Choosing Geographic Scales for the Analysis of Labour Market and Related Statistics

Approximately 74,000 excluding Bibliography and following sections

Sally-Ann Goold

December 2020

Submitted to the University of Hertfordshire in partial fulfilment of the requirements for the degree of PhD (Doctor of Philosophy)

Contents

List of tables

List of figures

Abstract

Researchers often only have access to aggregated statistics about people and businesses rather than individual data. This means that research into the relationships between labour market outcomes for geographic areas and demographic characteristics often has to rely on the analysis of statistics for areas rather than for individuals. This presents a problem as the results of the analysis of aggregated, areal data often lead to different results depending on the geographic scale of aggregation. When areal statistics are available for different geographic scales then researchers building statistical models have to choose which geographic scales to include in their models. When areal statistics are not available for different geographic scales then researchers have to consider whether the results of their research would have been different if areal statistics had been available for a different geographic scale to the one that they were forced to use.

That different geographic scales can give rise to different results is important if the results are to be used to inform policies (to improve labour market outcomes for example). That researchers and those using their research are aware that different geographic scales in statistical models can give rise to different results is important as it may focus attention on the importance of choosing which scales to use. It may also help explain differences between different results from similar research projects.

The specific aim of this project was to assess which geographic scales are the most appropriate and useful to include in the statistical modelling of selected UK labour market statistics and which geographic scales provide unhelpful or misleading information. The wider aim of this project was to develop an approach built using one set of labour market statistics that could subsequently be applied to other labour market statistics or other business or socioeconomic statistics in order to provide guidance to researchers on the effects of using different geographic scales for the analysis of areal data. The intention was to create transferable guidance on levels and methods of analysis rather than solely to analyse a single data set.

This project contributes to knowledge by providing some original information about which geographic scales to include in models of various labour market outcomes. Moreover, it contributes to professional practice by describing the different stages used in choosing the geographic scales to include in the modelling of labour market outcomes.

The research described in this report was conducted using multilevel modelling. The R statistical programming environment, R Cran Project (2019), was used to build the models and produce all the figures in the final report. Earlier model building was carried out using both MLwiN software and R. Whilst MLwiN produces user-friendly output which helps in understanding multilevel models, R was chosen for the main modelling as it allowed model building and the creation of charts in one language which could be documented and replicated easily in the form of R scripts, examples of which are included in the Annex to the report. The scripts did not contain functions written as part of the research. Instead, they contained sections of code that built models using parameters named 'Output_variable' and 'Predictor_variable' which could be set to each of the variables required for the models using an earlier section of the script. The data used by the R scripts were read in from csv files stored separately to the scripts rather than being contained in packages. The use of scripts rather than packages simply evolved as the code was written and was sufficient to produce and run the models required for the research. If the work were developed further, then the writing of packages containing code and data to make it easier for other researchers to run the models could be considered.

The data used in the research were all downloaded from official UK government statistics websites. The dataset used for the main section of model building described in chapters 4 and 5 of this thesis consists of outcome variables at local authority level for the 326 English local authority districts and unitary authorities in existence up until early 2019 together with predictor variables mainly at local authority level.

The research presented in chapter 5 of this thesis consisted of three stages, investigating the geographic scale of variation in the outcome variables, choosing the geographic scale to use for predictor variables, and choosing the geographic scales to include as levels in multilevel models. Many of the multilevel models contained one or more of The Europewide 'Nomenclature of Territorial Units for Statistics' (NUTS) geographic scales (Eurostat, 2018) as model level(s). This nomenclature provides a set of hierarchical areas for the collection and analysis of statistics. In the UK, the NUTS 1 areas are Scotland, Wales, Northern Ireland and the nine former government office regions in England. NUTS 2 areas in the UK generally consist of one or more counties depending on county population sizes. Single NUTS 3 areas in the UK can be either a single unitary authority, a group of local authorities or a single county depending on local population sizes, or a single London borough.

The amount of variation at different geographic scales is important as it helps to show how similar units within the same areas are to each other and how different units in different areas tend to be from each other. The geographic scale at which units within areas are similar to each other and units in different areas are different to each other is important in finding which geographic scales it is helpful to have in multilevel models. The main conclusions from investigating the geographic scale of variation in the outcome variables were that:

- for local authority unemployment rates there were higher proportions of variance at NUTS 1 and NUTS 3 areas levels than at NUTS 2 area levels;
- for local authority employment rates and workplace earnings there were broadly similar proportions of variance at NUTS 1, NUTS 2 and NUTS 3 area levels;
- for local authority mean hours and median hours variables there were negligible amounts of variance at NUTS 3 areas level;
- for job density there was a negligible amount of variation at NUTS 1;
- for the median residents' earnings variable there were equal proportions of variance at NUTS 2 and NUTS 3 areas levels and twice that proportion at NUTS 1 areas level.

The main finding from the investigation into the geographic scale to use for predictor variables in models of local authority level outcomes was that it was usually better to use local authority level predictor variables rather than predictors calculated at higher geographic scales and that it was unnecessary to use predictors calculated at multiple geographic scales. For that reason, the main modelling part of the project was devoted to multilevel models of local authority level outcomes using only local authority level predictors.

The research consisted of building a large number of models for each outcome variable using different predictor variables all measured or calculated at local authority level but within multilevel models that grouped the local authorities at different geographic levels. The models were then analysed to see which ones fitted the data better by comparing the AIC values for groups of the models that used the same outcome and predictor variables in different ways. This found the following models to be among the best for the various outcome variables:

- four-level random intercept models for models of unemployment rates, residents' earnings and workplace earnings;

- two-level random intercept models with grouping by NUTS 2 areas for models of mean hours worked and models of job density;

- a variety of models for employment rates depending on the predictor variable used.

An overall finding from the results was that there was often a choice to be made between complex, i.e. random coefficient, models with just two levels and simpler, i.e. random intercept, models with four levels. Given that this choice may have to be made, it was suggested that consideration should be given to what sort of information is sought from the model in order to help choose which geographic levels to include. To learn about influences coming from different geographic scales a random intercept model with many different levels is likely to be appropriate. However, to learn about different strengths of effects in different parts of a study area a random coefficient multilevel model with just two levels or a four-level model with random coefficients at just one level may be more useful.

The recommendations of this project include guidance to researchers on how to choose which geographic scales to include in models. The guidance is presented in the form of a set of steps. The steps cover choosing outcome variables that have distributions suitable for linear modelling, dealing with outliers, building null models to investigate the proportion of variance of the outcome variables that occurs at different geographic scales, considering the intended purpose of the model to determine whether a random coefficient model would be helpful and being aware that the geographic scales to use for random coefficient models may be different to those to use for null or random intercept models, comparing the AIC values of models that include different geographic scale levels to assess which fit the data better, and where appropriate checking for any spatial patterns in the random coefficients estimated by a model.

1. Introduction

1.1 Origins and development

This project arose from the Hertfordshire Business School's 2017 competition for a PhD fee-waver studentship into 'Understanding Spatial Effects in Business Research'. It became focused on how to choose the geographic scales to use when modelling labour market statistics and other business and socioeconomic statistics.

1.2 Research problem

Information about people and businesses is often available only for geographic areas rather than for individual people or businesses. This can be particularly true for survey data where information for areas has been obtained from just a sample of businesses or individuals rather than from all businesses or people in each area.

It is known that analysing areal data at different geographic scales can give different results. Therefore, if source information is available for different geographic scales then researchers and analysts need to decide which scale information to use. If source information is available for only one geographic scale they may not have a choice about which scale to use but could be left wondering whether their results would have been different if they had been able to use data for a different geographic scale.

1.3 Research philosophy

The research for this project was undertaken in line with the positivist research philosophy that originated in the natural sciences. In business research, positivism assumes that the social world exists externally and is viewed objectively; that research is value free; and that the researcher is independent and takes the role of an objective analyst, Burns and Burns (2008). The methods employed were all quantitative, based on the statistical analysis of secondary data that provide measurable facts about people and geographic areas.

1.4 Research aims and questions

The specific aim of the project was:-

To assess which geographic scales are the most appropriate and useful to include in the statistical modelling of selected UK labour market statistics and which geographic scales provide unhelpful or misleading information. The modelling will take account of both hierarchical and proximity effects present in the statistics.

The wider aim was:-

To develop an approach, tool or method built using one set of labour market statistics that could subsequently be applied to other labour market or other business or socioeconomic statistics in order to provide guidance to researchers on the effects of using different geographic scales for the analysis of areal data.

The project aimed to create transferable guidance on levels and methods of analysis rather than solely to analyse a single dataset.

1.5 Original contribution to knowledge

In order to give an empirical example of the research problem the project provided original information on the relationships between local authority level employment and unemployment rates, residential and workplace earnings, hours worked and job density and the factors that affect these labour market outcomes which operate at different geographic scales. The project thereby provided information on the geographic scales of operation of social, economic and environmental factors that influence residents' chances of being in employment, how long they work each week and how much they are paid. In parallel to developing the findings the project provided guidance and insight into ways in which the geographic scale of operation of relationships between business and socioeconomic factors may be analysed.

1.6 Contribution to professional practice

The demonstration that the choice of geographic scales of analyses of areal data can affect results of the analyses should encourage the careful consideration of which geographic scales to use in analyses. Also, the examples of ways in which geographic scales of operation can be analysed serve to provide guidance to researchers on how to choose appropriate geographic scales for their future analyses.

There is the potential for the findings from this project to influence the geographic scales at which areal statistics are made available to researchers and to influence the geographic scales used to inform policy decisions about interventions designed to improve employment chances within local authorities and other geographic areas. For example, the effect of the geographic scale of location identifiers in microdata on their results and usefulness to researchers may led to consideration of the importance of the choice of geographic scale used for microdata.

2. Literature Review

2.1 Introduction to the literature review

In order to provide relevant guidance to other researchers on which geographic scales are the most useful to include in models of areal labour market statistics it is important: 1) to review existing research into the concept that relationships between variables can differ when measured at different geographic scales; 2) to investigate at methodologies used by other researchers to explore the effects of modelling data at different scales and the effects of missing scales out of models; and 3) to find a number of variables commonly used by researchers who model areal labour market statistics. This chapter reports on each of these three aspects in turn.

2.2 Spatial data and relationships at different spatial scales

2.2.1 Spatial data

The defining characteristic of spatial data is that they have a location associated with them. Indeed, at their most basic spatial data are just locations. In business statistics these could be the locations of companies, employees or customers. Analysis of such data includes point pattern analysis exploring whether the locations appear to be random or whether they might have been generated by an underlying spatial process. Ways of conducting point pattern analysis have been described by Greig-Smith (1952) and Diggle (2013) for example.

Many spatial data comprise location information and attribute data. Successful analysis of spatial attribute data needs to take account of the locational information also as individuals, institutions and areas that are close to each other tend to be more similar to each other than those that are further apart (Tobler, 1978). Ignoring the locational information would mean ignoring the fact that individuals or elements that are close geographically are not statistically independent of each other. This should be avoided as assuming observations are independent when they are not increases the chances of type 1 errors increasing the risk that null hypotheses may be rejected when they are in fact true. A good introduction to the importance of this special nature of spatial data is given by Longley et al (2005).

In statistics the tendency for individuals within the same area to be more alike than those in different areas is described as within-area homogeneity. For spatially aggregated statistics this can be measured by the intra-area correlation (IAC). This can be calculated by dividing the variation between the values of summary statistics for different areas (area-level variation) by the variation in individual values for all individuals regardless of the area they are based in (individual-level variation). The IAC gives the proportion of the total variance that is related to geographic areas and provides an indicator of the strength of contextual effects (Lindstrom et al, 2003). It is possible to study the geographic scale at which people or entities tend to be similar by comparing IAC values calculated using different geographic scales. Geographic scales that generate higher IAC values indicate a stronger tendency for individuals within areas at that scale to be similar. Including such geographic scales in multilevel models may generate models that fit the data better than models which include geographic scales that generate lower IAC values (as lower IAC values indicate a weaker tendency for individuals within areas to be similar).

IAC values are calculated specifically for datasets where the individual observations are located in different geographic areas. They give a measure of the proportion of variance that is at area level. This is the same as the value of the Intraclass Correlation Coefficient (ICC) calculated for null and random intercept multilevel models (which is described in sections 2.3.2 and 2.3.3 below).

In geographic information science the tendency for individuals within areas to be similar is formalised by the concept of spatial autocorrelation which describes the degreea to which things that are spatially near to each other are also similar in attributes. Formal measures of spatial autocorrelation include the Moran Index (Moran, 1948, and de Smith et al, 2015) which calculates the correlation between spatial information and attribute information for geographic areas or observations. Given a spatial dataset which contains location and attribute data it can be useful to calculate a measure of spatial autocorrelation for each of the attribute variables of interest. This may provide useful descriptive information about the dataset by itself and also help to determine how important it is to include the locational information in further analysis and modelling. If there is a high positive degree of spatial autocorrelation for a variable, then further analysis and modelling involving that variable should take account of the locational aspect of the data. If there is negligible spatial autocorrelation, then it may not be necessary to take account of the locational information

for that variable. In geography the idea that things that are closer together are more similar than those things that are further apart was described by Tobler (1970). The geographic scale at which this is true can vary enormously and investigating the scale at which things are similar can be an important part of the analysis of spatial data as it can help inform the choice of the appropriate geographic scale for the analysis which is central to this project.

It is understandable that there may be a higher level of spatial autocorrelation at a finer geographic scale than at a courser geographic scale as the closer things are the reasons for them to be similar could be stronger. It is possible that in certain situations there might be a lower level of spatial autocorrelation at a finer scale than at a particular courser scale that is important for the variable in question. Both of these situations can be dealt with by calculating measures of spatial autocorrelation at different geographic scales and then choosing the geographic scale that displayed the higher positive spatial autocorrelation as the geographic scale to use for further analysis and modelling.

2.2.2 The implications for traditional statistical methods

In order to work successfully and produce trustworthy estimates of standard errors traditional statistical methods rely on the assumption that individual observations are independent. As spatial data tend not to be independent (as nearby things tend to be more similar than those further apart) most traditional statistical methods underestimate standard errors. This is because using lots of similar observations effectively means the number of independent observations is smaller than the total number of observations. The effective sample size is therefore overestimated which is what leads to the underestimated standard errors. Using underestimated standard errors can cause Type 1 errors where null hypotheses that variables are not related are falsely rejected. This can lead to false conclusions that relationships do exist between variables.

2.2.3 Proximity and hierarchy

People and places that are geographically closer to each other tend to be more similar than those that are further apart (Tobler, 1970, Longley et al, 2005, Tranmer and Steel, 2001a, 2001b). Also, people are affected by their surroundings and local opportunities. In order to account for these tendencies, the analysis of business and social statistics for geographic areas should consider taking account of the proximity of different areas for example by

incorporating a measure of distance such as 'as the crow flies', or a travel time or cost measure.

The similarity of people living near to each other is used to generate official and commercial area type classifications. The Office for National Statistics (ONS) has published residential area classifications for a variety of different geographic scales (ONS, 2019h). The largest geographic areas for which classifications are available are local authorities. Local authorities each have a subgroup, group and supergroup classification. The classification is hierarchical. Subgroups are the lowest level in the hierarchy, groups are the middle level and supergroups are the highest level. For example, Torbay on the Devon coast is classified as a 'Seaside Living' area at subgroup level, a 'Remoter Coastal Living' area at group level and a 'Countryside Living' area at supergroup level. Describing a different coastal area, Blackpool is classified as a 'Service Economy' area at subgroup, a 'Services, Manufacturing and Mining Legacy' area at group level and a 'Services and Industrial Legacy' area at supergroup level. Details of how the classification for 2011 Census output areas were created are given by Gale and Longley (2013). Details of the commercial ACORN consumer classification of areas are available from CACI (2019) and information about the MOSAIC consumer classification is available from Experian (2018). Area type information can provide contextual information that can be used in the analysis of business-related areal statistics.

People and places are also affected by factors related to their local government areas. This can be true at many geographic scales from NUTS 1 level (Eurostat, 2018) regional influences and policies through to local parish council decisions that might affect local businesses. The Europewide 'Nomenclature of Territorial Units for Statistics' (NUTS) provides a set of hierarchical areas for the collection and analysis of statistics (Eurostat, 2018). In the UK, NUTS 1 areas comprise Scotland, Wales, Northern Local and the nine former government office regions in England. NUTS 2 areas in the UK generally equate to one or more counties depending on county population sizes. Single NUTS 3 areas in the UK can consist of either a single unitary authority, a group of local authorities or a single county depending on local population sizes. In London each borough forms its own NUTS 3 area. Administrations make and implement policies that can affect the social, cultural and economic opportunities available to their residents (and those living somewhat further afield if transport links are good). In the UK local government is organised on a hierarchical

basis. For example, many of England's local electoral wards are nested within local authority districts which are in turn nested within counties and then NUTS 1 areas. In other areas a flatter structure of unitary authorities (UAs) exists, however, these can still be grouped by counties and are nested within NUTS 1 areas.

2.2.4 Cross-correlation and the modifiable area unit problem

When working with spatial data it can be found that the value of a dependent variable may depend on the values of independent variables for surrounding areas as well as those for the dependent variable's own location. This phenomenon is known as cross-correlation. It can lead to scale problems if it takes place at a different geographic scale to the scale of analysis. Flowerdew et al (2001) noted that it can be difficult to find the scale at which cross-correlation operates and that this means that regression models of the same variables measured at different geographic scales can have different regression coefficients. Scale problems are one part of the modifiable areal unit problem (MAUP). The other part is known as the aggregation problem whereby the analysis of areal statistics calculated from the same building blocks aggregated differently can give different results and is outside the scope of this research.

2.2.5 Simpson's paradox

The term Simpson's paradox can be used to describe the fact that it is possible for two variables to be positively correlated in two sub-groups of a population yet to be negatively correlated when the two sub-groups are amalgamated and the data for the entire population are analysed together. It is the reversal of the direction of the correlation that is significant. The name refers to an article by Simpson (1951) describing how the association between variables in contingency tables can be reversed when all cases are analysed together.

For geographic data Simpson's paradox can occur due to a phenomenon known as the modifiable area unit problem (MAUP) whereby statistical analysis can give different results depending on the size and shape of the geographical units chosen for the analysis. The MAUP is described by Openshaw in publications including a book of that name (Openshaw, 1984).

A good example of how the choice of geographic units for the analysis of spatial data can change the results is given by Wilson (2013). The example describes how different

researchers studying the relationships between concentrations of property foreclosure and crime data using different areal units were producing differing results. To investigate in detail the author modelled crime in North Carolina based on levels of foreclosures and factors such as disadvantage at both US Census tract level and US Census block level. US Census tracts are small areas with populations of between 1,200 and 8,000 people (United States Census Bureau, 2020) for which statistics are published. US Census blocks are subdivisions of tracts (United States Census Bureau, 2020) with smaller populations, or even with no residents, which form the building blocks of all higher level US Census geography areas (United States Census Bureau, 2011). Analysis at block level showed that higher residential stability was associated with lower crime rates whereas analyses for the same study area at tract level showed higher stability was associated with higher crime rates. This displayed a clear example of Simpson's paradox as aggregating the finer geographic scale Census block level data to the coarser geographic scale Census tract level had the effect of reversing the direction of the association between the crime and residential stability variables. The explanation offered for this was that local spatial effects happening at the finer block level were not being detected when the coarser geographic scale US Census tract level data were used for the analysis.

The importance of analysing spatial data for spatial autocorrelation at different levels before carrying out further analysis was noted by Wilson (2013). In this example there was found to be significant positive spatial autocorrelation at block level but not at tract level. Again, this suggested that the tract-level analysis was not picking up local spatial effects that were occurring at block level.

Separate scatter plots for crime and foreclosure data were shown for block and tract levels with regression lines with a positive gradient on the block level plot and a negative gradient on the tract level plot. Maps for block and tract-level data were also shown highlighting areas with combinations of high and low crime and high and low foreclosure levels. Different spatial patterns could be seen on the two maps reinforcing the analytical finding and suggesting a way of visualising the effects of analysing spatial data at different geographical scales. The article went on to report the results of nearest neighbour analysis to find the average distance between foreclosure properties. Nearest neighbour analysis involves finding the observed minimum distances between a set of points, e.g. crime scenes or woodland trees, and then comparing the distribution of the distances with that which would be expected if the points had been generated randomly, i.e. generated by a completely random spatial process (CRS). The comparison makes it possible to quantify

whether the points are likely to be clustered due to a non-random cause or to be randomly distributed (by a CRS process). Essentially, it provides a way of indicating whether points are clustered or randomly distributed in space. Nearest neighbour analysis has been used by many researchers and is discussed in depth by Anselin (2000) and in summary by Rossbacher (1986), for example. The research by Wilson (2013) suggested that strong clustering of foreclosure rates was present in the dataset.

Kernel density estimation was then used by Wilson (2013) to produce a surface to visualise the clusters of foreclosures. Kernel density estimation is a technique used to measure the density of point events by considering how many events there are in a circle of a certain radius around each possible point in a study area (Anselin, 2000, and page 337, Longley et al, 2005). By overlaying block and tract boundaries on the foreclosure density surface it was possible to see that the blocks were better than the tracts at capturing the spatial concentrations of foreclosures.

The main analytical considerations raised by the research were: that it demonstrated that Simpson's paradox had an effect on the results of the analysis; that using larger geographical units compromised the analysis; that it supported other authors who had suggested that the results of their work might have been different if they had used smaller geographical units. Finding the geography that best captures spatial effects in data was suggested as one way of mitigating Simpson's paradox. The main policy consideration of the research was that urban policies to address particular problems should be formulated at the same geographical scale as the problems themselves.

An article by Ma (2015) explores Simpson's paradox in the analysis of economic data when aggregated at different levels. One example was that of studies of unemployment rates in the USA during the 1980s recession and in the 2008/9 downturn. For each stratum of society, defined by educational attainment, unemployment rates were higher in 2009 than in the 1980s, whilst the overall unemployment rate was lower in 2009 than in the previous period as the proportions of people in each stratum had changed between the 1980s and 2008/9. An economic example with particular relevance to geographic scale was that of aggregating per capita GDP for different countries over time. Specifically, per capita GDP had increased between 1996 and 2001 for both developed and developing countries but as population growth had been stronger in developing countries than in other countries, overall per capita GDP had decreased. Essentially there were relatively more people living in lower GDP countries at the end of the period than in higher GDP countries meaning the

average GDP had decreased. This demonstrates that data analysed at a larger geographic scale, i.e. all countries, may show a different trend from data analysed at a finer scale. In effect it misses the detail happening at the finer scale.

Two important points to take forward from these examples of Simpson's paradox are the need to take account of the geographic scale at which individuals or entities tend to be similar to each other and that using coarser geographic scales may miss important relationships that exist between variables at finer geographic scales.

2.3 Methodologies to investigate effects of scale in models

2.3.1 Multilevel modelling

Multilevel modelling is known by a number of different names. One of the most self-explanatory is hierarchical modelling which references the hierarchical, or nested, nature of the data being modelled. An important reason for the use of multilevel modelling is that it allows the total variance in a dataset to be partitioned into the amounts of variance at different levels in the model. If the levels are nested geographic scales, then multilevel modelling can show how much of the variation in the data is due to each geographic scale. This can indicate which geographic scales may be important in the model. A leading authority on multilevel modelling in the UK is the Centre for Multilevel Modelling (CMM) based at the University of Bristol. It has a comprehensive website providing information and guidance (CMM, 2019a) and has developed the MLwiN software (CMM, 2019b). Multilevel modelling is suitable for modelling hierarchical areal data as by including the different hierarchical geographic scales as levels in the models the models can take account of and report on effects that occur at different levels in the hierarchy of areal units.

A major reason for the early development of multilevel modelling was the desire to analyse school effectiveness data. It provides a technique which can be used to analyse the effects of different influences on pupils' educational development and outcomes. The influences tended to be arranged hierarchically. At the heart of the system were individual pupils who formed the first level in most multilevel models. Pupils were typically nested within classes which were nested in schools which were nested in geographic areas such as local education authorities. A thorough introduction to multilevel modelling is given by Goldstein (1995); many topics are covered in great detail by Snijders and Bosker (2012); a

comprehensive introduction is provided by Kreft and de Leeuw (1998) and a very clear first introduction is provided by Robson and Pevalin (2016).

Despite frequent references to schools and pupils due to the early development and extensive use of multilevel modelling in school effectiveness studies (Goldstein et al, 2000, Openakker and Van Damme, 2000, Van den Noortgate et al, 2005, Moerbeek, 2004, and Van Landeghem et al, 2005), multilevel modelling can be used for any datasets where a hierarchical model can be used to explain the structure of the influences on the primary units. Statistically the models are designed to investigate the amount of variation in the data that occurs at each level in the model.

2.3.2 Null models

The most basic multilevel models are termed null models as they do not contain any explanatory variables other than group membership information. For example, a two-level null model would contain a variable measured at level one and information on which level two unit each level one unit belonged to. In business research the data could be incomes of individual employees and knowledge of which employees worked for which firm. A null multilevel model can be used to provide an initial assessment of how appropriate it is to proceed with multilevel modelling by providing information on how much of the variability in the data is between groups and how much is within groups for example by calculating the Intraclass Correlation Coefficient (ICC). For example, Lee (2000) notes that calculation of the ICC is often the first step in hierarchical linear modelling and the ICC value obtained can be used to gauge whether hierarchical rather than single level modelling should be considered.

2.3.3 Intraclass correlation coefficient

For random intercept multilevel models, i.e. for null models with just a random intercept and for models with a random intercept and one or more predictor variables with fixed slopes, the ICC is exactly the same as the proportion of variance explained at group level rather than at individual level. The proportion of variance at group level can be referred to as the Variance Partition Coefficient (VPC), Goldstein et al (2002) and Browne et al (2005). For random slope models the ICC is not simply the proportion of variance explained at group level as the varying slopes mean that a unique ICC value does not exist. For random

intercept models an ICC value can be calculated as described below in words in equation 1 and in mathematical terms in equation 2.

ICC = Variance of the Intercept / (Variance of the Intercept + Variance of the Residuals)   (1)

ICC = ρ=variance between groups/total variance = $\tau^2/(\tau^2 + \sigma^2)$ (2)

In equation 2, σ^2 is the population within group variance, a quantity which can be estimated by the observed within group variance, and τ^2 is the population between group variance, which cannot be estimated solely by the observed between group variance. τ^2 can instead be estimated by the observed between group variance less a 'correction factor']equal to the observed within group variance divided by ñ where, following Snijders and Bosker (2012, p.20),

$$\tilde{n} = \frac{1}{N-1}\left\{ M - \frac{\Sigma_j n_j^2}{M} \right\}$$ (3)

where N is the number of groups, M is the total sample size, and n_j is the number of units in group j.

The observed within group variance can be calculated as an average of within group variances weighted by the group sizes.

In situations where the groups in a multilevel model are geographic areas then the ICC for null and random intercept models measures the same concept as the Intra-Area Correlation described in section 2.2.1 above. Both the ICC and the IAC seek to determine the differences between groups compared with overall differences or alternatively the similarities in observations within each group compared to the overall differences both of which are effectively the same concept – the extent to which an individual is similar to others within the same area combined with the degree to which typical individuals in separate groups tend to be different to those in other groups.

Multilevel modelling is more relevant in those situations where a higher degree of the variation is between the groups, i.e. when the variation between group-level statistics

accounts for an appreciable proportion of the total variation of all individual-level statistics resulting in a high ICC value. In such situations the relationships between variables (both the random effect group means and the random effect group coefficients) are more likely to be different in different groups than it is in situations where there is little variation between group-level statistics.

Where there is not an appreciable proportion of variation between group-level statistics it is often suggested that there is little if any need to use multilevel modelling. Indeed, Nezlek (2008) simply says that "various authors" suggest that where ICC values are zero or low then multilevel modelling is not appropriate implying that this is a view widely held by many researchers (but not by Nezlek himself). In more detail, Lee (2000) suggests that an ICC value of 10% or more could indicate that multilevel modelling should be considered when analysing differences in educational achievement between schools. Lee (2000) reports typical ICC values of around 25% to 30%. Robson and Pevalin (2016) say that there is no agreed 'cut-off' ICC value to determine whether multilevel modelling is required and refer to Lee's (2000) suggestion of 10% or higher being large enough to indicate that multilevel modelling should be used rather than single level modelling. Hox (2002) says that an ICC of 10% seems reasonable within educational and organisational contexts where an ICC of 15% may be regarded as high. Snijders and Bosker (2012) quote a report by Hedges and Hedburg (2007) as giving ICC values for educational performance studies in American schools as ranging from 10% to 15%. In contrast to the idea that multilevel modelling is not required for data with a low or zero ICC value, Nezlek (2008) argues that multilevel modelling should be used whenever the data to be modelled have a hierarchical structure (and more than a very small number of groups at each level of the hierarchy). Nezlek (2008) says that the reason it is important to use multilevel modelling whenever the data have a hierarchical structure is that if hierarchical data are analysed in a way that does not take account of their hierarchical structure then the results of the analyses may not be accurate. Nezlek (2008) explains that by ignoring the hierarchical structure in data and not using multilevel modelling techniques to model the data then one is assuming that the relationships between different variables are the same across groups when in fact the relationships may be different in different groups. The reason why it is important to use multilevel modelling rather than relying on single level regression modelling is that units in different groups, in this project people or small areas within larger areas, are more likely to

be alike than units in different groups meaning the assumption of independence of observations that is required for single level regression is not met by data that have a hierarchical structure. If units within groups tend to be more similar than units in different groups, i.e. they are not independent, then the effective group size will be lower than the actual group size and if this is not accounted for then the standard errors in models of the data will be underestimated leading to a higher risk of Type 1 errors than indicated by the p-values generated by the models. Huang (2018) explains that this is still true even for very small ICC values by stating that even for a low ICC value of 0.01 the chance of a Type 1 error is 0.20 which is four times higher than the rate would be if the data were independent and the ICC was therefore zero. This shows that even for the relatively small ICC values found for some of the outcome variables used in this project when grouped by administrative and statistical areas it is still necessary to take account of the clustering of outcome values within these hierarchical administrative and statistical areas across England. Multilevel modelling is therefore appropriate to use for this project despite some of the ICC values being relatively small as it is a method of modelling that is ideally suited to modelling data that have a hierarchical structure and takes account of the fact that the units within each group are not independent of each other. The levels should be kept in multilevel models even if the ICC values are small. It is important as the levels themselves are important as they relate to the hierarchical structure of the data. This is in contrast to situations such as blocking in ANOVA models where blocking is left in models even if it is not important.

For very small numbers of groups, e.g. up to five, Nezlek (2008) suggests a separate single level regression analysis could be carried out for each group. That would allow the relationships between variables to be studied separately for each group. The relationships found to exist within each group could then be compared to see how similar they were to each other. If the relationships were found to be different for each group that would provide useful information on the separate relationships and also show that a single level regression analysis for all the data points together was not appropriate.

There is a history of using different measures to describe the degree of similarity between observations from the same higher-level unit. As early as the 1920s Fisher described methods to calculate the degree of similarity between observations in the same higher-level unit, or class, and used the term intraclass correlation coefficient to describe the similarity (Fisher, 1925). This early version of an ICC could take negative as well as positive

values. By the 1970s there were many different versions of ICC as evidenced by Shrout and Fleiss (1979) who provided guidelines on how to choose which of six different ways of calculating an ICC measure to use in different circumstances. In 2012 Wu et al compared five methods of calculating ICC for binomial variables. More recently, Koo and Li (2016) published guidance on how to choose which of a total of ten different ICC formulae to use for different conditions and said it was important for researchers to publish details of the version of ICC they had used and for readers to check whether such information had been reported.

2.3.4 ICC for proportions

If variables are a proportion at the level of interest, having been created from binary variables at a lower level, and have thus been modelled as binomial variables in a generalised linear model, e.g. using a logit link function, then their ICC values cannot be calculated using the simple formula of Variance at Group level / (Variance at Group level + Variance of Residuals). This is because the residuals from logistic generalised linear models are the differences between the observed and fitted logits rather than the differences between the observed and fitted values of the variable. The residuals produced by a such a generalised multilevel linear model have mean 0 and have been shown (Snijders and Bosker, 2012, p. 305, and Wu et al, 2012) to have variance equal to $\pi^2/3$ (=3.29). ICC values for such models can therefore be calculated using equation 4.

ICC = Group level variance / (Group level variance + 3.29)   (4)

2.3.5 Examples of the use of ICC

ICC values were used by Ashworth and Armstrong (2003) to compare the attitudes of general practitioners (GPs) grouped in partnerships to a) prescribing issues and b) management issues. They discovered low ICC values for prescribing issues suggesting that GPs' attitudes to prescribing were not related to which partnership they belonged to. However, they found high ICC values for management attitudes suggesting that GPs' attitudes to management issues were related to which partnership they belonged to. These findings mean that it would not be necessary to use multilevel modelling to analyse GPs' prescribing practices, but it would be helpful to use multilevel modelling to analyse GPs' management related attitudes (with GPs grouped within partnerships).

2.3.6 Random intercept models

Slightly more complex multilevel models contain group membership information and predictor variables at level one. Such a model has random intercepts, i.e. each group may have a different intercept, but each group has the same, i.e. fixed, coefficient for the predictor variables. Both null models and those with random intercepts and fixed coefficients are known as random intercept models.

2.3.7 Random coefficient models

More complex still are random slope models where both the group intercepts and the group coefficients can vary between groups. Random intercept and random slope models can contain predictor variables at all levels in the models and can also include terms to model the interactions between levels. They can measure whether the relationship between two variables varies according to a third context variable. For example, a model could investigate whether the relationship between wages and qualification levels was different in different types of company. Examples of multilevel model equations are given in sections 4.2.5.1 and 4.2.5.2.

2.3.8 The effects of missing out levels in a multilevel model

One of the most important steps in carrying out multilevel modelling successfully is defining the levels to be used in the model. The levels should match the inherent hierarchical structure of the data. Decisions need to be made about which levels to include in the model in order to model the data sufficiently well to study relationships between different variables whilst being subject to the constraints of data availability and reliability at different levels.

Research by Tranmer and Steel (2001a, 2001b) addressed this problem in a review of the effects of missing out a level in the analysis of 1991 UK Census data for individuals, enumeration districts (EDs) of around 500 people and local authority wards using two different approaches. The first approach was a theoretical approach in which variance and covariance matrices were created for a three-level model consisting of individual, ED and ward levels, and then compared with variance and co-variance matrices for three alternative models each based on two levels only. These calculations showed that missing

the ward level from the model had little effect on the individual level but added the ward-level effects to the ED level, i.e. missing out the top level from the model for data structured at three levels transferred the effects to the middle level. Missing out the ED (middle) level from the model caused its effects to be transferred to both the individual (bottom) level and the ward (top) level. Finally, missing out the individual level from the model caused the individual-level effects to be transferred to the ED (middle) level and did not change the ward (top) level effects. This final finding is of particular relevance to the work reported in this thesis as researchers often only have access to aggregate rather than individual-level data. It can be argued that this means an individual level will be missed out from most models used in this research. These findings suggest that any individual effects that there might have been will be transferred to the lowest level of geography that is included in the model and will not change the effects for any higher levels of geography that are included in the model. Therefore, if the focus of the research is at middle to higher geographic levels then little information is lost by not including individual data in the models.

