

## Exploration of unemployment in Wales by settlement type for small areas

### Introduction

This paper reports an exploratory piece of analysis using unemployment data from the Census of Population, 2011. The Census data are used to examine the distribution of unemployment within Wales. The analysis summarises comparisons of the distribution of unemployment:

- at Output Area (OA) and Lower Super Output Area (LSOA) level
- across a broad settlement size and context classification

The intention is to consider the strengths and weaknesses of the OA and LSOA data and what implications these may have for what data analyses are appropriate. Also to look at whether there are systematic differences between the settlement categories and whether these change depending on the scale of analysis.

The paper is a first contribution to this debate. While it highlights several key findings it is not the final definitive statement on these matters.

### Key Findings

- For analysing unemployment data by settlement type, it is not clear that OA level data is necessarily better than LSOA level data.
  - There are potential data quality issues for very small areas (Output Areas). In the case of unemployment analysis, these raise questions about whether the differences observed between OA and LSOA level data are real, random fluctuations or a product of those data quality issues.
  - The numbers of unemployed people at OA level can be very small and potentially unreliable e.g. more than half the OAs contained fewer than 10 unemployed people. A single household of two unemployed adults moving could have a significant impact on an OAs unemployment rate.
- Overall there is little systematic difference in the picture of unemployment by settlement size and context if the scale of analysis is changed from LSOA to OA.
  - There is a higher level of variation (i.e. more extremes) at OA level – though this must be considered alongside quality issues and the small numbers involved.
  - There are individual areas in the rural areas that have locally high unemployment rates. However, this is both relatively and absolutely less than for the more urban areas
  - The change from LSOA to OA scale does not particularly change the pattern of unemployment in the more rural areas. There is as much, or even more, change in the more urban areas.

- The broad picture at both LSOA and OA is:
  - The more rural areas tend to have lower unemployment rates.
  - They also tend to have relatively few areas with very high unemployment rates compared to the more urban areas.
  - There are still significant numbers of unemployed people in rural areas.

The paper focuses on an analysis of the distribution of unemployment in Wales at OA and LSOA level. Some of the issues relating to usefulness of analysis at a very small level will be relevant to other data variables or data sources, though this cannot be generalised. However a key conclusion from this work is that any assumption that analysis at a smaller level is always better than analysis at a higher level needs to be challenged – taking into account data quality and other issues surrounding the use of smaller areas.

### **Unemployment data from Census of Population 2011**

The data for these analyses are taken from published results from the 2011 Census of Population. The published figures report the employment status for people aged 16 to 74. All calculations of unemployed numbers and rates will exclude people under the statutory school age and people aged at least 75. The following categories are used:

- In work – employees and the self employed.
- Unemployed – not in work but available and looking for work.
- Economically Inactive – not in work and either not available or not looking for work. Includes retired people, students, the long term sick or disabled and people looking after home or family.

Together those in work and unemployed are known as “economically active”. The unemployment rate is expressed as a share of economically active people.

The measure of unemployment used in the Census is not related to whether people claim benefits or not.

### **Statistical geography**

For analysis of data at a local level this paper uses two National Statistics statistical geographies. These geographies were defined for the 2011 Census of Population.

- Output Areas (OA) – areas with around 300 people each at Census 2011. In Wales there are just over 10,000 Output Areas.
- Lower Super Output Areas (LSOA) – areas with around 1,600 people each at Census 2011. In Wales there are just over 1,900 Lower Super Output Areas.

The LSOA are built up of OA. Generally LSOA will contain from 4 to 8 OA. The areas can be added up to give local authority results.

To show the difference in scale of OA and LSOA, Table 1 counts the areas by the number of unemployed people in an area. This is given as a reminder that, while there are many good reasons to want to work with the more locally specific data, there is a price to pay in terms of the volatility of the results for individual areas.

**Table 1**

**OA and LSOA by number of unemployed people in an area**

|              | Number of areas |              | Share         |               |
|--------------|-----------------|--------------|---------------|---------------|
|              | OA              | LSOA         | OA            | LSOA          |
| Under 5      | 1,947           | 0            | 19.4%         | 0.0%          |
| 5 to 9       | 3,913           | 3            | 39.0%         | 0.2%          |
| 10 to 14     | 2,335           | 23           | 23.3%         | 1.2%          |
| 15 to 19     | 1,066           | 71           | 10.6%         | 3.7%          |
| 20 to 24     | 500             | 127          | 5.0%          | 6.7%          |
| 25 to 29     | 172             | 180          | 1.7%          | 9.4%          |
| 30 to 34     | 57              | 182          | 0.6%          | 9.5%          |
| 35 to 39     | 24              | 172          | 0.2%          | 9.0%          |
| 40 to 44     | 16              | 152          | 0.2%          | 8.0%          |
| 45 to 49     | 1               | 148          | -             | 7.8%          |
| 50 to 54     | 3               | 141          | -             | 7.4%          |
| 55 to 59     | 0               | 123          | 0.0%          | 6.4%          |
| 60 to 64     | 1               | 108          | -             | 5.7%          |
| 65 to 69     | 0               | 89           | 0.0%          | 4.7%          |
| 70 to 74     | 1               | 83           | -             | 4.3%          |
| 75 to 79     | 0               | 60           | 0.0%          | 3.1%          |
| 80 to 84     | 0               | 53           | 0.0%          | 2.8%          |
| 85 to 89     | 0               | 38           | 0.0%          | 2.0%          |
| 90 to 94     | 0               | 43           | 0.0%          | 2.3%          |
| 95 to 99     | 0               | 25           | 0.0%          | 1.3%          |
| At least 100 | 0               | 88           | 0.0%          | 4.6%          |
| <b>Total</b> | <b>10,036</b>   | <b>1,909</b> | <b>100.0%</b> | <b>100.0%</b> |

Source: Census of Population, 2011

Table 1, which shows that a fifth of all Output Areas in Wales contained fewer than 5 unemployed people and more than half the OAs contained fewer than 10 unemployed people at the time of the 2011 Census, is key to understanding why using the lower geographical scale may not always be the best thing to do. The other tables in this bulletin need to be considered in the light of the very small number of unemployed people that are found in a single OA. If an OA has, say, five unemployed people in it then a single household of two unemployed adults moving into the area would have a significant impact on the unemployment rate.

### **Data Quality issues**

In using any data about small areas it is important to understand the issues that affect the data at that level and there are a number of such issues that should be taken into account when using the Census of Population data on unemployment for Output Areas:

- A small number of households do not respond to the Census. There are measures in place to account for those households but this does include some imputation of results for these households.
- Census results are recorded for an individual's "usual address". This usual address is how individuals are allocated to the output areas. In some instances there may be people with complex living patterns e.g. students, those with a second address for work purposes, where there may be some question over what the most appropriate allocation of output area is for the analysis being undertaken.
- The Census is a self completion form and the employment status for some people may be mis-stated.
- In an area where very few people have a particular characteristic there may be a risk of disclosing information about individuals. In these areas results can be adjusted slightly for disclosure control.

Sources of error in any data collection can include mis-reporting, mis-allocation to an area, non-response and processes for handling that non-response and possible processing errors<sup>1</sup>. Huge effort goes into the Census programme to minimise any such errors, but they will not all be eliminated. However, as the Census attempts to capture information about all households it is not subject to sampling error in the same way that a smaller survey would be and therefore it provides, in many ways, a best case scenario for the use of OA level data.

At the OA level the numbers of people are much lower than for LSOA. If there are few people in an area then any errors can have a large proportionate effect. This raises the key question about whether the extra geographical precision of the OA level data compared to the LSOA level compensate for the lower accuracy of the OA level data? In practical terms, if we see differences between OA and LSOA level estimates is there any reason to prefer the OA level?

Throughout this paper these issues will be referred to as we try and assess if the Census unemployment results remain "good enough" at the OA level.

## **Rural classification**

Part of the purpose of this paper is to see if there are significant differences between unemployment in rural and urban areas and if this relationship changes at the OA and LSOA levels. Therefore we require a rural classification. Defining urban and rural areas is not a precise science. There are many plausible and feasible ways to do it. Different classifications will be preferred for different purposes. As background for this the "Statistical Focus on Rural Wales" has a useful chapter on "rural definitions and how to choose between them". Although now dated in terms of the available classifications this is still a useful summary of the principles involved. See link in "references".

For this paper we combine elements of the National Statistics Urban-Rural classification<sup>2</sup> with the statistical Built Up Area classification<sup>3</sup> from ONS

---

<sup>1</sup> For a further explanation about the types of errors that could occur see Annex 1

<sup>2</sup> <http://www.ons.gov.uk/ons/guide-method/geography/products/area-classifications/2011-rural-urban/index.html>

There is potential confusion because some local authorities in Wales have the same name as settlements (Cardiff, Swansea, Newport, Wrexham and Bridgend) however the areas are not the same. The boundaries of the 'settlement' represent the built environment and do not respect administrative boundaries. This paper is not concerned with local authorities and all references are to the settlements unless specifically mentioned.

The Built Up Areas are a way of defining "settlements" which were created for the 2011 Census. They are constructed from contiguous areas of built up (developed) land (residential, commercial, industrial, infrastructural). Parcels of developed land are identified and where the gap between two parcels is less than 200m, they are amalgamated. This is known as 'chain linking'. There are 427 built-up areas in Wales.

This chain linking can have a marked effect. For example the Built Up Area of Cardiff extends out to include the separate contiguous areas of Penarth, Pontypridd and Caerphilly.

Not all OAs or LSOAs are included in an identified Built Up Area. Some will be villages and smaller settlements that fall below the limit of detection for the Built Up Areas. Such areas will be assumed to be in the smallest size category.

2011 Census data for Built-Up Areas are produced by aggregating the statistics for one or more whole Output Area (OA) (on a best-fit basis) to form the total estimate. Each OA is either in a single Built Up Area or not in any. Published information from ONS identifies this for each OA. At the LSOA level it is possible for an LSOA to contain OAs in more than one Built Up Area. To allocate an LSOA to a Built Up Area thus needs a best fit for those areas with multiple settlements.

The following settlement size groups are used in this paper.

