated and is in fact simpler than the commonly-used transformation to a normal curve which necessitates the use of a look-up table. The resulting distribution is illustrated below in a histogram.

Each transformed domain has a range of 0 to 100, with a score of 100 for the most deprived LSOA. Ten percent of LSOAs have a score higher than 50. When transformed scores from different domains are combined by averaging them, the skewness of the distribution reduces the extent to which deprivation on one domain can be cancelled by lack of deprivation on another. For example, if the transformed scores on two domains are simply averaged, with equal weights, a (hypothetical) LSOA that scored 100 on one domain and 0 on the other would have a combined score of 50 and would thus be ranked at the 90th percentile. (Averaging the untransformed ranks, or after transformation to a normal distribution, would result in such a LSOA being ranked instead at the 50th percentile: the high deprivation in one domain would have been fully cancelled by the low deprivation in the other.) Thus the extent to which deprivation in some domains can be cancelled by lack of deprivation in others is, by design, reduced.

Histogram of a transformed domain



The Shrinkage Technique

Shrinkage is the term given to the procedure used to improve the quality of the small area data in indicators where populations are small. For some indicators, where occurrences and / or populations are small, data may be unreliable – more likely to be affected by measurement error or sampling error, with particular LSOAs scores being unrepresentatively high or low for certain indicators. Calculating a scores standard error measures the extent of the unreliability.

The Shrinkage technique involves moving unreliable or extreme scores towards another more robust score. For the WIMD 2005, the more robust score used was the Local Authority (LA) average score, as no suitable set of geographies below the local authority level exist and the average score for Wales would not be representative due to the large variations in existence across the country. Most scores will move only a small amount, although extreme scores (those with a very large or small standard error) will move significantly. The size of the movement will not depend on the size of the standard error alone, but also takes into account the amount of heterogeneity amongst the LSOAs within a LA.

Shrinkage of proportions

The 'shrunk estimates' of a LSOA level proportion is a weighted average of the original proportion for the LSOA and for that the corresponding LA. The weighted average is calculated on the logit scale, for technical reasons, principally because the logit of a proportion is more nearly normally distributed than the proportion itself.

The weightings used are determined by the relative magnitudes of within LSOA variability and between LSOA variability.

If the data for the LSOA *j* are r_j individuals with a particular attribute out of n_j in total in the LSOA, the empirical logit is

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

whose estimated standard error (s_j, say) is the square root of

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

[See, for example, D. R. Cox (1970), *Analysis of Binary Data,* pp. 33-34]. The corresponding counts *r* out of *n* at LA level give the LA-level logit,

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

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

$$z_j^* = w_j z_j + (1 - w_j) z,$$

where w_j is the weight given to the 'raw' LSOA-*j* data and $(1-w_j)$ the weight given to the overall LA figure. 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 LA, calculated as

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

Thus large LSOAs, where precision $1/s_j^2$ 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 LAs.

The final step is to back-transform the shrunk logit z_j^* using the 'anti-logit', to obtain the shrunken LSOA level proportion

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

for each LSOA.

Shrinkage for average points scores

The 'shrunk estimates' of a LSOA level average point score (whether for KS2, 3 or 4) is a weighted average of the two original average points scores for the LSOA and for the corresponding LA.

The weightings used are determined by the relative magnitudes of within LSOA variability and between LSOA variability.

If for LSOA $_j$ there are l_j individuals with individual point scores m_{lj} then the average point score

$$z_j = \frac{\sum_{lj} m_{lj}}{l_j}$$

for LSOA$_j$, z_j, is:

The corresponding values at a local authority level of *l* and m give the average point score for the local authority, z, is

$$z = \frac{m}{l}$$

The 'shrunk' LSOA-level average point score is then the weighted average

$$z_j^* = w_j z_j + (1 - w_j)z,$$

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 LA. 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 LA, calculated as

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

and s^2 is the within LSOA variance for the l_j individual point scores

$$s^2 = \frac{1}{l_j - 1}\sum_{lj} (m_{lj} - z_j)^2$$

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 LAs.

The Factor Analysis Technique

Factor Analysis Overview

Factor analysis is a method for assessing the extent to which a set of indicators may all be measuring the same underlying construct or factor. The premise behind a one-common-factor model is that the underlying factor is imperfectly measured by each of the indicators in the dataset but that indicators that are most highly correlated with the underlying factor will also be highly correlated with each other. By analysing the correlation between indicators it is therefore possible to make inferences about the common factor and as a result estimate a 'factor score' for each LSOA. This score is derived from a set of weights for each of the indicators in the data set that is generated by the process of factor analysis. This factor score can then be used as the domain index.