The second approach was a practical one that combined the use of individual and aggregate data. The individual data came from a 2% sample of individuals in the 1991 Census in the 'Reigate and Banstead with Tandridge' Standard Anonymised Record (SAR) district. The aggregate data were for enumeration districts (EDs) and wards. For confidentiality reasons only anonymised data were available in the individual data and the lowest geography identifier for each individual was that of SAR district. Local authority districts with populations of 120,000 or over form single SAR district in their own right whereas less populous districts are merged together to form composite SAR districts with suitably large populations to ensure confidentiality. Data from the individual and aggregated data sets were used to compare estimates of variance that would occur in a three-level model of individuals, EDs and wards with those for two alternative models one of which missed out the wards and the other missed out the EDs. Missing out the wards appeared to transfer the ward-level effects to the ED level of the model. Missing out the EDs transferred the ED-level effects to either up to the ward level or down to the individual level depending on which variable was used in the analysis. Missing out the individual level generally led to its effects being transferred to the ED level. Thus, the results from the practical approach were consistent with those for the theoretical approach.

In the field of school effectiveness studies an article by Opdenakker and Van Damme (2000) discussed the importance of specifying levels for multilevel analysis. The research was

based on a reference model using four levels: pupil, class, teacher and school. Six alternative models were created each of which missed out one or two of the four levels. The results from the six alternative models were compared with those for the four-level reference model. The comparisons showed firstly that missing out any of the middle or higher levels affected the proportion of variance at those levels that were included meaning that it was important to include all four levels; that missing out either of the higher levels, i.e. the school or teacher levels, transferred the effects of the missing level to the highest level that was included in the model; and that missing out one of the middle levels transferred its effects to both the level just above and the level just below.

Van den Noortgate et al (2005) discussed the effect of missing a level in the multilevel analysis of school effectiveness data. The article noted that the effects of missing a level in multilevel analysis was an important research topic as there was often not sufficient data to include all levels in a multilevel analysis. The article described the remodelling of the data from the 2000 Opdenakker and Van Damme article. It concluded that missing a top or middle level from a multilevel model often effected the results of the modelling. The article found that correctly choosing levels to include in a multilevel model was very important and that in order to study effects at a certain level it was important to include the levels immediately above and immediately below the level of interest.

Moerbeek (2004) wrote about the consequences of missing out a level in the analysis of social science data using the three levels of pupils, classes and schools in a three-level reference model. In general terms the results of missing out any of the three levels in an alternative two-level model was found again to transfer the effects of the missing level to the other levels. Not including the school (top) level transferred its effects to the class (middle) level. It also gives a falsely high standard error for the coefficient of the class level predictor, but it does not change the standard error of the coefficient for the pupil level predictor. Missing out the class level transferred its variance to both up to the school level and down to the pupil level. It also gives a falsely high standard error for the coefficient of the pupil level predictor, but it does not change the standard error of the coefficient of the school level predictor.

The Moerbeek (2004) article was followed by the publication of a comment on the article by Van Landeghem et al (2005). This commented on the specific conditions referred to in the original article and was not convinced that in general it could be argued that the changes in sizes of standard errors were as predictable as the original article claimed. The

comment suggested that researchers put resources into trying to obtain data for all relevant levels rather than try to counteract the effects of missing certain levels.

2.3.9 Geographically weighted regression

One way of taking account of the distances between points or between areas when working with spatial data is to use geographically weighted methods such as geographically weighted regression (GWR). This builds on ordinary least squares regression by weighting the predicator variables according to their distance from each point being modelled. The location of an area can be summarised by its geographic or population-weighted centroid. GWR generates a separate regression model for each point in the dataset. The output from a GWR model includes residuals, coefficients and their standard errors for each point. By studying these, for example by mapping the residuals, coefficients and standard errors, it is possible to see how the relationships between variables can differ across space. The coefficients can show how the strength, and even the direction, of the relationship between a predictor and response variable can be different across a study area. Hypothetically, they could show that in some parts of a country higher qualifications have a strong influence on incomes whereas in other parts they have a much weaker influence. Details of geographically weighted methods of analysis are provided by Fotheringham et al (2000) and examples are illustrated by Fotheringham (2016). A disadvantage of using GWR alone is that it simply works on the distances between places and does not take account of any hierarchical grouping of areas. It could not, for example, take account of the fact that two areas that are very close together have different education policies. This means that GWR by itself would not be sufficient to find the geographic scales that should be used to model labour market and related statistics. Although future research could explore combining GWR and multilevel modelling to find appropriate geographic scales, it is not necessary to use GWR in order to find which existing administrative and statistical area scales should be included in such models.

2.3.10 Research suggesting multilevel modelling can account for proximity effects

Moellering and Tobler (1972) carried out fairly early research into the appropriate geographic scales to study processes. They noted that human society was often arranged in nested hierarchies, e.g. government units, towns, counties and census areas, and that available data were often aggregated by spatial units. They used analysis of variance

techniques to examine the proportions of variance at each level in geographic hierarchies. They suggested that analysing data at different levels of a hierarchy was equivalent to carrying out the analyses at different geographic scales (on the grounds that the higher levels areas were physically larger than the lower levels that they contain).   Effectively they suggested analysing the amount of variance at different available levels to examine the effects of spatial scale in census (rather than survey) data where the data are available for fully nested hierarchical areas as a way of studying which geographic scales are appropriate to model underlying processes. This can now be done by using multilevel modelling techniques and software not available to Moellering and Tobler. It is noteworthy that they suggested using existing hierarchical areas as an appropriate way to study geographic scale as this supports the idea of using multilevel modelling for hierarchical areal data for existing hierarchical areas as the main approach for this project (rather than using geographically weighted techniques to search for new statistical areas for the analysis of data).

2.4 Labour market statistics

Labour market statistics cover a variety of subjects related to work and employment. The ONS describes labour market statistics as measures of work, jobs, the workforce, types of work done, payment and other benefits, and working patterns (ONS, 2019a). The *Journal for Labour Market Research* (SpringerOpen, 2020) publishes research about labour markets, employment, education, training and careers. Research in labour market statistics can include research using statistics for different geographic areas, industrial sectors and for people with different demographic characteristics. In order for this work to be particularly relevant to future researchers and modellers of labour market statistics it was important that the dependent and independent variables used in this work included some of those commonly used in published research.   It was also appropriate that established definitions of unemployment and employment were used. This chapter therefore includes a selection of references to published analysis and research by other researchers working in the field of labour market statistics research, in particular research based on official UK statistics. It also provides information on central definitions of unemployment and employment. It is important to note that although unemployment and employment are highly related they are not simply complementary as can be seen in the definitions given below.

2.4.1 Income and pay levels

ONS estimates of labour market outcome statistics including the median weekly earnings of full-time workers by local authority are available (ONS, 2019c). The estimates cover the whole of the UK except for a small number of local authorities for which not enough data were available to produce estimates. These weekly earnings estimates provide good outcome variables for models of labour market statistics. As the lowest geographic scale for which they are available is that of local authority the lowest level that could be used in multilevel models of earnings is local authority level. Similar statistics for larger local geographies including travel to work areas and NUTS 3 areas are also available (ONS, 2019d). The UK Census does not include questions on income and therefore the 2011 and earlier Censuses cannot provide information on earnings by smaller geographic areas across the UK. Depending on the focus of research income and pay levels can be regarded as dependent, labour market outcomes that depend on other socioeconomic variables such as available workers and jobs, or could in some circumstances be regarded as independent variables that could be used to model other labour market outcomes such as rates of unemployment or employment.

An ONS survey (ONS, 2015a) found rates of long-standing illness or disability (LSI) were higher for people with lower incomes. The rates of LSI ranged from a high of 45% for people earning under £10,000 to a low of 24% for those earning over £50,000. This link between income and a health-related variable and the fact that it is reported by the ONS provides evidence that future researchers into income levels may be interested in the use of health in models of income levels and that it is therefore appropriate to include health as a covariate when modelling income levels in this project.

An international example of modelling household income in Sweden by Gomez-Rubio et al (2008) included age, gender and education information for heads of households among the independent variables. This confirms that researchers, including those using data from outside the UK, make use of age, gender and education levels as covariates when modelling labour market statistics showing again that it is appropriate to include these demographic variables in this project.

Another international example of modelling income levels for geographic areas, this time in Mexico, is given by Tzavidis et al (2018) who modelled total household per capita income with predictors including age, education and socioeconomic class. This shows that it is relevant to include measures of age, education and socioeconomic class in this project. Socioeconomic class information can be included in models by using variables based on occupational group.

2.4.2 Highest qualification

There is evidence that people with higher qualifications tend to travel further to work than those with lower qualifications. For example when travel to work areas (geographic areas where the majority of people both live and work within the same area) are generated separately for people with different qualification levels a total of 416 areas emerged for people with low qualifications whilst only 53 areas emerged for people with higher qualifications, (ONS, 2016c). As this project is concerned with geographic scales for the analysis of labour market statistics it is therefore relevant that information on qualification levels is considered for use in its models. The smaller number of travel to work areas for people with higher qualifications indicates that people with different qualification levels travel different distances to work, on average. This implies that different geographic scales should be used to model employment for people with different levels of qualification. This is beyond the scope of this current project but could be informed by the findings from this project. Qualification levels are generally used as independent variables in models of dependent labour market outcomes. However, it could be argued that area-level qualification rates might depend on the area-level abundance of jobs paying different salaries. Bell and Blanchflower (2011) analysed youth unemployment by a number of different variables including educational attainment and thus provided evidence of the use of many of the socioeconomic variables used in this research, including highest qualification, to model labour market outcomes (including unemployment and hours worked for example). Felstead et al (2013) reporting on skills and employment in Britain, carried out analyses of the proportions of jobs requiring a degree on entry and the proportion of jobs requiring no qualifications on entry showing that qualifications held, and in particular having a degree or having no qualifications, were variables used by academics researching labour market outcomes. Phimister et al (2006) also included qualifications held as a variable in their research on low pay in urban, as opposed to rural, areas demonstrating that other researchers also used qualifications held as a variable in research

into labour market outcomes. Phimister et al (2006) also included a number of other variables including age, gender, housing tenure and industrial sector in their analyses of ways of escaping low pay demonstrating that other researchers used variables similar to Bell and Blanchflower (2011). Cribb et al (2015) report on the percentage of workers with higher education qualifications by industrial sector demonstrating the use of qualifications held as a variable in the analysis of labour market statistics.

2.4.3 Hours worked and balance of full-time and part-time work

Analysis of travel to work areas for full-time and part-time workers found that on average full-time workers travel further to work than part-time workers (ONS, 2016c). This implies that different geographic scales might be appropriate for models of employment-related statistics for people in full-time and part-time jobs. Alternatively, a summary area-level statistic for mode of employment, e.g. percentage of part-time workers could be a relevant variable for models of employment rates for example. This would not remove the problem of different geographic scales being more relevant for people with different characteristics but it might improve the fit of models built for all workers regardless of their different characteristics. Bell and Blanchflower (2011) showed that young people who are working part-time would generally prefer to work more hours whilst other people working part-time would prefer to work fewer hours. Cribb et al (2015) use full-time employment as a variable in their research into public sector pay showing that the mode of employment, full-time or part-time, is used by other researchers in their analysis of labour market statistics. In some models the median number of hours worked or the proportion of workers who work part-time could be the dependent variables whilst in others they could be independent variables depending on the focus of the research. In this thesis the number of hours worked is used as a dependent variable as it may indicate the availability of work in local areas whilst the proportion of workers who work part-time is used as an independent variable as it may help to explain dependent variables such income levels and employment rates by indicating the availability of full-time work.

2.4.4 Health

The ONS (2019b) analysis of people who have never worked showed large numbers of the people were counted as either long-term or short-term sick. This gives evidence of the link between sickness and whether a person is in employment. Further evidence is given by a survey in 2014 in which the proportion of unemployed people who reported a long standing illness or disability (LSI) was found to be higher, 33%, than the proportion of employed people who reported a LSI, 24% (ONS, 2015a). The survey report reported those findings to be in line with the 2013 Labour Force survey which had found that the unemployment rate was higher for people with a disability than for people who did not have a disability. The fact that these ONS analyses of employment and unemployment include health related variables indicates that researchers into employment and unemployment use health as a covariate. This shows that it is relevant for this project to include health related variables in models of unemployment rates. Murray et al (2019) analysed retirement ages using data from the 1946 British Birth Cohort. This is relevant as early retirement by individuals can be associated with lower area-level employment rates (and also with higher area-level unemployment rates). Their models included local authority area level unemployment rates, binary individual level health status (fair/poor health or excellent/good health for individuals at mid-life). Earlier research by Arrow (1996) provides more evidence that researchers sometimes use health as an independent variable in their studies of unemployment. In general health is likely to be an independent variable used to model dependent labour market outcomes. However, an area-level prevalence of certain occupations may have an influence on population health, perhaps more in the past than today, and area-level high unemployment or job insecurity or low pay can still have an effect on population health, so in some research area-level health might be used as a dependent variable that depends on good job opportunities. Shahidi et al (2016) studied the differences between self-reported health for employed and unemployed people across 23 European countries using data from the 2012 European Social Survey. They grouped good and very good health as one category and combined fair, bad and very bad as the other category creating a binary individual-level variable and used age, gender and years of schooling as control variables. Virtanen et al (2013) also studied statistical relationships between self-reported health status and unemployment. They also included gender and social economic position, in terms of occupational group, in their models. That they were analysing links between health and unemployment and using gender and occupational group in their research shows that researchers use these variables

together. This therefore provides further support to the inclusion of health, gender and occupational group variables in this project.

2.4.5 Age

The ONS (2019b) analysis of people who have never worked shows the proportions to be very different in different age groups. Bell and Blanchflower (2011) show that youth unemployment tends to be higher than adult unemployment and more sensitive to economic conditions. The differences between employment and unemployment rates for younger and older workers indicate that it is appropriate to take age into account in models of employment and unemployment rates. This provides both a reason to include age as an independent variable in models of unemployment rates in this project (as unemployment caused by having never worked is more prevalent in some age groups than others) and evidence that other researchers and analysts use age as a variable when studying employment and unemployment. In research relating to unemployment in different age groups in the UK during the great recession, which started in 2008, published by Bell and Blanchflower (2010) it was shown that younger people were particularly badly affected suffering an increase in unemployment and a decrease in employment. Phimister et al (2006) include age as a variable in their analysis of what happens to workers following periods of low pay showing that other researchers also use age in their research into labour market outcomes.

2.4.6 Gender

The high proportion of those looking after home or family who are female (ONS, 2019b) provides one example of the different employment patterns for males and females. Another example of employment analysis by gender is given by the inclusion of the comparison of male and female employment rates in analysis of employment by travel to work areas (ONS, 2016c).   Bell and Blanchflower (2011) analysed youth unemployment by gender and found it to be higher amongst males than amongst females.

Charts of employment, unemployment and economic inactivity rates by gender over time in ONS statistical bulletins (2019k) show that although the differences in employment and economic inactivity rates for males and females have reduced substantially in the past forty years there are still differences. These differences show that it is still relevant to consider

45

gender as a variable in models of employment and economic inactivity. It is outside the scope of this project, but future research could include building separate models for male and female unemployment rates.

A chart of separate male and female unemployment rates over time shows that although rates are generally very similar, particularly in recent years, there have been differences in the past. The chart shows that usually during periods of rapidly rising unemployment the unemployment rate grows more steeply for men than for women leading to periods when the male unemployment rate is distinctly higher than the female unemployment rate. That unemployment rates for males and females have changed at different speeds in the past shows that information on gender should be considered for inclusion in models of unemployment. More generally the fact that separate economic activity information is still provided for males and females in ONS statistical bulletins suggests that others researching employment statistics may still be using gender as a variable which gives another reason for including gender as a variable in this project.

A House of Commons Library (2019) briefing paper on labour market statistics at UK region level provides analysis of employment, unemployment, mode of employment and average earnings by gender providing evidence that employment status, mode of employment and gender are variables used by researchers analysing labour market statistics for different parts of the UK.

Gender was also used as a variable in research into periods of low pay by Phimister et al (2006) and by Cribb et al (2015) in their research into public sector pay in the UK. These examples demonstrate the use of gender by other researchers analysing labour market outcomes, especially pay levels.

2.4.7 Type of industry

There has long been research into the relationship between industrial diversity and the stability of unemployment rates to explore the idea that whilst greater industrial diversity may protect areas from economic downturns, greater industrial specialisation may give faster growth in employment, e.g. McLaughlin (1930), Jackson (1984), Tran (2011), Felix (2012), Feser et al (2014), Deller and Watson (2015), Trendle and Shorney (2004), Mason and Howard (2010), O'Donoghue (1999) and ONS (2017b).

There are a number of diversity indicators in use. The Herfindahl index is a simple indicator of the amount of industrial diversity in an area. It is calculated using the proportions of those employed in an area that are employed in each of the k industries or industrial sectors in an area. It is a measure of the chance of any two randomly selected employees in an area being employed in the same industry. As higher values indicate a higher chance of two employees being in the same industry, i.e. lower diversity, it is technically a measure of industrial specialisation rather than diversity so that lower values indicate a lower chance and thus indicate higher industrial diversity. The Herfindahl index is widely used in economic and industrial research where it is sometimes called the Herfinadalh-Hirschman Index (HHI). The use of different names for the index within economics was commented upon by Hirschman (1964).

The Herfindahl index, H, is calculated as follows:

$$H = \sum_{i}^{k} p_{i}^{2} \tag{5}$$

Where, p_i = proportion of employment that is in the ith industry

   k = number of industries

2.4.8 Occupational group

Average incomes are different for people in different occupational groups. For example, people in professional occupations have incomes approaching twice those of people working in sales and customer service roles (ONS, 2019g). Unemployment rates for economically active people in different occupational groups can also vary. For example, in the first quarter of 2018 the unemployment rate amongst economically active people in professional occupations was 1.4% compared with 5.1% for people who work in sales and customer services roles (ONS, 2018). These differences in incomes and unemployment rates suggest that the composition of a geographic area by occupational group could affect the area's average income and unemployment rate.

Variables based on occupational group information can also be used as measures of social class. As noted above, Tzavidis et al (2018) found social class to be a good predictor to use when modelling income levels in Mexico. This international evidence of a link between income and social class/occupation supports the use of occupational group information in models in this project. In their analyses of retirement ages using data from the 1946 British

Birth Cohort Murrary et al (2019) included categorical mid-life and young adult occupational class information (professional, managerial/technical, skilled manual, partly skilled and unskilled manual, missing data). As noted earlier there can be a relationship between age of retirement and area-level employment rates (and higher area-level unemployment rates) making their choice of covariables relevant to models of employment and unemployment rates. The links between occupational group and employment outcomes (income and chances of unemployment) mean that it would be appropriate for this project to include occupational group information when modelling employment outcomes. That the ONS publish information on income and unemployment by occupational group implies that the ONS and other researchers make use of occupational group when analysing labour market statistics. This is another reason to include occupational group information in models in this project.

2.4.9 Census Area type

ONS local authority level estimates of unemployment and income make use of information on 'types of area' in order to borrow strength from geographic areas for which more survey data are available to help make estimates for similar types of area for which less survey data are available (e.g. ONS, 2003, ONS, 2006, page 8, and ONS 2019e). The fact that the ONS methodology for modelling small area labour market statistics uses Census Area type suggests a link between similar areas and unemployment rates. It also suggests that it would be appropriate to include Census Area type, as given by ONS (2019h), in this project and moreover this would be relevant to other researchers modelling labour market statistics for different geographic areas.

2.4.10 Indices of Multiple Deprivation (IMD)

The English Indices of Multiple deprivation seek to identify geographic areas that contain high levels of deprivation. Along with overall multiple deprivation, separate information is available for various domains including income deprivation and employment deprivation. Deprivation indices rank small geographical areas according to deprivation scores calculated from a rich variety of relevant data. Full details of the 2015 English Indices of Deprivation are available from the Department for Communities and Local Government (2015a). Guidance on the use of the 2015 IMD (Department for Communities and Local Government (2015b) indicates that relative deprivation in different geographic areas

should be compared by comparing their IMD ranks or by comparing their IMD decile rather than by comparing their IMD scores. It notes that an area with an IMD rank twice that of another area should not be regarded as being half as deprived. An example of the use of IMD deciles in analysing life expectancy is given by ONS (2019i). Similarly, IMD deciles were used by the ONS (2019j) when analysing statistics on avoidable deaths. Oxford Consultants for Social Inclusion (2010) writing about the 2010 English IMD reiterated that it was important to use IMD ranks rather than scores when comparing geographic areas because the indices order areas according to levels of deprivation rather than giving absolute measures of deprivation. The 2015 IMD were updated by Oxford Consultants for Social Inclusion (2018) and the latest indices, 2019 IMD, were released in September 2019 after the data collection and data processing phases of this project. The 2019 IMD local authority level summary statistics were calculated and published for the revised set of local authorities in existence from April 2019 onwards rather than the set of local authorities in existence since the 2011 Census that was already in use for this project. This work therefore makes use of the 2015 IMD average ranks for local authorities for overall multiple deprivation and for employment deprivation. Future research involving labour market statistics could both make use of the 2019 IMD on their own and also compare them with the 2015 IMD to create an indicator of change in deprivation to be used in modelling labour market statistics.

2.4.11 Commuting distances

Information on the distances that people travel to work is available from the 2011 Census. For example, data showing the numbers of people travelling less than 2km, 2km to less than 5km, 5km to less than 10km, 10km to less than 20km, 20km to less than 30km, 30km to less than 40km, 40km to less than 60km, 60km and over, working mainly at home, and 'other', is available for Lower Super Output Areas (LSOAs) and higher geographies. At the extreme ends of the distribution the percentage of workers travelling 30km or more varies from less than 1% in some London LSOAs to around 40% in some parts of the St Albans and Sevenoaks local authorities. Different commuting patterns in different parts of England may mean that it would be appropriate to use different geographic scales when modelling labour market outcomes in different parts of the country. The differing patterns in commuting distances are drivers of different sized travel to work areas (TTWA) in different parts of the UK. That workers travelling to full-time or part-time jobs tend to travel different distances to work has been noted above. As commuting distances appear to vary

for different parts of the country it would seem appropriate to include a measure of commuting distances as a covariate in models of labour market outcomes. Future research may use travel to work areas as a geographic level in such models.

2.4.12 Rural/urban classification

One areal variable that can be used for modelling labour market outcomes is the ONS rural/urban classification of areas derived using 2011 Census data. A methodology report (DEFRA, 2013a) and a user guide (DEFRA, 2013b) are available online along with maps showing the classification for Census output areas, lower and medium super output areas and local authorities. The classification was primarily intended to identify rural settlements of less than 10,000 population. It is possible that living in a rural or urban area might affect access to employment and wage levels. There is evidence that other researchers are interested in analysis of labour market statistics by rural/urban classification of areas. For example, the ONS published an article on worker productivity in Great Britain by rural and urban area classification (ONS, 2017d), Somerset Intelligence (2019) produced a rural/urban analysis of unemployment and economic inactivity for local authorities in the area, and the Improvement and Development Agency (2010) produced a guide for local authorities on tackling unemployment in rural areas which noted that the DEFRA rural/urban classification that provided information at local authority district and unitary authority level was the most appropriate for use with other statistics that were available at this geographic scale and referred readers to a DEFRA technical guide (DEFRA, 2005) which provides details of the local authority level rural/urban information based on the 2001 Census which was the latest classification available at the time the Improvement and Development Agency's 2010 guide was published.   A measure derived from the rural/urban classification based on the 2011 Census was used in this project. Phimister et al (2006) demonstrate the use by other researchers of whether people live or work in rural or urban areas when analysing employment statistics. Similarly, Culliney (20016), Cartmel and Furlong (2000) and the Commission for Rural Communities (2012) show other researchers make use of whether people are from rural or urban areas in their research into pay, unemployment and employment respectively, showing the relevance of rural/urban indicators in research into labour market outcomes.

2.4.13 Country of birth and ethnicity

Research on labour market outcomes for individuals by nationality, race, ethnicity and class has been carried out by many including Kingston et al (2015), Freeman (2012) and Yu and Sun (2019). For example, in 2018, 77% of white people were in employment compared with 65% of people from all other ethnic backgrounds (gov.uk, 2020). Country of birth and also ethnicity indicators can be used as independent variables in models of labour market outcomes. This project has used the proportion of people born outside the UK as a dependent variable to give a measure of the diversity of the population at area level. Area-level country of birth variables could also be regarded as dependent variables in some models used to see if the proportion of people moving to different areas from outside the UK is dependent on the job opportunities in different areas.

2.4.14 Housing tenure

There is much research on the relationships between housing tenure and labour market outcomes, e.g. Ma et al (2008), Edward and Fisher (2009), Borg and Branden (2018), Robson (2003), Martin et al (2001) and Phimister et al (2006). The percentages of owner-occupied households and the percentage of households in social rented housing are often used as different predictor variables in models of socioeconomic outcomes, e.g. Tranmer and Steel (2001b).

2.4.15 Job density

Research by Clark and Summers (1982) is reported by Bell and Blanchflower (2011) as showing that teenage unemployment is due to a shortage of jobs. This evidence of a link between unemployment and a lack of jobs supports the idea of including job density in models in this project as a possible predictor of unemployment rates, and arguably also as a predictor of employment rates.

2.4.16 Unemployment

A widely accepted definition of unemployment is provided by the International Labour Organization (ILO). The ILO is a United Nations agency founded in 1919 under the League of Nations as part of the Treaty of Versailles following the first world war; in 1946 it became an agency of the new United Nations (glo.be, 2019 and ILO, 2015). The ILO aims to

bring together the interests of employers, workers and governments (ILO, 2019). The ONS reports the ILO definition of unemployment as being either "*without a job, have been actively seeking work in the past four weeks and are available to start work in the next two weeks*" or "*out of work, have found a job and are waiting to start it in the next two weeks*" (ONS, 2019a). It is this definition of unemployment that is used in the UK Labour Force Survey. The unemployment rate can then be defined as the number of people who are unemployed according to the ILO definition divided by the sum of the number of people who are employed and the number of people who are unemployed. This is consistent with the UN's Sustainable Development Indicator 8.5.2 and Sustainable Development Goal 8.5 concerning full and productive employment, decent work and equal pay (UN, 2019). This provides another reason why an unemployment variable based on this definition may therefore be of particular relevance and importance to other researchers and analysts with an interest in official unemployment statistics. Generally, unemployment is likely to be a dependent variable in models of labour market statistics.

2.4.17 Employment

The ONS (2019a) gives a definition of people in employment as those people aged 16 or over who are paid to work for at least an hour a week together with people who have a job but are temporarily away from work for example on leave or off sick. People in employment includes employees and those who are self-employed. The headline employment rate used in the UK is the number of people aged between 16 and 64 who are in employment divided by the total number of people in this age range (ONS, 2019a). Separate employment rates can also be calculated for different age groups and for men and women. Employment is likely to be used as a dependent variable in models of labour market statistics.

2.4.18 Economic activity and inactivity

The total number of people who are counted as economically active is given by adding together those people who are employed and those who are unemployed according to the ILO definition.

The number of people who are counted as economically inactive is all those aged 16 or over who do not have a job and have not looked for a job in the last four weeks and/or are

not able to start work in the next two weeks. Economically inactive people include students, people looking after a family or home, people who are sick or disabled, retired people and discouraged workers (ONS, 2019a). Discouraged workers are defined in ONS, 2019a, as "*A small subgroup of the economically inactive population who said their main reason for not seeking work was because they believed there were no jobs available.*".

A report on the characteristics of people who have never worked (ONS, 2019b) provides an insight into some of the reasons for people being economically inactive. The 2017/18 total of 3.6 million people over 16 in the UK who had never worked included 2 million full-time students; 180,000 students who were not in full-time education; 250,000 unemployed people; 510,000 looking after family or home (of whom 94% were female); 440,000 who were short-time or long-term sick; 40,000 retired people who had never had any paid work; and 190,000 others including those waiting to hear about a job, discouraged workers, those who say they do not need a job, and those who were not yet looking for a job.

It is likely that levels of economic activity or inactivity would be used as dependent variables in models of labour market and socioeconomic statistics.

3. Methodology Part 1 – Overview and Output Areas and microdata methodology

3.1 Introduction to the two methodology chapters

The central purposes of this work are assessing which geographic scales are the most useful to include in statistical models of labour market statistics and providing guidance to researchers on the effects of using different geographic scales in their models of areal labour market statistics. The approach taken to start to investigate the usefulness of different geographic scales in models of labour market statistics was to model a key labour market statistic, that of economic activity, at the building block geographic scale of 2011 Census Output Area (OA). Section 3.2 below notes why OA level was chosen as the starting point for the investigation. The data and methodology used for these first stage investigations is discussed in sections 3.2.1 - 3.2.4.6 of this chapter, and the results from these investigations are set out in sections 4.1 and 4.1.1 in the next chapter, *Methodology Part 2*.

It was envisaged from the beginning that as much as possible of the data used for the project would comprise statistics aggregated over geographic areas within the UK in order to make the research in this project consistent with that undertaken by the large numbers of researchers who only have access to aggregated areal statistics. That large numbers of researchers use aggregate areal statistics is evidenced by the ONS's statement that "The aggregate data are the most commonly and widely used component of the outputs from UK censuses.", UK Data Service Census Support, 2020a. Thus using aggregated areal data for this project helps to make the recommendations from this research relevant to large numbers of other researchers.

It must be acknowledged however that much of the variability in social economic data occurs at the level of the individual. It was therefore decided to carry out a small exploration using records for individuals to address any potential questions about whether too much information was lost by not using data for individuals. This was carried out using data from the 2011 Census microdata teaching data (ONS, 2018c). The methodology used is described in this chapter, *Methodology Part 1*, and the results obtained are set out in the next chapter, *Methodology Part 2*.

The experience and results obtained from the analyses of OA level data and the exploration of microdata for individuals helped to shape the major modelling part of this project. This involved the modelling of a number of labour market outcome variables measured at local authority level. A range of two-level and four-level multilevel modes were built for each local authority level outcome variable using a set of independent variables measured at different geographic scales. The variables used in these models were chosen to be representative of the variety of statistical topics used by other researchers in their models and analyses of labour market outcomes. Section 4.2.2.1 introduces these topics and links the variables used for this project, which are detailed in table 5 of section 4.2.2.2, to relevant published research described in section 2.4 of the literature review. Section 4.2.3 explains how multilevel modelling of areal statistics for hierarchical administrative takes account of both hierarchical and proximity considerations making it an appropriate methodology for the main modelling section of this project. The results of the multilevel models of local authority level labour market outcomes are set out in chapter 5 of this thesis, *Results for local authority areas*.

3.2 Output Area methodology

The choice of Census Output Area (OA) level to start modelling data was deliberate. OA level provided a ground level starting point for modelling aggregate areal data. One reason for this was that large amounts of data are available for OAs which can be aggregated up to provide data at any higher geographic scale. In addition, OA level models were not only useful in themselves but also provided a good basis for the development of further models as models that could be run successfully at such a fine geographic scale could be adaptable to run successfully using data from higher geographic scales.

3.2.1 Introduction to Output Areas

The 2011 Census is the main source of socioeconomic statistics for people at the lowest possible geographic scale across the whole of the UK. The ONS states that "Census aggregate data provide the most complete source of information about the demographic and socio-economic characteristics of the UK population that is available.", UK Data Service

Census Support, 2020a. This is because, by definition, the Census aims to collect information about all individuals in the UK and this in turn makes it possible to produce information about people at the finest geographic scale possible as, in theory, data for all individuals are available so there is no need to consider whether the sample size is large enough to produce reliable information for small areas. An important constraint that determines the finest geographic scale that can be used for the publication and dissemination of Census data is that of data confidentially for individuals. Census outputs have to meet strict criteria to ensure that information about individuals cannot be inferred from aggregate statistics for small geographic areas. More recent statistics about people are available for many topics, for example Labour Market Statistics from surveys such as the UK Labour Force Survey are available from the ONS and from Nomis, 2019. But they are available only for much larger geographic areas than OAs. This is especially true for statistics for smaller areas for the whole of the UK or the whole of England. A very useful feature of the sets of statistics available from the 2011 Census is that many of the same statistics are available for a whole range of geographic scales making the 2011 Census an ideal source for comparing multilevel statistical models.

The finest geographic scale for which 2011 Census statistics are available for England is that of 2011 Census Output Area (OA), (ONS, 2016aa). OAs were first used for data for England for the 2001 Census. They were generated from the raw data collected during that Census rather than being determined in advance of the Census. Martin et al (2001) and the ONS (2016) explain how OAs were created by using an algorithm to amalgamate unit postcode areas to form small contiguous areas that were both homogeneous with respect to housing tenure and housing type and that also respected physical barriers such as major roads and rivers. The aim was to create a set of the smallest possible areas for which a considerable number of statistics could be published whilst respecting confidentially of the people living in each area. Statistics for all other geographic scales are generated by grouping together data for the OAs that are within in each areal unit at the larger scales. The set of OAs for 2011 was created by making as few adjustments as possible to the 2001 OAs in order to provide a set of OAs that were as consistent as possible with the 2001 OAs whilst maintaining data confidentially and desirable population size when applied to the raw data collected for the 2011 Census. As a result, there are 171,372 OAs for England for the 2011 Census and, on average, each OA has a population of around 300 people.

3.2.2 Output Area variables used for modelling

As OAs form the building blocks for 2011 Census statistics they were a natural choice for the lowest level for the first set of multilevel models for the analysis of labour market statistics at a variety of different geographical scales. Statistics from the 2011 Census for each OA were downloaded from the ONS (ONS UK data service, 2017) for each of the 171,372 OAs in England giving information on:

- the numbers of people aged 16 or over by economic activity status (either economically active, i.e. working or looking for work, or economically inactive);
- the numbers of people reporting their general health to be either very good, good, fair, bad or very bad;
- the numbers of people with different levels of highest qualification.

These topics of economic activity, general health and highest qualification were chosen as representative of the topics used by other researchers in their models of labour market statistics. In particular economic activity is a key labour market outcome statistic that can be used to give an indication of the size of the active work force in an area. Section 2.4.18 of the literature review gives information on economic activity and inactivity and published analyses of the reasons why people are economically inactive. General health is used by other researchers as a predictor variable in models of labour market outcomes as shown in section 2.4.4 of the literature review of the thesis making it a suitable variable for inclusion in the project. Highest qualification is similarly often used by other researchers as a predictor variable in their models of labour market outcomes as shown in section 2.4.2 of the literature review. This makes it a suitable variable to use in this research. Just one outcome variable (economic activity) and two predictor variables (health and highest qualification) were selected for this early stage of modelling in order to keep the models straightforward and transparent. This was particularly important given the very large number of OA records in this stage of the project.

Along with these statistics information was downloaded identifying the administrative areas associated with each OA.   For the majority of the OA records this information was a county or Unitary Authority (UA) code identifying the OA's upper level local administration. A small number of the OA records had a merged Census district code indicating they were

part of the City of London or the Isles of Scilly as appropriate. These areas are merged with the City of Westminster unitary authority or the Cornwall unitary authority for some purposes when their low populations make the provision of separate statistics not possible due to practical or data protection reasons. In total there were 90 separate county/UAs which provided a set of geographic units at a scale in between OA and NUTS 1 areas that could be used in multilevel modelling. Each OA record also had a NUTS 1 areas code showing which of the nine regions of England it was within and providing a high level set of geographic units to use in multilevel modelling of the dataset.