- At least 100,000 people
- 25,000 to 99,999 people
- 10,000 to 24,999 people
- 2,000 to 9,999 people
- Under 2,000 people (including those in areas without a named Built Up Area)

As well as the overall size of a settlement we also want to consider the context in which the settlement lies. Whether it is close to other settlements or is more dispersed. The National Statistics classification divides Wales and England into three contexts; conurbations, sparse setting and others. There are no settlements in Wales that are classed as conurbations, so for Wales the settlements are categorised as being in the "sparsest" or "less sparse" contexts. In practical terms we can think of this as the difference between "accessible" and "remote" areas.

The classification used in this paper combines the size groups with the context. Note that in the sparsest context there are no settlements with more than 25,000 people.

There are three Built Up Areas with a population of at least 100,000 people in Wales. These are, in size order, Cardiff, Newport and Swansea. In all three cases the Built Up Area extends well beyond the traditional city boundaries as described above.

---

<sup>3</sup> ONS document: 2011 Built-up Areas - Methodology and Guidance. In the 2001 Census these areas were referred to as 'urban areas'

In the sparsest context there are four towns of at least 10,000 people. These are Holyhead, Newtown, Aberystwyth and Carmarthen. Aberystwyth is the largest with a usual resident population of around 18,000. Whether these towns should be treated as “urban” or “rural” is ambiguous because of their highly unusual geography. In some ways they are towns like others of a similar size. However, their separation from other large population centres and their importance within the local area in some circumstances it may be more appropriate to think of them as “rural”. In the following analysis and in the settlement classification considered appropriate for Wales these four towns are included as ‘rural’.

The settlement classification is shown in Map1 and Table 2. The two figures should be read in combination. The large land areas, but small number of people, in the more sparsely populated OA can give a false impression of which categories contain the most people. However, used together a balanced picture is presented.

Taken together the chart and table show the relative scale of the various settlement types. The balance between the larger and small places is important. Over half the people in Wales live in larger settlements of at least 25,000 people. However, just under a third live in smaller settlements of fewer than 10,000 people and nearly 20 per cent in settlements of under 2,000.

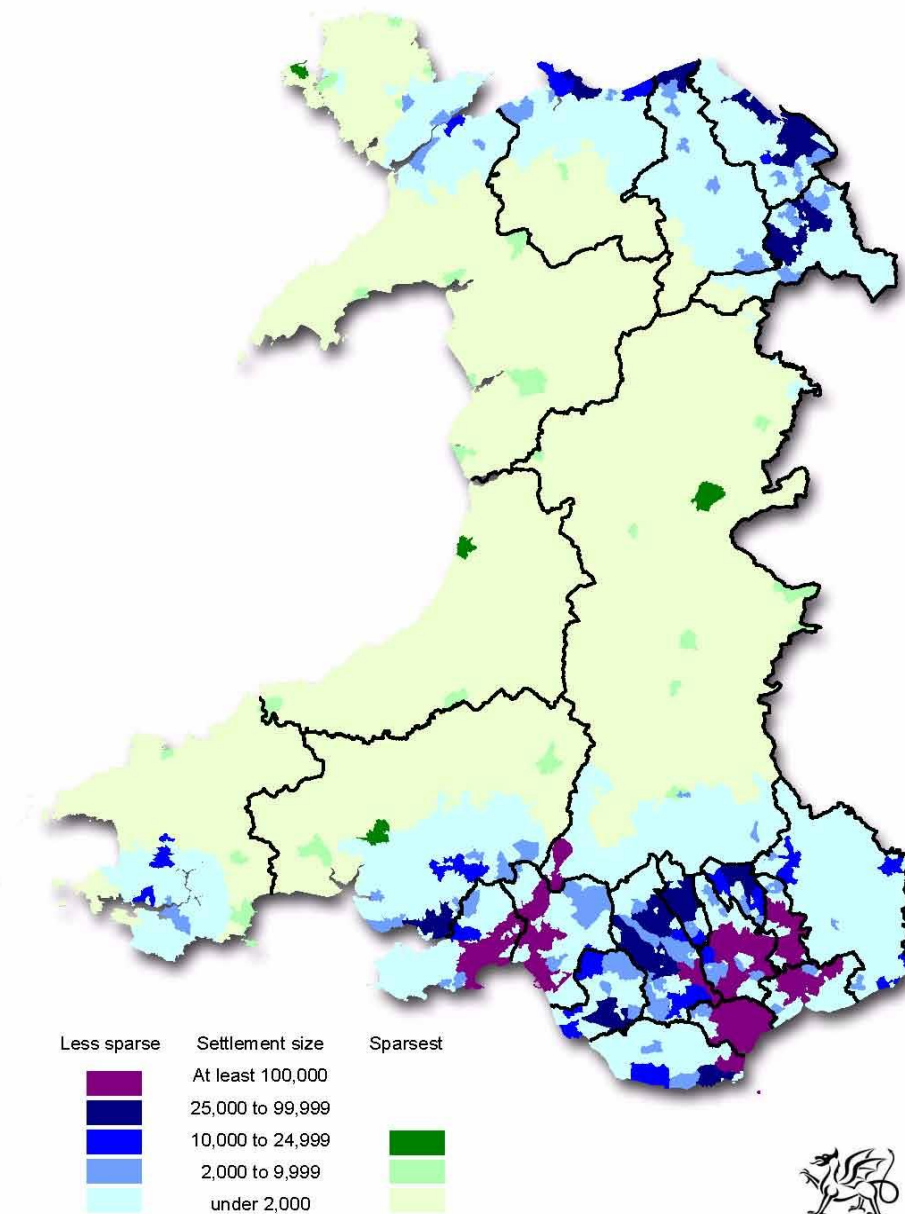
**Table 2**

**People and land in Wales by settlement classification at OA level**

|                            | People<br>(‘000) | Area<br>(sq km) | Share of Wales |               | Persons<br>per sq km |
|----------------------------|------------------|-----------------|----------------|---------------|----------------------|
|                            |                  |                 | People         | Land          |                      |
| <b>Less sparse context</b> |                  |                 |                |               |                      |
| At least 100,000           | 1,054            | 689             | 34.4%          | 3.3%          | 1,530                |
| 25,000 to 99,999           | 600              | 568             | 19.6%          | 2.7%          | 1,057                |
| 10,000 to 24,999           | 346              | 379             | 11.3%          | 1.8%          | 913                  |
| 2,000 to 9,999             | 337              | 679             | 11.0%          | 3.3%          | 496                  |
| Under 2,000                | 302              | 5,874           | 9.9%           | 28.3%         | 51                   |
| <b>Total</b>               | <b>2,639</b>     | <b>8,189</b>    | <b>86.2%</b>   | <b>39.4%</b>  | <b>322</b>           |
| <b>Sparsest context</b>    |                  |                 |                |               |                      |
| 10,000 to 24,999           | 57               | 64              | 1.9%           | 0.3%          | 900                  |
| 2,000 to 9,999             | 90               | 286             | 2.9%           | 1.4%          | 314                  |
| Under 2,000                | 277              | 12,242          | 9.0%           | 58.9%         | 23                   |
| <b>Total</b>               | <b>424</b>       | <b>12,591</b>   | <b>13.8%</b>   | <b>60.6%</b>  | <b>34</b>            |
| <b>Wales</b>               |                  |                 |                |               |                      |
| At least 100,000           | 1,054            | 689             | 34.4%          | 3.3%          | 1,530                |
| 25,000 to 99,999           | 600              | 568             | 19.6%          | 2.7%          | 1,057                |
| 10,000 to 24,999           | 403              | 443             | 13.2%          | 2.1%          | 911                  |
| 2,000 to 9,999             | 427              | 965             | 13.9%          | 4.6%          | 442                  |
| Under 2,000                | 579              | 18,115          | 18.9%          | 87.2%         | 32                   |
| <b>Total</b>               | <b>3,063</b>     | <b>20,780</b>   | <b>100.0%</b>  | <b>100.0%</b> | <b>147</b>           |

Source: Census of Population, 2011 with Built Up Areas and National Statistics rural classification

**MAP 1 Settlement classification for Wales at OA level, 2011**



**Source: Knowledge and Analytical Services**

© Crown Copyright and database right 2014.  
Welsh Government. Licence number 100021874.

  
Llywodraeth Cymru  
Welsh Government  
082.14-15  
Cartographics, FCS

## Unemployed people in a settlement category

The first step is to consider the number of unemployed people in each settlement category and the average unemployment rate. Table 3 shows this for the Output Area level. The table shows for each settlement category:

- People of all ages ("Total")
- People aged 16 to 74 who are economically active – that is in work, self employed or unemployed ("Active").
- People aged 16 to 74 who are unemployed ("Unemp")
- Average unemployment rate – unemployed people divided by economically active people ("Unemp rate").

The number of people and the share of the Wales total are shown for the first three items show above.

**Table 3**

### Unemployment by classification of settlements at OA level

|                     | People ('000) |        |       | Unemp<br>rate | Share of Wales total |        |       |
|---------------------|---------------|--------|-------|---------------|----------------------|--------|-------|
|                     | Unemp         | Active | Total |               | Unemp                | Active | Total |
| Less sparse context |               |        |       |               |                      |        |       |
| At least 100,000    | 36            | 476    | 1,054 | 7.5%          | 37%                  | 34%    | 34%   |
| 25,000 to 99,999    | 22            | 276    | 600   | 7.9%          | 22%                  | 20%    | 20%   |
| 10,000 to 24,999    | 11            | 154    | 346   | 7.4%          | 12%                  | 11%    | 11%   |
| 2,000 to 9,999      | 11            | 155    | 337   | 7.0%          | 11%                  | 11%    | 11%   |
| Under 2,000         | 7             | 146    | 302   | 4.8%          | 7%                   | 10%    | 10%   |
| Total               | 87            | 1,207  | 2,639 | 7.2%          | 90%                  | 86%    | 86%   |
| Sparsest context    |               |        |       |               |                      |        |       |
| 10,000 to 24,999    | 2             | 23     | 57    | 7.9%          | 2%                   | 2%     | 2%    |
| 2,000 to 9,999      | 3             | 40     | 90    | 6.6%          | 3%                   | 3%     | 3%    |
| Under 2,000         | 6             | 133    | 277   | 4.2%          | 6%                   | 9%     | 9%    |
| Total               | 10            | 196    | 424   | 5.1%          | 10%                  | 14%    | 14%   |
| Wales               |               |        |       |               |                      |        |       |
| At least 100,000    | 36            | 476    | 1,054 | 7.5%          | 37%                  | 34%    | 34%   |
| 25,000 to 99,999    | 22            | 276    | 600   | 7.9%          | 22%                  | 20%    | 20%   |
| 10,000 to 24,999    | 13            | 177    | 403   | 7.5%          | 14%                  | 13%    | 13%   |
| 2,000 to 9,999      | 14            | 195    | 427   | 6.9%          | 14%                  | 14%    | 14%   |
| Under 2,000         | 13            | 278    | 579   | 4.5%          | 13%                  | 20%    | 19%   |
| Total               | 97            | 1,403  | 3,063 | 6.9%          | 100%                 | 100%   | 100%  |

Source: Census of Population, 2011 with Built Up Areas and National Statistics rural classification

Table 3 looks at unemployment in three important ways.