Factor analysis has only been applied to three domains: Health, Education and Geographical Access to Services. Factor analysis is used in these domains because they contain indicators that measure, on potentially different metrics and with different levels of accuracy, a number of forms of that deprivation and therefore cannot otherwise easily be combined. The main reasons why Factor Analysis has been used are:

- Because the indicators are on different metrics and have different levels of accuracy, and so cannot simply be summed
- To ascertain the factor that underlies the indicators within the Domain
- To help take into account the problem of 'double counting' within a Domain

In the Employment and Income domains we can identify individuals who are or are not deprived in terms of the domain definition. The number of deprived people can then simply be summed and divided by a suitable denominator to create an area rate. This is not possible in the other five domains. These deprivations tend to present themselves in different ways at different times. Thus, for example, an individual is 'health deprived' if they die prematurely or are long-term sick. While the long-term sick may be more likely to die prematurely than others, these events do not occur to the same people at the same time. Typically such domains include data on people at different ages and stages e.g. in the education domain, lack of qualifications in the adult population as well as poor results at school level. Instead we hypothesise that there is an underlying factor at the local area level (e.g. health deprivation) that makes these different states likely to exist together in the same area. This underlying factor cannot be measured directly but can be identified through its effects on specific individual measures (e.g. premature death, long-term sickness, low birth-weight children etc.). We have therefore collected a number of indicators that measure, with different levels of accuracy, the effects of this underlying factor. By looking at the relationship between all these indicators the underlying factor can be identified and quantified.

Factor analysis also takes some account of the problem of 'double-counting' within domains. The Health, Education and Access domains potentially contain indicators that overlap with each other. For example, in the Health domain, it is possible for an individual to have had cancer and also potentially to be included in the limiting long-term illness indicator. Combining data using other methods such as 'z scores' more directly double-weights these cases by taking them all into account. Factor analysis, however, takes some account of this overlap in that an indicator may have a lower weight if the contribution it makes has already been taken into account.

The choice of maximum likelihood estimation method

WIMD 2005 follows the methodology applied by Oxford University for WIMD 2000 as well as the Indexes for the other three UK countries.

In Principal Components Analysis all variance in an indicator is analysed including measurement error (*error variance*) and the indicators' imperfect measurement of the underlying construct or constructs (*specific variance)*. This is because it does not attempt to separate *common variance* (i.e. variance shared between three or more indicators) from *unique variance* (i.e. specific variance and error variance). It assumes that an indicator is perfectly reliable and measured without error. It was therefore not appropriate to use the Principal Components method. The appropriate technique, where it is suspected that indicators are not perfectly reliable or measured without error, is *common factor analysis* of which Maximum Likelihood Factor (ML) analysis is a type.

Principal Factoring (PF) has, in the past, been the favoured method of common factor analysis but this was probably because of its relative computational simplicity. With the advent of high-powered computers more sophisticated methods, such as ML factor analysis, are now easily accomplished. PF has a number of disadvantages in comparison to ML factor analysis. The PF solution depends on the scale of measurement of the input indicators (i.e. depends on whether or not they have been standardised), which means that there is not one but an infinity of PF solutions, the choice among, which is arbitrary. The factor model itself is intrinsically scale free, and thus any procedures for its estimation should be scale invariant. ML is scale invariant. ML also treats the correlation matrix as a sample correlation matrix and attempts to explain variance in the *population* correlation matrix. This treatment of the data as a sampled dataset is consistent with the proposal, made throughout this project, that even 'census' indicators should be seen as a sample from a super-population.

Communality

This is the proportion of a variable's variance explained by a factor structure. A variable's commonality must be estimated prior to performing a factor analysis. A communality does not have to be estimated prior to performing a principal component analysis. Communality estimates are estimates of the proportion of common variance in a variable. *Prior communality estimates* are those which are estimated prior to the factor analysis. Common methods of prior communality estimation are to use (1) an independent reliability estimate, (2) the squared multiple correlation between each variable and the other variables, (3) the highest off-diagonal correlation for each variable, or (4) iterate by performing a sequence of factor analyses using the final communality estimates from one analysis as prior communality estimates for the next analysis. *Final communality estimates* are the sum of squared loadings for a variable in an orthogonal factor matrix.

The default setting for communality prior estimates, Square Multiple Correlation, was used for WIMD 2005 calculations.

Calculation Process

The indicators were first transformed to the standard normal distribution. The transformed indicators were then entered into a one common factor Maximum Likelihood factor analysis. Fuller's regression method was used to derive factor scores from the resulting solution. The process was undertaken in SAS Enterprise Guide and the following details the settings used.

- The normally transformed values for each of the domain indicators were entered as the analysis variables;
- Maximum likelihood factor analysis was chosen as the factoring method, for the reasons described above;
- The smallest eigenvalue was set to 1 because this is a commonly used indicator showing that sufficient factors have been extracted to reasonably explain the 'common variance' between the indicators.
- For prior communality estimates the method chosen was Squared Multiple Correlation with all other columns, as described above;
- For the rotation method, no rotation was selected as we are only looking for a single factor solution and rotation only applies if there two or more factors.