The specific OA-level dependent/outcome variable used in multilevel models was the OA-level percentage of people aged 16 to 64 who were economically active, 2011. The specific OA-level independent/predictor variables used in the multilevel models were: the OA-level percentage of people who described their general health as good or very good, 2011; and the OA-level percentage of people who reported their highest qualification as being at NVQ level four (first degree or equivalent) or higher, 2011. These particular variables were chosen as representative of the outcome and predictor variables commonly used by other researchers in their models of labour market statistics. Further details of these and other variables used by other researchers in their models of labour market statistics are discussed in section 2.4 of the literature review of this thesis.

3.2.3 NUTS 1 areas level variables processed for the analysis

For each OA record three variables for NUTS 1 areas statistics were added to the OA dataset. These were: the NUTS 1 area percentage of people who are economically active, 2011; the NUTS 1 area percentage of people who described their general health as good or very good, 2011; and the NUTS 1 area percentage of people who had reported their highest qualification as being at NVQ level four (first degree/equivalent) or higher, 2011.

3.2.4 Analysis techniques used for Output Area data

3.2.4.1 Separate OA level and NUTS 1 areas level models

The initial, exploratory, models built using the OA level dataset were separate single level regression models at OA level and at NUTS 1 areas level. The OA level model modelled the percentage of people in each OA who were economically active as the dependent variable using the percentage of people in each OA who had reported their health to be either good

or very good as the independent, i.e. predictor, variable. The NUTS 1 areas level model used the percentage of people who were economically active in each NUTS 1 area as its dependent variable and the percentage of people in each NUTS 1 area who had reported their health to be good or very good as the independent variable. The NUTS 1 areas level model only had nine data points whilst the OA level model had over 171,000 data points. At this early stage of modelling data at different geographic scales a simple approach of modelling the percentages of people who were economically active as if they were Normally distributed variables rather than binomially distributed proportions was used. One advantage of this was that it made the models more transparent and thus easier to illustrate and compare.

3.2.4.2 Separate models for each NUTS 1 area

Next a set of nine regression models, one for each NUTS 1 area, were built using the OA level data. These models were created in order to see whether the coefficients were different in each of the nine regression areas. Again, the dependent variables were modelled as if they were Normally distributed rather than binomially distributed. The intercept and coefficient values for each of the models were compared in a table which included the adjusted R^2 value for each model. This table provided evidence that the relationship between the same variables in different parts of a study area, i.e. different NUTS 1 areas/regions of England, could be different. The different R^2 values showed the goodness of fit of the models varied among the different NUTS 1 areas. Figure 3 in section 4.1.1 below shows the nine regression lines plotted in different colours on the same axes to provide an illustration of the different relationships between the OA-level variables in different NUTS areas.

3.2.4.3 Multilevel models

The next step in the analysis of the OA dataset was to build a number of multilevel models of the OA-level percentages of 16 to 74 year olds who were economically active. Initially these were built as linear models which treated the percentages as if they were Normally distributed dependent variables. This was done for simplicity and to correspond to the single level regression models already built using this dataset. These models were subsequently replaced by generalised multilevel linear models of the OA proportions of the people who were economically active. These were built in R using the logit link function.

An example of the R code is included in Annex 1, example 1. It was necessary to use generalised multilevel linear models take account of the fact that the percentages were measures of the proportions of individuals in each OA that had one or the other of the individual-level binary outcomes of economically active or economically inactive. To model area level proportions such as these the proportions should be assumed to have a binomial rather than Normal distribution. For models where the dependent variable is a binary variable the amount of variation at individual level is not available from the model as the model residuals are not simply the difference between the observed and model values for the dependent variable. Instead an estimate must be made. For models that use the logit link it has been shown (Snijders and Bosker, 2012, p. 305, and Wu et al, 2012) that an acceptable accepted estimate to use is $\pi^2/3$. This can be approximated as 3.29.

Consideration of extra-binomial dispersion

Although it should be assumed that areal proportions have a binomial distribution rather than a Normal distribution, it is possible, or even likely, that areal level proportions may exhibit overdispersion when modelled due to a lack of sufficient predictor variables being included in the model (Skrondal and Rabe-Hesketh, 2007). Browne et al (2005) note that there are two methods for modelling overdispersion in binomial dependent variables created by the aggregation of binary individual-level variables. Often extra-binomial dispersion, which can be overdispersion or under-dispersion, is investigated in a model using the multiplicative approach. In this approach the variance in the assumed binomial variable is calculated without restricting it to that predicted by the sample proportion. A scale factor is then generated by comparing the calculated variance with the predicted variance. A scale parameter of less than one shows under-dispersion and a scale parameter larger than one shows overdispersion. The larger the absolute size of the scale parameter the greater the overdispersion. A limitation of the multiplicative approach to modelling extra-binomial dispersion is that it only evidences the existence, direction and extent of extra-binomial dispersion it does not attempt to give any information on the causes of the extra-binomial dispersion.

Browne et al (2005) proceeded to use the alternative, additive approach to modelling overdispersion rather than the multiplicative method. The additive method for multilevel modelling of binomial variables that display overdispersion does not use a scale factor to measure the size of the overdispersion. Instead it works by adding a pseudo level to the

model which uses the same group identifiers as the first level in the model. The pseudo level forms the second level in the model. This means that the second level units in the model all contain only one lower level unit. In such a model the first level units can be assumed to have a variance comprised of the expected variance for a binomial distribution, i.e. $\pi^2/3$ where the logit link function is used and 1 where the probit link function is used (Snijders and Bosker, 2012, p. 305, and Wu et al, 2012), and the extra variance due to the overdispersion which is shown in the model output as the variance of the second level in the model. The extra variance indicates that there are influences on the binomial variable, at the level of the binomial variable, that are not included in the model. Adding appropriate predictor variables to the model should reduce the extra variance and improve the model. A limitation of the additive approach is that it cannot be used to model binomial under-dispersion. This is unlikely to be a problem in this project as the clustered nature of the data mean it is very likely that there may be factors missing from the model that cause apparent overdispersion, which is in fact more likely to be just an indication that the model is either under specified or poorly specified than of genuine overdispersion. It is not very likely that there will be any under-dispersion. Skrondal and Rabe-Hesketh (2007) approached observed overdispersion in binomial variables by adding random effects to explain the apparent overdispersion. They stress the importance of adding appropriate variables to better specify the model rather than simply introducing an overdispersion parameter to quantify observed overdispersion that is more likely to be due to an underspecified model rather than true overdispersion.

Using the additive approach to overdispersion of adding a pseudo level can lead to very complex models and computational problems if the number of units in the lowest level is very large. This is because the pseudo level adds an extra parameter to be estimated for each of the units in the lowest level. In the case of the Output Area level models the number of units and hence the number of extra parameters to be estimated would be extremely large. Whilst the additive approach to overdispersion was used successfully in the local authority level models described in section 4.2 and reported on in chapter 5, it was not practical to use the technique for the OA level models described here in chapter 3 and reported on in section 4.1 of the next chapter as experimental models including a pseudo level were found to be too complex to calculate successfully when the additive approach to over dispersion was trialled. For that reason, the OA level models reported on in section 4.1, at this first stage of the modelling for this research, all assume the outcome

variable to be binomially distributed without taking account of possible overdispersion. The assumption was made as it was the only practical way to model the OA level data.

3.2.4.4 Null models using OA data

The first multilevel models built using the OA level data were null models, i.e. than contained no predictor variables other than codes identifying the associated higher level geographic areas for each OA. First a two-level null model of the proportion of people who were economically active in each OA was built with NUTS 1 areas as the higher level in the model. This was followed by a two-level model of OA proportions with the county/unitary authority area as the higher level. Finally, a three-level null model was built for the OA proportions with county/unitary authority as the middle level and NUTS 1 areas as the highest level in the model.

For each of these models the proportion of variance that was due to each level was calculated and the Akaike Information Criterion (AIC) value produced by the model was noted. The AIC is a measure of relative model fit. For different models of the same dataset the model with the lower AIC fits the data better. If the difference between the AIC values is less than two then the model with the higher AIC value fits the data almost as well as the model with the lower AIC value (Fabozzi et al, 2014). Therefore, if the difference between the AIC values is two or more the model with the lower AIC value can be taken to have a better fit than the model with the higher AIC value. More information on the calculation and use of AIC values is given on page 58 of Robson and Pevalin (2016) and by Fabozzi et al (2014).

The AIC values are all given in the results chapter. Annex 1, example 2, contains an extract of the R code used to build the three null models and extract the ICC values. The ICC percentages and the AIC values from the three null models are shown and discussed in chapter 4 of this thesis.

3.2.4.5 Random intercept models using OA data

Separate three-level random intercept models of the OA-level proportions of people who were economically active were produced using the proportions of people in good or very good health and the percentage of people with an NVQ level 4 or higher qualification as

independent, predictor variables, one in each model. The R code used to build these models is shown in Annex 1, example 3, and the resulting ICC percentages and AIC values are shown in the chapter 4 of this thesis. The difference between these codes and that for the three-level null model shown in Annex 1, example 2, is simply the addition of either the OA level proportion of people in good/very good health variable or the OA level proportion of people with an NVQ level four or higher qualification variable as appropriate to the fixed effects part of the model. The random effects part of the model is unchanged.

3.2.4.6 Random coefficient models using OA data

R code to produce random coefficient three-level models of the OA level proportions of people who were economically active is at Annex 1, example 4. The only difference to the code for the random coefficient models shown above is the inclusion of the independent variables in the random effects parts of the model (shown in bold) as well as in the fixed effects part of the model. The AIC values for these models are reported in chapter 4 of this thesis. As the models are random coefficient models, as described more fully in the methodology for the microdata models shown below (see section 3.3.3.3), it is not possible for unique ICC values to be produced. Examples of the R code used to build these models is given in Annex 1, example 4. Due to the large number of OAs each model took a considerable time to runFor example a random coefficient model of OA-level economic activity rates with grouping by county/unitary authorities and NUTS 1 areas with the health predictor variable took 25 minutes to run and a similar model with the NVQ level four or higher predictor variable took 40 minutes to run. Although these times are not excessive in themselves for individual models, they show that it would be time consuming to run different OA-level models that each included different geographic levels for each of a large number of predictor variables in turn. However, the OA-level models that were run were useful in themselves and facilitated the development of the essential R code to run multilevel models for data at other geographic scales. The code to run the models was developed incrementally so it is difficult to quantify how long the code for each model took to write. The initial code to run OA-level models was developed over about a month towards the beginning of the research. It was subsequentially used as the basis of the code to run local authority level models although these were developed further particularly in regard to the use of a pseudo level to account for binomial overdispersion and to experiment with models that had random intercepts at all levels of the model whilst allowing for random coefficients at just a subset of these levels.

Subsequent multilevel models for areal data were all built using local authority level rather than OA level data. The local authority level models were then used to assess which higher geographic scale levels should be included in models of labour market and related socioeconomic statistics as described in chapter 5 of this thesis.

3.3 Microdata methodology

Microdata is the term used to describe data consisting of records for individual people from censuses and official surveys. Ensuring the privacy of data from individuals drives the way in which such data are made available. The terms and conditions under which microdata are made available depend on the degree of anonymity in the data and the sensitivity of the information. For example, data on incomes requires the strictest access conditions. Generally, the more detailed the geographic information held in a dataset the more secure the access arrangements need to be.

3.3.1 Exploration of microdata variables

In keeping with the ethos of using information that was widely available to other researchers it was envisaged that as much as possible of the data used for the project would consist of statistics aggregated to geographic areas within the UK. This was in order to help make the research consistent with that undertaken by large numbers of researchers who only have access to aggregated areal statistics and thereby help to make any recommendations relevant to such researchers. However, as much variability in social economic data is to be found at the level of the individual, in order to address any questions about whether too much information could be lost by not using data for individuals, a small exploration was carried out using records for individuals. An appropriate data source for this was 2011 Census microdata teaching dataset which is the only microdata resource publicly available from the 2011 Census (ONS, 2018c) without restrictions. It was created to give students experience in using microdata and to provide a publicly available set of microdata to allow data exploration and the development of methodology. It provides a 1% sample of data from the 2011 Census. The largest of the individual teaching datasets for the UK is that for England and Wales (ONS, 2018c) and this is what was used in this section of the project. Similar datasets are available separately for

Scotland (National Records of Scotland, 2018) and Northern Ireland (Northern Ireland Statistics and Research Agency, 2018).

A zipped file containing the 2011 Census teaching dataset for England and Wales as a csv file, a user guide, details of variables and a data quality note for the religion variable were downloaded from the ONS website (ONS, 2018c). It was noted that the citation required for the dataset was 'Source: Office for National Statistics licensed under the Open Government Licence v.1.0.' R scripts were created to load and describe the data.

The dataset consisted of 569,741 records and 18 variables which are listed in table 1 below.

Table 1: Variables included in microdata teaching set

Variable	Description
Geographic identifier	9-digit region codes – (nine starting with 'E' to identify each of the nine English regions and one code, starting with 'W', for all records for Wales).
Type of residence	Whether the residence was a communal establishment or not.
Family composition	Type of family lived in.
Population base code	Code to identify whether individuals were usual residents, students living away during term time, or short-term residents.
Age	
Sex	
Marital status	
Schoolchildren/full-time students	A binary code to indicate records for school children and full-time students.
Country of birth code	Values used are: UK; Non-UK; 'No code required as record is for a student or schoolchild living away during term time'.

Variable	Description
General health	No code required for students or
Ethnic group	schoolchildren living away from home
Religion	during term time
Economic activity	
Occupation	Full information was not provided for
Industry	these variable for some groups of people
Hours worked per week	such as those aged under 16.
Approximate socioeconomic grade	

To start to investigate and to illustrate the variables contained in the microdata a collection of charts was produced to give base line information. These showed usual residents by economic activity, usual residents by occupation, usual residents by industry and usual residents by hours worked per week. They helped to inform the choice of variables for this work not just for the microdata modelling but in particular for the more extensive modelling of local authority level data. Figure 1 below shows usual residents by economic activity.
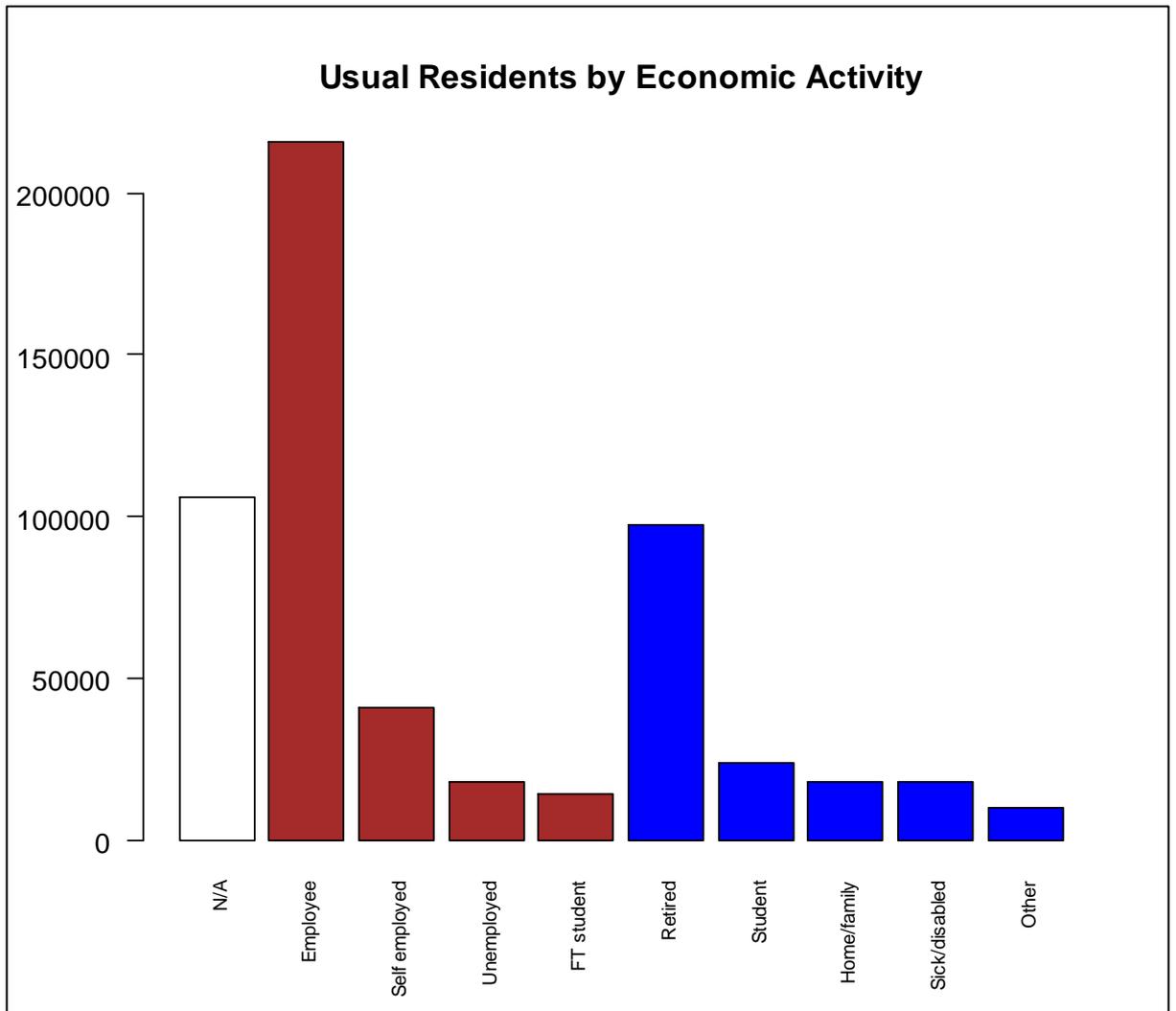
Figure 1 above shows usual residents from the 1% microdata sample by economic activity. The bar on the left labelled N/A was coloured white to blend into the background as it represents those aged under 16 or students or schoolchildren living away during term-time who are not usually included in labour market statistics. The bars shown in brown show categories that are classed as economically active (employee, self-employed, unemployed, and full-time students). The bars shown in blue show categories classed as economically inactive (retired, student, looking after home or family, long-term sick or disabled, and other). The chart shows employees to be the largest economically active group and retired people to be the largest economically inactive group. The colour coding was intended to make it clear that being unemployed or a full-time student are counted as being economically active. The chart illustrated the fact that unemployment accounts for only a

small proportion of those people who are not employed or self-employed. This gave evidence to suggest that modelling unemployment rates alone would not capture all the relationships between the numbers of people working and those not working. This showed that it would be beneficial for the project to model employment rates and economic activity rates as well as unemployment rates.

Details of the occupational and industrial sector codes used in the microdata records are shown below in tables 2 and 3.

Table 2: Definitions of microdata occupation codes

Code	Label	Definition
-9	NA	No code required (people under 16, people who have never worked, and students or schoolchildren living away during term-time)
1	Manager	Managers, Directors and Senior Officials
2	Professional	Professional Occupations
3	Technical	Associate Professional and Technical Occupations
4	Admin	Administrative and Secretarial Occupations
5	Skilled	Skilled Trades Occupations
6	Service	Caring, Leisure and Other Service Occupations
7	Sales	Sales and Customer Service Occupations
8	Process	Process, Plant and Machine Operatives
9	Elementary	Elementary Occupations

Table 3: Definitions of microdata industrial sector codes (by sector size)

Code	Definition
-9	No code required (people under 16, people who have never worked, and students or schoolchildren living away during term-time)
4	Wholesale and retail trade; Repair of motor vehicles and motorcycles
2	Mining and quarrying; Manufacturing; Electricity, gas, steam and air conditioning system; Water supply
8	Real estate activities; Professional, scientific and technical activities; Administrative and support services
11	Human health and social work activities
10	Education
6	Transport and storage; Information and communication
3	Construction
5	Accommodation and food storage activities
9	Public administration and defence; compulsory social security
12	Other community, social and personal service activities; Private households employing domestic staff; Extra-territorial organisation and bodies
7	Finance and insurance activities; Intermediation
1	Agriculture, forestry and fishing

3.3.2 Microdata used for modelling

As the only geographic indicator provided in the microdata teaching set was that of NUTS 1 there was limited scope for building multilevel models from the microdata. Only two-level models could be built with individuals forming level one and NUTS 1 areas forming level two. The models all used the same outcome variable. This outcome variable, at individual level, was whether each of the usual residents aged 16 to 74 was unemployed or not. The chance of any usual resident in this age group being unemployed is not the same as the unemployment rate given by the standard definition of headline unemployment which was presented earlier in section 2.4.1.6. It is instead a basic measure of the prevalence of unemployment among those in this age range that was used to give a straightforward labour market outcome suitable for modelling at this early stage of the modelling for this project. The only predictor variable used in the microdata models was a measure of whether each usual resident aged 16 to 74 had reported their health as being very good, good or fair (as opposed to being bad or very bad) created from the general health variable to provide a simple, binary predictor variable for health status. To avoid possible double counting and for compatibility with published Census results, the analysis of the data was restricted to records for usual residents. These accounted for 98.5% of the records

3.3.3 Modelling techniques used with microdata

Two-level generalised linear models of unemployment were built with individuals at level one and NUTS 1 areas at level two. These used only the records for usual residents aged 16 to 74 to give results that would broadly reflect those of working age. The modelling compared those with an economic status of unemployed with all those in this age group. The models were built using R with the logit link function.

3.3.3.1 Null model

The first model built was a null model of unemployment with no predictor variables other than the regional code to group the individual records at region (NUTS 1 areas) level. The R commands to build the null, random intercept and random coefficient models and summary output from the model are included at Annex 1, example 5, together with an automated ICC command and its results. As the automated ICC command only gave a value to three decimal places 'manual' calculations of ICC were also carried out to give

more precise values and to demonstrate how the ICC can be calculated for models of binary variables.

Whilst the R code is shown in Annex 1, some pertinent definitions and details are set out here.

Unemployed – a binary variable which was created from the economic activity status variable, coded as 1 if economic status = unemployed, and 0 for any other economic activity status

(1|Region) – the variable **Region** contained the group identifiers in the individual level records that were used for the level two groups; the **1** indicates that the intercepts are random between groups, i.e. different groups can have different intercepts

family= binomial("logit") – the models are of a binary variable and use the logit link function

Region (intercept) variance – variance that can be accounted for by the different intercept values for different Regions

ICC – Intraclass Correlation Coefficient – proportion of overall variance that can be accounted for by grouping in the model

AIC – Akaike Information Coefficient, a measure of relative model fit. For different models of the same dataset the model with the lower AIC fits the data better. More information on the use and choice of the AIC is given in section 4.2.6 below, and also on page 58 of Robson and Pevalin (2016) and by Fabozzi et al (2014).

Manual calculation of ICC for the null model

The Intraclass Correlation Coefficient (ICC) is a measure of the proportion of the overall variation in a multilevel model that can be attributed to the grouping structure of the model. The amount of variance at each of the grouping levels in a multilevel model can always be extracted from the model. For example, the amount of variance due to groups is 0.0257 in the model described above. However, as noted in section 3.2.4.3 above when discussing the generalised linear multilevel models of the Output Area level data, it is not possible to obtain the variance for the lowest level in these models from the models

themselves. Instead the estimated value of $\pi^2/3$ (which can be approximated as 3.29) was used, as discussed by (Snijders and Bosker, 2012, p. 305, and Wu et al, 2012).

The calculation of the ICC value then becomes:-

ICC = group level variation / (group level variation + 3.29)

In the null model above this gives, ICC = 0.0257 / (0.0257 + 3.29) which is equal to 0.008 as a proportion to three decimal places, consistent with the automated ICC function value shown in Annex 1, or 0.78% when expressed as a percentage to two decimal places as shown in chapter 4 of this thesis.

3.3.3.2 Random intercept model

In order to see the effect of adding a predictor variable at level one to the model, the next model of unemployment using the microdata teaching set was a random intercept model with an individual level predictor variable to indicate whether each person had reported their general health to be very good, good or fair, all of which statuses were coded as 1 in a new 'Healthy' variable, or to be poor or very poor, both of which statuses were coded as 0 in the 'Healthy' variable. The R code used to run the model and output results from the random intercept model are included in Annex 5, example 5. The difference between the code for the random intercept model and the null model is simply the addition of the independent predictor variable 'Healthy', highlighted in bold. As this variable is not in the (…..|Region) part of the code, which defines the Random effects part of the model, its coefficients do not take random (i.e. different) values for different groups of individuals. The intercept values of the Healthy variable do however take random (different) values for different groups of individuals. Information for the Healthy variable is shown in the Fixed effects part of the model output shown in Annex 1, example 5. Using the automated ICC command gave a value of 0.008 the same as it did for the null model. However, by calculating the ICC value manually as shown below a more precise value equivalent to 0.81% was obtained which is higher than the ICC value for the null model. This showed that a slightly higher proportion of the variance was at group level in the random intercept model meaning that the addition of the Healthy variable to the Fixed part of the two-level model had reduced the variability at individual level by adding a variable that helped to model or explain some of the variation. This means that the Healthy variable should be

included in the model as its inclusion helped to explain some of the variability and produce a better fitting model.

Manual calculation of ICC for the random model

Using the formula *ICC = group level variation / (group level variation + 3.29)* for the random intercept model gives, ICC = 0.02691 / (0.02691 + 3.29) which is equal to 0.008 as a proportion to three decimal places, again consistent with the automated ICC function, 0.81% when expressed as a percentage to two decimal places.

3.3.3.3 Random coefficient model

The third and final model run using the microdata teaching set was a random coefficient model. This was built to explore the difference made to the model by allowing the coefficients for the predictor variable used in the random intercept model to vary for the different level two units, i.e. regions. The code to run this model and the resulting model output are included in Annex 1, example 5. The difference between this model and the random intercept model is that the Healthy independent predictor variable is in the Fixed effects and the Random effects part of the model as highlighted in bold in the top line of the code. Information for this variable consequently now appears in both the Fixed effects and Random effects parts of the model output.

The random coefficient model did not appear to be appreciably better than the random intercept model at modelling the unemployment variable. This meant that there was no advantage in terms of model fit in allowing the coefficients to vary between groups so the random coefficient model should not be used unless there were specific reasons for using it such as to see how the strength of the relationship between the outcome and predictor variable differed across the whole study area. As the random coefficient model requires a greater number of parameters to be fitted there would be an increased risk of type 2 errors if the random coefficient mode was used rather than the random intercept model. A warning message arose from the model to say that it had a singular fit and it can be seen that there was in fact no variation in the regional coefficients (variance for the region grouping is shown to be 0.00000 in the Random effects part of the output). In the documentation for the R lme4 package used to run the model, Bates et al (2020) explain that the singular fit warning message means that the variance of one or more of the random effects in the model is either zero or close to zero. This suggests that the model is overfitted and as a result may have poor power (a relatively high chance of a Type 2 error)

and/or a higher chance of computational problems meaning that the model is at risk of non-convergence (Bates et al, 2020). Possible solutions to the issue of singular fit models include creating less complex models (Bates, et al, 2020). In this project the AIC value for such models is reported in the results along with the warning of a singular fit or non-converging model. These AIC values are shown for comparison with the AIC values of other similar models to compare the fit of the models. However, this project suggests that models that produce a singular fit or that do not converge without a warning message are best avoided on the grounds that they may be overfitted. It is recommended that instead similar but simpler models are selected from those reported in the same tables in the results as the singular fit or non-converging model. This advice, which is particularly relevant when looking at the results for local authority models shown in chapter 5 of this thesis, is based on the historic convention that overfitted models should be avoided. There is now some new research in machine learning which suggests that for modelling based on neural networks overfitting of models is not a problem and can sometimes be an advantage (Belkin et al, 2019).

The automated ICC command produces a warning message rather than a value. This is always the case for random coefficient models. This is because for random coefficient models the ICC is not simply the proportion of variance explained at group level as the varying slopes mean that a unique ICC value does not exist (page 63, Kreft and de Leeuw, 1998). Therefore, a manual ICC calculation was not carried out for the random coefficient model.

3.3.3.4 Comparison of AIC values for the three microdata models

One way of comparing the model fit of different models of the same data is to compare their AIC values. The lower the AIC value the better the relative fit of the model. For the null model of unemployment using the microdata teaching set the AIC value was 147,707. For the random intercept model the AIC value was lower at 147,451. For the random coefficient model the AIC value was 147,448, only marginally lower than the AIC value for the random intercept model (only just over the value of two regarded as showing a difference in fit between two models, Fabozzi et al, 2014). That the AIC value for the random intercept model was considerably lower than the AIC value for the null model can be taken to mean that the random intercept model of unemployment with 'Healthy' as a predictor variable is appreciably better at fitting the data than the null model. That the AIC

value for the random coefficient model was only marginally lower than that for the random intercept model indicates that it was not much better, if at all, than the random intercept model at fitting the data. The combination of the warning about a singular fit and the only marginally better AIC value can be taken together to mean that the random coefficient model is no better than the random intercept model. More information about singular fitting models is given in section 3.3.3.3 of this thesis. Considering the evidence from the singular fit warning together with the only marginally better AIC value and taking them together to mean the random coefficient model was not an improvement on the random intercept model led to the 'tool' of considering both warning messages and differences in AIC values together as a way of choosing which models to recommend for modelling data for all the model building exercises described in the rest of this thesis. The combination of a warning message plus a similar or worse AIC value became a criterion to use when choosing which of a set of similar models to recommend. When the different models included data for different geographic scales then this criterion in effect helped to choose which geographic scales to include in the model of the dependent labour market statistic of interest. This criterion is often used in chapter 5 of this thesis to compare results from the local authority level modelling part of this project where there are often a number of different models of the same dependent variable some of which include different geographic scales.

3.3.4 Evidence from the literature about missing the lowest level on future methodology

As set out in section 2.3.8 of the literature review there is research to show that if the lowest level is missed out of a multilevel model then the effects from the missing level are transferred to the lowest level that is included in the mode but not to any higher levels, e.g. Tranmer and Steel (2001a, 2001b). Thus if the individual level is missed out of a model, for example because individual level data are not available, then any individual effect will be transferred to the lowest level that is included in the model, such as local authority level, and that the higher geographic scales, e.g. NUTS 1 areas, NUTS 2 areas or NUTS 3 areas, that this project is focussed on, would not be affected. This suggested that not having individual level data in models would not cause a loss of information about which of the higher geographic scales to include in models. This helped to answer the question raised in

3.1 above about whether too much information would be lost by not using data for individuals by suggesting that not using data for individuals would not cause too much relevant information about higher geographic levels to be lost.

4. Methodology Part 2 – Results from Output Area and microdata investigations and methodology for local authority areas

4.1 Results from Output Area analyses and Microdata investigations

4.1.1 Results from Census Output Areas analyses

The first analysis of areal labour market statistics used data for the 171,372 Census Output Areas in England in the 2011 Census. Further details of the methodology used for modelling OA level data are given in sections 3.2.1 to 3.2.4.6. Initially, exploratory single level linear regression models were built separately for OA level data and for NUTS 1 areas level data. In each model the percentage of people who were economically active was modelled using the percentage of people whose health was either good or very good as the predictor variable. For simplicity at this stage of the work, the models used linear modelling treating the outcome variables as if they were Normally distributed variables rather than generalised linear modelling that should be used for areal outcomes (on the grounds that they represent a proportion of individuals in an area and therefore have a binomial distribution).

Figure 2: Scatter plots and regression lines at OA level shown in blue and NUTS 1 areas level shown in red, England 2011
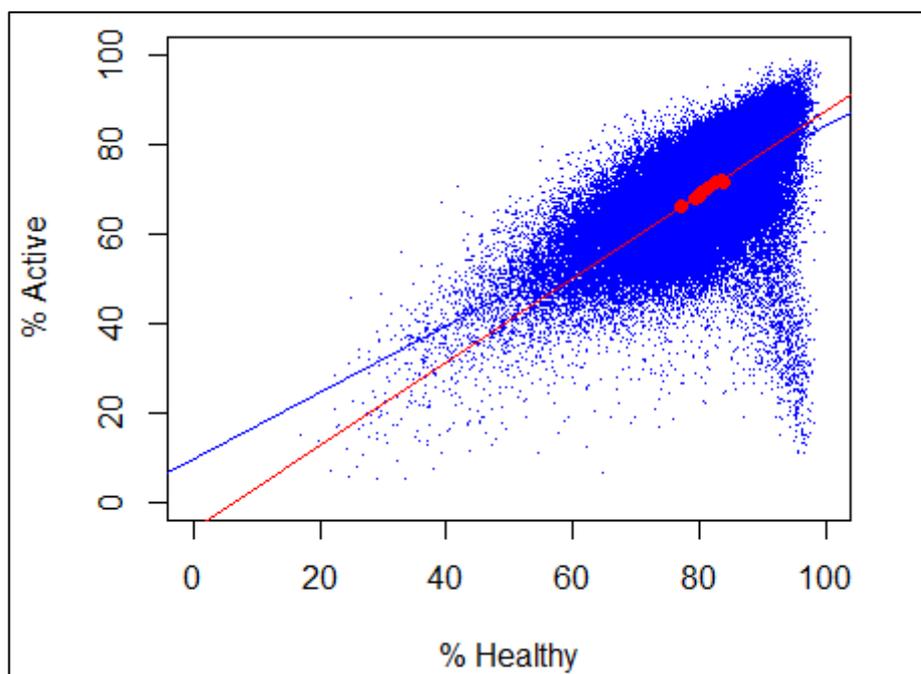
Figure 2 shows the two regression lines drawn on scatter plots of OA and NUTS 1 area data. The red regression line which indicates the relationship between the NUTS 1 areas percentage of people who were economically active and the percentage of people who reported that their health was either good or very good shows a steeper slope than the blue regression line representing the same relationship at OA level. This chart provides the first illustration in this work of the fact that the relationships between the same variables can be different at different geographic scales. It is the first evidence in this work that researchers need to consider which geographic scales to use when modelling labour market and related statistics.

Next separate OA level linear regression models were built for each NUTS 1 area in England. The slopes and intercepts for each NUTS 1 area are shown in table 4 below along with adjusted R2 values which indicate the proportion of variance that can be explained by each model. The corresponding regression lines are shown in figure 3 below.

Table 4: Separate regression models for each NUTS 1 area

NUTS 1 Area Code	NUTS 1 Area Name	Intercept	Coefficient for % of people who were 'healthy'	Adjusted R^2
E12000001	North East	15.67	0.66	0.40
E12000002	North West	8.17	0.76	0.44
E12000003	Yorkshire & the Humber	11.33	0.72	0.38
E12000004	East Midlands	11.78	0.72	0.37
E12000005	West Midlands	11.96	0.71	0.35
E12000006	East of England	8.68	0.76	0.42
E12000007	London	3.82	0.81	0.37
E12000008	South East	12.30	0.72	0.34
E12000009	South West	10.71	0.73	0.36

London stands out in this table as having a much lower intercept and a higher coefficient. The North East has the highest intercept and the lowest coefficient. The East Midlands and

West Midlands have similar values to each other. The intercepts and coefficients for the separate regional models demonstrate that the relationship between the percentage of people who are economically active and the percentage of the population who are in good health is different in different NUTS 1 areas across England. This showed that multilevel modelling using NUTS 1 areas as one of the levels may provide more useful information about the relationships in the data than single level modelling alone can provide.