- The simple counts.
- The unemployment rate.
- The share of the Wales total.

Having the three measures together allows a fuller picture of unemployment to be painted than any single measure on its own. This is a good reminder that any one measure on its own is incomplete.

These measures all look at the group of settlements as a whole. They show the total number of unemployed people in a group or the average unemployment rate for a group, for example. What they do not show is how areas vary within a group. That question will be addressed in a later section.

The category of settlements of under 2,000 people stands out as different from the others. The average unemployment rate is lower for these settlements than the Wales average and the other settlement sizes. The settlements of at least 2,000 people have reasonably consistent unemployment rates in the range from 7 to 8 per cent. The smallest settlements have an average rate of 4.5 per cent. The smallest settlements in the sparsest context have a slightly lower average unemployment rate than those in the less sparse context.

If unemployment rates were consistent across different types of areas of Wales we might expect 19 per cent of the unemployed people to be in the smallest settlements but they actually have 13 per cent. Whilst this is lower than expected, this still means there are 13,000 unemployed people in these settlements. Particularly in the sparsest context, the impact of unemployment may be more in these areas because of problems accessing the job market and the usual range of support services.

Table 4 extends this idea of considering the “expected” number of unemployed people in an area. Assume that every settlement type has the Wales average unemployment rate. The table shows both the actual and expected number of unemployed people and the difference between them. Again, the table considers each settlement size band as a group.

**Table 4**

**Unemployed people by settlement classification at OA level**

| Unemployed people by settlement classification at OR level | Unemployed people ('000) |        |      | Relative diff | Unemp rate |
|------------------------------------------------------------|--------------------------|--------|------|---------------|------------|
|                                                            | Actual                   | Expect | Diff |               |            |
| Less sparse context                                        |                          |        |      |               |            |
| At least 100,000                                           | 35.6                     | 32.8   | 2.8  | 8%            | 7.5%       |
| 25,000 to 99,999                                           | 21.7                     | 19.0   | 2.7  | 14%           | 7.9%       |
| 10,000 to 24,999                                           | 11.4                     | 10.6   | 0.8  | 8%            | 7.4%       |
| 2,000 to 9,999                                             | 10.9                     | 10.7   | 0.2  | 2%            | 6.9%       |
| Under 2,000                                                | 7.1                      | 10.1   | -3.0 | -30%          | 4.8%       |
| Total                                                      | 86.7                     | 83.2   | 3.5  | 4%            | 7.2%       |
| Sparsest context                                           |                          |        |      |               |            |
| 10,000 to 24,999                                           | 1.8                      | 1.6    | 0.2  | 14%           | 7.8%       |
| 2,000 to 9,999                                             | 2.7                      | 2.8    | -0.1 | -5%           | 6.4%       |
| Under 2,000                                                | 5.5                      | 9.1    | -3.6 | -40%          | 4.1%       |
| Total                                                      | 10.0                     | 13.5   | -3.5 | -26%          | 5.1%       |
| Wales                                                      |                          |        |      |               |            |
| At least 100,000                                           | 35.6                     | 32.8   | 2.8  | 8%            | 7.5%       |
| 25,000 to 99,999                                           | 21.7                     | 19.0   | 2.7  | 14%           | 7.9%       |
| 10,000 to 24,999                                           | 13.2                     | 12.2   | 1.0  | 8%            | 7.5%       |
| 2,000 to 9,999                                             | 13.5                     | 13.4   | 0.1  | 1%            | 6.8%       |
| Under 2,000                                                | 12.6                     | 19.2   | -6.6 | -34%          | 4.5%       |
| Total                                                      | 96.7                     | 96.7   | 0.0  | 0%            | 6.9%       |

Source: Census of Population, 2011 with

Built Up Areas and National Statistics rural classification

"Expect" uses the Wales average unemployment rate of 6.9%

"Diff" is the actual minus the expected value

"Relative diff" is "diff" as a share of the expected value

The differences are generally small in terms of the number of unemployed people. However, the relative differences can be quite large. This raises practical questions of what the impact of this might be in terms either of how it effects people in the area or on the provision of services. Is the relative small difference in the number of unemployed people a key driver or is it more the percentage increase or decrease in unemployment?

**Range of unemployment rates within a settlement category**

The previous tables have looked at the settlement size bands as individual units. In this section we start to look at the variation within a size band.

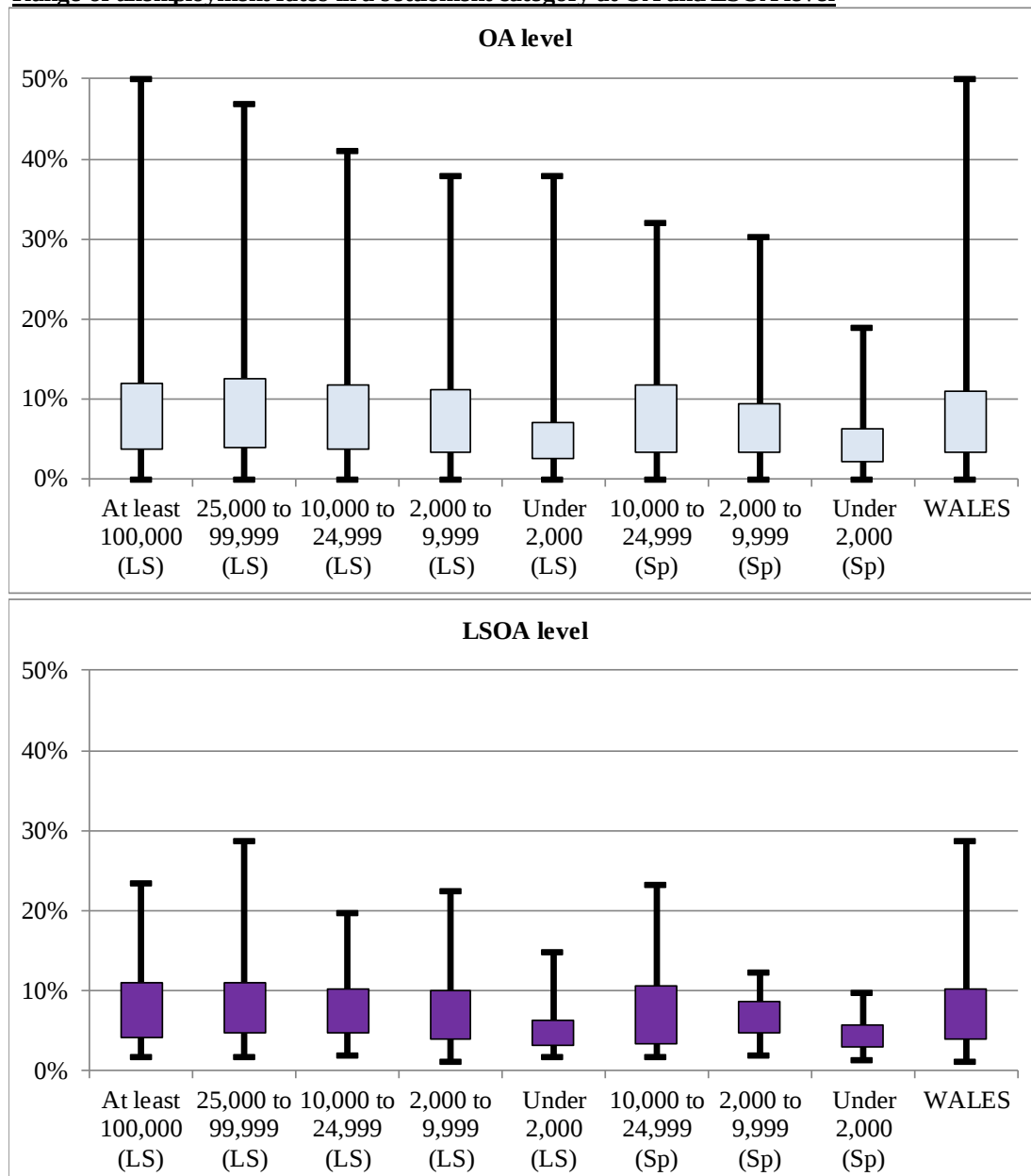
Chart 1 examines the range of unemployment rates occurring in a settlement size band. The chart uses both the OA and LSOA level. The vertical line runs from the

minimum to the maximum unemployment rate for a size band. In any band 60 per cent of the areas (OA or LSOA) lie within the coloured box. There are 20 per cent of the LSOA above the box and 20 per cent below. Table 5 presents the detailed figures.

The chart makes the higher variation at the OA level clear. Given the small number of unemployed people in almost all OA this is not surprising. If there are small errors or misallocations for an individual OA then these can have a significant effect on the unemployment rate of that OA. This relates back to the data quality issues discussed earlier.