Figure 3: Regression lines from separate models for each NUTS 1 area in England, 2011
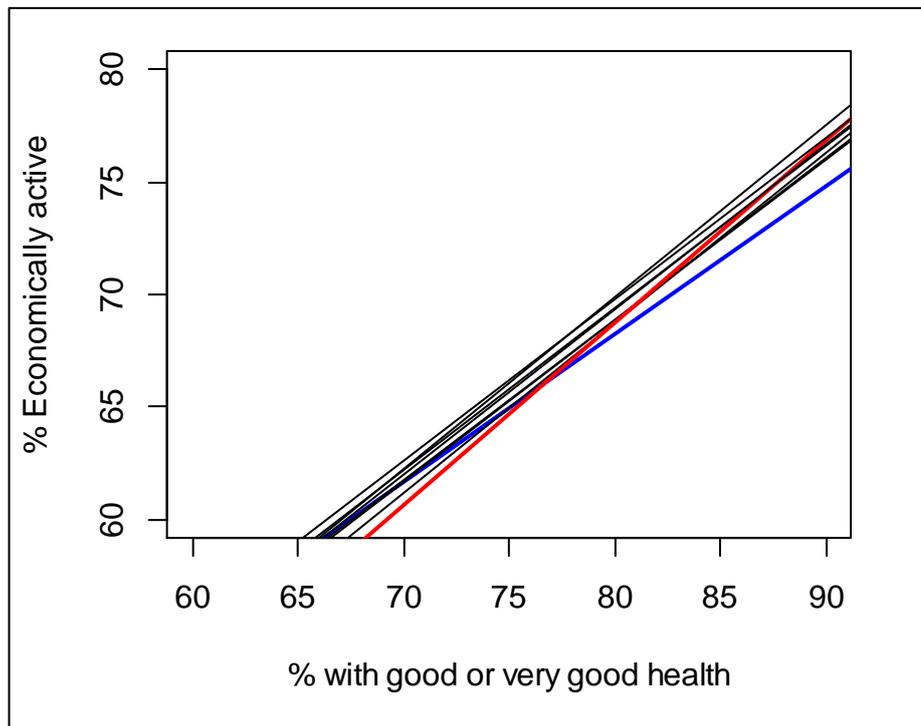
The red regression line for London has the steepest slope. The blue regression line for the North East has the shallowest slope.

In order to investigate the effect of adding a middle level between NUTS 1 areas and OAs to models of economic activity rates, each OA was mapped to its county for two-tier administrative areas, or to its unitary authority for single-tier administrative areas. This allowed a level equivalent to upper tier local authorities to be used in multilevel models. Initially linear multilevel models were created of the OA level proportions of 16-74 year olds who economically active. These were later superseded by generalised multilevel linear models of the proportions who were economically active. This was in order to take account of the fact that proportions representing individuals within areas have binomial distributions and should therefore be modelled using generalised linear modelling. The

results below are for generalised multilevel linear models using the logit link function. The models show the proportion of variance at each geographic level which is known as the Variance Participation Coefficient (VPC), Goldstein et al, 2002, and Browne et al, 2005. The VPC is calculated by dividing the amount of variance at each geographic level by the total amount of variance. VPC values are therefore the same as the Intraclass Correlation Coefficient (ICC) values as although the ICC measures the correlation between two observations in different groups (see sections 2.3.2 and 2.3.3) it can be calculated by dividing the variation in the data at group level by the total variation in the model.

Akaike Information Criterion (AIC) values for the models are also shown for reference. The AIC value for a model is a measure of relative model fit. The fit of two models of the same dataset can be compared by comparing the AIC values. As the AIC is a measure of 'lack of fit' then the model with the lower AIC value has the better fit of the two models. A difference of two or more between AIC values can be taken as a true indication that the model with the lower AIC value fits the data better. More information on the calculation and use of AIC values is given on page 58 of Robson and Pevalin (2016) and by Fabozzi et al (2014).

Null two-level model, OA and NUTS 1 areas (9 groups)

NUTS 1 areas

Variance Partition Coefficient (VPC) = 0.2%

AIC = 2,632,062

Null two-level model, OA and County/UA (90 groups)

County/UA VPC = 0.7%

AIC = 2,553,116

Null three-level model, OA, County/UA and NUTS 1 areas

VPC:

County/UA level   = 0.5%

NUTS 1 areas level = 0.2%

AIC = 2,553,102

The proportion of variation at NUTS 1 areas level is very small in both the two-level null models that include NUTS 1 area levels and in the three-level null model. The proportion of variance in the two-level model of 0.7% is reduced to 0.5% when the NUTS 1 areas level is added to the null model. This provides an example of how missing a level from the top of a multilevel model (in this case missing the NUTS 1 areas level from a model) causes the variance at that level to be transferred down to the next level that is included in the model as is discussed in the literature about missing levels from multilevel models details of which are given in section 2.3.8 of the literature review, e.g. Tranmer and Steel (2001a, 2001b), Opdenakker and Van Damme (2000), Van den Noortgate et al (2005) and Moerbeek (2004).

Random intercept three-level model, OA, County/UA and NUTS 1 areas, with the OA level proportion of people with good or very good health as the predictor variable

VPC:

County/UA level   = 0.3%

NUTS 1 areas level = 0.0%

AIC = 2,203,750

Random intercept three-level model, OA, County/UA and NUTS 1 areas, with the OA level proportion of people with an NVQ level 4 or higher qualification as the predictor variable

VPC:

County/UA level   = 0.4%

NUTS 1 areas level = 0.1%

AIC = 2,339,782

Including the general health indicator predictor variable or the highest qualification predictor variable in the three-level model reduces the proportion of variance at County/UA level and at NUTS 1 areas levels and reduces the AIC value of the model both of which provide evidence that these models fit the data better than the null model. AIC values give a measure of the lack of fit of models. When comparing models of the same outcome variable the model with the lowest AIC value can be assumed to fit the data better than the others.

Random coefficient three-level model, OA, County/UA and NUTS 1 areas, with the OA level proportion of people with good or very good health as the predictor variable

Note that this model produced a singular fit (see sections 3.3.3.3 and 5.10 for information on singular fitting models). A VPC value was not calculated as variances at each level are not unique for random coefficient models (Kreft and de Leeuw, 1998, page 63).

AIC = 2,164,929 which is much lower than the corresponding random intercept model IAC value of 2,203,750 shown above. This shows that the random coefficient model fits the data better than the corresponding random intercept model.

Random coefficient three-level model, OA, County/UA and NUTS 1 areas, with the OA level proportion of people with an NVQ level 4 or higher qualification as the predictor variable

A VPC value was not calculated as variances at each level are not unique for random coefficient models (Kreft and de Leeuw, 1998, page 63).

AIC = 2,294,547 which is much lower that the AIC value of the corresponding random intercept model. Given that a difference of two or more is considered sufficient to show that one model fits the data better than another, then this AIC value shows that the random coefficient model fits the data better than the corresponding random intercept model.

4.1.2 Results from analyses of microdata

The purpose of this project is to assess which geographic scales to include in statistical models and how to make that choice. This relates mainly to the aggregation of data at different geographic scales to create and model statistics for areas rather than to building models using separate records for individuals. However areal statistics, even those for small areas, combine information about lots of different people and can hide much of the variety that exists within each geographic area. In order to see how data for individual people might be used as part of this work, and to help determine whether it was necessary to make extensive use of data for individuals in order to satisfy the aims of the work, some limited modelling was carried out using a set of data for a 1% sample of individuals from the 2011 Census for England and Wales known as the teaching set (ONS, 2018c). A major limitation of this dataset from the point of view of assessing the geographic scales to use in

models was that the only geographic indicator it contained was the NUTS 1 area code for each person.

Two-level models of unemployment were built with individuals at level one and NUTS 1 areas at level two. These were restricted to usual residents aged 16 to 74 and compared the numbers with an economic status of unemployed with all those in this age group. The models were built using R with the logit link function. The proportions of variance at NUTS 1 areas level and at residual individual level are shown below for each model along with the AIC value for each model.

Null two-level model of unemployment

VPC values:

Individual level =   99.2%

NUTS 1 areas level = 0.78%

AIC value = 147,707.1

Random intercept two-level model of unemployment,

predictor variable: of health status = very good, good or fair

VPC values:

Individual level =   99.19%

NUTS 1 areas level = 0.81%

AIC value = 147,450.8

Random coefficient two-level model of unemployment,

predictor variable: health status = very good, good or fair

VPC values: not calculated as unique values not possible for random coefficient models

AIC value = 147,447. This model resulted in a singular fit so should not necessarily be recommended despite having a marginally lower AIC value than the corresponding random intercept model fit (see section 3.3.3.3 above for information on singular fitting models) .

The overwhelming finding from these models is that the proportion of variance at NUTS 1 areas level is tiny compared to the proportion of variance at individual level. This suggests that using NUTS 1 areas geographic scale as the only geographic scale in models of

unemployment status is not particularly helpful. In turn this implies that it is not particularly helpful to use datasets for individuals where the only geography indicator in the dataset is the NUTS 1 areas code. These results therefore suggested that there was little benefit in using individual data for this work unless more detailed geographic indicators were available within the dataset. Using individual level data with more detailed geographic information would mean that there would have needed to be very tight data security restrictions on the access to individual data that might be of use for this work. This was a contributing factor in deciding not to make extensive use of data for individuals for this project. Other reasons included the slow speed of running models with very large numbers of level one units and more importantly the fact that many researchers and analysts are likely only to have access to areal labour market and related statistics rather than to data for individuals. These considerations meant that it would be more helpful to other researchers to demonstrate relevant methods of determining which geographic scales to include in their models of labour market and related statistics if the data used for this project were areal statistics rather than individual data. Subsequent results in this work therefore relate to areal statistics.

4.1.3 Lessons learnt for methodology from the microdata models

The main finding from the three models built using the microdata was that using only the NUTS 1 areas geographic level to group the data for individuals in multilevel models of unemployment explained only a very small proportion of the variance in the model, less than 1%. This suggested that using only NUTS 1 areas as the only higher level in multilevel models was not particularly helpful. Put in the terms of the aims of this project, the NUTS 1 geographic scale on its own was not particularly appropriate or useful to include in statistical models of a selected labour market statistic. This finding together with the fact that the only geographic scale in the microdata teaching set was NUTS 1 areas code showed that there would be little to gain in terms of which geographic scale to include in statistical models by building further models using the microdata teaching set. This influenced the future methodology of the project as it was therefore decided not to build any further models using this dataset. Any microdata sets with more geographic detail would require very tight data security restrictions making access more difficult not only for this project but also for other researchers. There are two classes of microdata which require security restrictions: Safeguarded Data; and Secure Data. Safeguarded Data are available under licence to users who register and agree to data management conditions.

Such users can then download data to use at their own workplace whilst keeping the data secure. Secure Data have to be accessed through the Secure Research Service separately for England and Wales, or Northern Ireland or Scotland. There are two levels of Secure Data. The first comprises 'regional data' that can be downloaded for use at one's place of work but for which the region (NUTS 1 area) is the only geographic information provided for each record. The second comprises 'grouped local authority data' that includes local authority level geographic identifiers for records from local authorities with populations over 120,000 and a grouped local authority identifier for records from local authorities with smaller populations. The grouped local authority data are classed as personal data by the Statistics and Registration Act and have to be assessed only in a secure location rather than by download. Further details about microdata and assess arrangements are available from UK Data Service Census Support, 2020b. In order to help make this research more relevant to other researchers who often only have access to aggregated areal labour market statistics it was decided that all subsequent modelling would be based on aggregated areal data rather than data for individuals.

4.2 Local Authority methodology

Labour market and other socioeconomic statistics for local authority areas, local authority districts or unitary authorities are often used for the publication, analysis and modelling of statistics. For local authority administrations they give information on what is happening in the area they are responsible for, where they need to provide services and implement policies. For individuals they give measures for areas that people can relate to, where they might live, work or study. For researchers they provide statistics that can be compared with other local authority-based statistics. For companies they give information on potential workers and customers. For the providers of official statistics they are often the lowest geographic scale statistics that can be released reliably from surveys or without loss of confidentiality from Censuses.

4.2.1 Local authorities

The base geographic units for the local authority level models in this project were defined to be the 326 local authorities and unitary authorities in England as at 2011, the time of the last Census. For consistency, and to ease the process of mapping, the local authority codes

as contained in the standard 2011 boundary files were adopted as the standard for this project. This meant that where necessary records for later statistics for the following local authorities were recoded to match their 2011 codes: East Hertfordshire, Gateshead, Northumberland, St Albans, Stevenage and Welwyn Hatfield.

In April 2019, approximately half-way through this project, there was a reorganisation of local government administration in some parts of England. This meant that 14 local authorities were merged to form five new authorities. The decision to continue to use data for the pre-April 2019 set of local authorities was based on data availability. Although statistics for the post-April 2019 set of local authorities started to published in the summer of 2019, all local authority statistics up to March 2019, which is the vast majority of statistics available at the time of data collection for the project and includes all 2011 Census data and statistics based on them, e.g. Census area classifications, are for the pre-April 2019 set of local authorities. At a suitable point in the future it will be appropriate for researchers to switch to using the new set of local authorities as more data become available however it was not appropriate to do so during this project which uses data from the 2011 Census and more recent data all published for the pre-2019 boundaries. The methodology used and lessons learnt from this project will be appropriate for future models built using data for the new boundaries. The release of statistics for the 2021 Census might be the point at which it becomes appropriate for the majority of academic researchers using data for a number of years to switch to the set of local authority boundaries that are in use at that time. Information on changes to local authority codes can be obtained from the ONS UK Geographies (2019) webpages and the ONS Open Geography Portal (2019)**.**

The City of London and the Isles of Scilly
The figure of 326 local authorities counts the City of London and the Isles of Scilly as separate local authorities. For a number of official datasets, and for the standard local authority geographic boundary files, they are merged with the borough of Westminster and the unitary authority of Cornwall respectively. Where statistics were available for all 326 local authorities the City of London and the Isles of Scilly were included in models as separate units. Where statistics for them were not available they were either merged or excluded from the models depending on data availability. Even when data or results are

available for the City of London and the Isles of Scilly they are not displayed on maps of England. However, this does not affect the appearance of maps of the whole of England as the geographic areas concerned are relatively small.

4.2.2 Variables processed for local authorities and larger areas

4.2.2.1 Choice of topics for inclusion in local authority level models

In order for the findings of this project to be useful and relevant to other researchers building models of labour market statistics the variables chosen for inclusion in the models of local authority labour market statistics were chosen to be representative of the statistics used by other researchers in their published research on labour market issues. The literature review for this project therefore included looking at published research to see which labour market topics were typically included in other researchers' models of labour market to ensure that the variables modelled in this project were relevant. Section 2.4 of the literature review chapter of this thesis sets out a number of topics that are typically included by researchers in their models of labour market statistics. The topics presented there cover a selection of labour market outcomes and factors that are used as explanatory or predictor variables in models of labour market outcomes. These topics are outlined briefly below. To make this research relevant to other researchers each of the variables used in this research was chosen to represent one of these topics. The research topics and the precise details of the data used to create the variables used in this research are set out in table 5 below to show which variable relates to which topic.

Two key labour market outcome topics are levels of unemployment and of employment. Definitions of these are given in sections 2.4.16 and 2.4.17 of the literature review. These sections also explain the importance of these variables. It is because both are important, and crucially as they describe different concepts, that measures of both unemployment and employment were included as variables in models in this project. Details of the variables used to represent employment and unemployment are shown in items 1 and 2 of table 5 below.

Income and pay levels are also labour market outcome topics that are modelled by other researchers, as described in section 2.4.1 of the literature review. This made it relevant to

include indicators of earnings in this project. Earnings pertaining to a geographic area can relate to the earnings of those workers who live in the area, described as residents' earnings, or the earnings of those workers who work in an area, described as workplace earnings. As these can differ and as both are important it was decided to include measures of both in the variables used in this project. Details of the earnings variables used in the local authority level models are shown in items 4 and 5 of table 5 below.

Another labour market outcome topic related to a variable used in the local authority level models in this project is that of job density. Job density is a measure of the number of jobs in an area. For example, it can be defined as the number of jobs per working age resident. It indicates how rich an area is in job opportunities. Section 2.4.15 in the literature review notes that unemployment can be due to a lack of jobs. This provided evidence that the abundance of jobs in an area is a relevant topic to include in models of labour market statistics. In this project job density was modelled as an outcome variable using a measure described in item 3 of table 5 below as job density could be taken as an outcome dependent on the characteristics of an area and the people living within it. Job density could alternatively be used as a predictor variable that affects the employment or unemployment rate in a geographic area.

The final labour market outcome topic chosen for inclusion in the local authority level models for this project was a measure of the number of hours worked per week. This topic was included as the average number of hours worked per week in an area can be an indicator of the availability of sufficient full-time jobs to meet demand in an area. For example, section 2.4.3 of the literature review reports research that found that part-time workers in different age groups would generally prefer to work either more or fewer hours depending on their age. The details of the variables considered are given in item 6 of table 5 below.

The predictor variables used in the local authority-level models in this project are defined in items 7 to 23 of table 5 below. Each of these predictor variables was chosen to be included in the local authority levels models for this project to represent one of the topics

used by other researchers in their models of labour market statistics as described in section 2.4 of the literature review.

Items 7 and 8 detailed in table 5 below are both area level measures of educational attainment. One measures the proportion of highly educated adults in an area and the other measures the proportion of adults with no qualifications in an area. These were chosen for inclusion as predictor variables in the area level models of labour market outcomes as highest qualification obtained is often used by researchers in models of unemployment as qualifications held by individuals can affect their chances of being in employment and the employment opportunities available to them. Section 2.4.2 of the literature review also notes that there is research published that shows the distance travelled to work tends to be related to people's educational attainment with those with higher qualifications travelling further to work on average than those with lower qualifications. The qualification levels required to enter certain professional level jobs would naturally indicate a link between educational attainment and earnings which would make it appropriate to include measure of highest qualification in models of labour market statistics.

Item 9 in table 5 is a measure of part-time workers as a percentage of all workers. This was included in this research as the proportion of workers in part-time work can be an indication of the availability of full-time jobs. Section 2.4.3 of the literature review notes research by Bell and Blanchflower (2011) into whether those with part-time jobs would prefer to work more hours or fewer hours and research by the ONS (2016c) into commuting distances for full-time and part-time workers.

Item 10 in table 5 describes a variable measuring health, that of the percentage of people reporting their general health to be bad or vary bad. Health is often used by other researchers as a predictor variable of labour market outcomes. Examples of this are given in section 2.4.4 in the literature review. The frequent use of health status in research into labour market outcomes made it an important variable to include in this research.

Item 11 in table 5 below gives details of an area-level measure of population age – median areal age. Section 2.4.1 shows that age is used in published research into unemployment as the rates of unemployment can be different in different age groups and people in different age groups can be affected to a greater or lesser degree by changes in labour market conditions. This makes it relevant to include a measure of age in this research. Using areal statistics means only limited information on the ages of people can be included in models. Median age was chosen as a representative measure.

Gender is included as item 12 in table 5 below. Gender balance is used in this research as an area level measure of a variable that is known from other research to have relationships with labour market outcomes. Section 2.4.6 of the literature review gives examples of published research into the relationships between gender and various labour market outcomes.

Section 2.4.7 in the literature review briefly introduces the idea of industrial diversity and refers to a number of researchers who have used measures of industrial diversity to model unemployment and in particular how the resilience of areal unemployment rates to external labour market shocks may be related to industrial diversity. This shows that industrial diversity is used in models of labour market statistics. A measure of industrial diversity is therefore included in this research. The details of the variable used are shown in item 13 in table 5 below.

Items 14 and 15 in table 5 below are measures of the percentages of workers in different roles. These measures are included as variables in this research as representative of measures of socioeconomic class that are used by other researchers in models of income levels and unemployment for example. A reference is given in section 2.4.1 of the literature review to the use by Tzavidis et al (2018) of socioeconomic class in models of income. A reference to the use of socioeconomic class by Virtanen et al (2013) in models of unemployment is given in section 2.4.4 of the literature review.

As outlined above there has been research by others into the effects of industrial diversity on labour market outcomes and into the effects of socioeconomic class (which is measured by occupation level) on labour market outcomes. Item 16 in table 5 below describes a measure of occupational diversity which was included in this research to combine the concept of diversity indicators such as industrial diversity indicators and socioeconomic class.

Census area classifications are included in this research as listed in item 17 of table 5 below. Section 2.4.9 of the literature review notes that the ONS uses Census area classifications to help to produce local authority level unemployment and income estimates. This shows the relevance of including Census area types in models of labour market outcomes.

Item 18 and 19 of table 5 below show measures of multiple deprivation and employment deprivation respectively that were included in this research. Section 2.4.10 in the literature review gives details of the English Indices of Multiple Deprivation that were the source of these measures. They are included in this research due to the close links between variables used to create the indices and labour market outcomes. Indeed, the indices draw on various labour market outcomes to create both the overall index of multiple deprivation and the index of employment deprivation.

Item 20 in table 5 details the variable used in this research as a representative measure of the proportion of workers with a long journey to work each day. Section 2.4.11 of the literature review comments on differing commuting rates in different parts of England. Section 2.4.2 of the literature review comments on how the distances travelled to work tend to be different for workers with different qualification levels. These sections give examples of some of the ways that other researchers include distance travelled to work in their analysis of labour market statistics which supports the inclusion of a measure of commuting in this research.

Item 21 of table 5 below shows that the percentage of people living in rural areas was included in this research. It was included as other researchers do use information on whether people live in rural or urban areas in their research into labour market outcomes. Some examples of research by others using rural/urban information in their analyses of worker productivity, unemployment and economic activity are given in section 2.4.12 of the literature review.

Item 22 of table 5 below gives details of the country of birth variable used for this research. The variable was included in this research as there is research by others that shows that there can be relationships between labour market outcomes and individuals' country of birth and ethnic background. Section 2.4.13 of the literature review gives some examples research using these variables.

Item 23 of table 5 shows that the percentage of households living in social rented accommodation was used as a predictor variable in this research. This was included as housing tenure is often used by other researchers in their research into labour market outcomes. Examples of such research are given in section 2.4.14 of the literature review of this thesis.

4.2.2.2 Details of variables chosen for local authority level models

Table 5 shows the candidate variables that were chosen for possible inclusion in local authority level models to represent the topics described in section 4.2.2.1 above. As noted individually for each topic in section 4.2.2.1 above details of other researchers' work using these topics are given in section 2.4 of the literature review. The variables come from a variety of UK government surveys and analyses of official statistics including the 2011 Census. Despite the variety of original sources, the majority of the statistics were accessed via the Nomis website (Nomis, 2019). This gave access to Annual Population Survey (ONS, 2012) statistics for qualification levels and type of occupation; the Annual Survey of Hours and Earnings (ONS, 2016a) to estimate the percentage of people working part-time; the Business Register and Employment Survey (ONS, 2016b) for industrial sector information; and 2011 Census data relating to general health and distance travelled to work. Age and gender variables were obtained from ONS mid-2017 population estimates (ONS, June

2018). Official analyses of 2011 census data provided the Census area classifications (ONS, 2017) and the percentage of people living in rural areas (ONS, 2014). The deprivation variables came from the Department for Communities and Local Government Indices of Deprivation (2015a). Fuller details and definitions for each variable are shown in table 5 below.

Table 5: Candidate Outcomes and Predictor Variables for Local Authority Level Modelling Section

Research topic	Geography of data obtained	Data used to create variable
Outcome Variables		
1.Employment rate	Local authority and NUTS 1	Percentage of 16 to 64 year olds in employment in 2018. *Source: ONS, 2019 LI01 Local labour market indicators by unitary and local authority (based on employment data from Oct 2017-Sept 2018 Annual Population Survey, population data from mid-2017 population estimates).*
2.Unemployment rate	Local authority and NUTS 1	Percentage of economically active people aged 16 and over who are unemployed in 2018. *Source: ONS, 2019 LI01 Local labour market indicators by unitary and local authority (uses ONS model-based estimates of unemployment using APS and Job Seeker's Allowance data). This variable is consistent with the UN's Sustainable Development Indicator 8.5.2 and Sustainable Development Goal 8.5 concerning full and productive employment, decent work and equal pay (UN, 2019).*
3.Job density	Local authority and NUTS 1	Number of jobs per resident aged 16 to 64 in 2017. *Source: ONS, 2019 LI01 Local labour market indicators by unitary and local authority (job data include employees from December BRES and self-employed people).*
4.Resident earnings	Local authority and NUTS 1	Median weekly gross earnings of full-time workers resident in area. *Source: ONS, 2019 LI01 Local labour market indicators by unitary and local authority (data from Annual Survey of Hours and Earnings (ASHE)).*
5. Workplace earnings	Local authority and NUTS 1	Median weekly gross earnings of full-time workers by workplace. *Source: ONS, 2019 LI01 Local labour market indicators by unitary and local authority (data from Annual Survey of Hours and Earnings (ASHE)).*
6. Mean hours worked/week (workplace)	Local authority and NUTS 1	Mean paid hours worked per week, 2018, by workplace. Median hours also available. *Source: ONS (October 2018) Work Geography table 7.9a, data from the Annual Survey of Hours and Earnings.*

Research topic	Geography of data obtained	Data used to create variable
Predictor variables		
7. High qualifications rate	Local authority, NUTS 3, NUTS 2, NUTS 1	Percentage of people aged 16-64 who have an NVQ level four or higher qualification, 2018. *Source: ONS Crown Copyright Reserved [from Nomis on 18 July 2019]. Data from the Annual Population Survey*
8. No qualifications rate	Local authority, NUTS 3, NUTS 2, NUTS 1	Percentage of people aged 16-64 who have no qualifications, 2018. *Source: ONS Crown Copyright Reserved [from Nomis on 18 July 2019]. Data from the Annual Population Survey*
9. Part-time workers rate (residents)	Local authority, NUTS 3, NUTS 2, NUTS 1	Part-time workers as a percentage of full-time and part-time workers, 2018. *Source: ONS Crown Copyright Reserved [from Nomis on 18 July 2019]. Data from Annual Survey of Hours and Earnings Resident Analysis.*
10. Poor health rate	2011, Local authority, NUTS 3, NUTS 2, NUTS 1; also 2014, NUTS 1	Percentage of people reporting their general health to be either Bad or Very Bad, 2011 Census. *Source: ONS data accessed via Nomis, 19 July 2019. (Statistics for NUTS 1 areas also available for 2014 from the ONS Integrated Household Survey.)*
11. Median age	Local authority, NUTS 1	Median age, 2017. *Source: ONS mid-2017 population estimates.*
12. Gender balance	Local authority, NUTS 3, NUTS 2, NUTS 1	Females as a percentage of males and females, 2017. *Source: ONS mid-2017 population estimates.*
13. Herfindahl index of industrial diversity	Local authority, NUTS 3, NUTS 2, NUTS 1	A Herfindahl index of diversity (see section 2.4.7) was calculated based on 18 industrial groupings, 2017. *Data source: ONS Crown Copyright Reserved [from Nomis on 17 July 2019]. Statistics from open access Business Register and Employment Survey data.*
14. Managerial/professional rate	Local authority, NUTS 3, NUTS 2, NUTS 1	Managers and professionals as a percentage of all workers, April 2018 to March 2019. *Source: ONS Crown Copyright Reserved [from Nomis on 17 July 2019], data from the Annual Population Survey.*
15. Process/plant machine/elementary workers rate	Local authority, NUTS 3,	Process workers, plant machine and elementary job workers as a percentage of all workers, April 2018 to March 2019. *Source: ONS Crown Copyright*

Research topic	Geography of data obtained	Data used to create variable
	NUTS 2, NUTS 1	*Reserved [from Nomis on 17 July 2019], data from the Annual Population Survey.*
16. Herfindahl index of occupational diversity	Local authority, NUTS 3, NUTS 2, NUTS 1	A Herfindahl index of diversity (see section 2.4.7) was calculated based on nine occupational groups, April 2018 to March 2019. *Data source: ONS Crown Copyright Reserved [from Nomis on 17 July 2019], data from the Annual Population Survey.*
17. Census area classifications	Local authority (OA, LSOA and health area also available)	2011 Census area supergroup, group and subgroup classifications. *Source: ONS, 2017, the 2011 Area Classification for Local Authorities, revised 15 September 2017*
18. IMD average rank	Local authority (LSOA also available)	Average rank for IMD 2015 score. *Source: Department for Communities and Local Government, 2015a, the English Indices of Deprivation 2015, File 10 Local Authority District Summaries*
19. IMD employment domain average rank	Local authority (LSOA also available)	Average rank for IMD 2015 employment domain score. *Source: Department for Communities and Local Government, 2015a, the English Indices of Deprivation 2015, File 10 Local Authority District Summaries*
20. Commuting rate	Local authority, NUTS 3, NUTS 2, NUTS 1	Workers aged 16 to 74 travelling 30km or more to work each day as a percentage of all such workers, 2011. *Source: ONS Crown Copyright Reserved [from Nomis on 20 July 2019]. Data from 2011 Census, QS702EW - Distance travelled to work.*
21. Rural living rate	Local authority, NUTS 3, NUTS 2, NUTS 1	Percentage of population living in rural areas (including hub towns). *Source: ONS, 2014, the 2011 rural/urban classification.*
22. Country of birth	Local authority, NUTS 3, NUTS 2, NUTS 1	Percentage of people born outside the UK, 2017-18. *Source: Data download from ONS Bulletin, Population of the UK by country of birth and nationality July 2017 to June 2018.*
23. Housing tenure	Local authority, NUTS 3, NUTS 2, NUTS 1	Social rented tenure as a percentage of all households, 2011. *Source: ONS Crown Copyright Reserved [from Nomis on 2 October 2019]. Data from 2011 Census, QS405UK - Tenure – Households.*

Notes on boundary data

Digital boundary data for local authorities in the UK were downloaded from the UK Data Service (2019). The file infuse_dist_lyr_2011_clipped was the version of the boundary used. This contains boundary data for local authorities in the UK as at 2011 for compatibility with the 2011 Census. For England this means 324 authorities when the City of London and Westminster are combined as one and the Isles of Scilly and Cornwall and combined as one. As noted above, when the City of London and the Isles of Scilly are not merged with the larger authorities then there are 326 local authorities in England as at 2011. For mapping purposes maps produced with or without separate data for the City of London and the Isles of Scilly would not appear different from each other as the two local authorities cover such small areas. A reference list of local authorities was also downloaded from the same source, UK Data Service (2019). For completeness boundary information and a list were also downloaded for the four separate areas of City of London, Westminster, Isles of Scilly and Cornwall. These four separate areas are known as 'merging districts'. Their boundary information is stored in the file named 'infuse_merging_dt_2011_clipped'.

Any maps produced using these boundary datasets must show the following copyright notice. *Contains National Statistics data © Crown copyright and database right [2019]. Contains NRS data © Crown copyright and database right [2019]. Source: NISRA : Website: www.nisra.gov.uk. Contains OS data © Crown copyright [and database right] (2019).*

In addition, references to downloading the data should use the following text. *Office for National Statistics; National Records of Scotland; Northern Ireland Statistics and Research Agency (2011). 2011 Census: boundary data (United Kingdom) [data collection].*
UK Data Service. SN:5819 UKBORDERS: Digitised Boundary Data, 1840- and Postcode Directories, 1980-. Retrieved from http://census.ukdataservice.ac.uk/get-data/boundary-data.aspx. Contains public sector information licensed under the Open Government Licence v3.

4.2.3 Modelling techniques used for local authority models

As stated in the research aims of this project it was initially thought that this project would build models to take account of both the geographic proximity of places and the hierarchical nature of administrative geography in the UK. It was expected that multilevel modelling would be used as the primary method as it is ideally suited to the analysis of data that are available for the hierarchical levels of UK administrative and statistical areas. Building multilevel models was therefore expected to be the main focus of the model building phase of the project. It was thought that some specifically geographically weighted methods might be needed to explore the idea that some socioeconomic and business related processes operate depending on the distances between people and places rather than on whether they are located within the same administrative areas, e.g. employment chances for residents of one local authority may depend on job opportunities that are close by despite being in neighbouring local authorities. However, the suggestion by Moellering and Tobler (1972) that analysing data for hierarchical areas takes sufficient account of geographic scales implied that geographically weighted methods of analysis were not required for this project. The models built for this project using local authority level dependent variables were therefore all multilevel models using hierarchical administrative areas to take account of both hierarchical and distance effects.

For research projects which aim to search for a set of new geographic areas for the analysis of data (which is outside the scope of this project) geographically weighted techniques such as geographically weighted regression, described in 2.3.9 of the literature review, may be appropriate. They are less appropriate for the study of which existing areas are the most helpful to include in models that is at the core of this project. There is research by Browne et al (2001), Fielding and Goldstein (2006), Dong and Harris (2015), Dong et al (2016), Chen and Truong (2012) and Park and Kim (2014) that could be drawn on to incorporate geographically weighted techniques into multilevel modelling for such research.

4.2.4 Choice of local authority level dependent variables

Once the topics for this research had been chosen, as discussed in section 4.2.2.1, and data from which variables could be created had been selected, as set out in section 4.2.2.2, histograms were plotted for the proposed outcome variables. The histograms are shown and discussed in sections 4.2.4.1 and 4.2.4.2 below.