**Chart 1**

**Range of unemployment rates in a settlement category at OA and LSOA level**



Source: Census of Population, 2011 with Built Up Areas and National Statistics rural classification

In the chart (LS) refers to the less sparse context and (Sp) to the sparsest context

**Table 5****Range of unemployment rates in a settlement category at OA and LSOA level**

|                       | <u>Minimum</u> | <u>Lower<br/>10%</u> | <u>Lower<br/>20%</u> | <u>Upper<br/>20%</u> | <u>Upper<br/>10%</u> | <u>Maximum</u> |
|-----------------------|----------------|----------------------|----------------------|----------------------|----------------------|----------------|
| <b>OA level</b>       |                |                      |                      |                      |                      |                |
| At least 100,000 (LS) | 0.0%           | 2.6%                 | 3.6%                 | 12.0%                | 15.5%                | 50.0%          |
| 25,000 to 99,999 (LS) | 0.0%           | 2.7%                 | 3.9%                 | 12.4%                | 16.1%                | 46.9%          |
| 10,000 to 24,999 (LS) | 0.0%           | 2.7%                 | 3.7%                 | 11.6%                | 14.6%                | 41.0%          |
| 2,000 to 9,999 (LS)   | 0.0%           | 2.5%                 | 3.3%                 | 11.1%                | 14.7%                | 37.9%          |
| Under 2,000 (LS)      | 0.0%           | 1.7%                 | 2.5%                 | 7.0%                 | 9.2%                 | 38.0%          |
| 10,000 to 24,999 (Sp) | 0.0%           | 2.0%                 | 3.3%                 | 11.6%                | 16.4%                | 32.1%          |
| 2,000 to 9,999 (Sp)   | 0.0%           | 2.1%                 | 3.3%                 | 9.4%                 | 12.1%                | 30.3%          |
| Under 2,000 (Sp)      | 0.0%           | 1.4%                 | 2.1%                 | 6.2%                 | 7.9%                 | 19.0%          |
| <b>WALES</b>          | <b>0.0%</b>    | <b>2.3%</b>          | <b>3.3%</b>          | <b>10.9%</b>         | <b>14.4%</b>         | <b>50.0%</b>   |
| <b>LSOA level</b>     |                |                      |                      |                      |                      |                |
| At least 100,000 (LS) | 1.8%           | 3.4%                 | 4.2%                 | 11.0%                | 13.7%                | 23.4%          |
| 25,000 to 99,999 (LS) | 1.7%           | 3.9%                 | 4.7%                 | 11.0%                | 13.1%                | 28.7%          |
| 10,000 to 24,999 (LS) | 1.9%           | 3.9%                 | 4.6%                 | 10.2%                | 12.7%                | 19.7%          |
| 2,000 to 9,999 (LS)   | 1.1%           | 3.2%                 | 3.9%                 | 10.0%                | 12.4%                | 22.5%          |
| Under 2,000 (LS)      | 1.7%           | 2.7%                 | 3.1%                 | 6.2%                 | 7.5%                 | 14.9%          |
| 10,000 to 24,999 (Sp) | 1.8%           | 3.2%                 | 3.4%                 | 10.6%                | 14.1%                | 23.3%          |
| 2,000 to 9,999 (Sp)   | 1.9%           | 4.1%                 | 4.6%                 | 8.6%                 | 10.1%                | 12.4%          |
| Under 2,000 (Sp)      | 1.3%           | 2.4%                 | 3.0%                 | 5.7%                 | 6.4%                 | 9.8%           |
| <b>WALES</b>          | <b>1.1%</b>    | <b>3.2%</b>          | <b>3.9%</b>          | <b>10.0%</b>         | <b>12.5%</b>         | <b>28.7%</b>   |

Source: Census of Population, 2011 with Built Up Areas and National Statistics rural classification

**Unemployment rate quintiles**

For further analysis we will concentrate on the unemployment rate quintiles for Wales. At the Wales level the areas are ranked by the unemployment rate. We split the areas in five groups with as near as possible the same number of areas in them.

For this paper we will name the groups

- Highest – with the highest unemployment rates
- High
- Medium
- Low
- Lowest – with the lowest unemployment rates

Different ways of building the quintiles are possible. For example, instead of splitting in groups with equal number of areas we could have groups with (broadly) equal numbers of unemployed people. However, for simplicity this paper sticks with the traditional equal number of areas approach.

The lower and upper boundaries of the unemployment quintiles at OA and LSOA level are shown in Table 6. The table also shows the number of unemployed people in each category and the share of the Wales total unemployed. Note that the boundary of the lowest quintile will correspond to the “lower 20%” value in Chart 1 and Table 5 while the highest boundary will correspond with the “upper 20%”.

**Table 6**  
**Unemployment rate quintiles at OA and LSOA level**

| Unemployment rate quantiles at OA and LSOA level |                   |              |                    |                   |             |
|--------------------------------------------------|-------------------|--------------|--------------------|-------------------|-------------|
|                                                  | Unemployment rate |              | Number<br>of areas | Unemployed people |             |
|                                                  | Min               | Max          |                    | Working           | Unemp       |
| OA                                               |                   |              |                    |                   |             |
| Highest                                          | 15.8%             | 15.8%        | 2,007              | 37                | 38%         |
| High                                             | 8.8%              | 9.8%         | 2,007              | 23                | 24%         |
| Medium                                           | 6.0%              | 6.6%         | 2,007              | 17                | 18%         |
| Low                                              | 5.0%              | 4.6%         | 2,007              | 12                | 13%         |
| Lowest                                           | 0.5%              | 2.7%         | 2,008              | 7                 | 7%          |
| <b>Total</b>                                     | <b>0.5%</b>       | <b>15.8%</b> | <b>10,036</b>      | <b>97</b>         | <b>100%</b> |
| LSOA                                             |                   |              |                    |                   |             |
| Highest                                          | 10.1%             | 28.7%        | 381                | 32                | 33%         |
| High                                             | 7.3%              | 10.1%        | 382                | 23                | 24%         |
| Medium                                           | 5.4%              | 7.3%         | 382                | 18                | 19%         |
| Low                                              | 3.9%              | 5.4%         | 382                | 14                | 14%         |
| Lowest                                           | 1.1%              | 3.9%         | 382                | 10                | 10%         |
| <b>Total</b>                                     | <b>1.1%</b>       | <b>28.7%</b> | <b>1,909</b>       | <b>97</b>         | <b>100%</b> |

Source: Census of Population, 2011

For the LSOA level data this shows that the range of unemployment rates in the lower four quintiles is quite uniform at around 2 percentage points. However, the highest quintile has a much larger range from around 10 per cent to 29 per cent. The picture is similar for the OA, but the ranges for each quintile are not quite so uniform. At the OA level for the 20 per cent of OA with the highest unemployment rates the values range from just under 11 per cent to 50 per cent.

The practical implication of this is that if an area increases or decreases its unemployment rate by a single percentage point then it can move a long way up or down the rankings if it started outside the highest quintile. However, in the highest quintile such a change would probably have only a small effect on the rank of the area.

Another practical concern from this table is the share of the unemployed people that fall into each quintile. Since the quintiles are based on equal numbers of areas split by unemployment rates, there are more unemployed people in the highest quintile than in any other. However, note that the highest quintile does not contain most of the unemployed people in Wales. At the OA level it has 38 per cent of the total and 33 per cent at LSOA level.

The majority of the unemployed people in Wales live in areas that are outside the highest quintile for unemployment rate. This means that they are in the region where small changes in the rate can mean a large change in ranking. This has implications in thinking about the balance between local area interventions and general interventions for all unemployed people in Wales.

Maps 2 and 3 show the distribution of the unemployment quintiles across Wales at OA and LSOA levels. The maps are intended to show how the distribution of unemployment changes between the OA and LSOA level. They show where the areas of high or low unemployment occur at both scales.

Overall the two maps do seem to be picking up largely the same pattern, though, of course, there are differences. A problem with such maps is that it is so much easier to see differences in the smallest settlements (where the individual areas have large land areas) than in the larger settlements where the land areas are typically small (see Table 2). For example it is very difficult to see what is happening in the settlement of Cardiff but easy for the local authority of Powys.

For example, the OA map (Map 2) clearly shows an OA on the Powys/Ceredigion border in the highest OA quintile. This is the area around Ponterwyd (OA code W00002902 in LSOA W01000543). It shows up clearly because it includes a large area of nearly empty land, including the summit of Pumlumon. The OA identified had fourteen unemployed people recorded by the 2011 Census with an unemployment rate of nearly 13 per cent. The other OA in the LSOA all have unemployment rates below 5 per cent – well below the Wales average of 6.7 per cent. The LSOA as a whole has an unemployment rate of just over 4 per cent.

This raises another question. Is it important that there is a minor hot spot for unemployment in this area which contains around 300 people, especially when most of its neighbours are in the lower unemployment rate quintile? From table 1 we might expect around 10-15 unemployed people in the area.

The greater variation at the OA compared to LSOA level is clear in maps as it was in Chart 1. Again this leads to the theme running through this paper about whether the data quality issues for the lower level geography outweigh any advantages from using OA level data.

**MAP 2**      Unemployment quintiles at the Output Area level, 2011



**Source: Knowledge and Analytical Services**

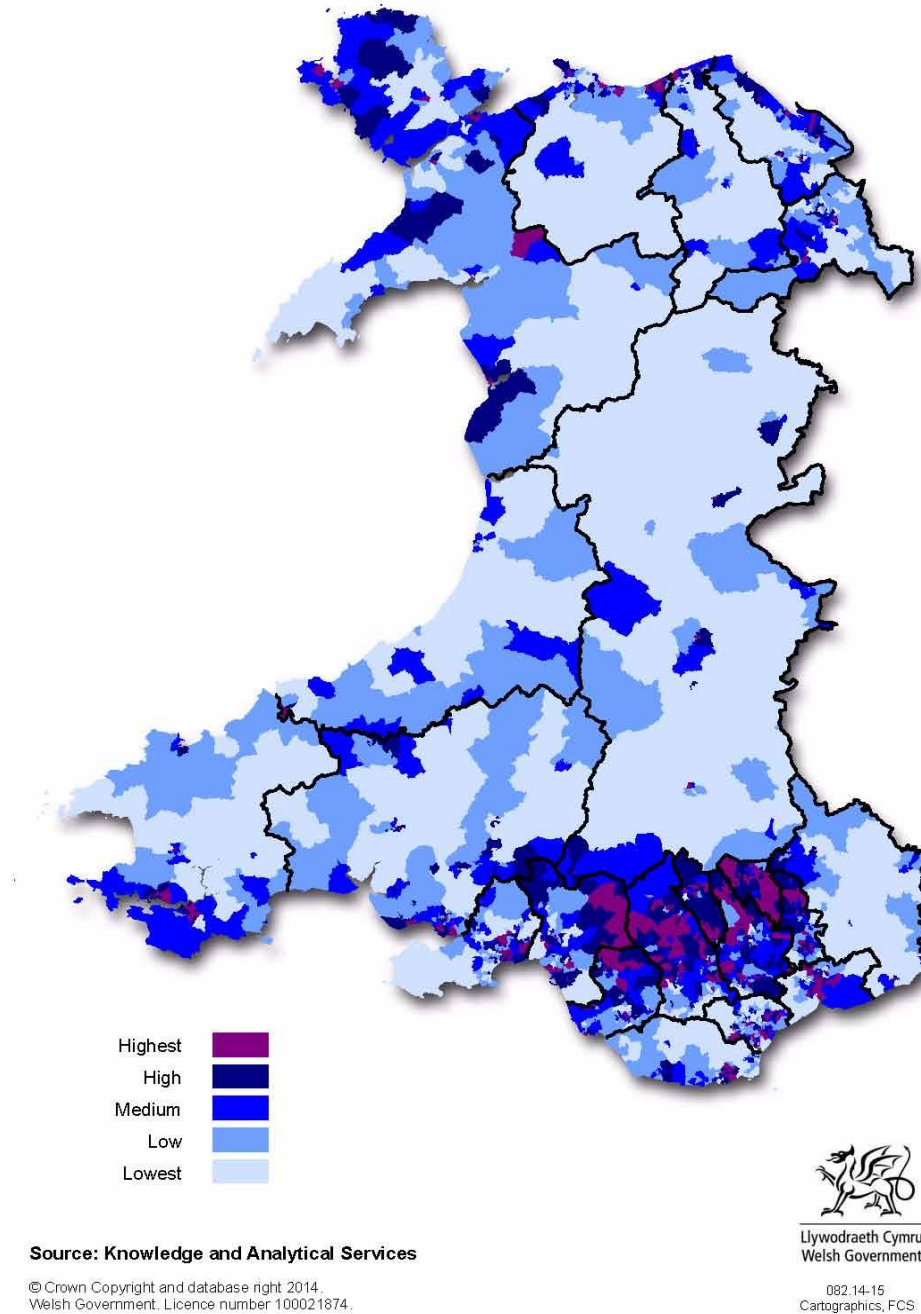
© Crown Copyright and database right 2014.  
Welsh Government. Licence number 100021874.



Llywodraeth Cymru  
Welsh Government

082.14-15  
Cartographics, FCS

**MAP 3      Unemployment quintiles at LSOA level, 2011**



### **Unemployment quintiles by settlement category at OA and LSOA level**

The next step is to examine the distribution of these unemployment quintiles across the settlement categories. Table 7 shows this distribution for the OA and LSOA levels. For each settlement category the table shows the share of the areas in that group that fall into each of the quintiles. If the distribution was entirely uniform each of these values would be 20 per cent. Therefore it is the differences from 20 per cent

for the individual entries that are interesting. In the table the settlements of 10,000 to 24,999 people in the less sparse and sparsest context have been combined because of the small size of the latter. All settlements of at least 25,000 people are in the less sparse context.

**Table 7**

**Share of areas in a quintile at OA and LSOA level by settlement category**

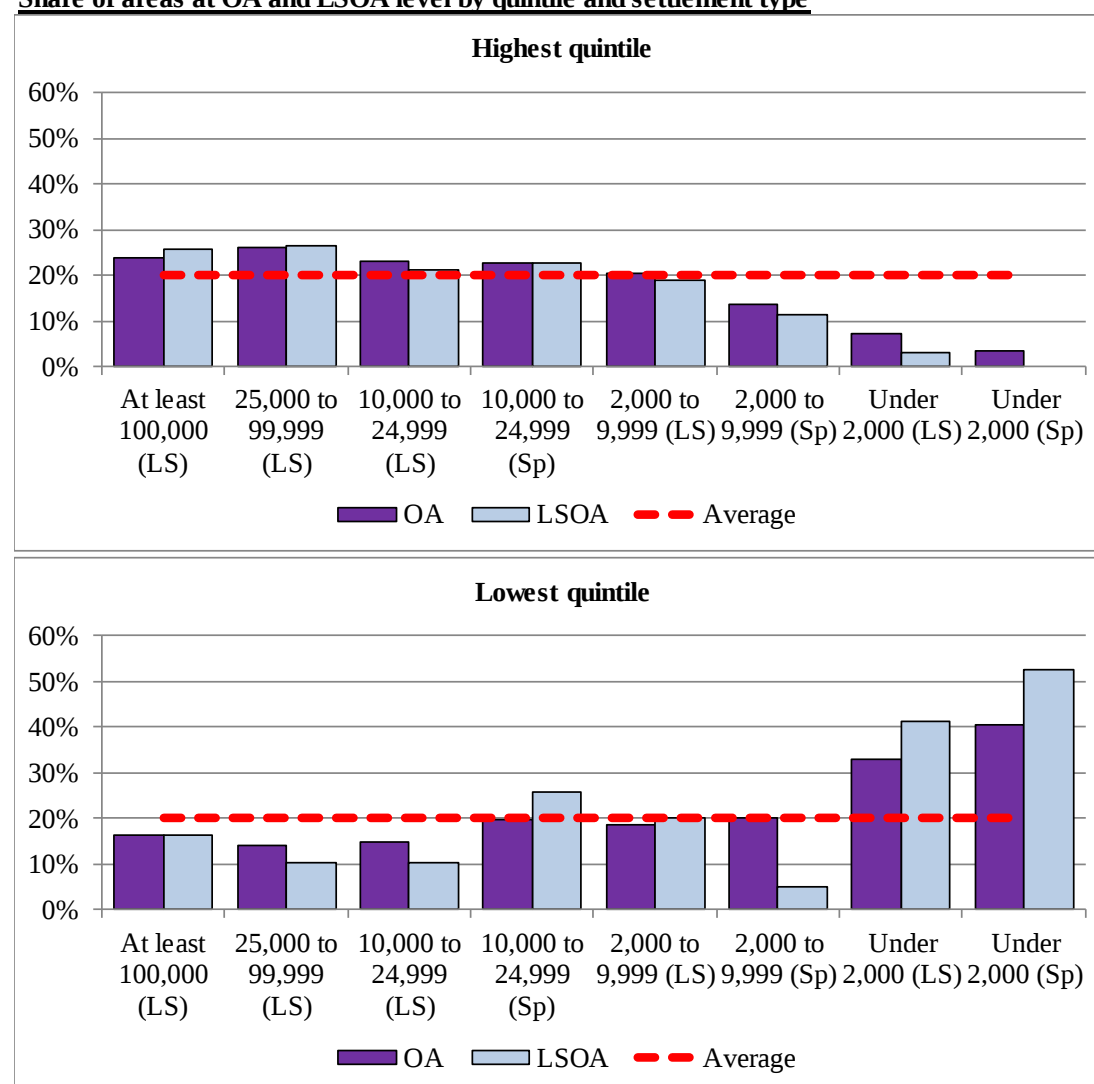
| Share of areas in a quintile at OA and LSOA level by settlement category |          |       |        |       |        |                  |
|--------------------------------------------------------------------------|----------|-------|--------|-------|--------|------------------|
|                                                                          | Quintile |       |        |       |        | Total<br>(=100%) |
|                                                                          | Highest  | High  | Medium | Low   | Lowest |                  |
| <b>At least 100,000</b>                                                  |          |       |        |       |        |                  |
| OA                                                                       | 23.8%    | 21.2% | 20.1%  | 18.5% | 16.4%  | <b>3,404</b>     |
| LSOA                                                                     | 25.8%    | 19.8% | 19.2%  | 18.6% | 16.5%  | <b>666</b>       |
| <b>25,000 to 99,999</b>                                                  |          |       |        |       |        |                  |
| OA                                                                       | 26.0%    | 23.7% | 19.3%  | 17.0% | 14.0%  | <b>1,981</b>     |
| LSOA                                                                     | 26.4%    | 26.4% | 20.2%  | 16.5% | 10.5%  | <b>382</b>       |
| <b>10,000 to 24,999</b>                                                  |          |       |        |       |        |                  |
| OA                                                                       | 23.0%    | 23.3% | 19.1%  | 19.0% | 15.5%  | <b>1,324</b>     |
| LSOA                                                                     | 21.5%    | 24.7% | 22.3%  | 19.1% | 12.4%  | <b>251</b>       |
| <b>2,000 to 9,999 - less sparse</b>                                      |          |       |        |       |        |                  |
| OA                                                                       | 20.7%    | 21.0% | 20.7%  | 19.0% | 18.7%  | <b>1,113</b>     |
| LSOA                                                                     | 18.9%    | 23.9% | 18.9%  | 18.0% | 20.3%  | <b>222</b>       |
| <b>2,000 to 9,999 - sparsest</b>                                         |          |       |        |       |        |                  |
| OA                                                                       | 13.8%    | 22.9% | 24.8%  | 18.5% | 20.1%  | <b>319</b>       |
| LSOA                                                                     | 11.5%    | 16.4% | 27.9%  | 39.3% | 4.9%   | <b>61</b>        |
| <b>Under 2,000 - less sparse</b>                                         |          |       |        |       |        |                  |
| OA                                                                       | 7.4%     | 11.1% | 20.6%  | 27.8% | 33.1%  | <b>978</b>       |
| LSOA                                                                     | 3.1%     | 9.8%  | 19.0%  | 27.0% | 41.1%  | <b>163</b>       |
| <b>Under 2,000 - sparsest</b>                                            |          |       |        |       |        |                  |
| OA                                                                       | 3.4%     | 9.8%  | 19.5%  | 26.8% | 40.5%  | <b>917</b>       |
| LSOA                                                                     | 0.0%     | 4.9%  | 18.9%  | 23.8% | 52.4%  | <b>164</b>       |

Source: Census of Population, 2011 with Built Up Areas and National Statistics rural classification  
For the larger settlements the results combine less sparse and sparsest context

Chart 2 summarises these results for the highest and lowest quintiles.

**Chart 2**

**Share of areas at OA and LSOA level by quintile and settlement type**



Source: Census of Population, 2011 with Built Up Areas and National Statistics rural classification

LS - less sparse context

Sp - Sparsest context

The table and chart show that there are differences between the settlement categories. However, the differences are generally consistent between the OA and LSOA levels. The higher share in the lowest quintile for the smallest settlements is present at both scales. The share in the lowest quintile is higher for the smallest settlements in the sparsest context – the more remote areas.