4.2.4.1 Histograms of proposed outcome variables

Figure 4: Histogram of employment rate, 16-64 year olds, by local authority, England, 2018

Employment rate, England, 16-64 year olds

Percentage of 16 to 64 year olds in employment
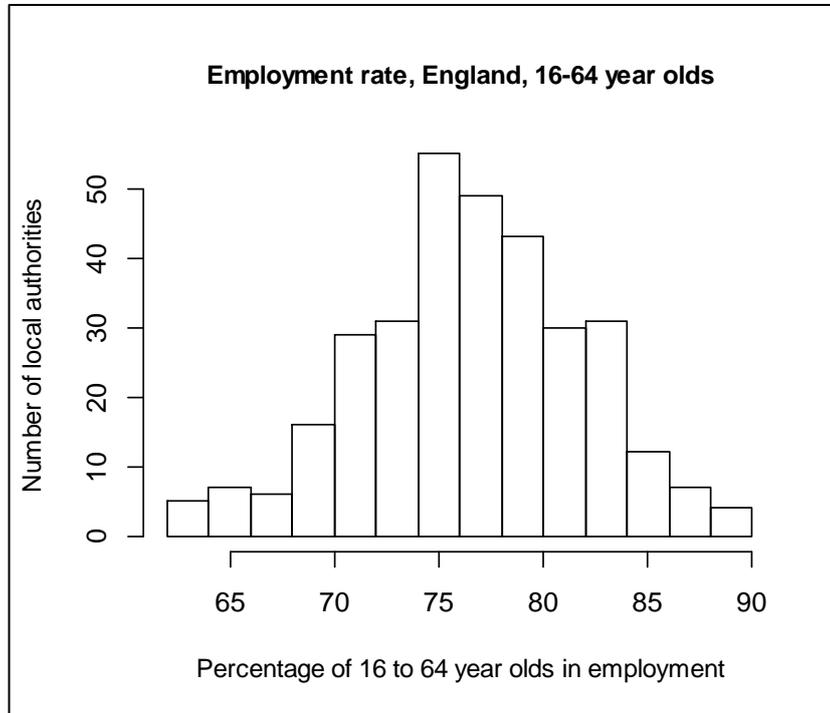
Figure 5: Histogram of unemployment rate, unemployed as a percentage of economically active, aged 16 and over, by local authority, England, 2018

Unemployment rate, England, 16-64 year olds

% of economically active 16 to 64 year olds who are unemployed
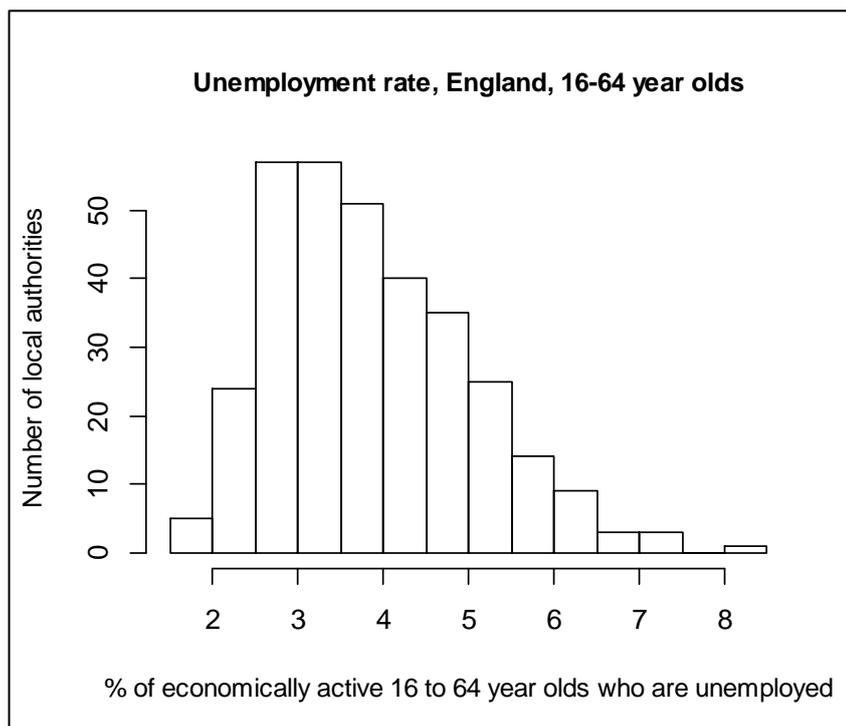
Figure 6: Histogram of residents' weekly gross earnings, by local authority, England, 2019

Median residents' weekly earnings, England
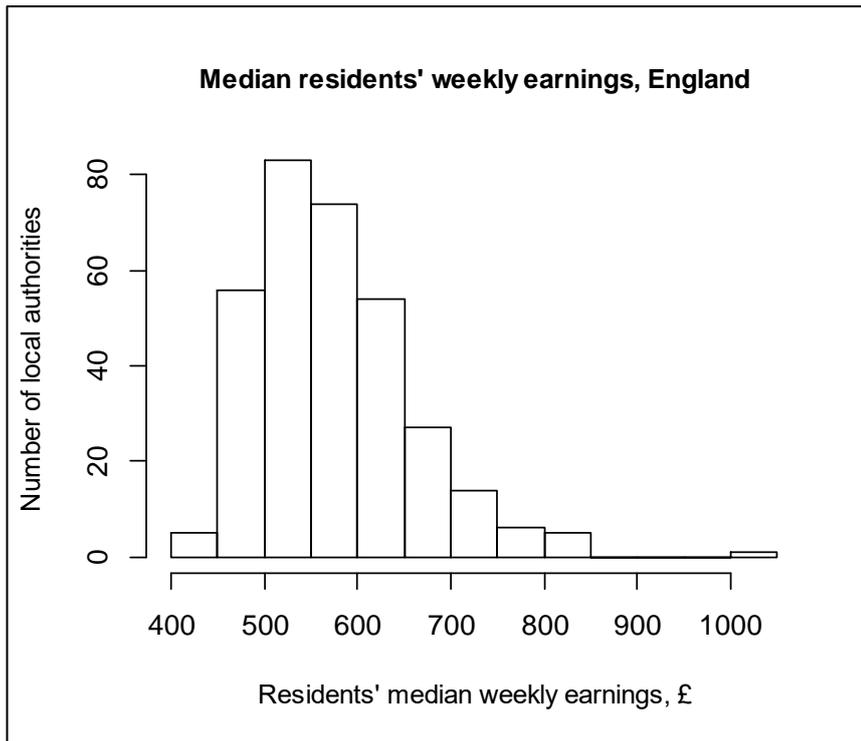
Figure 7: Histogram of workplace weekly gross earnings, by local authority, England, 2019
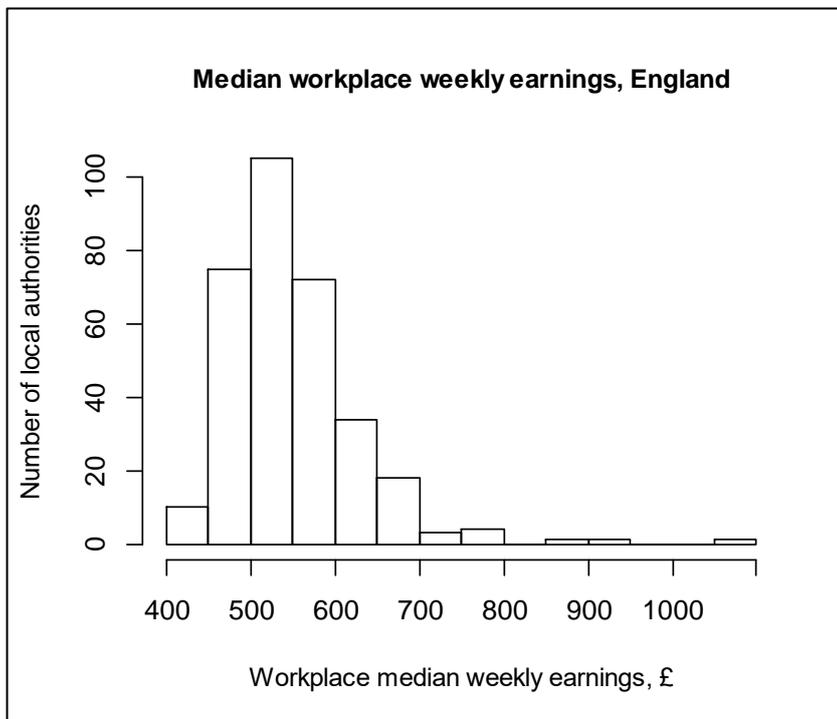
Median workplace weekly earnings, England
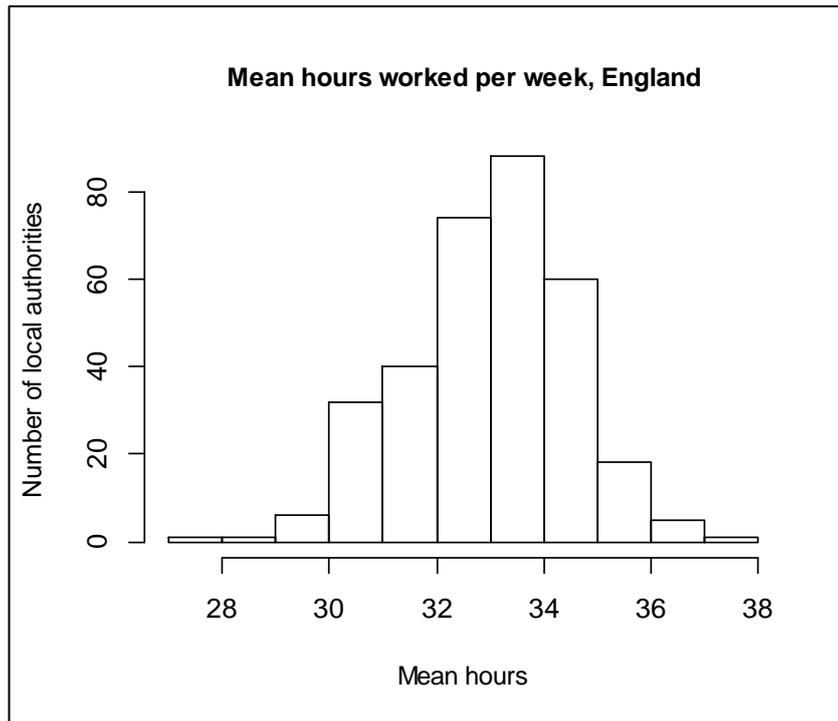
For a linear model to be reliable the residuals, the differences between the actual and fitted values, should be Normally distributed in order for the assumptions needed for a linear model to be met. That the histogram of the mean number of hours worked per week outcome shows a Normal distribution is a very reliable indication that the residuals will also have a Normally distribution.

Figure 9: Median number of hours worked per week, by local authority, England, 2018

Median hours worked per week, England

[Histogram: x-axis labelled "Median hours" ranging from 30 to 38; y-axis labelled "Number of local authorities" ranging from 0 to 150. Bars show a small count near 33, around 30 near 34, around 25 near 35, a tall bar exceeding 150 near 36, and about 85 near 37.]
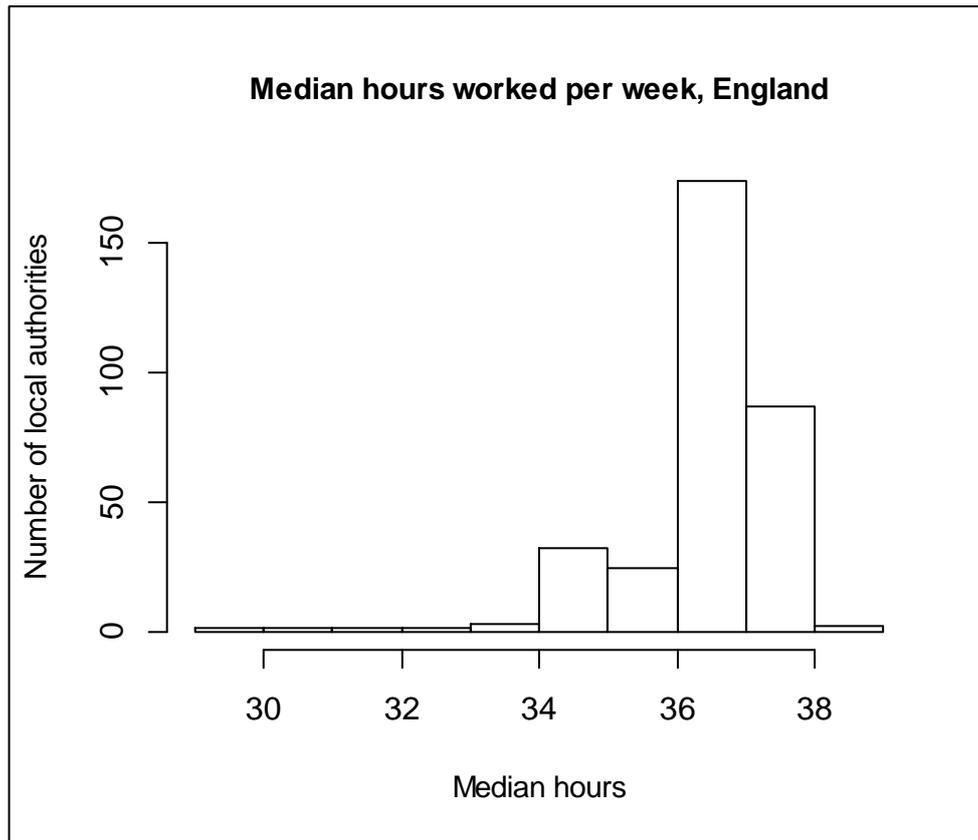
The histogram for the median number of hours worked per week shows an irregular distribution rather than a Normal distribution. This showed that the median number of hours is not a particularly suitable outcome variable for linear modelling. Between them figures 8 and 9 recommend that researchers wanting to model weekly earnings by local authority should focus on building models of the mean number of hours worked per week rather than the median number of hours worked per week.

Figure 10: Histogram of job density variable by local authority, 2017

Job density, all local authorities, England
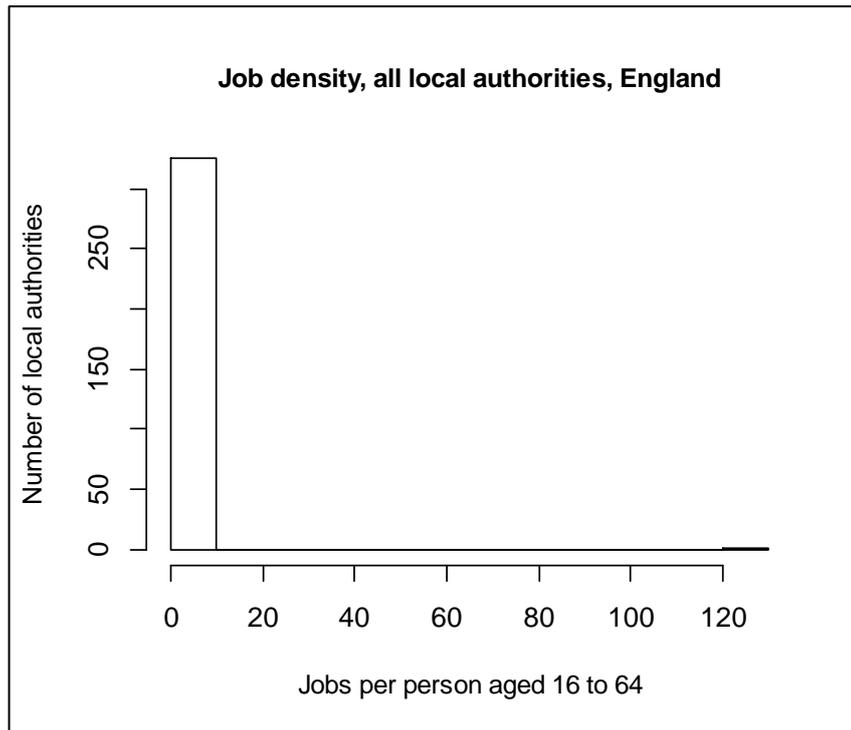
This histogram shows that the job density variable does not have a normal distribution. It illustrates the fact that most of the values are under ten and that one, perhaps, is over 120. As might be expected, subsequent inspection of the data confirmed the one very high job density to be for the City of London in line with its historical role as a workplace for many people but home to only a small population of residents. The data was also found to contain very high job density values for Westminster and Camden. To look at the distribution of the other local authorities in detail a histogram was produced that omitted the three local authorities with outlying job density values.

Job density, three outliers removed, England

The histogram of the job density values with the three outliers removed reveals a relatively Normal distribution for the rest of the local authorities. This raised questions about whether the outliers should be removed for subsequent analysis and modelling of the job density data which are answered at the end of section 4.2.4.2 below.

4.2.4.2 Suitability of outcome variables for modelling

The histograms of the potential local authority level outcome variables shown in figures 4 to 11 indicated which of the variables would be suitable for linear or generalized linear modelling. Models were therefore built using generalized linear modelling with the logit link function and a pseudo level at local authority level for the employment rate and unemployment rate variables (as they are each binomial variables which may display overdispersion) and linear modelling for the weekly earnings by place of residence, weekly earnings by place of work and mean number of hours worked per week (as they are all approximately Normally distributed).

The histogram of the median number of hours worked per week, figure 9, showed a skewed distribution especially when compared with the histogram for the mean number of hours worked per week, figure 8. This indicated that more reliable models would be built for the mean number of hours worked per week than for the median number of hours worked per week.

The histogram for the job density variable, figure 10, showed a very uneven distribution which gave the appearance of almost all of the local authorities having a job density of less than ten jobs per person and just one having a job density of over 120 jobs per person. Investigation of the numbers showed the City of London to have far higher job density than all other local authorities and Camden and Westminster to have higher job densities than the remaining local authorities. To investigate the distribution for the remaining local authorities a new variable was created by removing the job density values for the remaining local authorities. This enabled models of the local authorities without outlying job densities to be modelled directly using the full dataset. The histogram for this new variable, figure 11, showed a Normal distribution indicating that it would be appropriate to build linear models for the remaining set of local authorities. In order to check whether including or excluding the three local authorities with outlying job density values might make a difference to the results and conclusions that could be drawn, two sets of models were initial produced so that their results could be compared. On observing the results for the complete dataset it was deduced that modelling the reduced set of local authorities provided more reliable models of job density. Therefore, only results for the set of data excluding the outliers are shown in this thesis.

4.2.5 Calculating ICC/VPC values for null models

4.2.5.1 Null models for the local authority level unemployment and employment rates

In order to gauge which geographic scales were responsible for larger proportions of variance in local authority level unemployment and employment rates, a number of null models of the local authority level proportions of economically active people aged 16 years or over who were unemployed and of the local authority level proportions of people aged 16 to 64 who were employed were built that used different geographic scales as levels in the different models. The proportions of variance at different geographic levels were calculated for each model. An example of the R code used to build the models is shown in Annex 1, example 6. As described in the microdata and OA sections of this methodology chapter (sections 3.2 and 3.3 respectively), the percentages of variance at individual level are estimates produced using the estimated variances at individual level of $\pi^2/3$ for models using the logit link function and one for models using the probit link function. To take account of possible binomial overdispersion a pseudo level (as described in section 3.2.4.3

above) at local authority level was included in the local authority level models of unemployment rates and of employment rates. An example of a model equation for a null model of an extra-binomially distributed outcome such as an area unemployment or employment rate using the logit link function and containing a pseudo level to account for binomial overdispersion is shown as equation 6 below.

Binomial response with pseudo level to account for overdispersion

– Null model for local authorities grouped by NUTS 3 areas

Where $y_{ijk} \sim$ Binomial(n_{ijk}, p_{ijk}), the model can be written as:

$$\text{logit}(p_{ijk}) = \beta_0 + v_K + u_{jk} \tag{6}$$

Where the v_K and the u_{jk} are the random effects.

This is based on the assumptions that: $v_k \sim N(0, \sigma^2_{NUTS3})$ and $u_{jk} \sim N(0, \sigma_u^2)$ and that these random effects are uncorrelated.

i = Level one = local authorities

j = Level two = the pseudo level which is also the local authority level (each level two unit will contain one level one unit).

K = Level three NUTS 3 areas

The proportions of variance for null and random intercept models are the same as the ICC values for each model. The ICC values are presented and discussed in sections 5.3 (null models) and 5.5 (random intercept models) in chapter 5 of this thesis. Important for the methodology for the rest of the project were the general findings that there was an appreciable proportion of the variance at each of the NUTS areas levels that were included in the models of unemployment. Although the proportions of variance may be relatively small which would often be regarded as grounds for not using multilevel modelling there is also an argument that because the data are hierarchical (local authorities are nested within NUTS 3 areas which are nested within NUTS 2 areas which are nested in NUTS 1 areas) then multilevel modelling should be used (see section 2.3.3 of the literature review and Nezlek,

2008). Given that the hierarchical nature of the data is grounds for using multilevel modelling then the relative size of the proportions of variance at each of the group levels in the models can be taken to indicate which of the group levels should be used to model the data. Therefore in this case the fact that relatively large amounts of the proportions of variance were at each of the NUTS areas levels showed that all three of the NUTS area levels should be included in the multilevel models for the remaining local authority level dependent variables.

4.2.5.2 Null models for the local authority level Normally distributed dependent variables

Null, four-level models were built of the local authority level mean number of hours worked per week, median weekly earnings for residents of each local authority, median weekly earnings for those whose workplace was in each local authority, the difference between median residential and median workplace earnings for each local authority, and job density by local authority (excluding the three local authorities with outlying job density values, namely the City of London, Camden and Westminster). The R code used to build each of these models is shown in Annex 1, example 7. An example of a model equation for a multilevel model for an interval response variable such as area-level hours worked or weekly earnings is shown as equation 7 below.

Interval response – Null 2-level model for local authorities grouped by NUTS 3 areas

$$y_{ij} = \beta_0 + u_j + e_{ij} \tag{7}$$

based on assumptions that: $u_j \sim N(0, \sigma^2)$ and $e_{ij} \sim N(0, \sigma^2_e)$ and that these random effects are uncorrelated,

where,

y_{ij} = value of response for i^{th} local authority in j^{th} NUTS 3 area

β_0 = overall mean of responses regardless of local authority

u_j = random effect = group residual = difference between mean for NUTS 3 area j and overall mean

e_{ij} = random effect = individual residuals

= difference between i^{th} observation in NUTS 3 area j and mean for NUTS 3 area j

4.2.6 Models with independent variables at different geographic scales

Building on from null models which use dependent variables and area identifiers at different levels but no independent variables, random intercept models of local authority employment and unemployment rates were built to investigate the effect on model fit of including independent variables at different geographic scales.   The idea behind this was to test whether it was necessary or helpful to include independent variables at geographic scales larger than local authority level (in order take account of opportunities influences that effect larger geographic areas) and whether it was sufficient to include predictor variables for one geographic scale only or whether predictor variables at more than one geographic scale were necessary in order to build well-fitting models.   For both local authority employment and local authority unemployment rates five separate models were built for each of 18 different independent variables. For each dependent/independent variable combination a model was built using the independent variable at local authority level, at NUTS 3 areas level, at NUTS 2 areas level, at NUTS 1 areas level and a model with four independent variables – all measures of the same statistic, one at local authority level and one at each of the NUTS 3, NUTS 2 and NUTS 1 areas levels.

The first step for each independent variable was checking that suitable data were either available at each geographic scale or could be calculated from local authority level data. For some independent variables that were created at area rather than individual level, e.g. Census area type, it was not possible to aggregate the local authority data to create meaningful values for the larger geographic areas. For example, median values for larger areas cannot be simply calculated from median values for each of the smaller areas that they contain. Similarly, area-level classifications such as Census area type cannot readily be combined for different areas to give an area type on the same basis for a larger area. For each of the local authority, NUTS 3 areas, NUTS 2 areas and NUTS 1 areas geographic scales an industrial diversity indicator and an occupational diversity indicator were calculated to use as an independent variables. R code written to calculate the local authority level industrial diversity indicator is shown at Annex 1, example 8. In order to compare the different models of each dependent/independent variable combination a single model fit measure was adopted, that of the Akaike Information Criterion (AIC). The merits of using the AIC rather than the alternative Baysian Information Criterion (BIC) are weighed by Fabozzi et al 2014. The AIC gives a figure for the lack of fit for each model adjusted to take some account of the number of parameters in the model (page 58, Robson and Pevalin, 2016). If the AIC values for two models are compared the lower value indicates a better

fitting model that does not contain too many extra model parameters in line with the principle of choosing a parsimonious model. Generally, an absolute decrease of two or more tends to be taken to identify a better fitting model (Fabozzi et al, 2014). However, given the very large AIC values of most of the models in this project some consideration was also given to the relative differences rather than just the absolute differences. Although AIC values tend to be positive rather than negative there is no reason why an AIC value should not be negative. The equation to calculate the AIC value of a model is:

AIC = (2 * number of parameters) – (2 * ln(L))   (8)

where L is the maximised value of the likelihood. The AIC value will therefore be negative if the natural log of the maximised likelihood is greater than the number of parameters to be estimated by the model.

In the models of job density in this thesis the AIC values calculated are all negative. The important point to note is that when comparing the AIC values of two models to see which fits the data better it is the difference between the two values that is important rather than whether they are negative or positive. Where both AIC values are negative then it will be the model with the most negative AIC value that fits the data better rather than the model with the AIC value that is closest to zero. This is relevant when considering the results shown in section 5.7.6 of this dissertation.

 All the models were written using R code as shown in Annex 1, example 7. The dependent variables, referred to as the output variable in the model, were set before running the code to create the model. In that way the model building code could be used for both the employment and unemployment dependent variables although separate, similar, code sections were written for each independent variable at this stage of the project. From this exercise the way of writing the codes for subsequent models for later stages of the project way adapted to write more generic code in which both the dependent and independent variables could be set and so that the same lines of code could be re-used to save duplicating large amounts of similar code. At this stage of the project it was however useful to write the separate code for each independent variable as it focussed attention on ensuring the appropriate data were used for each independent variable at each geographic scale where possible. The AIC values generated from the models are shown and discussed in section 5.4 of this thesis. From the point of view of the methodology used for the rest of the project, the important finding was that using independent data at local authority scale was generally more helpful than using independent data at larger geographic scales. This

finding led to all of the models in the subsequent stages of the project using local authority level independent variables rather than independent variables for larger geographic scales.

4.2.7 VPC values for random intercept models

Section 4.2.5 above described how null models were built for each of the local authority level dependent variables and their variance partition coefficient values were calculated and compared. For the next stage in the project random intercept modes were built for each of the local authority level dependent variables. This was done in order to see if models which included a predictor variable were better fitting than the equivalent null models. For each dependent variable / independent variable combination a four-level random intercept model was built with local authority as the lowest level, NUTS 3 areas as the next level, then NUTS 2 areas level, and finally NUTS 1 areas level as the highest level in the model. For each of these models the VPC values for each level of the model were calculated. These random intercept models were built to see if geographic scales with relatively large proportions of variance in the null models had similarly large proportions of variance when each of the independent variables was added. The VPC values are all presented and discussed in sections 5.3 and 5.5 of this thesis. Comparing the VPC values for null models and random intercept models made it possible to see if the geographic scales which might be chosen by researchers based on evidence from VPC values for null models would be different if they used VPC values from random intercept models instead to make their choices. This exercise was not extended to include random coefficient models as VPC values are not generally calculated for random coefficient models as the variable coefficients mean that such models do not have unique proportions of variance at each level (Kreft and de Leeuw, 1998, page 63, and Nalborczyk, 2017).

Each of the four-level random intercept models was built using R. An example of the code used to build the models for Normally-distributed variables and extract the VPC (ICC) values is shown in Annex 1, example 9.

4.2.8 Possible higher level models

The VPC values for residents' earnings and for workplace earnings calculated as described above in section 4.2.7 showed very high proportions of variance for the NUTS 1 areas level,

the highest level in the four-level models. The literature on missing levels from multilevel models discussed in section 2.3.8 of the literature review, e.g. Tranmer and Steel (2001a, 2001b), Opdenakker and Van Damme (2000) and Moerbeek (2004), suggests that if the highest level is missed out from a multilevel model then effects due to that level are transferred to the highest level that is included in the model. This might mean that the high proportion of variance that appears to be at NUTS 1 areas level for the two earnings variables may in fact be partly due to a level higher than NUTS 1 areas level that is not included in the four-level models. This might mean that there is a higher geographic scale that future researchers could find it helpful and appropriate to include in their models of earnings. It is reasonable to argue that earnings tend to be higher in London for reasons such as London weighting, the London Living wage (Mayor of London/London Assembly, 2020, and London Data Store, 2020) and relatively high housing and other living costs in the capital and surrounding counties.   An initial candidate 'missing level' might therefore consist of just two areas such as London and "the rest of England". Five-level models were built for residents' earnings and for workplace earnings to start to explore this idea. The results, in terms of VPC values, are reported in section 5.5.8 of this thesis. The very high proportion of variances at the new London/not London higher level in both the residents' earnings and workplace earnings models suggested that the higher level should be included in models of earnings. To provide an alternative higher level arrangement, a split of the local authorities into "London and the South East" and "the remaining regions" was also generated and used to build a five-level model. The R code used to create the higher level identifiers and to build the five-level models is shown in Annex 1, example 10. An alternative explanation of why there is a high proportion of variance at NUTS 1 areas levels is that it is simply a matter of London workplace incomes being higher than those in all other NUTS 1 areas, e.g. average workplace earnings in London in 2019 were £713, which was 24% higher than the average for the whole of England, £575, and 41% higher than the average for workplaces in the North East NUTS 1 area, £507 (ONS, 2019f). The differences in incomes between London and the rest of England could be modelled by introducing a dummy variable into models that takes the value of one for all local authorities that are in London and zero for all local authorities that are not in London. A modification of this would be for the dummy variable to take the value of one for local authorities that are in London and for those that are 'near' to London with the definition of near being experimented with to try to find the geographic area which needs to take the value of one for the dummy variable in order to improve the fit of models of income levels.

4.3 Summary of methodology and results shown in chapter 4

The earlier stages of this research used microdata for individuals and Output Area level data for small areas to start to investigate the proportions of variance in dependent labour market variables such as employment and unemployment rates that occurred at different geographic scales. OA level data in particular also provided a good environment to start to develop multilevel models of employment and unemployment rates. There were limitations however in how much progress could be made using microdata due to the very limited geographic information available (only NUTS 1 area identifiers). OA level data provide good information for very small areas which are the building blocks for producing all other Census and related statistics. However, the very fine geographic scale means that very large numbers of observations are needed to cover the whole of England which can make models very complex and very slow to run. For example, experiments with random coefficient multilevel models at OA level led to models that were over complex involving very large numbers of parameters which tended not to converge or to give a model with a singular fit (see section 3.3.3.3 above). Similarly, adding a pseudo level at OA level to account for binomial overdispersion proved impractical as it created a very large number of groups with only one member which would not converge successfully without generating warning messages.

Knowledge gained from the research using microdata and OA level data was then used in the investigation and modelling of local authority level data. The modelling of local authority data started with calculating ICC/VPC values for null models. Next random intercept models were built using independent variables at different geographic scales. The AIC values of the models with independent variables at different scales were compared and it was found that in general those models with independent variables at local authority scale fitted the data better than those models with independent variables at larger geographic scales. This informed the decision to use independent variables at local authority rather than larger scales for the rest of the modelling. A series of four-level multilevel models were then built for each dependent variable with each independent variable in turn, each at local authority level and their VPC values were studied to see which geographic scales accounted for the highest proportions of variance for each combination of dependent and independent variables. These models only included one independent variable at a time in order to study the geographic scale effects for each independent variable separately. As there were very many separate independent variables

the modelling process was not developed further to include models with more than one independent variable. Future development of this research could include building models with more than one independent variable to see how this affects the proportions of variance at each geographic scale. Finally, a comprehensive set of null models and random intercept and random coefficient models for each combination of dependent and independent variables was built using R. The AIC values were extracted from each model and compared. The results are shown in section 5.7 and Annexes 2 to 6. The R codes used to build the models for Normally distributed and binomially distributed dependent variables are shown in Annex 1, example 11.

4.4 Local authority random intercept and random coefficient models for all dependent and independent variables

Although this stage of the research builds the most substantial set of models, it is the culmination of the research that has gone before. No new elements are introduced in the codes, instead the codes are used to build null, random intercept and random coefficient models for all combinations of dependent and independent variables. For the outcome variables with Normal distributions null models, models with random intercepts at all levels and models with random intercepts and random coefficients at all levels were built first and their AIC values were compared. As many of the four-level models with random coefficients at all levels either failed to converge or, more commonly, generated a warning about the model having a singular fit (see section 3.3.3.3 above) additional four-level models were built with random intercepts at all levels and random coefficients at NUTS 1 areas level only or at NUTS 2 areas level only. The AIC values for all these models are shown in section 5.7 and Annexes 2 to 6 of this thesis.

The code to build multilevel models for the binomially distributed dependent variables includes a pseudo level at local authority level. The code with the pseudo level worked well to produce null and random intercept models. However, the code with pseudo level in the random coefficient models led to warning/error messages for all combinations of dependent and independent variables for all levels of geographic scale rather than producing models. An example of the message is as follows, *“Error: number of observations (=325) < number of random effects (=650) for term (1 + Predictor_variable | LAD13CD:NUTS315CD); the random-effects parameters are probably unidentifiable”*. The message is simply pointing out including a pseudo level in random coefficient models has the effect of including a random intercept and a random coefficient for each observation which means the model specified would have twice as many random effects as it has observations which does not give a usable model. Models with random coefficients at all levels are not therefore included in the results shown in Annex 5 for employment and Annex 6 for unemployment models as it did not prove possible to build models with random coefficient models at all levels including at the pseudo level that is used to account for binomial overdispersion. Instead a set of models were built which had random intercepts at all levels and random coefficients only at some levels. Specifically, for each

combination of outcome and predictor variable: one model was built with random coefficients at NUTS 3 areas level, NUTS 2 areas level and at NUTS 1 areas level but not at the pseudo level; one model was built with random coefficients at NUTS 2 areas level only; and one model was built with random coefficients at NUTS 1 area level only. The results for these are shown in Annexes 5 and 6.

5. Results for local authority areas

5.1 Introduction to the results for local authorities

This chapter presents the results from a variety of models of labour market statistics aggregated at the geographic scale of the 326 English local authority districts and unitary authorities in existence up until early 2019 (see section 4.2.1 in Chapter 4, *Methodology Part 2*, for details of why this set of administrative areas was chosen). The results comprise findings from different phases in the exploration and analysis of the statistics for English local authorities. Each phase provided findings which were both of merit in themselves and which informed subsequent phases of the analysis.

Section 5.3, *VPC for null models*, reports the results of an investigation into the proportions of variation in the outcome variables that can be attributed to different geographic scales. This was carried out by calculating and comparing the variance partition coefficient (VPC) values of null four-level models. VPC values provide a fundamental way of deciding which geographic scales to include in multilevel models. Those scales that are related to greater proportions of variance are likely to be important to include, and those related to much smaller proportions of variance are likely to be less helpful to include in such models.

Section 5.4, *Predictors at different scales*, provides results from four-level random intercept models of employment and unemployment rates which use predictor variables at different scales. These results were used to assess which geographic scales to use for predictor variables in multilevel models in the next phases of the work.

Section 5.5, *VPC values for random intercept models*, shows results from the calculation of variance participation coefficients (VPC) values for null models and random intercept models with predictor variables at local authority level. This section expands the use of VPC values for null models presented in section 5.3 by including predictor variables in order to see if the geographic scales at which larger proportions of variance are found are the same in random intercept models with predictor variables as they are in null models. Only local authority level predictors were included in the models as the results reported in section 5.4 showed that almost always local authority level predictors created models that fitted the data at least as well, and usually better, than models with predictors measured at higher geographic scales. Whilst the VPC values from null models provide an indication of

whether it is likely to be useful to carry out multilevel modelling rather than single level regression modelling, the VPC values from random intercept models provide more information on which geographic scales to include in multilevel models that include a variety of different predictor variables. Random coefficient models do not have unique VPC values due to their variable coefficients (Kreft and de Leeuw, 1998, page 63). Therefore section 5.5 does not include any random coefficient models.

Section 5.6 presents a summary of results shown in sections 4.1.1 to 5.5.8.