At the LSOA level the share of areas in the highest quintile broadly increases with settlement size. The share in the lowest quintile decreases with settlement size, again with the lowest share in the settlements between 25,000 and 100,000. The pattern is more irregular at the OA level. However, the largest two settlement size groups still have the highest share of OA in the highest quintile and the smallest size group has the highest share in the lowest quintile.

It is striking that there is much more of a difference between OA and LSOA levels for the *lowest* rather than the *highest* quintiles. For unemployment, the table confirms the picture given by the average unemployment rates for a settlement category. The

larger settlements tend to have more unemployment and the more rural areas less. Notice however, that it also confirms the wide mix of results. Examples of particularly high rates in rural areas and low rates in urban ones can be found.

There is a slight systematic difference between the OA and LSOA levels for the more rural areas. Their share of areas in the higher quintiles is slightly increased and in the lower slightly decreased. However, the scale is not great and needs to be considered alongside the question of whether the OA or LSOA paint the more reliable picture.

A simple split between urban and rural areas may be necessary in order to summarise the complex distribution. However, the table confirms just how wide the variation is within a category.

### Matching OA and LSOA unemployment rates

For every OA in Wales we can find which unemployment rate quintile it is in. We can also find which quintile the LSOA of which the OA is part is in. Obviously all the OA in an LSOA will be in the same LSOA quintile. Table 8 shows the cross tabulation of the OA and LSOA quintiles.

**Table 8**

**OA by unemployment rate quintile at OA and LSOA level, 2011**

| Quintile<br>for LSOA | Quintile for OA |              |              |              |              | Total         |
|----------------------|-----------------|--------------|--------------|--------------|--------------|---------------|
|                      | Highest         | High         | Medium       | Low          | Lowest       |               |
| <b>Output Areas</b>  |                 |              |              |              |              |               |
| Highest              | 1,272           | 442          | 177          | 55           | 23           | 1,969         |
| High                 | 524             | 741          | 465          | 205          | 96           | 2,031         |
| Medium               | 167             | 554          | 603          | 489          | 238          | 2,051         |
| Low                  | 42              | 225          | 528          | 698          | 527          | 2,020         |
| Lowest               | 2               | 45           | 234          | 560          | 1,124        | 1,965         |
| <b>Total</b>         | <b>2,007</b>    | <b>2,007</b> | <b>2,007</b> | <b>2,007</b> | <b>2,008</b> | <b>10,036</b> |

Source: Census of Population, 2011

The diagonal cells (in a box) of this matrix are where the OA and LSOA are in the same quintile. The lower left cells (in orange) are where the OA is in a lower quintile than the LSOA. The upper right cells (in blue) are where the OA is in a higher quintile than the LSOA.

The table shows that all the combinations occur for some OA. There are 23 that are in the lowest OA quintile and the highest LSOA quintile and 2 that go from highest OA quintile to lowest LSOA quintile. However, these extremes are not representative.

Table 9 summarises the results from Table 8. The colour coding of Table 8 shows how the categories in Table 9 are constructed.

- 2+ Lower – OA has a lower unemployment rate than the LSOA it is in and the gap is at least two quintiles (Medium to Highest, for example)
- 1 Lower – OA has a lower unemployment rate than the LSOA it is in and the gap is one quintile (Medium to High, for example)
- Same – OA and LSOA are in the same quintile
- 1 Higher – OA has a higher unemployment rate than the LSOA it is in and the gap is one quintile (Medium to Low, for example)

- 2+ higher – OA has a higher unemployment rate than the LSOA it is in and the gap is at least two quintiles (Medium to Lowest, for example)

**Table 9**

**Summary comparison of OA quintile with LSOA quintile**

|              | OA            | Share       |
|--------------|---------------|-------------|
| 2+ Lower     | 794           | 8%          |
| 1 Lower      | 1,923         | 19%         |
| Same         | 4,438         | 44%         |
| 1 Higher     | 2,166         | 22%         |
| 2+ higher    | 715           | 7%          |
| <b>Total</b> | <b>10,036</b> | <b>100%</b> |

Source: Census of Population, 2011

Overall less than half the OA fall into the same quintile as the LSOA they are in. However, around 85 per cent either stay the same or move by a single quintile. Roughly the same number of OA (7-8%) move at least two quintiles higher or lower.

Table 10 shows the share of OA in a settlement category that falls in these summary bands. The table also shows the total number of OA in a settlement category.

**Table 10****Share of OA by settlement class and difference in OA and LSOA quintiles**

| Share of OA by settlement class and difference in OA and LSOA quintiles | Difference between OA and LSOA quintile |          |      |         |          | Total<br>(=100%) |
|-------------------------------------------------------------------------|-----------------------------------------|----------|------|---------|----------|------------------|
|                                                                         | 2+ higher                               | 1 Higher | Same | 1 Lower | 2+ Lower |                  |
| Less sparse context                                                     |                                         |          |      |         |          |                  |
| At least 100,000                                                        | 6%                                      | 20%      | 47%  | 19%     | 8%       | 3,404            |
| 25,000 to 99,999                                                        | 5%                                      | 20%      | 44%  | 21%     | 10%      | 1,981            |
| 10,000 to 24,999                                                        | 7%                                      | 21%      | 42%  | 20%     | 10%      | 1,138            |
| 2,000 to 9,999                                                          | 7%                                      | 24%      | 42%  | 19%     | 8%       | 1,113            |
| Under 2,000                                                             | 11%                                     | 25%      | 42%  | 16%     | 6%       | 978              |
| Total                                                                   | 6%                                      | 21%      | 44%  | 20%     | 8%       | 8,614            |
| Sparsest context                                                        |                                         |          |      |         |          |                  |
| 10,000 to 24,999                                                        | 8%                                      | 21%      | 44%  | 18%     | 9%       | 186              |
| 2,000 to 9,999                                                          | 12%                                     | 21%      | 37%  | 23%     | 8%       | 319              |
| Under 2,000                                                             | 11%                                     | 25%      | 46%  | 14%     | 4%       | 917              |
| Total                                                                   | 11%                                     | 23%      | 44%  | 16%     | 5%       | 1,422            |
| Wales                                                                   |                                         |          |      |         |          |                  |
| At least 100,000                                                        | 6%                                      | 20%      | 47%  | 19%     | 8%       | 3,404            |
| 25,000 to 99,999                                                        | 5%                                      | 20%      | 44%  | 21%     | 10%      | 1,981            |
| 10,000 to 24,999                                                        | 7%                                      | 21%      | 42%  | 20%     | 10%      | 1,324            |
| 2,000 to 9,999                                                          | 8%                                      | 23%      | 41%  | 20%     | 8%       | 1,432            |
| Under 2,000                                                             | 11%                                     | 25%      | 44%  | 15%     | 5%       | 1,895            |
| Total                                                                   | 7%                                      | 22%      | 44%  | 19%     | 8%       | 10,036           |

Source: Census of Population, 2011 with Built Up Areas and National Statistics rural classification

The table shows that there is a small tendency for the more rural areas to move into higher quintiles at the OA rather than LSOA level. The scale of the change is not great and the table shows that there is variation across all the categories.

This returns to the question of whether the differences between OA and LSOA are improvements or just increases in the amount of noise and random fluctuation.

### **Areas in the highest quintile at OA or LSOA level**

We are often particularly interested in the areas with the highest unemployment rates - the "hotspots". An OA might be in the highest quintile for both OA and LSOA, for one but not the other, or for neither. Table 11 shows the share of OA in a settlement type that fall into these categories.

The areas of interest are where an OA is in the highest quintile at one scale but not the other. The matched areas, either in the highest quintile at both scales or neither scale, are of less interest as they seem to be more securely identified.

**Table 11****Share OA in a settlement class by whether the OA is in OA or LSOA highest quintile**

| Settlement<br>classification | OA in highest unemployment quintile |         |           |         | Total<br>(=100%) |
|------------------------------|-------------------------------------|---------|-----------|---------|------------------|
|                              | Both                                | OA only | LSOA only | Neither |                  |
| Less sparse context          |                                     |         |           |         |                  |
| At least 100,000             | 17%                                 | 7%      | 9%        | 68%     | 3,404            |
| 25,000 to 99,999             | 16%                                 | 10%     | 10%       | 64%     | 1,981            |
| 10,000 to 24,999             | 14%                                 | 10%     | 8%        | 69%     | 1,138            |
| 2,000 to 9,999               | 12%                                 | 8%      | 7%        | 72%     | 1,113            |
| Under 2,000                  | 3%                                  | 4%      | 2%        | 91%     | 978              |
| Total                        | 14%                                 | 8%      | 8%        | 70%     | 8,614            |
| Sparsest context             |                                     |         |           |         |                  |
| 10,000 to 24,999             | 15%                                 | 8%      | 7%        | 70%     | 186              |
| 2,000 to 9,999               | 5%                                  | 8%      | 5%        | 82%     | 319              |
| Under 2,000                  | -                                   | 3%      | none      | 97%     | 917              |
| Total                        | 3%                                  | 5%      | 2%        | 90%     | 1,422            |
| Wales                        |                                     |         |           |         |                  |
| At least 100,000             | 17%                                 | 7%      | 9%        | 68%     | 3,404            |
| 25,000 to 99,999             | 16%                                 | 10%     | 10%       | 64%     | 1,981            |
| 10,000 to 24,999             | 14%                                 | 9%      | 8%        | 69%     | 1,324            |
| 2,000 to 9,999               | 11%                                 | 8%      | 7%        | 74%     | 1,432            |
| Under 2,000                  | 2%                                  | 4%      | 1%        | 94%     | 1,895            |
| Total                        | 13%                                 | 7%      | 7%        | 73%     | 10,036           |

Source: Census of Population, 2011 with Built Up Areas and National Statistics rural classification

Overall the table shows that 7 per cent of the OA are in the highest quintile at the OA level but not at the LSOA level. The smallest settlements have the lowest share of OA in this category. The highest share is for the settlements between 25,000 and 99,999 people, rather than the largest category.

A partial explanation of this may be that the Heads of the Valleys area, which is generally considered to be an area of high deprivation, has several of the settlements in this size band, but none in the largest group.

Notice for the smaller settlements that there is an imbalance between the share of OA in the highest quintile for OA level only compared to LSOA level only. The share in OA only highest quintile is higher than that for LSOA only. By contrast these two categories are relatively balanced for the larger settlements.

### **Mismatches between OA and LSOA**

In this section we consider the mismatches between the highest quintiles at the OA and LSOA level. Establishing how common it is for an OA in the highest quintile *not* to be in an LSOA in the highest quintile. Table 12 divides LSOA into those that are in the highest quintile and the rest. For those outside the highest quintile we split the LSOA into those with none, one or two or more OA in the highest OA quintile.

**Table 12****LSOA inside and outside the highest quintile**

|                          | LSOA in<br>highest | OA in highest quintile |            |              | Total        |
|--------------------------|--------------------|------------------------|------------|--------------|--------------|
|                          |                    | 2 or more              | One        | None         |              |
| <b>Number of LSOA</b>    |                    |                        |            |              |              |
| Highest                  | 381                | 0                      | 0          | 0            | <b>381</b>   |
| High                     | 0                  | 153                    | 184        | 45           | <b>382</b>   |
| Medium                   | 0                  | 19                     | 129        | 234          | <b>382</b>   |
| Low                      | 0                  | 2                      | 38         | 342          | <b>382</b>   |
| Lowest                   | 0                  | 0                      | 2          | 380          | <b>382</b>   |
| <b>Total</b>             | <b>381</b>         | <b>174</b>             | <b>353</b> | <b>1,001</b> | <b>1,909</b> |
| <b>Share of quintile</b> |                    |                        |            |              |              |
| Highest                  | 100%               | none                   | none       | none         | <b>100%</b>  |
| High                     | none               | 40%                    | 48%        | 12%          | <b>100%</b>  |
| Medium                   | none               | 5%                     | 34%        | 61%          | <b>100%</b>  |
| Low                      | none               | 1%                     | 10%        | 90%          | <b>100%</b>  |
| Lowest                   | none               | none                   | 1%         | 99%          | <b>100%</b>  |
| <b>Total</b>             | <b>20%</b>         | <b>9%</b>              | <b>18%</b> | <b>52%</b>   | <b>100%</b>  |

Source: Census 2011

Table 12 shows that OA in the highest quintile are not confined to LSOA in the highest quintile. They can occur anywhere in Wales. However, they are much more common in those LSOA with relatively high unemployment rates. From the set of 764 LSOA in the low and lowest quintiles there are 2 that have more than a single OA in the highest quintile and 40 with a single highest quintile OA. By contrast in the high LSOA quintile 45 out of 382 LSOA have no OA in the highest OA quintile.

Table 13 extends the previous table by considering also the settlement categories. For brevity, the table shows the share of the LSOA in a settlement category that is either in the highest LSOA quintile or outside it. Those LSOA outside the highest quintile are split by the number of OA in them that are in the OA level highest quintile.

The Table confirms the relatively low unemployment rates at either OA or LSOA for the smallest settlements. Instead of the expected 20 per cent of LSOA in the highest LSOA quintile the smallest settlements have 2 per cent. An additional 4 per cent are outside the highest LSOA quintile but have 2 or more OA in the highest OA quintile and a further 13 per cent have a single OA in the highest quintile. This leaves just over 80 per cent of the LSOA that are outside the highest LSOA quintile and have no OA in the highest OA quintile.

The larger settlements categories are more mixed. The share of LSOA in a category outside the LSOA highest quintile and with no OA in the highest OA quintile varies between 40 and 52 per cent. Similarly the shares with one or two or more in the highest OA quintile are higher than for the smallest settlements.

This shows that while there is change because of the change of scale it is not particularly a small settlement (or rural) phenomenon. It can happen across the settlement types and tends to happen more in the larger settlements.

**Table 13****Share of LSOA inside and the highest quintile by settlement type**

| Settlement<br>category | LSOA in<br>highest | OA in highest quintile |     |      | Total<br>(=100%) |
|------------------------|--------------------|------------------------|-----|------|------------------|
|                        |                    | 2 or more              | One | None |                  |
| Less sparse context    |                    |                        |     |      |                  |
| At least 100,000       | 26%                | 8%                     | 16% | 50%  | 666              |
| 25,000 to 99,999       | 26%                | 13%                    | 21% | 40%  | 382              |
| 10,000 to 24,999       | 21%                | 12%                    | 26% | 41%  | 216              |
| 2,000 to 9,999         | 19%                | 11%                    | 19% | 51%  | 222              |
| Under 2,000            | 3%                 | 5%                     | 15% | 77%  | 163              |
| Total                  | 22%                | 10%                    | 19% | 49%  | 1,649            |
| Sparsest context       |                    |                        |     |      |                  |
| 10,000 to 24,999       | 23%                | 11%                    | 14% | 51%  | 35               |
| 2,000 to 9,999         | 11%                | 8%                     | 26% | 54%  | 61               |
| Under 2,000            | none               | 3%                     | 12% | 85%  | 164              |
| Total                  | 6%                 | 5%                     | 16% | 73%  | 260              |
| Wales                  |                    |                        |     |      |                  |
| At least 100,000       | 26%                | 8%                     | 16% | 50%  | 666              |
| 25,000 to 99,999       | 26%                | 13%                    | 21% | 40%  | 382              |
| 10,000 to 24,999       | 22%                | 12%                    | 24% | 42%  | 251              |
| 2,000 to 9,999         | 17%                | 10%                    | 21% | 52%  | 283              |
| Under 2,000            | 2%                 | 4%                     | 13% | 81%  | 327              |
| Total                  | 20%                | 9%                     | 18% | 52%  | 1,909            |

Source: Census of Population, 2011 with Built Up Areas and National Statistics rural classification

**Extreme mismatches**

To illustrate what is happening with the mismatches it may be useful to consider some specific examples in detail. These examples are picked as extreme cases rather than as a representative sample. They are helpful in demonstrating the extremes of what can happen, but should not be taken as being typical.

Table 14 considers 4 LSOA with low unemployment rates. The first two LSOA are in the lowest LSOA quintile but have an OA in the highest OA quintile. The second two are in the lowest two LSOA quintiles but have two OA each in the highest OA quintile. These are the only LSOA with these characteristics. Not shown in the table are an additional 38 LSOA in the lowest quintile with a single OA in the highest OA quintile.

At the other end of the scale, Table 15 shows LSOA in the highest LSOA quintile that have only a single OA in the highest OA quintile. There are 13 LSOA meeting this description. Table 15 is a selection of these to show a mix of larger and smaller settlements.

Table 14

**LSOA with generally low unemployment but outlying OA**

|                                                                          |                 | Economically active aged 16 to 74 |            |              |              |
|--------------------------------------------------------------------------|-----------------|-----------------------------------|------------|--------------|--------------|
| Codes                                                                    | Usual residents | In work                           | Unemployed | Total        | Unemp rate   |
| <b>W01000618 - Solva and Brawdy (under 2,000 sparsest)</b>               |                 |                                   |            |              |              |
| W00003321                                                                | 237             | 93                                | 12         | 105          | <b>11.4%</b> |
| W00003318                                                                | 392             | 195                               | 9          | 204          | <b>4.4%</b>  |
| W00003320                                                                | 238             | 94                                | 4          | 98           | <b>4.1%</b>  |
| W00003322                                                                | 155             | 75                                | 3          | 78           | <b>3.8%</b>  |
| W00003319                                                                | 620             | 427                               | 6          | 433          | <b>1.4%</b>  |
| W00003323                                                                | 235             | 90                                | 1          | 91           | <b>1.1%</b>  |
| <b>Total</b>                                                             | <b>1,877</b>    | <b>974</b>                        | <b>35</b>  | <b>1,009</b> | <b>3.5%</b>  |
| <b>W01000106 - Porthmadog and Morfa Bychan (2,000 to 9,999 sparsest)</b> |                 |                                   |            |              |              |
| W00000540                                                                | 169             | 71                                | 10         | 81           | <b>12.3%</b> |
| W00000541                                                                | 316             | 142                               | 6          | 148          | <b>4.1%</b>  |
| W00000542                                                                | 240             | 98                                | 4          | 102          | <b>3.9%</b>  |
| W00000544                                                                | 303             | 125                               | 4          | 129          | <b>3.1%</b>  |
| W00000543                                                                | 233             | 101                               | 2          | 103          | <b>1.9%</b>  |
| W00000538                                                                | 178             | 69                                | 1          | 70           | <b>1.4%</b>  |
| W00000539                                                                | 293             | 127                               | 0          | 127          | <b>0.0%</b>  |
| <b>Total</b>                                                             | <b>1,732</b>    | <b>733</b>                        | <b>27</b>  | <b>760</b>   | <b>3.6%</b>  |
| <b>W01000351 - Wrexham (25,000 to 99,999 less sparse)</b>                |                 |                                   |            |              |              |
| W00001886                                                                | 245             | 94                                | 15         | 109          | <b>13.8%</b> |
| W00001884                                                                | 297             | 105                               | 14         | 119          | <b>11.8%</b> |
| W00001889                                                                | 298             | 151                               | 11         | 162          | <b>6.8%</b>  |
| W00001890                                                                | 148             | 94                                | 4          | 98           | <b>4.1%</b>  |
| W00001885                                                                | 308             | 174                               | 6          | 180          | <b>3.3%</b>  |
| W00001887                                                                | 234             | 108                               | 3          | 111          | <b>2.7%</b>  |
| W00001891                                                                | 195             | 99                                | 2          | 101          | <b>2.0%</b>  |
| W00001888                                                                | 450             | 243                               | 4          | 247          | <b>1.6%</b>  |
| <b>Total</b>                                                             | <b>2,175</b>    | <b>1,068</b>                      | <b>59</b>  | <b>1,127</b> | <b>5.2%</b>  |
| <b>W01001885 - Cardiff (at least 100,000 less sparse)</b>                |                 |                                   |            |              |              |
| W00009713                                                                | 388             | 117                               | 19         | 136          | <b>14.0%</b> |
| W00010179                                                                | 153             | 50                                | 7          | 57           | <b>12.3%</b> |
| W00009676                                                                | 276             | 82                                | 10         | 92           | <b>10.9%</b> |
| W00009711                                                                | 359             | 209                               | 7          | 216          | <b>3.2%</b>  |
| W00010183                                                                | 294             | 171                               | 5          | 176          | <b>2.8%</b>  |
| W00010207                                                                | 306             | 172                               | 3          | 175          | <b>1.7%</b>  |
| W00010155                                                                | 241             | 153                               | 2          | 155          | <b>1.3%</b>  |
| <b>Total</b>                                                             | <b>2,017</b>    | <b>954</b>                        | <b>53</b>  | <b>1,007</b> | <b>5.3%</b>  |