Section 5.7 reports results from a different method of providing evidence about which geographic scales to include in multilevel models of labour market statistics that can be used for random coefficient models as well as for random intercept models, as AIC values can be compared for random coefficient as well as for random intercept models. This method consisted of building a large number of random intercept and random coefficient models for each outcome variable using different predictor variables, all measured or calculated at local authority level as recommend by the results in section 5.4, but deployed within multilevel models that grouped the local authorities at different geographic levels. The models were then examined to see which ones fitted the data better by comparing the AIC values for groups of the models that used the same outcome and predictor variables in different ways. For each outcome variable separate null, random intercept and random coefficient models were built for four different combinations of levels: a four-level model (local authority, NUTS 3 areas, NUTS 2 areas and NUTS 1 areas) and three two-level models with local authority level as level one and either NUTS 3 areas, or NUTS 2 areas, or NUTS 1 areas as level two.

5.2 Local authority level outcome variables

Seven outcome variables were considered for local authority level modelling as discussed in section 4.2.4 of the methodology chapter. Histograms of their distributions are shown in section 4.2.4.1 together with discussion on how they were used to determine exactly which local authority level dependent variables to model for this section of the project. As a result the following local authority level outcome variables were modelled and it is the results of these models that are presented in this chapter: unemployment rate for economically active people aged 16 or over; employment rate for all people aged 16 or

over; mean hours worked per week; median earnings for local authority residents; median earnings by local authority workplace; job density for local authorities excluding three London boroughs with exceptionally numbers of jobs per resident.

5.3 VPC for null models

5.3.1 Variance partition coefficient values for the unemployment rate outcome variable

Null models of the proportion of economically active people aged 16 years or over who were unemployed were produced using different geographic scales and the proportions of variance at different geographic levels were calculated. The results are set out below. As noted in the methodology chapters (in sections 3.2.4.3 , 3.3.3.1 and 4.2.5.1) percentages of variance at individual level shown in these tables are estimates produced using the estimated variances at individual level of $\pi^2/3$ for models using the logit link function and one for models using the probit link function.

Table 6: VPC values for local authority level unemployment for models with different levels

	Local Authority level grouping only	Local Authority and NUTS 3 areas grouping	Local Authority, NUTS 3 areas and NUTS 2 areas grouping	Local Authority, NUTS 3 areas, NUTS 2 areas and NUTS 1 areas grouping	Local Authority, NUTS 3 areas, NUTS 2 areas and NUTS 1 areas grouping
Link function	**Logit link**	**Logit link**	**Logit link**	**Logit link**	**Probit link**
Estimated individual level VPC	97.2%	97.3%	97.2%	97.2%	98.1%
Local Authority level VPC	2.8%	1.0%	1.0%	1.0%	0.7%
NUTS 3 areas level VPC	n/a	1.7%	0.7%	0.7%	0.5%
NUTS 2 areas level VPC	n/a	n/a	1.1%	0.3%	0.2%
NUTS 1 areas level VPC	n/a	n/a	n/a	0.8%	0.5%

In both the logit and probit null models of unemployment that use local authority, NUTS 3, NUTS 2 and NUTS 1 areas level there is a relatively large amount of the proportion of variance at each four of these geographic levels. This implies that there could be factors that affect local authority unemployment rates that operate at all of these levels and therefore that researchers should consider including grouping and predictor variables at all four of these geographic scales in their models of unemployment. Section 2.3.3 of the literature review and Nezlek (2008) note that though the proportions of variance may be relatively small which would often be regarded as grounds for not using multilevel modelling there is also an argument that because the data are hierarchical (local authorities are nested within NUTS 3 areas which are nested within NUTS 2 areas which are nested in NUTS 1 areas) then multilevel modelling should be used. Given that the hierarchical nature of the data is by itself a reason for using multilevel modelling then the relative size of the proportions of variance at each of the group levels in the models can be taken to indicate which of the group levels which be used to model the data. In this case the fact that relatively large amounts of the proportions of variance were at each of the NUTS areas levels implies that NUTS 3 areas level, NUTS 2 areas level and NUTS 1 areas level should all be included in the model.

Alternatively, in both of these models the proportion of variance that is at NUTS 2 areas level is approximately half that associated with any of the other levels which might tempt a researcher to miss the NUTS 2 areas levels out of models. The expectation arising from literature discussed in section 2.3.8 of the literature review, e.g. Tranmer and Steel (2001a, 2001b), Opdenakker and Van Damme (2000), Van den Noortgate et al (2005), and Moerbeek (2004), on the effect of missing a level from a multilevel model is that missing the NUTS 2 areas level from either of these models would transfer the variance due to NUTS 2 areas effects to either or both of the NUTS 3 and NUTS 1 areas levels and leave the proportion of variance at the local authority level unchanged. To demonstrate this a new probit model was built that missed out the NUTS 2 areas level.

Local authority, NUTS 3 and NUTS 1 areas levels model using the probit link function
The new probit model gave the following proportions of variance.
Estimated Individual level VPC = 98.2%
Overdispersion/LA level variance VPC = 0.7%

NUTS 3 areas level VPC = 0.6%

NUTS 1 areas level VPC = 0.6%

These VPC percentages show that missing the NUTS 2 areas level out of the model has caused some of the variance at NUTS 2 areas level to be transferred down to the NUTS 3 areas level and some to be transferred up to the NUTS 1 areas level.

5.3.2 Summary of results from the variance partition coefficient values for unemployment

The variance partition coefficient (VPC) values for the four-level null model of the unemployment rate variable showed there to be appreciable variation at all four geographic levels in the model (local authority and lower, NUTS 3 areas, NUTS 2 areas and NUTS 1 areas). A three-level null model that missed out the NUTS 2 areas level, on the grounds that it had the lowest VPC, gave a higher VPC for the NUTS 3 areas level and an unchanged VPC at the local authority and lower level. This was consistent with research by others on the effects of missing out a level in a multilevel model discussed in section 2.3.8 of the literature review, e.g. Tranmer and Steel (2001a, 2001b), Opdenakker and Van Damme (2000), Van den Noortgate et al (2005) and Moerbeek (2004), that showed that missing out a level causes the variance at that level to be transferred either up or down (or both) to the next levels in the model (in this case the NUTS 3 areas level) but not to other levels (in this case the local authority or lower level). These results together suggested that multilevel models of unemployment rates should include all four levels even though the NUTS 2 level had a smaller VPC than other levels. This is because missing out the NUTS 2 areas level would have an effect on the NUTS 3 areas level variance estimate of the model.

5.3.3 Variance partition coefficient values for the employment rate outcome variable

The null models built for the proportion of all people aged 16 – 64 years in employment using different geographic scales as the levels in the models showed the following proportions of variance at different geographic scales.

Local authority (LA) level model only using the logit link function

Estimated Individual level VPC = 97.3%

Overdispersion/LA level VPC = 2.7%

Local authority and NUTS 3 areas levels model using the logit link function

Estimated Individual level VPC = 97.3%

Overdispersion/LA level VPC = 2.1%

NUTS 3 areas VPC = 0.6%

Local authority, NUTS 3 and NUTS 2 areas levels model using the logit link function

Estimated Individual level VPC = 97.3%

Overdispersion/LA level variance VPC = 2.1%

NUTS 3 areas level VPC = 0.2%

NUTS 2 areas level VPC = 0.5%

The proportion of variance at NUTS 2 areas level is higher than that at NUTS 3 areas level. This implies that there may be factors that operate at NUTS 2 areas levels that have a great effect on a person's chance of being at work than the factors that operate at NUTS 3 areas levels, e.g. job opportunities in the wider NUTS 2 area may play an important part in the chances of people being in employment. An implication of this is that researchers should include grouping by NUTS 2 areas and predictor variables at NUTS 2 areas levels in their models.

Local authority, NUTS 3, NUTS 2 and NUTS 1 areas levels model using the logit link function

Estimated Individual level VPC = 97.2%

Overdispersion/LA level variance VPC = 2.1%

NUTS 3 areas level VPC = 0.2%

NUTS 2 areas level VPC = 0.2%

NUTS 1 areas level VPC = 0.3%

Local authority, NUTS 3, NUTS 2 and NUTS 1 areas levels, <u>probit </u>link function

Estimated Individual level VPC = 97.0%

Overdispersion/LA level variance VPC = 2.3%

NUTS 3 areas level VPC = 0.2%

NUTS 2 areas level VPC = 0.2%

NUTS 1 areas level VPC = 0.3%

The VPC values for the NUTS 3, NUTS 2 and NUTS 1 area levels are all similar. This suggests that there may be factors effecting the chances of a 16-64 year old being in employment that operate at each of these geographic scales. This provides some evidence that researchers should try including grouping and predictor variables at each of these levels in their models as well as predictor variables at local authority level. Both probit and logit link models were built simply to see whether the choice of link function had an appreciable affect on the VPC values that would change any conclusions drawn from the values. Both models produced very similar VPC values suggesting that the choice of link function did not make a difference to any conclusions. Subsequent models therefore were all built just using the logit link function as there appeared to be no advantage in building both logit and probit versions of each model.

The proportion of variance due to overdispersion or at LA area level is approximately ten times higher than that at each of the NUTS areas 3, 2, and 1 levels. This is different to the finding from the models of unemployment for economically active people aged 16 or over where the proportion of variance at LA level was broadly similar to the proportion of variance at the NUTS areas 3 level and at the NUTS 1 areas level. This implies that the reasons for the proportions of 16-64 year olds that are employed may be more closely related to conditions within the local authority than are the reasons for the proportions of

economically active people aged 16 or over who are unemployed. Part of the reason for this may be that those over 16 year olds who are economically active may find it easier to travel outside of their home local authority to work than the general population of 16 to 64 year olds who either choose or are restricted in some way to remaining within their home local authority which limits suitable job opportunities. There is evidence of transport deprivation being a barrier to economically inactive people entering employment. In Northern Ireland for example, Department for Employment and Learning and Department of Enterprise, Trade and Investment (2015) list accessible transport issues as a structural barrier to labour market participation (page 32) and include 'addressing issues of transport deprivation' as an area for further exploration (page 37). Titheridge et al (2014) report on links between poverty and transport deprivation (including lack of access to a car and level and cost of bus services) and job seekers attending interviews and young people entering jobs, training or work (pages 15-16). Any such suggestion would need to be explored in ways outside of this work before any claims could be made. This project can only suggest that the mobility of economically active people aged over 16 might be a factor that could be considered by future research into the reasons underlying local authority employment and unemployment rates. A conclusion that can be drawn that is central to this work is that the proportion of variance at different geographic scales and hence the geographic scales that researchers should consider including in their models can vary for seemingly similar variables.

5.3.4 Summary of results from the variance partition coefficient values for employment

Comparing the VPC values for two-level, three-level and four-level models shows there to be variance at each of the NUTS areas levels that could be missed by using models with only two or three levels instead of four levels. The effects of missing out a level in a multilevel model and the importance of including NUTS 1 area levels in models of employment rates is illustrated by comparing models with and without grouping by NUTS 1 areas on the VPC value for NUTS 2 areas. The NUTS 2 areas VPC is approximately halved for the four-level null model that includes grouping by NUTS 1 areas compared to the three-level null model which does not include the NUTS 1 areas level.

5.3.5 Mean hours worked per week

A null, four level model for the mean number of hours worked per week by local authority produced the following variance partition coefficients.

Local Authority and Lower level VPC = 87.71%

NUTS 3 areas level = 0.09%

NUTS 2 areas level = 10.33%

NUTS 1 areas level = 1.88%

These results show approximately 88% of variance in mean hours worked being due to local authority area level and lower level effects, 10% being due to NUTS 2 areas level effects, 2% being due to NUTS 1 areas level effects and only a negligible amount of the variance being due to NUTS 3 areas level effects. This suggests that it may be important to include NUTS 2 areas level grouping and predictors in models of mean hours worked by those in different local authorities but not important to include NUTS 3 areas level groupings or variables.

5.3.6 Summary of results from the variance partition coefficient values for mean hours worked outcome variable

The calculation of VPC values for mean hours worked per week for each of the four geographic scales ('local authority and lower', NUTS 3 areas, NUTS 2 areas and NUTS 1 areas) showed that only a very tiny proportion of the variance could be attributed to the NUTS 3 areas scale. This showed that NUTS 3 areas were not an appropriate geographic scale for studying the mean number of hours worked per week. This suggests that it would not be particularly helpful to have statistics on hours worked, and possibly other related variables, at NUTS 3 areas level. A higher proportion of variance was attributed to NUTS 2 areas meaning that it would be more helpful to future researchers to have statistics about the numbers of hours worked for NUTS 2 areas instead.

5.3.7 Median earnings for local authority residents

A null, four-level model for median weekly earnings for residents by local authority showed the following proportions of variance at each of the geographical scales in the model.

Local Authority and Lower level VPC = 35.16%

NUTS 3 areas level = 13.90%

NUTS 2 areas level = 13.83%

NUTS 1 areas level = 37.11%

These results show approximately a third of the variance in median residential incomes to be due to local authority level and lower level differences, just over a third due to NUTS 1 areas level differences and the remaining almost a third split between NUTS 2 area levels differences and NUTS 3 areas level differences. These results imply there that there could be important factors affecting the median earnings of people living in local authorities that operate at each of the geographic scales in the model (LA, NUTS 3, NUTS 2 and NUTS 1). This shows that it could be important for researchers to include grouping and predictor variables at each of these levels in their models of the median earnings of residents in local authorities.

It is interesting that the proportion of variation at NUTS 1 area level is high, being similar to that at local authority level and larger than that at NUTS 2 areas levels and NUTS 3 areas level combined. Further research might be appropriate to see how much of the NUTS 1 areas level variation is due to different income levels in the NUTS 1 London area compared with other NUTS 1 areas, i.e. to see whether there is a 'London effect' that has a substantial influence on the overall variation in income levels across local authorities in England.

5.3.8 Summary of results from variance partition coefficient values for residents' earnings

The VPC values for different levels in the four-level null model of median earnings for people who live in each local authority show around a third of the variation to be at NUTS 1 level. This is much higher than the proportion at this level for other outcome variables, e.g. 2% for mean hours, less than 1% for unemployment and less than 1% for employment. This suggests that the earnings of people are very much influenced by pay levels for people in the surrounding NUTS 1 area. This is a logical finding as some residents of each local authority commute to work in other parts of their NUTS 1 area. The NUTS 1 area effect on earnings appears to be bigger than the NUTS 1 areas effects on hours worked which are themselves bigger than the NUTS 1 areas effects on unemployment rates and particularly

on employment rates. This suggests that income levels are affected by the NUTS 1 area that people live in, whilst the chances of being employed depend far more on factors at a local level (including individual level). There are also fairly large VPC values for residents' earnings for the NUTS 2 areas level and for the NUTS 3 areas level. Together the VPC values indicate that there are effects on residents' earnings that operate at NUTS 1, NUTS 2 and NUTS 3 scales. This implies that researchers should try including all three of these levels in their models of residents' earnings.

5.3.9 Median earnings for local authorities by workplace

In addition to statistics on the earnings of people by place of residence the ONS produces statistics on the earnings of people by place of work. In local authorities with large proportions of people commuting out to work, or large proportions travelling in to work, or both, there can be substantial differences between the incomes of people living in a local authority and the incomes of people working in the local authority. Comparison of separate models of the earnings of those living in local authorities and the earnings of those working in local authorities may help to explain the differences in income levels for residents and for workers.

A null, four-level model of the median earnings for local authorities by workplace found the following proportions of variance at the four different geographic scales.

Local Authority and Lower level VPC = 41.81%

NUTS 3 areas level = 17.03%

NUTS 2 areas level = 16.85%

NUTS 1 areas level = 24.31%

The proportion of variance in median earnings by workplace is by far larger at local authority and lower levels than at any other geographic level. This indicates, as might be expected, that local authority level effects have a greater effect on the earnings of those working in each local authority than they do on the median earnings of those who live in each local authority some of whom may work in the local authority but others of whom may work elsewhere. This suggests that changes within a local authority (e.g. new jobs or different jobs) may affect the earnings of those working in the local authority more strongly than they effect the earnings of the residents of the local authority. The proportion at NUTS1 areas level is also large, but not as large as in the model of median earnings of residents in each local authority.

As people often live in one local authority and work in another local authority workers are at the same time nested in both their home local authority and in their workplace local authority, rather than in a strict hierarchy of local authorities (as would be the case if each person worked in the local authority that they lived in). If data for individual workers were available with geographic indicators to show both their home local authority and their workplace local authority, then cross-classified multilevel models of individual incomes could be built that simultaneously take account of where people lived and where they worked. However, as the data used for this project were aggregate local authority level data there was not enough information available to build cross-classified models. Instead, separate multilevel models were built for the two separate variables of median income for people living in each local authority and median income for people working in each local authority.

5.3.10 Summary of results from variance partition coefficient values for workplace earnings

The VPC calculations for the median earnings for people who work, as opposed to reside, in each local authority showed about a quarter of the variation in earnings to be at NUTS 1 areas level. This was slightly lower than the same proportion for residents' earnings but still much higher than for all of the other outcome variables. This confirmed the finding for residents' earnings that there is a large NUTS 1 (English regional) effect on earnings.

5.3.11 Local authority differences between median residential earnings and median workplace earnings

Clearly the difference between the median earnings of residents in a local authority and the median earnings of those whose workplace is within the local authority depends on the relative numbers of people commuting in and out of the local authority and on whether both sets of commuters tend to earn more than they would in the local authority that they live in. A null, four level model of the excess of residential earnings over workplace earnings produced the following proportions of variance.

Local Authority and Lower level VPC = 92.75%

NUTS 3 areas level = 1.05%

NUTS 2 areas level = 3.60%

NUTS 1 areas level = 2.60%

That over 90% of the variation is at local authority and lower levels shows that differences in the earnings of residents and workers are mainly due to local authority and lower level effects (including individual level effects). This suggests that models of this statistic should concentrate on including predictors measured at local authority level rather than at higher geographic levels. It is interesting that only 1% of the variation is found at NUTS 3 areas level. This implies that the NUTS 3 area around each local authority has little effect on whether the earnings of residents or workers are higher, possibly this might indicate that those commuting for higher pay do so to areas outside their immediate NUTS 3 area. Researching this is outwith the scope of this project.

5.3.12 Job density – excluding outliers

A null, four level model of job density values for local authorities excluding the three outliers produced the following variance partition coefficients.

Local Authority and Lower level VPC = 82.29%

NUTS 3 areas level = 3.91%

NUTS 2 areas level = 13.80%

NUTS 1 areas level = 0.00%

5.3.13 Summary of results from variance partition coefficient values for job density outcome variable

The VPC calculations for the set of local authorities with the outliers removed showed that a negligible proportion of variation was at NUTS 1 areas level. They also implied that NUTS 2 areas level effects had a greater influence than NUTS 3 areas effects.

5.4 Predictors at different scales

The next stage of the main modelling section of the project consisted of building multilevel models of the proposed local authority level outcome variables using predictor variables at different geographic scales. These models sought to show which level predictor variables appeared to be better suited to modelling the local authority level outcome variables. The criterion used to compare the different models was the AIC which gives a measure of model fit moderated by the number of parameters to be estimated by each model. A discussion on the choice of AIC or BIC (Bayesian Information Criteria) is given by Fabozzi et

al (2014) tends to favour the AIC. Tables were produced to show the AIC values for a number of outcome variables modelled by predictor variables at different scales.

5.4.1 Tables of AIC values for models with predictors at different geographic values

In order to determine which geographic levels to use for predictor variables in models of labour market statistics a number of multilevel models were built of the proportions of people aged 16 to 64 who were employed and of the proportions of economically active people aged 16 or over who were unemployed using predictor variables at different geographic scales.

Tables 7 - 11 provide a single model fit statistic, namely the AIC value for a large number of random intercept multilevel models of the proportion of 16 to 64 year olds who were employed. Each column in each table relates to a different predictor variable measured (or calculated) at different geographic levels. The AIC value for the null, four level (local authority (LA), NUTS 3 areas, NUTS 2 areas, and NUTS 1 areas) is repeated at the top of each column. This gives a base level model fit statistic to compare the other AIC values in each column against.

5.4.2 Employment rate

Table 7: AIC values for Four level models of the proportion of 16-64 year olds who are employed

Geographic scale for predictor Variable	Predictor = Proportion of people with an **NVQ level four or higher** qualification	Predictor = Proportion of people with **no qualifications**	Predictor = Proportion of people with **Bad or Very Bad Health – published data**	Predictor = Proportion of people with **Bad or Very Bad Health – Calculated from 2011 LA data**
Null model for reference	6,395	6,395	**6,395**	**6,395**
Local Authority level predictor	6,674	*6,121 boundary (singular) fit*	*6,302 boundary (singular) fit*	*6,302 boundary (singular) fit*
NUTS 3 areas level predictor	6,380	*6,338 boundary (singular) fit*	2011 Health data was not published at NUTS 3 level	*6,374 boundary (singular) fit*
NUTS 2 areas level predictor	6,393	*6,370 boundary (singular) fit*	2011 Health data was not published at NUTS 2 level	*6,390 boundary (singular) fit*
NUTS 1 areas level predictor	6,396	**6,391**	*2014 NUTS 1 data 6,388 boundary (singular) fit*	*6,388 boundary (singular) fit*
Predictor variables **at four levels** (LA, NUTS 3,2,1 areas)	**6,377**	*6,120 boundary (singular) fit*	*2011 LA data and 2014 NUTS 1 areas data only 6,297 boundary (singular)*	*Error: couldn't evaluate grouping factor LAD13CD:(LAD13CD:(NUTS315CD:(NUTS215CD:NUT115CD))) within model frame: try adding grouping factor to data frame explicitly if possible*

Note: In this and following tables AIC values that are particularly low indicating a better fit than comparable models are shown in bold for models that converged without producing a singular fit as these models stand out as being particularly suitable to use.

Table 8: AIC values for Four level models of the proportion of 16-64 year olds who are employed, continued

Geographic scale for predictor Variable	Predictor = **Part-time**	Predictor = **Median Age rescaled to Age/100**	Predictor = **Gender balance (proportion female)**	Predictor = **Industrial Diversity** Herfindahl Index (18 groups)
Null model for reference	6,395	**6,395**	**6,395**	6,395
Local Authority level predictor	6,396	6,371 Note that data had to be rescaled	6,393 *Model failed to converge*	6,390
NUTS 3 areas level predictor	6,397	Data not available	6,392 *Model failed to converge*	**6,377**
NUTS 2 areas level predictor	6,393	Data not available	6,397 *Model failed to converge*	6,388
NUTS 1 areas level predictor	*6,395*	*6,395 Model failed to converge*	*6,396 Model failed to converge*	*6,386 boundary (singular) fit*
Predictor variables **at four levels** (LA, NUTS 3,2,1 areas)	6,397	LA and NUTS 1 predictors 6,373	*6,396 Model failed to converge*	*6,377 Model nearly unidentifiable, large eigenvalue ratio*

Table 9: AIC values for Four level models of the proportion of 16-64 year olds who are employed, continued

Geographic scale for predictor Variable	Predictor = Occupational Diversity Herfindahl Index (9 groups)	Predictor = Managerial/professional rate	Predictor = Process/plant machine/elementary workers rate
Null model for reference	6,395	6,395	6,395
Local Authority level predictor	**6,393**	**6,388**	6,382
NUTS 3 areas level predictor	**6,392**	6,393	**6,372**
NUTS 2 areas level predictor	6,396	6,393	6,393
NUTS 1 areas level predictor	6,396	6,396	6,395
Predictor variables **at four levels** (LA, NUTS 3,2,1 areas)	6,396	6,386	6,377

Table 10: AIC values for four level models of the proportion of 16-64 year olds who are employed, continued

Geographic scale for predictor Variable	Predictor = **Commuting Rate, proportion of workers traveling 30 km or more**	Predictor = Proportion of people living in **rural areas** including hub towns	Predictor = **Housing Tenure, Social Rented**	Predictor = **Country of birth outside the UK**
Null model for reference	6,395	6,395	6,395	6,395
Local Authority level predictor	6,383	6,380	**6,360**	**6,361**
NUTS 3 areas level predictor	**6,376**	**6,377**	6,367	6,385
NUTS 2 areas level predictor	6,388	6,387	6,385	6,396
NUTS 1 areas level predictor	*6,390*	6,393	6,394	6,397
Predictor variables **at four levels** (LA, NUTS 3,2,1 areas)	6,381	*6,380 Model failed to converge*	*6,364*	**6,361**

Geographic scale for predictor Variable	Predictor = **2011 Census Areas Super Group**	Predictor = **2011 Census Areas Sub Group**	Predictor = **IMD 2015 Average Rank** for LSOAs in each LA	Predictor = **Employment Deprivation 2015 Average Rank** for LSOAs in each LA
Null model for reference	6,395	6,395	6,395	6,395
Local Authority level predictor	6,359	6,342	6,284	6,290

Note: The Census Area type data used for this research classify local authorities by Census Area Super Group and by Census Area Sub Group. Census Area types are not available for larger geographic areas. Due to the categorical nature of the data and the fact that the classifications were created to describe small areas it is probably not practical nor helpful to try to create 'average type' data for larger geographical areas.

Although it would be numerically straightforward to calculate average IMD 2015 and average Employment Deprivation 2015 ranks for larger geographical areas this is not something that is usually done by other researchers and has not been done for in this research. It would not be in keeping with the purpose of areal deprivation indicators in general which aim to identify deprived areas at as local a level as possible in order to allocate resources or inform policy.

5.4.3 Summary of results for predictors of employment rates

Comparing the AIC values of four-level models of employment with predictor variables measured at different geographic scales found that for about half of the predictors the models with predictors just at local authority level fitted the data better than models with predictors at different geographic levels. For the other predictors, models with predictors at all levels fitted the data about as well or slightly better than models with predictors just at local authority level. This finding suggests that researchers modelling employment at local authority level may reasonably safely build models that use just local authority level data to get models that fit the data as well as might be expected by models with predictors at a number of different levels. This was an important finding as it suggested that it was not generally necessary to obtain or calculate predictor variables at different scales. This

finding could spare future modellers of employment rates from spending valuable resources trying to obtain predictor variable data for larger geographic areas.   This is in line with a similar finding for models of unemployment given below.

5.4.4 Unemployment rate

Table 12: AIC values for Four level models of unemployed people aged 16 or over as a proportion of all economically active people aged 16 or over (i.e. unemployed + employed)

Geographic scale for predictor Variable	Predictor = Proportion of people with an **NVQ level four or higher** qualification	Predictor = Proportion of people with **no qualifications**	Predictor = Proportion of people with **Bad or Very Bad Health – published data**	Predictor = Proportion of people with **Bad or Very Bad Health – Calculated from 2011 LA data**
Null model for reference	5,097	5,097	5,097	5,097
Local Authority level predictor	**5,042**	4,907	4,961	4,961
NUTS 3 areas level predictor	5,078	5,046 *boundary (singular) fit*	Predictor not published for NUTS 3 areas	5,062
NUTS 2 areas level predictor	5,098	5,074 *boundary (singular) fit*	Predictor not published for NUTS 2 areas	5,095 *boundary (singular) fit*
NUTS 1 areas level predictor	5,099	5,096	5,097 *boundary (singular) fit*	5,087 *boundary (singular) fit* Lower than Null model
Predictor variables **at four levels** (LA, NUTS 3,2,1 areas)	5,042	4,895	LA and NUTS 1 only, *boundary (singular) fit* 4,948	*Error: couldn't evaluate grouping factor LAD 13CD:(LAD13CD:(N UTS315CD: (NUTS215CD:NUT1 15CD))) within model frame: try adding grouping factor to data frame explicitly if possible*

Geographic scale for predictor Variable	Predictor = **Part-time**	Predictor = **Median Age**	Predictor = **Gender balance** (proportion female)	Predictor = **Industrial Diversity** Herfindahl Index (18 groups)
Null model for reference	5,097	5,097	5,097	5,097
Local Authority level predictor	**5,092**	**5,002**	5,078	5,067
NUTS 3 areas level predictor	5,094	data not available	5,086	5,068
NUTS 2 areas level predictor	5,086	data not available	5,098	5,092
NUTS 1 areas level predictor	5,094	5,094 *Model failed to converge*	5,098 *Model is nearly unidentifiable*	5,087
Predictor variables **at four levels** (LA, NUTS 3,2,1 areas)	5,088	LA and NUTS 1 only model, 5,003	5,081 *Model is nearly unidentifiable*	5,049 *Model is nearly unidentifiable*

Part-time models show the proportion of people working part-time in wider geographic areas has an effect on, or is correlated with, unemployment rates for economically active people aged 16 or over. Comparing with the same predictor used to model the employment rate for people aged 16 to 64 shows the effect of part-time working variable has an effect at further distances on economically active people aged 16 or over than on all people aged 16 to 64.

Industrial Diversity models show there appear to be effects at all four separate areas levels as a predictor variable at each level produces a model that fits the data better than the null model. Best model (in terms of lowest AIC value) is one with Industrial Diversity predictors at all four geographic area levels.

Table 14: AIC values for Four level models of unemployed people aged 16 or over as a proportion of all economically active people aged 16 or over (i.e. unemployed + employed), continued

Geographic scale for predictor Variable	Predictor = **Occupational Diversity** Herfindahl Index (9 groups)	Predictor = **Managerial/professional rate**	Predictor = **Process/plant machine/elementary workers rate**
Null model for reference	5,097	5,097	5,097
Local Authority level predictor	**5,083**	**5,047**	**5,035**
NUTS 3 areas level predictor	5,096	5,075	5,057
NUTS 2 areas level predictor	5,099	5,098	5,098
NUTS 1 areas level predictor	5,099	5,099	5,099
Predictor variables **at four levels** (LA, NUTS 3,2,1 areas)	5,085	**5,047**	**5,028**

Table 15: AIC values for four-level models of unemployed people aged 16 or over as a proportion of all economically active people aged 16 or over (i.e. unemployed + employed), continued

Geographic scale for predictor Variable	Predictor = **Commuting Rate,** proportion of workers traveling 30 km or more	Predictor = Proportion of people living in **rural areas** including hub towns	Predictor = **Housing Tenure, Social Rented**	Predictor = **Country of birth outside the UK**
Null model for reference	5,097	5,097	5,097	5,097
Local Authority level predictor	**5,067**	**5,017**	**4,998**	**5,046**
NUTS 3 areas level predictor	5,074	5,053	5,040	5,072
NUTS 2 areas level predictor	5,092	5,080	5,080	5,094
NUTS 1 areas level predictor	5,096	5,094	5,094	5,098
Predictor variables **at four levels** (LA, NUTS 3,2,1 areas)	**5,066**	**5,012**	**4,998**	5,049

Table 16: AIC values for four-level models of unemployed people aged 16 or over as a proportion of all economically active people aged 16 or over (i.e. unemployed + employed), continued

Geographic scale for predictor Variable	Predictor = **2011 Census Areas Super Group**	Predictor = **2011 Census Areas Sub Group**	Predictor = **IMD 2015 Average Rank** for LSOAs in each LA	Predictor = **Employment Deprivation 2015 Average Rank** for LSOAs in each LA
Null model for reference	5,097	5,097	5,097	5,097
Local Authority level predictor	4,974	4,934	4,836	4,862

Note: An explanation of why this table does not include predictors for areas larger than local authorities is given at the foot of table 11 above.

5.4.5 Summary of results for predictors of unemployment rates

Four-level random intercept models of local authority level unemployment built using predictor variables at different geographic scales showed that, for most predictor variables the models which included predictors at different scales did not fit the data much better, if at all, than four-level models which only used a local authority level predictor variable. This is in line with the findings for predictors for employment described above. Together these findings suggest that for models of other labour market and related statistics it is sufficient for researchers to focus on building multilevel models with predictors at local authority level only rather than to expend resources on obtaining or calculating predictor variables at a number of different geographic scales. The following phases of the work therefore just used local authority level predictor variables.

5.5 VPC values for random intercept models

The results for four-level random intercept models, shown here in tables 17 to 22 below, expand upon the results for four-level null models included in section 5.3 above. These additional results make it possible to see if the geographic scales that were found to have relatively large proportions of variance when included in four-level null models have similarly large proportions of variance when included in four-level random intercept models. This makes it possible to see if the geographic scales which might be chosen by researchers based on evidence from VPC values for null models would be different if they were to use evidence provided by VPC values for random intercept models. VPC values are

not generally calculated for random coefficient values as the variable coefficients mean that such models do not have unique proportions of variance at each level (Kreft and de Leeuw, 1998, page 63, and Nalborczyk, 2017).

All the models for this section of the work were built using R. All the models in this section are four-level models with level 1 using local authority data, level 2 being NUTS 3 areas, level 3 being NUTS 2 areas and level 4 being NUTS 1 areas level. The tables below each include a null model and a number of random intercept models.

For each model the total variance is shown. For the null models this is the total unconditional variance, i.e. the total variance when no attempt has been made to use predictor variables to model or 'explain' the variance in the outcome variables. The variance for each of the random intercept models with a predictor variable is different from the unconditional variance in that the variance is the variance that exists in the model after the predictor variable has been included in the model. One would expect the conditional models, the ones with the predictor variables, to be better fitting models than the null model for each outcome variable as they use more information to model the outcome variables, i.e. the conditional variances would be expected to be smaller than the unconditional variance. The total variances shown in the tables below show that this is generally, but not always, the case. Predictor variables that lead to models with a much lower total variance than the unconditional variance of the null model suggest themselves to be worthwhile predictor variables for the outcome. Those predictor variables that lead to similar or higher variances suggest themselves not to be very helpful to include in models of the outcome variable.

All the outcome variables are at local authority level. As the results that were included in section 5.4 above showed that models with local authority level predictor variables generally fitted the data better than models with predictor variables at different levels, the predictor variables in all the models in section 5.5 below are local authority level variables.

5.5.1 Residents' earnings

Table 17: VPC values and total variances for null and random intercept models of Residents' Earnings

	Null	NVQ4+	No Qual-ifications	Poor Health	Part-time	Age
LA/lower	33%	**52%**	37%	35%	35%	30%
NUTS 3	13%	11%	10%	10%	13%	13%
NUTS 2	13%	12%	11%	**22%**	13%	13%
NUTS 1	40%	**25%**	41%	**33%**	39%	**45%**
Total variance	7,946	3,817	5,931	5,212	7,602	8,486

	Female	30Km+	Industrial Diversity	Occupat-ional Diversity	Managers /profess-ionals	Process /Plant
LA/lower	33%	31%	34%	**51%**	44%	40%
NUTS 3	13%	12%	14%	11%	10%	10%
NUTS 2	13%	13%	13%	11%	13%	12%
NUTS 1	40%	**45%**	40%	**28%**	33%	38%
Total variance	7,992	8,287	7,908	3,765	4,248	5,034

	Rural prop-ortion	Social Renting	Born outside the UK	IMD 2015	Employ-ment Depriva-tion
LA/lower	32%	27%	28%	**21%**	29%
NUTS 3	13%	12%	12%	**6%**	**7%**
NUTS 2	13%	**17%**	10%	**21%**	**20%**
NUTS 1	42%	45%	**50%**	**52%**	44%
Total variance	8,205	8,675	8,612	6,523	5,331

For the null model and for the random intercept models of residents' earnings for most of the predictor variables, there is a much higher proportion of variance at NUTS 1 areas level than at NUTS 2 areas level or at NUTS 1 areas level and also a higher proportion of variance at NUTS 1 areas level than at local authority/lower areas level. This suggests that NUTS 1 areas should be included in all models of residents' earnings. Literature discussed in section 2.3.8 of the literature review, e.g. Tranmer and Steel (2001a, 2001b), Opdenakker and Van Damme (2000), Van den Noortgate et al (2005) and Moerbeek (2004), suggests

that missing a higher level from a model would cause the variance due to the missing higher level to be transferred to the highest level that is included in the model. Therefore, the particularly high proportions of variance at NUTS 1 areas level also suggest that there could potentially be a missing higher level, above NUTS 1 areas level, that it would be helpful to include in models of residents' earnings. This possibility is investigated in section 4.2.8 above.