Source: Census 2011

**Table 15****LSOA in highest unemployment quintile but only 1 OA in highest quintile**

| W01001942 - Cardiff                                        |                 |                                   |            |       |            |
|------------------------------------------------------------|-----------------|-----------------------------------|------------|-------|------------|
| Codes                                                      | Usual residents | Economically active aged 16 to 74 |            |       | Unemp rate |
|                                                            |                 | In work                           | Unemployed | Total |            |
| W01001942 - Cardiff                                        |                 |                                   |            |       |            |
| W00008803                                                  | 293             | 94                                | 32         | 126   | 25.4%      |
| W00008804                                                  | 353             | 157                               | 18         | 175   | 10.3%      |
| W00008807                                                  | 456             | 230                               | 26         | 256   | 10.2%      |
| W00010124                                                  | 253             | 118                               | 9          | 127   | 7.1%       |
| Total                                                      | 1,355           | 599                               | 85         | 684   | 12.4%      |
| W01001015 - Maesteg                                        |                 |                                   |            |       |            |
| W00005382                                                  | 310             | 82                                | 26         | 108   | 24.1%      |
| W00005370                                                  | 292             | 107                               | 13         | 120   | 10.8%      |
| W00005381                                                  | 297             | 127                               | 13         | 140   | 9.3%       |
| W00005379                                                  | 280             | 94                                | 8          | 102   | 7.8%       |
| W00005373                                                  | 296             | 130                               | 10         | 140   | 7.1%       |
| Total                                                      | 1,475           | 540                               | 70         | 610   | 11.5%      |
| W01000043 - Valley                                         |                 |                                   |            |       |            |
| W00000220                                                  | 387             | 139                               | 26         | 165   | 15.8%      |
| W00000221                                                  | 270             | 92                                | 10         | 102   | 9.8%       |
| W00000222                                                  | 257             | 109                               | 10         | 119   | 8.4%       |
| W00000226                                                  | 303             | 131                               | 11         | 142   | 7.7%       |
| Total                                                      | 1,217           | 471                               | 57         | 528   | 10.8%      |
| W01001042 - Pontycymer (2,000 to 9,999 less sparse)        |                 |                                   |            |       |            |
| W00005531                                                  | 355             | 137                               | 24         | 161   | 14.9%      |
| W00005525                                                  | 268             | 99                                | 12         | 111   | 10.8%      |
| W00005524                                                  | 370             | 157                               | 18         | 175   | 10.3%      |
| W00005527                                                  | 345             | 166                               | 13         | 179   | 7.3%       |
| Total                                                      | 1,338           | 559                               | 67         | 626   | 10.7%      |
| W01001358 - Pen Twyn, Caerphilly (under 2,000 less sparse) |                 |                                   |            |       |            |
| W00007122                                                  | 326             | 109                               | 20         | 129   | 15.5%      |
| W00007121                                                  | 238             | 84                                | 14         | 98    | 14.3%      |
| W00007120                                                  | 355             | 94                                | 15         | 109   | 13.8%      |
| W00007104                                                  | 324             | 149                               | 21         | 170   | 12.4%      |
| W00007118                                                  | 297             | 120                               | 4          | 124   | 3.2%       |
| Total                                                      | 1,540           | 556                               | 74         | 630   | 11.7%      |

Source: Census 2011

The tables show a few examples where there is an outlying OA in an LSOA. These extremes are not common. The key question posed by the two tables is how do we interpret individual OA that are out of step with the LSOA that they are in? What is the importance of an isolated OA in the highest quintile in an area of generally relatively low unemployment compared to one in an area of generally high unemployment? This question is particularly important given the small number of people involved, meaning that the analysis is at risk from the sort of minor misspecifications that can happen with any data source.

The LSOA level results are not immune to these problems, particularly for areas with lower unemployment rates. However, the larger size of the LSOA gives a measure of protection against the minor misallocations when we want to find areas with particularly high unemployment rates.

### **Conclusions from this analysis**

The paper has shown an exploratory analysis of the distribution of unemployment in Wales as measured in the 2011 Census of Population. It compares the distribution of unemployment at the OA and LSOA level.

Within settlement groups, settlements or LSOA we can identify changes between the LSOA and OA level. Places that have very high unemployment at one scale may not do so at the other. However, the overall impression is very much one of consistency. The overall picture of the distribution of unemployment is quite consistent at OA and LSOA level. It does not identify large scale and systematic differences.

Particularly the general picture of rural areas having generally lower unemployment rates is consistent at OA and LSOA level – and extends to the local authority level as well. We can identify some examples of rural areas that have relatively high unemployment rates when examined at OA, but not LSOA, level. This is also true, and involves more OA and LSOA, for the more urban areas.

The key issue in deciding which level of analysis is appropriate is data quality. The quality of the data at OA level will, almost inevitably, be worse than the LSOA level. Therefore, for what reason would we assume that the OA level estimates were to be preferred?

In the specific example of the analysis of Census unemployment data we have noted the general similarity of the distribution at OA and LSOA level. It does suggest that the differences are as likely to be from random noise as they are to be genuine.

In this case the general data quality issues that need to be considered for all OA level analysis combine with the results of the analysis that suggest the patterns are much the same anyway. It is therefore, not clear that for analysing unemployment data by settlement type that OA level data is necessarily better than LSOA level data. In fact, it could be argued that LSOA level analysis could be preferred as being more robust.

### **Wider conclusions**

The analysis in this paper focusses on a single issue. Therefore, it is not a definitive statement. However, it does raise some key issues about the use of OA level data that have general relevance.

A key conclusion from this work is that we need to challenge the *assumption* that analysis at a smaller level is always better than analysis at a higher level. Under some

circumstances it may be so. However, there are enough data quality and other issues surrounding the use of smaller areas to make it important that the assumption is tested.

On balance there are advantages and disadvantages of using different levels of data.

The great advantage of working with high level aggregates – say local authority level – is that the results are based on large amounts of data and will be robust. They give a robust picture of the variation *between* categories.

What is lost is a sense of the variation *within* categories. This is a sensible driver to want to examine smaller geographical areas.

The goal should be to assess what the appropriate level of data is for various types of analysis. This gives a justification for the choices rather than just assuming that smaller (or bigger) is better.

This makes more work for the analyst, of course. However, the payback is that the results are more appropriate for the users, within the limitations of the data.

## References and further information

References for the following:

National Statistics Area and Rural-Urban classifications, Office for National Statistics  
<http://www.ons.gov.uk/ons/guide-method/geography/products/area-classifications/index.html>

Built Up Areas, Office for National Statistics  
<http://www.ons.gov.uk/ons/guide-method/geography/beginner-s-guide/census/built-up-areas---built-up-area-sub-divisions/index.html>

Statistical Focus on Rural Wales  
<http://wales.gov.uk/statistics-and-research/statistical-focus-rural-wales/?lang=en>

Welsh Index of Multiple Deprivation 2014: A guide to analysing deprivation in rural areas  
<http://wales.gov.uk/statistics-and-research/welsh-index-multiple-deprivation/?lang=en>

For further information on this work, rural definitions or rural deprivation, please contact.

Stuart Neil  
Agriculture and Rural Affairs Statistician  
029 2082 6822  
[Stats.agric@wales.gsi.gov.uk](mailto:Stats.agric@wales.gsi.gov.uk)

There is always a trade off between the precision of the local geography and the quality and accuracy of the estimates. There is thus always a decision to make about what level of analysis is appropriate. How far can we go before the estimates stop being “good enough” to use? This is not a simple question answered by looking at a single quality measure. However, there are some key issues that always need to be considered.

- Number of people in an area. If there are few people in an area then any errors can have a large proportionate effect. This is particularly true when rates are calculated with the multiplicative effect of errors in both denominator and numerator.
- Misallocation of people to an area. This may arise from mistakes in an address or problems with postcodes, for example. Also from problems trying to allocate complicated cases to a single area. The smaller an area is the closer everyone in the area is to a boundary, making misallocation more likely.
- Miss-response of people in an area. With any response there is a chance that the question might be misunderstood and miss answered. This may be simple error, confusion over what the question means, or trying to force a complex reality into the simple question framework.
- Non-response. For a non-response we do not have the true data but have to estimate. This is similar to when we have to estimate the values from those not selected in a sample to give population estimates. If the non-respondents are a random selection from the population, then non-response is simply a matter of reducing the effective sample size. However, if particular groups, or regions, are more likely to respond than others then a bias is introduced.
- Processing errors in the data. For the Census a number of processes need to be run to produce the final estimates at LSOA, OA or whatever level. Individuals are allocated to their “usual residence”. There are a number of special populations such as armed forces, prisoners, people living in communal establishments and students for which this is not always straightforward. Students, in 2011 were allocated to their term time address. Where households have not responded to the Census values are imputed for the missing households. Finally where there are small numbers of people with particular characteristics in an area there are disclosure control techniques applied to prevent identification of individuals. Survey design. With the Census the intention is to cover the whole population. The issue is with how well domestic households can be identified. The vast majority are simple. However, there are more complicated cases where a dwelling is part domestic and part commercial and so on, some of which may be miss-specified.
- Sampling error. This is not an issue for the Census because it attempts to cover all households. However, it is mention because in most other application it will be crucial. In very simple terms the sampling error from a survey decreases as the number of respondents increases. The practical implication that survey error will be higher for OA than LSOA level estimates.

Each of the bullets above may produce some people who are not allocated correctly. With a source like the Census huge effort has gone into minimising the errors. However, it is not possible to ensure that there are zero errors. Other sources, particularly those that are sample based , will usually have more issues with the data quality. Thus the Census in many ways provides a best case scenario for the use of OA level data.

## **Contact**

Tel: 029 2082 6822

Email: [stats.agrci@wales.gsi.gov.uk](mailto:stats.agrci@wales.gsi.gov.uk)

Produced by Knowledge and Analytical Services, Welsh Government

26 November 2014



All content is available under the Open Government Licence v3.0 , except where otherwise stated.

<http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/>