For the four random intercept models of residents' earnings that use the poor health, social renting, IMD 2015 and employment deprivation predictor variables there is a higher proportion of variance at NUTS 2 areas level than there is for the null model or for the other random intercept models. This suggests that researchers should include NUTS 2 areas as a geographic scale in multilevel models of residents' earnings that include these predictor variables and, as these predictors are all associated with deprivation, possibly in multilevel models of earnings that include other predictors associated with deprivation.

The two predictor variables for which the overall patterns in the proportions of variance are different are the proportion of people with qualifications at NVQ level 4 or higher (equivalent to having a degree or higher level qualification) and the occupational diversity index predictor variable. Models with each of predictors had a much higher proportion of variance at local authority or lower level and a lower proportion of variance at NUTS 1 areas level when compared with the null model or the majority of the other random intercept models. For these two models the proportion of variance being very high at local authority level or lower level may be due to strong effects on earnings that come from an individual's qualifications rather than due to local authority area level effects. The lower proportions of variance being at NUTS 1 areas level may mean that including the proportion of highly qualified people in models helps fit the model more closely to the data in a way that specifically reduces the amount of residual variance at NUTS 1 areas level.

Although the NUTS 1 areas level has a higher proportion of variance than the NUTS 2 areas level and the NUTS 3 areas level, there is also an appreciable proportion of variance at these two geographic scales. The proportions at NUTS 2 areas level and NUTS 3 areas levels are very similar to each other in most models. That there tends to be appreciable amount of variance at each of the NUTS geographic scales included in these models suggests that researchers should try including all three of these geographic scales in their models of residents' earnings.

5.5.2 Workplace earnings

Table 18: VPC values and total variances for null and random intercept models of Workplace Earnings

	Null	NVQ4+	No Qual-ifications	Poor Health	Part-time	Age
LA/lower	40%	59%	44%	44%	47%	44%
NUTS 3	16%	16%	15%	15%	14%	17%
NUTS 2	16%	11%	14%	20%	15%	17%
NUTS 1	27%	14%	27%	21%	24%	22%
Total variance	6,469	4,330	5,312	5,566	5,409	5,810

	Female	30Km+	Industrial Diversity	Occupat-ional Diversity	Managers /profess-ionals	Process/P lant
LA/lower	44%	14%	41%	**61%**	48%	46%
NUTS 3	15%	17%	18%	14%	15%	15%
NUTS 2	15%	16%	14%	10%	14%	16%
NUTS 1	27%	26%	27%	15%	22%	24%
Total variance	5,709	6,373	5,861	4,053	5,424	5,717

	Rural prop-ortion	Social Renting	Born outside the UK	IMD 2015	Employ-ment Depriva-tion
LA/lower	41%	43%	**56%**	40%	43%
NUTS 3	17%	18%	20%	14%	14%
NUTS 2	17%	14%	11%	18%	18%
NUTS 1	24%	26%	**13%**	28%	25%
Total variance	6,190	6,053	4,271	6,391	5,848

The main pattern in the VPC values for models of the earnings of people who work in each local authority, shown in table 18, is that there is an appreciable proportion of variance at all of NUTS 1 areas level, NUTS 2 areas level and NUTS 3 areas level. This suggests that those researching earnings by local authority of workplace should include all of these three NUTS geographic scales in their models.

Comparing the VPC values in table 18 for models of earnings by workplace with those in table 17 for models of earnings by place of residence the most striking difference is that the proportions of variance at NUTS 1 areas level are lower for the models of workplace earnings than for the models of residents' earnings. It is also noticeable that the

proportions of variance at local authority and lower level are higher for models of workplace earnings than for models of residents' earning. This suggests that effects at local authority or lower level have greater association with earnings at workplaces within each local authority than they do with the earnings of those who live in each local authority.

5.5.3 Mean hours worked

Table 19: VPC values and total variances for null and random intercept models of Mean Hours Worked

	Null	NVQ4+ *'sing. fit'*	No Qual-ifications	Poor Health	Part-time	Age
LA/lower	87%	87%	89%	87%	90%	88%
NUTS 3	0%	0%	0%	0%	0%	0%
NUTS 2	10%	11%	6%	10%	6%	10%
NUTS 1	3%	2%	5%	4%	3%	2%
Total variance	2.362	2.340	2.278	2.378	2.103	2.343

	Female	30Km+	Industrial Diversity	Occupat-ional Diversity	Managers /profess-ionals	Process /Plant
LA/lower	90%	87%	88%	85%	84%	87%
NUTS 3	0%	1%	0%	1%	1%	0%
NUTS 2	7%	10%	9%	11%	11%	11%
NUTS 1	3%	2%	3%	3%	3%	2%
Total variance	2.307	2.337	2.342	2.381	2.340	2.357

	Rural prop-ortion	Social Renting	Born outside the UK	IMD 2015	Employ-ment Depriva-tion
LA/lower	87%	87%	87%	87%	87%
NUTS 3	0%	0%	0%	0%	0%
NUTS 2	9%	11%	5%	10%	10%
NUTS 1	4%	0%	**8%**	2%	4%
Total variance	2.374	2.247	2.293	2.358	2.377

Table 19 above shows the proportions of variance in null and random intercept four-level models of local authority level mean hours worked. Compared with the proportions of variance at all NUTS geographic scales in models of earnings outcomes, shown in tables 17

and 18 above, the proportions of variance at each of the NUTS geographic levels in the models of mean hours worked are lower. In particular the proportions of variance at NUTS 3 areas level are negligible for the null model and for most of the random intercept models of mean hours worked. This suggests that it would not be of help to researchers to include NUTS 3 area levels in their models of hours worked. In the models of mean hours worked the proportions of variance at NUTS 2 areas level are consistently higher than the proportions of variance at NUTS 1 areas level. This implies that it is more important for researchers to include the NUTS 2 areas geographic scale in their models of hours worked than to include the NUTS 1 areas level. It shows that an important amount of detail may be missed if researchers rely on using NUTS 1 areas geographic indicators rather than including NUTS 2 areas in their models. The model with the proportion of people with qualifications at NVQ level four or higher produced a singular fit (see section 3.3.3.3 above).

5.5.4 Job density

Table 20: VPC values and total variances for null and random intercept models of Job Density (excluding outliers)

	Null	NVQ4+	No Qual-ifications	Poor Health	Part-time	Age
LA/lower	82%	87%	87%	91%	81%	76%
NUTS 3	4%	2%	2%	0%	3%	4%
NUTS 2	14%	11%	11%	9%	16%	20%
NUTS 1	0%	**9%**	0%	0%	0%	0%
Total variance	0.0365	0.0323	0.0353	0.0321	0.0366	0.0371

	Female	30Km+	Industrial Diversity	Occupat-ional Diversity	Managers /profess-ionals	Process /Plant
LA/lower	82%	83%	76%	86%	88%	87%
NUTS 3	4%	4%	9%	2%	2%	1%
NUTS 2	14%	14%	15%	10%	10%	12%
NUTS 1	0%	0%	0%	**2%**	0%	0%
Total variance	0.0356	0.0365	0.0357	0.0335	0.0337	0.0353

	Rural prop-ortion	Social Renting	Born outside the UK	IMD 2015	Employ-ment Depriva-tion
LA/lower	81%	74%	75%	89%	93%
NUTS 3	4%	4%	3%	1%	0%
NUTS 2	14%	23%	13%	10%	7%
NUTS 1	0%	0%	**8%**	0%	0%
Total variance	0.0367	0.0378	0.0377	0.0346	0.0317

Table 20 shows the proportions of variation at different geographic scales for four-level null and random intercept models of job density. The null model and most of the random intercept models show negligible amounts of variance at NUTS 1 areas level. This suggests that it would not normally be helpful to include NUTS 1 areas as a level in multilevel models of local authority level job density. The exceptions are those random intercept models of job density that include the proportion of people with an NVQ level 4 or higher qualification, the proportion of people born outside the UK or, to a lesser extent, the local authority level occupation diversity index as their predictor variable. These exceptions

show that the relationships between these predictor variables and job density can be different in different NUTS 1 areas.

For the null model and for all of the random intercept models, the proportion of variance at NUTS 2 areas level is a number of times higher than the proportion of variance at NUTS 3 areas level. This implies that regardless of which predictor variables are used in random intercept models it is more important for researchers to include the NUTS 2 areas level in multilevel models of job density than to include the NUTS 3 areas level. This is a useful finding as NUTS 2 areas data may be more readily available than NUTS 3 areas data and grouping by NUTS 2 areas will lead to less complex models with fewer parameters to be estimated than models with grouping by NUTS 3 areas.

5.5.5 Employment rates

Tables 21 and 22 show VPC values for models of employment rates and unemployment rates both of which can be assumed to have binomial distributions. For binomial outcomes it is appropriate to use generalised linear models rather than linear models. Generalised linear modes were built in R using the logit link function and the additive approach to account for over-dispersion of binomial variables. For generalised multilevel linear models, the proportion of variance at the lowest level in the models cannot be obtained from the models. As noted earlier, e.g. in sections 2.3.1 and 3.3.3.1, in order to calculate the proportion of variance at each level in such a model an estimate has to be used for the amount of variation at the lowest level. For models using the logit link function this can be taken to be $\pi^2/3$ (Hox, 2002, Snijders and Bosker, 2012, Sommet and Morseli, 2017, and Wu et al, 2012) and this value was used in the calculations of proportions of variance shown in tables 21 and 22. However, the tables do not include the proportion of variance due to this lower level which is why the percentages in the table below do not add up to 100%. The fact that all the values in table 21 and 22 are much lower in general than those in tables 17 to 20 above is due to the use of generalised linear modelling and the logit link function and there being very strong individual level characteristics that affect individuals' employment statuses that significantly outweigh any area level characteristics.

Table 21: VPC values and total variances for null and random intercept models of Employment Rates

	Null	**NVQ4+**	**No Qual-ifications** *'sing. fit'*	**Poor Health** *'sing. fit'*	**Part-time**	**Age**
LA/lower	2.1%	2.0%	1.8%	1.8%	2.1%	2.0%
NUTS 3	0.2%	0.1%	0.1%	0.0%	0.2%	0.1%
NUTS 2	0.2%	0.2%	0.0%	0.0%	0.2%	0.2%
NUTS 1	0.3%	0.4%	0.2%	0.1%	0.3%	0.2%
Total variance	3.383	3.379	3.358	3.354	3.384	3.375

	Female	**30Km+**	**Industrial Diversity**	**Occupat-ional Diversity**	**Managers /profess-ionals**	**Process /Plant**
LA/lower	2.1%	2.1%	2.1%	2.1%	2.1%	2.1%
NUTS 3	0.2%	0.1%	0.1%	1.1%	0.1%	0.1%
NUTS 2	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%
NUTS 1	0.3%	0.1%	0.3%	0.3%	0.3%	0.3%
Total variance	3.382	3.374	3.381	3.382	3.380	3.378

	Rural prop-ortion	**Social Renting**	**Born outside the UK**	**IMD 2015**	**Employ-ment Depriva-tion** *'sing. fit'*
LA/lower	2.1%	2.0%	1.9%	2.1%	1.7%
NUTS 3	0.2%	0.1%	0.1%	0.2%	0.0%
NUTS 2	0.2%	0.2%	0.2%	0.2%	0.0%
NUTS 1	0.3%	0.2%	0.5%	0.3%	0.1%
Total variance	3.375	3.367	3.385	3.348	3.350

Table 21 above shows that for the null four-level model of employment rates and for the majority of the random intercept models with different predictor variables, there are similar, small amounts of variance at all of the NUTS areas levels. This suggests that it is important for researchers to include all the NUTS areas geographic scales in their models of employment. Although the differences are small the proportions of variance at NUTS 1 areas levels are often slightly higher than the proportions of variance at NUTS 2 areas level and at NUTS 3 areas level. This suggests that it is most important for models of local authority employment rates to include NUTS 1 areas level geographic scales. This is particularly true for modes of employment rates that use the proportion of people born outside the UK as a predictor variable. For three of the predictor variables, the proportion

of people with no qualifications, the proportion of people with poor health and the employment deprivation variable, the models generated a warning from R that they had a singular fit because there was no variance at one of the geographic levels (see section 3.3.3.3 above for information about singular fitting models). As noted above the percentages in table 21 were generated by using the additive method of accounting for binomial overdispersion in which a pseudo level is added to the model at the same level as the lowest level used in the model. For local authority level outcomes that means an additional local authority level is added to the model. If this extra level is omitted from the model, i.e. no account is taken of overdispersion then the proportions of variance at NUTS 3 areas level are higher than those shown in the above table, generally around 1%. This shows that models that do not allow for overdispersion have a greater proportion of variance at NUTS 3 areas level than models that use the additive approach to overdispersion and would therefore indicate quite strongly that it was more important to include NUTS areas 3 level in models of employment than to include NUTS areas 2 level or NUTS areas 1 level. This also shows that missing out the pseudo level from models of employment rates transfers some of the variance that would be shown at the pseudo level to the NUTS 3 areas level suggesting the NUTS 3 areas level to be more important in modelling than the other levels if no account is taken of overdispersion.

5.5.6 Unemployment rates

Table 22: VPC values and total variances for null and random intercept models of Unemployment Rates

	Null	NVQ4+	No Qual-ifications	Poor Health	Part-time	Age
LA/lower	1.0%	0.8%	0.9%	0.7%	1.0%	0.8%
NUTS 3	0.7%	0.5%	0.4%	0.2%	0.7%	0.5%
NUTS 2	0.3%	0.4%	0.1%	0.3%	0.3%	0.1%
NUTS 1	0.8%	1.1%	0.7%	0.8%	0.7%	0.6%
Total variance	3.386	3.389	3.364	3.358	3.381	3.356

	Female	30Km+	Industrial Diversity	Occupat-ional Diversity	Managers /profess-ionals	Process /Plant
LA/lower	1.0%	1.0%	0.9%	1.0%	0.9%	0.9%
NUTS 3	0.6%	0.6%	0.6%	0.6%	0.5%	0.4%
NUTS 2	0.3%	0.3%	0.4%	0.4%	0.4%	0.4%
NUTS 1	0.8%	0.5%	0.7%	0.9%	0.9%	0.8%
Total variance	3.379	3.368	3.378	3.385	3.380	3.398

	Rural prop-ortion	Social Renting	Born outside the UK	IMD 2015	Employ-ment Depriva-tion
LA/lower	0.8%	0.8%	0.9%	0.5%	0.6%
NUTS 3	0.5%	0.4%	0.6%	0.2%	0.1%
NUTS 2	0.2%	0.2%	0.3%	0.2%	0.2%
NUTS 1	0.6%	0.3%	0.7%	0.5%	0.6%
Total variance	3.360	3.347	3.373	3.335	3.340

Table 22 for models of unemployment rates shows that for the null model, and for most of the random intercept models, the proportions of variance at NUTS 3 areas level and at NUTS 1 areas level were approximately twice the proportions of variance at NUTS 2 areas level. This implies that researchers should include NUTS 3 areas level and NUTS 1 areas level in their models of unemployment rates. The low proportions of variance at NUTS 2 areas level suggest that it is less important to include NUTS 2 areas as a level in models of unemployment rates. In common with table 21 above for employment models, the VPC values shown in table 22 for unemployment models were calculated using the additive approach to accounting for the overdispersion of binomial variables. The overall finding

from this table, that NUTS areas 3 and NUTS areas 2 are particularly important to include in models of unemployment, is essentially the same as the finding that came from a similar set of models of unemployment that did not take account of overdispersion. That suggested that it was most important to include NUTS 3 areas level but also important to include NUTS 1 areas level.

5.5.7 Summary of results for VPC values for random intercept models

The VPC values for four-level random intercept models generally show the same patterns as those for four-level null models. The exceptions that arise most frequently are for those models that use the proportion of people with NVQ level 4 or higher qualifications or the proportion of people born outside the UK as their predictor variables.

For both residential and workplace earnings there is appreciable variance at all NUTS areas levels included in the four-level models. This implies that researchers should include grouping by all three of these levels in their models of earnings. Section 2.3.3 of the literature review and Nezlek (2008) explain why multilevel modelling is being used despite the overall amount of the proportion of variance due to grouping being relatively small. Researchers could also consider looking for influences on earnings that operate at each of these geographic scales.

As the proportion of variance is fairly high for the NUTS 1 areas level for both earnings outcomes (residential and workplace) this may indicate a missing higher level. Researchers reading this finding may wish to explore the options for including a higher than NUTS 1 areas level in their models of earnings or incomes. Given the concept of London-weighting whereby public sector and associated salaries are higher for employees working in London than for those working outside London a starting point for a higher level geography may be a London v. 'all other NUTS 1 areas' split. Although the creation of new geographies for modelling and reporting statistics are outside the scope of this work some tentative VPC results for higher level geographies are shown in section 5.5.8 below.

The VPC values for models of mean hours worked show there to be appreciable variance at the NUTS 2 areas level, a small amount of variance at the NUTS 1 areas level and negligible variance at NUTS 3 areas level. Researchers reading this may wish to look for effects and statistics at NUTS 2 areas levels to include in their models of hours worked. The results in this section of the work suggest they should not spend resources on obtaining NUTS 3

areas data and do not need to include NUTS 1 areas grouping or data in models of hours worked. (Section 2.3.3 of the literature review and Nezlek (2008) explain why multilevel modelling is being used despite the overall amount of the proportion of variance due to grouping being relatively small.)

The findings in the section for job density are similar to those for mean hours worked. There is negligible variance at NUTS 1 areas level, little variance at NUTS 3 areas level and most variance at NUTS 2 areas level. This implies that researchers of job density should include NUTS 2 areas grouping in multilevel models or try obtaining and using NUTS 2 areas data.

The VPC values for models of employment and unemployment rates are much lower than for models of normally distributed labour market outcomes. The ratios of variance at NUTS 1, NUTS 2 and NUTS 3 areas levels show that most of the variance is at NUTS 3 areas level and NUTS 1 areas level for both the employment and unemployment outcome variables. Researchers should therefore include NUTS 3 areas level and NUTS 1 areas level in their models of employment and unemployment.

5.5.8 A possible higher level

The VPC values for residents' earnings and for workplace earnings are high for the NUTS1 areas level. This could mean that there is a missing higher level that it would be useful for researchers to include in models of earnings. As earnings are known to be higher in London due to London weighting and possibly due to the London Living Wage (Mayor of London/London Assembly, 2020, and London Data Store, 2020), one possibility is that the missing higher level is a split of NUTS 1 areas into London as one group and all the other NUTS areas as a second group. A small number of multilevel models were built to explore this idea briefly. A five-level null model of residents' earnings with London/not London as the fifth level gave the following proportions of variance at each level.

NUTS 3 areas – 6.9%
NUTS 2 areas – 6.7%
NUTS 1 areas – 5.2%
London/Not London – 64.4%.

The high proportion of variance at the London/Not London level suggests that this might be an important level to include in models of residents' earnings. It has a much higher proportion of variance than the NUTS 1 areas level in this model suggesting it to be more important to include in models than the NUTS 1 areas level.

A similar five-level model for workplace earnings gave the following proportions of variance at each level.

NUTS 3 areas – 9.8%

NUTS 2 areas – 9.0%

NUTS 1 areas – 1.6%%

London/Not London – 56.6%

Again, there was a high proportion of variance in this model at the London/Not London level which supports the idea that incorporating a London/Not London level into models of earnings might be worth exploring.

It is usually recommended that there be far more than one member in any group in a multilevel model. This point is discussed, for example by Robson and Pevalin (2016, page 27). This recommendation would mean that a London/Not London split would not be appropriate to use as a level in multilevel models. However the high proportion of variance potentially at this level suggests that it might be helpful for researchers to incorporate information on which local authorities are in the London NUTS 1 area into their models in either as a higher level or as a dummy variable.   A dummy variable could take the value of one for all local authorities that are in the London NUTS 1 area and zero for all local authorities, or could take the value of one for local authorities that are in the London NUTS 1 area and those that are close to London. Including different local authorities in the group that are regarded as close to London and comparing the fit of the different models that result could provide a different way of finding the geographic area to use take account of higher incomes in local authorities in and around London as opposed to those in the rest of England.

5.6 Summary of results presented in sections 4.1.1 to 5.5.8

There were a number of findings in sections 4.1.1 to 5.5.8 that stood out as being particularly interesting as they were unexpected or engendered further thought or modelling.

The models of OA-level proportions of people who were economically active showed there was a higher proportion of variance at county/unitary authority level than at NUTS 1 areas level, i.e. there was more variation in economic activity rates within NUTS 1 areas than between NUTS 1 areas. These was surprizing as it suggested there was less difference between regions (NUTS 1 areas) of England than might be expected.

The microdata models were interesting in that they confirmed the relative lack of importance of NUTS 1 areas on labour market outcomes by showing that only a very small proportion of variance in the chance of being unemployed was at NUTS 1 areas level (when compared to the variance at individual level).

An outstanding result from the models of mean hours worked per week was that these models showed there was much more variance at NUTS 2 areas level than at either NUTS 1 areas level or NUTS 3 areas level. This suggested that the number of hours, which could be regarded as a proxy for abundance of available work, varies most at a geographic scale in between local, NUTS 3 areas, and regional, NUTS 1 areas. The amount of variation in median earnings was greatest at NUTS 1 areas level. It was so high that it raised a question about whether there was another, higher level that should be added to the model, e.g. the North/South divide or London verses the rest. The effect was more pronounced for residents' earnings than for workplace earnings.

It was surprizing that job density did not vary noticeably at all at NUTS 1 areas level. This suggests that once the outlying job density values for the City of London, Westminster and Camden are removed from the assessment there is much less difference in job density, or availability, between English regions/NUTS 1 areas than might be expected.

Finally, it was surprizing that the multilevel models of local authority employment rates with predictor variables at NUTS 1, NUTS 2, NUTS 3 and local authority level did not generally fit the data better than models with predictor variables solely at local authority level. This was useful information as it suggested that it was not necessary to include predictors at geographic scales higher than local authority if multilevel modelling was used to take account of the hierarchical geographic structure of local authority and NUTS geographic scales.

These findings can be generalised and suggest wider implications of this research than which geographic scales researchers should use in their models of labour market statistics. The first generalisation is that NUTS 1 areas are not very useful for modelling labour market and related socioeconomic statistics. Considering just NUTS 1 areas masks the difference within NUTS 1 areas, for example at NUTS 2 or county/unitary authority level, that are important.

Whilst NUTS 1 scale over-generalises, NUTS 3 areas are also less useful than might have been expected. In contrast, the less well-known NUTS 2 areas are more useful than might have been expected. The generalisation from this finding is that research should be carried out at NUTS 2 areas level or should at least include analysis at NUTS 2 areas level. A wider implication of these findings is that policies developed and implemented at NUTS 2 areas level are likely to be more effective than those at a larger (NUTS 1/region) or finer (NUTS 3 or local authority level). This suggests the importance of the upper tier of local authority administration which consists of counties and unitary authorities and is therefore closer to NUTS 2 areas level than it is to local authority or NUTS 1 areas level.

More central to this research, the results presented in sections 4.1.1 to 5.5.8 above helped to determine the geographic scale to use for outcome variables for the rest of this project; to show how much variation is at each geographic scale; and to decide which geographic scale to use for predictor variables for the rest of this project. These results are summarised belowin order to show why the models presented in the rest of this theses focus on local authority level outcome and predictor variables.

154

5.6.1 Summary of results from Output Areas analyses shown in section 4.1.1

- Linear models of the percentage of people who were economically active regressed on the percentage of people with good or very good health produced different regression lines depending on whether the model was built at Output Areas (OA) level or at NUTS 1 areas level. This gave the first evidence that researchers should consider which geographic scales to use when modelling labour market and related statistics.

- Separate OA-level models for each NUTS 1 area of the percentage of people who were economically active regressed on the percentage of people reporting good or very good health produced different regression lines for each NUTS 1 area. This suggested that multilevel models which included NUTS 1 areas as one of the levels would be beneficial as they would allow separate parameters to be estimated for the different intercepts and coefficients in different NUTS 1 areas (as indicated by the different regression lines produced by the different OA-level models for each NUTS 1 area).

- VPC values for the OA-level null models of the proportions of people who were economically active showed there to be some variation at County/unitary authority level and some variation at NUTS 1 areas. This suggested that NUTS 1 areas level was not the only geographic scale at which there were effects on the OA-level proportions of people who were economically active and that finer geographic scales than NUTS 1 areas might usefully be included in models of labour market and related statistics.

- VPC values for random intercept OA-level models of the proportion of people who were economically active supported the above finding for null models and showed that there was a higher proportion of variance at county/UA level than at NUTS 1 areas level suggesting that local areas (e.g. counties and unitary authorities) had a greater effect on OA outcomes than did NUTS 1 areas.

5.6.2 Summary of results from microdata analyses shown in section 4.1.2

- As the only geographical data in the open access microdata from the 2011 Census were NUTS 1 area codes it was only possible to build two-level models so it was not possible to assess whether finer geographic scales could usefully be included in models of labour market and related socioeconomic statistics.

- A two-level null model built to model the numbers of adult usual residents who were unemployed showed that there was a small proportion of variance in the chance of an individual being unemployed that was at the NUTS 1 area level. This finding was also true for a random intercept model that used individuals' health (very good, good, or fair versus poor or very poor) as the predictor variable for the chance of an individual being unemployed.

- These findings suggested that little information could be gained about which geographic scales to include in models of labour market and other socioeconomic data by modelling the available open access microdata for the 2011 census and also showed that NUTS 1 areas by themselves were not particularly helpful to include in models of labour market statistics.

5.6.3 Summary of findings from VPC values for models of local authority level outcome variables

VPC values for null models

- A logit null model for the local authority level proportions of variance at different geographic levels showed that 97.2% of the variance in the proportions of adults who were unemployed was at individual level, 1.0% was at local authority level, 0.7% was at NUTS 3 areas level, 0.3% was at NUTS 2 areas level and 0.8% was at NUTS 1 areas level. A probit model produced a broadly similar result. Section 2.3.3 of the literature review and Nezlek (2008) note that although the fact that these percentages are fairly low there is still grounds for using multilevel modelling supplied by the hierarchical nature of the areal data. Rather than to determine whether or not to use hierarchical modelling by considering just the proportion of variance that is due to group effects overall this project uses the relative amount of the proportion of variation that is at each of the

levels in the model to indicate which of the NUTS areas level to include in the multilevel model. On that basis, these results suggested that NUTS 3, NUTS 2 and NUTS 1 areas should all be included in multilevel models of local authority level unemployment. The results from this section also showed that missing out a level, specifically NUTS 2 areas level, had an effect on the amount of variance reported as being at NUTS 3 areas which confirms the findings in literature discussed in in section 2.3.8 of the literature review, e.g. Tranmer and Steel (2001a, 2001b), Opdenakker and Van Damme (2000), Van den Noortgate et al (2005) and Moerbeek (2004), that missing out a level from a model transfers the effects to an adjacent level that is included in the model.

- A logit null model of the proportion of people aged 16 to 64 who were employed suggested that 97.2% of variation was at individual level, 2.3% was at local authority level, 0.2% was at NUTS 3 areas level, 0.2% was at NUTS 2 areas level and 0.3% was at NUTS 1 areas level. Results for a comparable probit model were very similar. The similar proportions of variance at NUTS 3 areas level, NUTS 2 areas level and NUTS 1 areas level suggest that there are effects on local authority level employment that operate at each of these three geographic scales meaning that are equally important.

- A four-level null model of mean hours worked per week suggested that the proportion of variation at NUTS 2 areas level was much higher, approximately 10%, than the proportions at NUTS 3 areas level and at NUTS 1 areas level. This suggested that it would be particularly helpful to make use of the NUTS 2 areas geographic scale in models of hours worked per week.

- A four-level null model of the median earnings of local authority residents showed approximately a third of the variation in local authority level earnings to be at local authority level, just over a third to be at NUTS 1 areas level and the remaining almost a third to be split equally between NUTS 3 areas level and NUTS 2 areas level. The amount of variation at NUTS 1 areas level was particularly high and raised questions about whether London salaries might be responsible for the large amount of variation reported as being at NUTS 1 areas level.

- A similar four-level null model of local authority level median earnings by place of work rather than by place of residence showed a larger proportion of the workplace earnings variation to be at local authority level and a smaller proportion to be at NUTS 1 areas level. This suggests that local authority level conditions have a greater effect on the incomes of those working within a particular local authority than on the incomes of those living in that local authority.

- A four-level null model of job density values for local authorities (excluding the three London boroughs with exceptionally high job densities) showed around 82% of the variation in job density to be at local authority level, about 4% to be at NUTS 3 areas level and about 14% to be at NUTS 2 areas level, and none of the variation to be at NUTS 1 areas level (possibly as the three nationally significantly highly dense London boroughs had been removed from the data).

VPC values for random intercept models

- In section 5.5 VPC values were calculated for random intercept models for unemployment rates, employment rates, residents' earnings, workplace earnings, mean hours worked and job density in order to see how the proportions of variance at different geographic scales compared with those for null models of the same outcome variables.

- For most of the random intercept models of unemployment the proportions of variance at NUTS 3 areas level and at NUTS 1 areas level were approximately twice those at NUTS 2 areas level. This was broadly in line with the finding from the comparable null model.

- For most of the random intercept models of employment there were similar, small proportions of variance at each of the NUTS 3 areas, NUTS 2 areas and NUTS 1 areas levels. This finding is in line with that from the comparable null model. Together the findings from the random intercept unemployment model and the random intercept employment model suggest that the NUTS 2 areas level is relatively more important to include in models of employment than in models of unemployment.

- For most of the random intercept models of local authority level mean hours worked per week the proportions of variance at NUTS 3 areas level

were negligible suggesting that NUTS 3 areas had little effect on local authority level hours worked per week Again for most of the random intercept models the proportion of variance at NUTS 2 areas level was appreciably greater than the proportion of variance at NUTS 1 areas level suggesting that NUTS 2 areas had a bigger effect on hours worked at local authority level than did NUTS 1 areas. This suggested that it was important to include NUTS 2 areas in models of hours worked per week. These findings for random intercept models of hours worked per week are in line with those for the null model of hours worked per week.

- For most of the random intercept models of residents' earnings a much higher proportion of the variance was at NUTS 1 areas level than at NUTS 3 areas level or NUTS 2 areas level. This was in line with the finding for the null model of residents' earnings. As the proportions were so high they raised a question about whether the models were in fact missing a higher level, at a coarser geographic scale than NUTS 1 areas, that might be responsible for some of the variance that was reported as being at NUTS 1 areas level.

- For most of the random intercept models of workplace earnings the proportions of variance at NUTS 1 areas level, whilst higher than those at NUTS 2 and NUTS 3 areas levels, were lower than the proportions of variance at NUTS 1 areas level in the random intercept models of residents' earnings. These findings suggest that the NUTS 1 areas geographic scale has a lesser effect on the earnings of people working in each local authority than on the earnings of people living in each local authority. This is broadly in line with the findings for the null models of residents' earnings and of workplace earnings.

- Most of the random intercept models of job density (for local authorities excluding the three outliers) showed there to be a negligible proportion of variation at NUTS 1 areas level. For most of the models the NUTS 2 areas had a larger proportion of variation than the NUTS 3 areas. These finding suggest that it is important to include NUTS 2 geographic scale in models of job density and not at all important to include NUTS 1 areas geographic scale in such models. These findings are in line with those for the null model of local authority job density statistics.

5.6.4 Summary of findings from models using predictors at different geographic scales

- Four-level models were built to model local authority level unemployment and local authority level employment in order to explore the effect on model fit of using predictor variables calculated at different geographic scales in models of local authority level labour market statistics. Each of the models, presented in section 5.4, used either a single predictor variable at one of local authority level, NUTS 3 areas level, NUTS 2 areas level or NUTS 1 areas level or used four predictor variables comprising the same statistic calculated at each of the four geographic levels. The AIC values were extracted from each model and compared to see how the model fit differed for the different models. This was in order to see if any of the single geographic scales for predictor variables stood out as producing models with much better fit than the others and to see whether using multiple geographic scales for the predictor variable produced a model that fitted the data with substantially better fit sufficient to warrant the extra complexity of using more than one geographic scale for the predictor variable.

- The models of local authority level employment with predictors at all four levels did not generally fit the data significantly better than those with predictors at either local authority level or NUTS 3 areas level. This was important as it suggested that it would not generally be necessary for researchers to put resources into obtaining predictors at geographic scales courser than local authority or NUTS 3 if they were not readily available. Also, it would not normally be necessary to build complex models with predictor variables at more than one geographic scale and that simpler, more parsimonious, models could be used.

- In slightly more detail, for models of employment the predictor variables which lead to lower AIC values when the predictor variables were calculated at local authority level were: the proportion of people with an NVQ level 4 or higher qualification; the proportion of people with no qualifications; the proportion of people with bad or very bad health (although this produced a singular fit, see section 3.3.3.3 above); and the percentage of households that were living in socially rented accommodation. These are all variables that are closely related to people

and their own socioeconomic characteristics. Again for employment, the predictor variables which gave clearly lower AIC values when the predictor variables were calculated at NUTS 3 areas level were the Herfindahl index of industrial diversity (calculated using a classification with eighteen industrial sectors); the proportion of workers who were managers or professionals, the proportion of workers who were process, plant machine operators or in elementary jobs; and the proportion of workers commuting 30km or more to work each day. These are all variables related to types of work and jobs. Together these findings suggest that predictor variables closely related to individuals' socioeconomic circumstances should be calculated at local authority level whilst those related to types of jobs and commuting patterns may be better calculated at the slightly courser NUTS 3 areas level (presumably as people often work slightly outside the local authorities that they live in).

- For four-level models of local authority level unemployment, the models that fitted the data best, as determined by having lower AIC values than others, were for most predictor variables those where the predictor variable was calculated at local authority level.

- Exceptions where the models with the predictor variable measured at all four levels gave an appreciably lower AIC value were the models with the following predictor variables: the proportion of people with no qualifications; the Herfindahl index of industrial diversity (although the model gave a 'nearly unidentifiable' warning'); and the model with the proportion of workers in process, plant machinery or elementary jobs.

- Another exception was the percentage of people working part-time predictor variable for which the model with the predictor variable calculated at NUTS 2 areas level gave the lowest AIC value showing that it fitted the data better than those with the predictor calculated at other levels.

- Overall, the findings from the models of unemployment showed that there was generally no advantage in using predictor variables calculated at coarser geographic scales than local authority scale. This is in contrast to the findings from the models of employment which suggested that there was an advantage in calculating those variables related to types of jobs and commuting at NUTS 3 areas level rather than at local authority level. For both outcome variables those predictor variables related more closely to people's individual characteristics produced better fitting models when the predictor variables were calculated at local authority level.

- Although only two outcome variables were modelled for this part of the research they seemed to provide good evidence that generally it was not necessary to use predictor variables calculated at more than one geographic scale in models of key labour market statistics and that it was generally better to use predictor variables calculated at local authority level if they were available (and at NUTS 3 areas level if they were not). As a consequence of this the next stage of the project focussed on models with only local authority level predictor variables.

5.7 Comprehensive set of models of local authority level outcomes using local authority level predictors

The main finding from the investigation into the geographic scale to use for predictor variables in models of local authority level outcomes, summarised in 5.6.4 above, was there was not normally any great advantage in including predictors calculated at more than one geographic scale in models when compared with just using local authority level predictors. The next, most extensive single stage of the modelling part of this project was therefore devoted to building multilevel models of local authority level outcomes that had a number of levels, e.g. local authority level and one or more of NUTS 3 areas level, NUTS 2 areas level or NUTS 1 areas level, but that only had predictors calculated at local authority level. The models were then examined to see which ones fitted the data better by comparing the AIC values for groups of the models that used the same outcome and predictor variables in different ways.

For each combination of outcome variable and predictor variable various random intercept and random coefficient models were built as set out below. These combinations were chosen to compare the fits of more complex four-level models with those for simpler two-level models and also to compare the fits of simpler random intercept models with those of more complex models that had random coefficients at one or more levels.

For outcome variables with a Normal distribution (residential earnings, workplace earnings, mean hours worked and job density) these comprised:

- A four-level random intercept model (local authority, NUTS 3 areas, NUTS 2 areas and NUTS 1 areas);

- Three two-level random intercept models (local authority level and one of NUTS 3 areas, NUTS 2 areas or NUTS 1 areas);

- A four-level random coefficient model (local authority, NUTS 3 areas, NUTS 2 areas and NUTS 1 areas);

- Three two-level (local authority level and one of NUTS 3 areas, or NUTS 2 areas or NUTS 1 areas) random coefficient models;

- Two four-level models with random intercepts at all levels and random coefficients at one of either NUTS 2 areas level or NUTS 1 areas level.

For the employment rate and unemployment rate outcomes (which are assumed to have either binomial or extrabinomial distributions) the models all included a pseudo level at local authority level to account for extrabinomial variation. For these outcomes the models built comprised:

- A four-level model with random intercepts at all levels and no random coefficients;

- Three two-level models (local authority and one of NUTS 3 areas level, NUTS 2 areas level or NUTS 1 areas level) all with random intercepts at both levels and random coefficients at specified levels excluding the pseudo level;

- A four-level model with random intercepts at all levels and random coefficients at NUTS 3, NUTS 2 and NUTS 1 levels but not at pseudo level;

- A four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level only;

- A four level with random intercepts at all levels and random coefficients at NUTS 1 areas level only.

As noted above these models were built in order to compare the model fit that could be achieved by grouping the local authority levels using different geographic scales as the higher levels in the models. For the earnings, hours worked and job density outcomes the tables can be read vertically to compare the AIC values of the four-level models with those for the three two-level models to see whether the four-level model has a better fit (lower AIC value) or whether any of the simpler two-level models has a fit that is either better or almost as good (in which case the simpler two-level model would be the preferred choice). The tables for these outcome variables can also be read horizontally to compare whether the random intercept or random coefficient versions of each model has the best fit (lowest AIC value). If the AIC values are similar then the simpler random intercept model would normally be chosen as it would have fewer parameters to estimate making it computationally easier to run and more likely to converge without generating a singular fit (see section 3.3.3.3). For the binomially, or extra-binomially, distributed outcome variables of employment rate and unemployment rate the tables can just be read vertically to compare the fit of more complex models (those with more level and/or that include random coefficients) with simpler models that have fewer levels and do not have any random coefficients.

- The resulting AIC values for the models of resident earnings are set out in section 5.7.3 below to give an example of the range of models and results produced for each output variable. The AIC values for models of the other output variables are shown in Annexes 2 to 6. Comments on the results of the models for each output variable are given below in the following sections:
- Resident earnings (section 5.7.3)
- Workplace earnings (section 5.7.4)
- Mean hours (section 5.7.5)
- Job density (section 5.7.6)
- Employment rate (section 5.7.8)
- Unemployment (section 5.7.9).

In discussing the results from the comprehensive set of models for local authority outcomes the idea of parsimonious models arises as do the issues of models that fail to converge or produce 'singular fits'. Comments on these concepts are given below in

sections 5.8 to 5.10.   A summary of the findings from the tables of results for the comprehensive set of models is given above in section 5.7.1. An overview of patterns and findings from all the results presented in chapters 4 and 5 is given at the start of chapter 6, *Discussion of results.*

5.7.1 Summary of the results and recommendations from section 5.7

Residential Earnings

For residential earnings, the four-level null model had the lowest AIC value however it failed to converge. The lowest AIC for a null model of residential earnings that did converge was for the two-level model with grouping by NUTS 2 areas. For all but one predictor variable the four-level random intercept models fitted the data better than any of the two-level models. For many of the predictor variables the four-level model with random coefficients at NUTS 1 areas fitted the data as well as the four-level random intercept model. The only exception was the social housing variable. For this predictor variable the four-level model failed to converge making the random intercept and random coefficient models with grouping by NUTS 2 areas the models which (equally) fitted the data the best.

Recommendations:

- All researchers of residents' earnings should build four-level random intercept models.
- Researchers wishing to investigate whether there are differing relationships between residential earnings and predictor variables in different part of England should build: four-level models with random intercepts at all levels and random coefficients at NUTS 1 areas level; and then build two-level random coefficient models with grouping by NUTS 2 areas; and then compare the AIC values of the two models to see which model fits the data best.

Workplace Earnings

For workplace earnings, the four-level null model fitted the data better than any of the two-level null models. For most predictor variables the four-level random intercept models and four-level models with random intercepts at all levels and random coefficients at one of NUTS 1 areas level or NUTS 2 areas level fitted the data better than any of the two-level models. There were a few exceptions where the two-level random coefficient model with grouping by NUTS 2 areas fitted the data best.

Recommendation:

- All researchers of workplace earnings should initially build four-level random intercept models. They should then build four-level models with random intercepts at all levels and random coefficients at NUTS 1 areas level and also build two-level random coefficient models with grouping by NUTS 2 areas and compare the AIC values of the three models to see if either of the models containing random coefficients fit the workplace earnings data appreciably better than the initial four-level random intercept model.

Earnings in general

These results suggest that in general models of earnings should be four-level models with grouping by all of NUTS 3 areas, NUTS 2 areas and NUTS 1 areas. Exceptionally two-level models with grouping by NUTS 2 areas can produce a better fit to the data, especially if the four-level model fails to converge. Two level models with grouping by NUTS 3 areas or NUTS 1 areas alone are not advised as they did not produce the best fitting models for any of the predictor variables in this study.

Recommendation:

- All researchers of areal earnings of any sort should initially build four-level random intercept models. They should then build four-level models with random intercepts at all levels and random coefficients at NUTS 1 areas level and compare the AIC values of the two models to see whether allowing a model to have random coefficients at NUTS 1 areas level produces a model that fits the earnings data better than the four-level random intercept model.

Mean hours worked

For the null models of mean hours worked it was the two-level null model with grouping by NUTS 2 areas that fitted the data best. For most predictor variables it was the random intercept two-level model with grouping by NUTS 2 areas that fitted the data best. For a few predictor variables two-level random coefficient models with grouping by NUTS 2 areas fitted the data equally well or better. These results suggested that researchers interested in modelling mean hours worked should build random intercept, and possibly random coefficient, two-level models with grouping by NUTS 2 areas.

Recommendations:

>   All researchers building models of mean hours worked should first build a random intercept two-level model with grouping by NUTS 2 area levels. They should then build a two-level random coefficient model with grouping by NUTS 2 areas level and compare the AIC values of the two models to see whether the random coefficient model fits the means hours data any better than the random intercept model for the particular predictor variables that they are including in their models.

Job density

Of the models of job density it was the null model with grouping by NUTS 2 areas that had the lowest (most negative, see section 4.2.6) AIC value showing that it fitted the data better than the other null models. For models incorporating predictor variables it was predominantly two-level models with grouping by NUTs 2 areas that fitted the data better with either the random intercept or random coefficient model with grouping at this geographic scale fitting the data best. This suggested that researchers interested in modelling job density should experiment with building random intercept and random coefficient models with grouping at the NUTS 2 areas geographic scale.

Recommendation:

- All researchers building models of job density should build two-level models in
  which local authorities are grouped by NUTS 2 areas.   Initially they should build a
  random intercept two-level model with grouping by NUTS 2 areas and then they
  should build a random coefficient two-level model with grouping by NUTS 2 areas
  and compare the AIC values of the two models to see whether the random
  coefficient model fits the job density data appreciably better than the random
  intercept model.

Employment

The four-level null model for local authority level employment rates had an AIC value that
was very slightly lower than those for the two-level null models with grouping by NUTS 2
areas or by NUTS 1 areas suggesting that they fitted the data better. For the majority of
predictor variables, the random intercept models that had four levels or had two levels
with grouping by NUTS 2 areas fitted the data similarly well and better than the remaining
two-level models. However, this was not true for all predictor variables as two-level
models with grouping by NUTS 1 areas fitting the data better than, or equally as well as, the
other models for a small number of predictor variables. The overall findings from the
models of employment was that there were predictor variables that can influence, or that
correlate with, local authority level employment rates that act at a variety of different
geographic scales depending on the predictor variable.

Recommendations:

- Researchers of employment rates should focus on building random intercept
  models and not put resources into building random coefficient models.
- Researchers of employment rates should initially build four-level random intercept
  models.
- Researchers may then experiment with building two-level random intercept
  models with grouping by NUTS 2 areas level or by NUTS 1 areas level to see

whether for the particular predictor variable(s) in their models either of these two-level models fit the data appreciably better than the four-level model did.

- Researchers should not use NUTS 3 areas level alone to build models of employment rates.

Unemployment

For models of local authority level unemployment the results were more uniform than those for models of employment rates. For the null models and for models for all the predictor variables, the random intercept four-level models and the four-level models with random intercept at all levels and random coefficients at one of NUTS 2 areas level or NUTS 1 areas level had the lowest AIC values showing that they fitted the data better than any of the two-level models. This suggested that grouping all three geographic scales, NUTS 3 areas, NUTS 2 areas and NUTS 1 areas, should be included in models of unemployment.

Recommendations:

- Researchers of unemployment rates should initially build four-level random intercept models of local authority level unemployment rates.

- Researchers could then allow the coefficients at either NUTS 2 areas level or at NUTS 1 areas level in the four-level models to vary to see if this produced a better fitting model (as indicated by the AIC values of the different models).

- Researchers of employment rates and researchers of unemployment rates should not assume that the geographic scales appropriate for either one of these labour market outcomes will necessarily be the most suitable geographic scales to use for the other one of these key labour market outcomes.

Table 23: Resident Earnings at LA level: AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level NVQ4+ and No qualifications predictors

	Resident Earnings – Null Models	**Resident Earnings** – Random Intercept models with LA level **NVQ4+** predictor variable	**Resident Earnings** – Random coefficient models with LA level **NVQ4+** predictor variable	**Resident Earnings** – Random Intercept models with LA level **No qualifications** predictor variable	**Resident Earnings** – Random Coefficient models with LA level **No qualifications** predictor variable
Four Level model	*3,618 Model failed to converge*	3,476	*3,478 boundary (singular) fit*	**3,410**	*3,412 boundary (singular) fit*
Grouping by NUTS 3 areas	3,697	3,514	3,511	3,492	*3,485 boundary (singular) fit*
Grouping by NUTS 2 areas	**3,659**	3,492	3,490	3,441	3,441
Grouping by NUTS 1 areas	3,674	3,506	3,498	3,450	*3,449 boundary (singular) fit*
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level	3,476			**3,411**	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level	**3,469**			*3,412 Model failed to converge*	

The resident earnings four-level null model gave a lower AIC value than the other resident earnings null models but it failed to converge. Of the null models that did converge the two-level model with grouping by NUTS 2 areas gave the lowest AIC value indicating that it fitted the data better than the others that converged.

Of the models using the NVQ level 4 or higher qualification predictor to model resident earnings the four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level appeared to fit the data best.

Moving attention to the models using the 'No qualifications' predictor to model resident earnings the four-level model with random intercepts at all levels and the four level model with random intercepts at all levels and random coefficients at NUTS 2 areas level fitted the data best. Three of the random coefficient models gave singular fit models (see section 3.3.3.3 above), one failed to converge and another gave an AIC value that was the same as the AIC value for the corresponding random intercept model.

Table 24: Resident Earnings at LA level: AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level Proportion with Bad/Very Bad Health and Proportion working Part-time predictors

	Resident Earnings – Null Models	Resident Earnings – Random Intercept models with LA level **Bad/Very Bad Health** predictor variable	Resident Earnings – Random coefficient models with LA level **Bad/Very Bad Health** 'predictor variable	Resident Earnings – Random Intercept models with LA level **working Part-time** predictor variable	Resident Earnings – Random Coefficient models with LA level **working Part-time** predictor variable
Four Level model	*3,618 Model failed to converge*	**3,475**	*3,474 Model failed to converge*	**3,606**	*3,604 Model failed to converge*
Grouping by NUTS 3 areas	3,697	3,553	*3,551 Model failed to converge*	3,683	3,671
Grouping by NUTS 2 areas	**3,659**	3,496	**3,472**	3,641	*3,635 boundary (singular) fit*
Grouping by NUTS 1 areas	3,674	3,537	3,526	3,651	3,645
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level		*3,476 Model failed to converge*		*3,608 boundary (singular) fit*	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level		**3,475**		**3,608**	

Of the models using the proportion of people with bad or very bad health predictor to model resident earnings the two-level random intercept model with grouping by NUTS 2 areas had the lowest AIC value although the value was only slightly lower than those for

the four-level model with random intercepts at all levels and the four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level.

Observing the AIC values of the models using the proportion of people working part-time predictor to model resident earnings, it was the four-level model with random intercepts at all levels and the four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level that fitted the data best. The AIC values for these two models were much lower than the others indicating that these were clearly the best fitting models.

Table 25: Resident Earnings at LA level: AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level Age and proportion Female predictors

	Resident Earnings – Null Models	**Resident Earnings –** Random Intercept models with LA level **Age** predictor variable	**Resident Earnings –** Random coefficient models with LA level **Age** 'predictor variable	**Resident Earnings –** Random Intercept models with LA level **proportion female** predictor variable	**Resident Earnings –** Random Coefficient models with LA level **proportion female** predictor variable
Four Level model	3,618 *Model failed to converge*	**3,600**	*3,601 boundary (singular) fit*	**3,606**	*3,598 Model failed to converge*
Grouping by NUTS 3 areas	3,697	3,687	3,691	3,685	3,666
Grouping by NUTS 2 areas	**3,659**	3,647	3,645	3,646	*3,640 Model failed to converge*
Grouping by NUTS 1 areas	3,674	3,663	3,660	3,661	3,661
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level	*3,600 Model failed to converge*			**3,608**	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level	*3,578 Model failed to converge*			**3,608**	

Of the models using the local authority level age predictor to model resident earnings the four-level random intercept model fitted the data best as evidenced by its AIC value which was much lower than those for any other of the models that converged successfully without generating any warning messages.

Observing the AIC values of the models using the proportion of people who were female as the predictor to model resident earnings, it was the four-level random intercept model and the two four level models with random intercepts at all levels and random coefficients at either NUTS 2 areas level or NUTS 1 areas level that fitted the data best as indicated by their AIC values which were much lower than those for the other successfully converging models.

For both of these predictor variables there was a large difference between the lowest AIC values and those for other models indicating that the fit of the best fitting models was much better than that of the other models. This showed that it was important to include all of the three NUTS areas geographic scales as levels in the models.

Table 26: Resident Earnings at LA level: AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level Proportion travelling 30km + and Industrial Diversity Indicator predictors

	Resident Earnings – Null Models	Resident Earnings – Random Intercept models with LA level **Proportion travelling 30km** predictor variable	Resident Earnings – Random coefficient models with LA level **Proportion travelling 30km** 'predictor variable	Resident Earnings – Random Intercept models with LA level **Industrial Diversity Indicator** predictor variable	Resident Earnings – Random Coefficient models with LA level **Industrial Diversity Indicator** predictor variable
Four Level model	3,618 *Model failed to converge*	**3,595**	*3,600 Model failed to converge*	**3,607**	*3,575 boundary (singular) fit*
Grouping by NUTS 3 areas	3,697	3,685	3,670	3,685	*3,661 boundary (singular) fit*
Grouping by NUTS 2 areas	**3,659**	3,639	3,635	3,648	*3,631 boundary (singular) fit*
Grouping by NUTS 1 areas	3,674	3,651	*3,652 boundary (singular) fit*	3,662	3,625
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level		*3,697 boundary (singular) fit*		*3,593 boundary (singular) fit*	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level		*3,597 Model failed to converge*		*3,586 Model failed to converge*	

Of the models using the local authority level proportion of people travelling 30 km or more to work to model resident earnings, it was the four-level random intercept model that fitted the data best. It is notable that its AIC value was much lower than the next lowest showing that it fitted the data much better than the other models. This indicates that using a four-level was very important when using this predictor variable in models of residents' earnings.

Observing the AIC values of the models using the local authority level industrial diversity indicator as the predictor to model resident earnings, it was again the four-level random intercept model that appeared to fit the data best with three of the four random

coefficient models producing a singular fit (see section 3.3.3.3 above). Again this showed the importance of including all four geographic scales in models of residents' earnings.

Table 27: Resident Earnings at LA level: AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level Proportion with Occupational Diversity Indicator and proportion managers and professionals predictors

	Resident Earnings – Null Models	**Resident Earnings –** Random Intercept models with LA level **Occupational Diversity Indicator** predictor variable	**Resident Earnings –** Random coefficient models with LA level **Occupational Diversity Indicator** 'predictor variable	**Resident Earnings –** Random Intercept models with LA level **proportion managers and professionals** predictor variable	**Resident Earnings –** Random Coefficient models with LA level **proportion managers and professionals** predictor variable
Four Level model	*3,618 Model failed to converge*	**3,462**	*3,464 boundary (singular) fit*	**3,464**	*3,472 boundary (singular) fit*
Grouping by NUTS 3 areas	3,697	3,502	*3,495 Model failed to converge*	3,524	3,519
Grouping by NUTS 2 areas	**3,659**	3,481	*3,481 boundary (singular) fit*	3,487	3,489
Grouping by NUTS 1 areas	3,674	3,498	3,501	3,501	*3,498 boundary (singular) fit*
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level		*3,462 boundary (singular) fit*		*3,466 Model failed to converge*	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level		**3,463**		**3,463**	

Of the models using the local authority level occupational diversity indicator as the predictor to model resident earnings, it was the four-level model with random intercepts at all levels and the four-level model with random intercepts at all levels and random

coefficients at NUTS 1 areas level that fitted the data best. For this predictor four out of the six models that included random coefficients either produced a singular fit or failed to converge (see section 3.3.3.3 above for information about models that have a singular fit or fail to converge).

Observing the AIC values of the models using the proportion of managers and professional as the predictor to model resident earnings, it was yet again the four-level model with random intercepts at all levels and the four-level models with random intercepts at all levels and random coefficients at NUTS 1 areas level that fitted the data best.

Table 28: Resident Earnings at LA level: AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level Proportion plant/process workers and proportion of people in rural areas inc. hub towns predictors

	Resident Earnings – Null Models	Resident Earnings – Random Intercept models with LA level **Proportion plant/process workers** predictor variable	Resident Earnings – Random coefficient models with LA level **Proportion plant/process workers** 'predictor variable	Resident Earnings – Random Intercept models with LA level **proportion of people in rural areas inc. hub towns** predictor variable	Resident Earnings – Random Coefficient models with LA level **proportion of people in rural areas inc. hub towns** predictor variable
Four Level model	*3,618 Model failed to converge*	**3,497**	*3,481 boundary (singular) fit*	**3,608**	*3,619 boundary (singular) fit*
Grouping by NUTS 3 areas	3,697	3,570	3,534	3,691	*3,690 boundary (singular) fit*
Grouping by NUTS 2 areas	**3,659**	3,525	*3,506 boundary (singular) fit*	3,651	*3,653 boundary (singular) fit*
Grouping by NUTS 1 areas	3,674	3,534	3,510	3,668	3,667
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level		3,499 *boundary (singular) fit*		3,601 *boundary (singular) fit*	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level		**3,499**		**3,610**	

Of the models using the proportion of plant and process workers as the predictor to model resident earnings, it was again the four-level level model with random intercepts at all levels and the four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level that fitted the data best.

Observing the AIC values of the models using the proportion of people who live in rural areas including hub towns as the predictor to model resident earnings, it was yet again the

four-level model with random intercepts at all levels and the four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level that fitted the data best. The large difference in the AIC values for this model and the other models with this predictor variable that converged shows that it fitted the data much better than the other models. As had been seen for models with a number of other predictor variables many of the random coefficient models produced singular fits when this predictor variable was used showing that random coefficient models are often over complex for modelling residents' earnings (see section 3.3.3.3 above for information about singular fitting models).

Table 29: Resident Earnings at LA level: AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level social housing and proportion of people born outside the UK

	Resident Earnings – Null Models	Resident Earnings – Random Intercept models with LA level **social housing** predictor variable	Resident Earnings – Random coefficient models with LA level **social housing** predictor variable	Resident Earnings – Random Intercept models with LA level **proportion non-UK born** predictor variable	Resident Earnings – Random Coefficient models with LA level **proportion non-UK born** predictor variable
Four Level model	*3,618 Model failed to converge*	*3,579 Model failed to converge*	*3,578 Model failed to converge*	**3,575**	*3,580 boundary (singular) fit*
Grouping by NUTS 3 areas	3,697	3,678	*3,669 Model failed to converge*	3,643	3,636
Grouping by NUTS 2 areas	**3,659**	3,624	3,622	3,625	3,621
Grouping by NUTS 1 areas	3,674	3653	3,650	3,629	3,627
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level		*3,579 Model failed to converge*		*3,571 Model failed to converge*	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level		**3,576**		*3,575 Model failed to converge*	

The models using the proportion of social housing in each local authority as the predictor to model resident earnings produced somewhat different results to those for a number of other predictor variables. For this predictor it was the four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas that had the lowest AIC value indicating that it fitted the data better than other models.   It is notable that the four-level random intercept model using this predictor variable and three of the models that included random coefficient models all failed to converge.

The AIC values of the models using the proportion of people who born outside the UK as the predictor to model resident earnings are in line with those for a number of other

predictor variables. That is to say that it was the four-level random intercept model that had the lowest AIC value (indicating that it fitted the data better than the other models using this predictor variable) and half of the models that included random coefficients either produced a singular fit or failed to converge (see section 3.3.3.3 above about models that have a singular fit or fail to converge).

Table 30: Resident Earnings at LA level: AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level IMD 2015 and employment deprivation predictors

	Resident Earnings – Null Models	**Resident Earnings** – Random Intercept models with LA level **IMD 2015**predictor variable	**Resident Earnings** – Random coefficient models with LA level **IMD 2015** predictor variable	**Resident Earnings** – Random Intercept models with LA level **employment deprivation 2015** predictor variable	**Resident Earnings** – Random Coefficient models with LA level **employment deprivation 2015** predictor variable
Four Level model	*3,618 Model failed to converge*	**3,409**	*3,416 Model failed to converge*	**3,427**	*3,419 boundary (singular) fit*
Grouping by NUTS 3 areas	3,697	3,557	3,560	3,539	3,530
Grouping by NUTS 2 areas	**3,659**	3,440	*3,432 Model failed to converge*	3,453	3,432
Grouping by NUTS 1 areas	3,674	3,495	3,4917	3,500	3,493
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level		**3,411**		**3,429**	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level		**3,411**		*3,429 Model failed to converge to converge*	

Of the models using the average IMD 2015 rank for each local authority as the predictor to model resident earnings, in common with models using many other predictor variables, it was the four-level random intercept model that fitted the data best in that it had the very lowest AIC value. However both the four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level and the four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level had AIC values that

were almost as low indicating that they fitted the data just as well as the four-level random intercept only model.

Of the models using the average 2015 employment deprivation rank for each local authority as the predictor to model resident earnings, it was the four-level random intercept model and the four-level model with random intercepts at all levels and random coefficients as NUTS 2 areas level that fitted the data best.

5.7.3 Summary of results from AIC values of models of resident earnings

For random intercept models of residents' earnings the four-level random intercept model usually fitted the data better than any of the two-level random intercept models. However for many of the predictor variables the four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level fitted the data equally well.

The four-level null model of residents' earnings failed to converge and most of the four-level models with random coefficients at all of the levels either failed to converge or produced a singular fit (see section 3.3.3.3 above). The two-level random coefficient models that did converge and had the lowest AIC values were often, but not always, those with grouping by NUTS 2 areas. These results are in line with those from the VPC calculations shown in sections 5.3.7 and 5.5.1 of this chapter which showed there to be fairly large proportions of variation at all three NUTS areas levels under consideration. Together these findings suggest that future researchers of residents' earnings should all build four-level random intercept models initially. If they wish to build random coefficient models to investigate whether there are differing relationships between the outcome and predictor variables in different part of England then they should build a) four-level models with random intercepts at all levels and random coefficients at NUTS 1 areas level, and b) two-level random coefficient models with grouping by NUTS 2 areas, and then compare the AIC values for all the models. The models with the lower AIC values, assuming they do not produce singular fits or fail to converge, will be the better fitting and thus the more appropriate to use.

5.7.4 Summary of results from AIC values of models of workplace earnings

AIC values indicating relative fits of models of workplace earnings are shown in Annex 2. For most predictor variables for workplace earnings the four-level random intercept models and four-level models with random intercepts at all levels and random coefficients at either NUTS 1 or NUTS 2 level fit the data better than two-level random intercept models suggesting there are effects at NUTS 3 areas level, NUTS 2 areas level and NUTS 1 areas level that influence local authority workplace earnings. Most four-level models with random coefficients at all levels either failed to converge or gave a singular fit (see section 3.3.3.3 above). The two-level random coefficient models that fitted the data better, as evidenced by having the lowest AIC value, were the two-level models with grouping by NUTS 2 areas.

The results shown in Annex 2 Table 5 typify the findings shown in many other tables with the four-level random intercept models and the four-level models with random intercepts at all levels and random coefficients at either NUTS 2 areas level or NUTS 1 areas level fitting the data best and all the other random coefficient models either failing to converge or having singular fits.

All researchers building models of workplace earnings should build four-level random intercept models initially. Those wishing to investigate possible different relationships between workplace earnings and predictor variables in different parts of the country should then build a) four-level models with random intercepts at all levels and random coefficients at NUTS 1 areas level, and b) two-level random coefficient models with grouping by NUTS 2 areas. By comparing the AIC for the different models researchers will be able to see which models fit the data better.

That the main findings for residents' earnings and for workplace earnings are similar indicates that any models of earnings should initially be built as four-level random intercept models. Subsequently adding random coefficients to the model at NUTS 1 areas would allow the model to take account of the large amount of variation at NUTS 1 areas level indicated by the VPC values reported in section 5.3.7 and 5.5.1 above.

As random coefficient models do not always give better fitting models than the corresponding random intercept models researchers who wish to build random coefficient models should therefore always also build random intercept models as well and compare the fit of the two models and use the fit as one of the criteria for choosing which of the two models to use.

5.7.5 Summary results from AIC values of models of mean hours worked

AIC values indicating relative fits of models of mean hours worked are shown in Annex 3. Four-level models of mean hours worked built using predictor variables at local authority and lower level often failed to converge or produced a singular fit implying that they may be over complicated (see section 3.3.3.3 above). Even when such four-level models did converge successfully, their AIC values were often found to be higher than those for the corresponding two-level random intercept model grouping by NUTS 2 areas. For example, the AIC values reported in Annex 3 Table 1 tend to be lower for models with grouping by NUTS 2 areas. These findings indicate that researchers building multilevel models of mean hours worked should include the NUTS 2 areas geographical level in their models of mean hours worked.

Models of mean hours that included random coefficients very often failed to converge or produced a singular fit. These findings agree with the analysis of VPC values in that they emphasise the importance of NUTS 2 areas when studying the mean hours worked outcome variable. The findings therefore recommend that researchers building models of mean hours worked should build two-level models that include grouping by NUTS 2 area levels.

5.7.6 Summary of results from AIC values of models of job density for dataset excluding City of London, Westminster and Camden

AIC values indicating relative fits of models of job density are shown in Annex 4. The AIC values for the job density models were all negative. Section 4.2.6 of chapter 4, *Methodology Part 2,* notes that it is possible for AIC values to be negative. It is also noticeable that the four-level null model and most of the four-level random intercept and four-level models with random intercepts at all levels and random coefficients at either NUTS 2 areas level or NUTS 1 areas level failed to converge.

In general, the two-level models with grouping by NUTS 2 areas gave the most negative AIC values showing that models of job density fitted the data better when they were built as two-level models with local authorities grouped by NUTS 2 areas. These findings lead to a recommendation that researchers should build two-level models of job density in which local authorities are grouped by NUTS 2 areas.

5.7.7 Employment rate

The remaining results in this chapter relate to the AIC values for null and random intercept models of local authority employment rates (AIC values shown in Annex 5) and of unemployment rates (AIC values shown in Annex 6). As discussed in sections 3.2.4.3 and 4.2.5.1 of this thesis, to take account of binomial overdispersion in the distributions of these two outcome variables a pseudo level can be added to models at the level of the level one observations. For the local authority models a pseudo level at local authority level was included in the models. This approach to modelling these outcome variables worked well for null and random intercept models. However all of the random coefficient models that were built using this method for local authority employment and unemployment rates gave a strong warning message that pointed out that the number of random effects for which parameters would need to be estimated would be twice the number of observations which would not produce a useable model. Therefore, instead of models with random coefficients at all levels including the pseudo level the f tables of AIC values shown in Annexes 5 and 6 include random intercepts models and models with random intercepts at all levels and random coefficients at specified levels but not at the pseudo level.

5.7.8 Summary of results from AIC values of models of employment rate

AIC values indicating relative fits of models of employment rates are shown in Annex 5. For most predictor variables the four-level random intercept models and the two-level random intercept models with grouping by NUTS 2 areas level fitted the data almost equally well and noticeably better than the other models. For some predictor variables the two-level model with grouping by NUTS 1 areas level was one of the two best fitting models. The two-level models with grouping by NUTS 3 areas level tended to have the highest AIC values showing them to be the least well fitting of any of the models. This suggests that factors that affect the chances of people being employed work at a higher geographical scale than the NUTS 3 area that they live in. This resonates with the idea of there being links between access to transport and being in employment that was raised in section 5.3.3. As two-level models with grouping by NUTS 3 area tend to fit the employment outcome data worse that other multilevel models it is recommended that NUTS 3 areas level alone should not be used for modelling employment rates, with the exception of models using the employment deprivation indicator for which any of the two-level models fit the data equally well.

The results imply that researchers should generally aim to use either all four levels in their random intercept models of employment if practical or to use two-level models with grouping by NUTS 2 areas level.

Logistic generalised multilevel models that included a pseudo level at local authority level in order to implement the additive approach to working with overdispersion of binomial variables and had random coefficients for all levels including the pseudo level increased the number of coefficients to be estimated to twice the number of observations in the model making the models extremely over-specified and impractical to run. Some of the models of local authority employment rates built with random intercepts at all levels and random coefficients at one or more levels not including the pseudo level ran successfully without producing warning messages, however most such models produced singular fits (see section 3.3.3.3 above) suggesting that it is not generally helpful to researchers building models of employment rates to include random coefficients at any levels in their models. It is therefore recommended that researchers focus on building random intercept models of employment rates and do not put resources into building random coefficient models of employment rates.

5.7.9 Summary of results from AIC values of models of unemployment rate

AIC values indicating relative fits of models of workplace earnings are shown in Annex 6. Four-level random intercept models of local authority unemployment rates and four-level models with random intercepts at all levels and random coefficients at one of either NUTS 2 areas level or NUTS 1 area level generally fitted the data better than any of the two-level random intercept models. These results come from logistic generalised multilevel models that use the additive approach to working with over-dispersed binomial variables. They are different to the results from an early experimental set of models that were built without taking account of overdispersion. The random intercept models that did not take account of overdispersion showed that when only random intercept models were compared then it was the four-level random intercept models that gave the lowest AIC values indicating that they fitted the data better than the two-level random intercept models in common with the models that included a pseudo level to take account of binomial overdispersion. However, when random coefficient models that did not allow for overdispersion were built

they indicated that random coefficient models with grouping by NUTS 3 areas gave lower AIC values than random intercept models. These values suggested that if account was not made for overdispersion then random coefficient models with grouping by NUTS 3 areas fitted the data better than random intercept four-level or two-level models.

The findings for the models using the proportion of people who were female as the predictor to model unemployment rates showed the four-level random intercept model to fit the data better than the other models. This implies that it may be unwise to build two-level models of unemployment rates with this predictor using either grouping by NUTS 3 areas or grouping by NUTS 2 areas, or to include any random coefficients. This is in contrast to the equivalent models for employment where the two-level model with grouping by NUTS 2 areas fitted the data as well as the four-level model suggesting that grouping by NUTS 2 areas should be included in models of employment rates. These findings emphasise that different geographic scales may be appropriate for models of unemployment and models of employment.

Overall, the findings for unemployment tend to show that four-level models tend to be better fitting than two-level models. This is different to the findings for employment which do not show a clear pattern of four-level models fitting the data better than two-level models.

There are two fundamental recommendations that come from the findings from the models of unemployment presented in Annex 6. Firstly, it is recommended that researchers build four-level models of local authority level unemployment rates that use grouping by NUTS 3 areas level, grouping by NUTS 2 areas level and grouping by NUTS 3 areas level rather than building two-level models of unemployment rates that only use grouping by one of these NUTS areas geographic scales. Secondly, it is recommended that researchers building models of unemployment rates and of employment rates do not assume that the geographic scales appropriate for one of these labour market outcomes will necessarily be the most suitable geographic scales for use in the other labour market outcome. This recommendation is made as the AIC values shown in Annex 5 and Annex 6 suggest different geographic scales for these two labour market outcomes. These different

results suggest that the geographic scale at which factors affect employment and unemployment operate at subtly different geographic scales at least for some of the working age population.

5.8 Parsimonious models

Reference is made in sections 4.2.6, 5.6.4, 5.7 and 7.3.3 to choosing parsimonious models rather than more complex models. Parsimonious models are those which model the data well but with as small a number of independent variables as are necessary to give a good model fit rather than models that fit the data marginally better but that include more variables than are necessary to give a good fit to the data. The advantages of parsimonious models include computational efficiency; transparency; and importantly predictive power when applied to new datasets. Computational efficiency may be less of a concern than it was in the past as models are run on more and more powerful computers however transparency and predictive power are still important. Transparency is important as it enables a model to be explained clearly to those that may use its results including policy makers and a wider public that may be affected by policies informed by the model. Predictive power when a model is applied to new datasets is important as it makes the model more useful in a greater number of real-world and rapidly changing situations. A less parsimonious model that fits the data very well for the dataset it was built for or trained on but is very complex, involving large numbers of variables meaning that large numbers of parameters have to be estimated may be useful for the particular dataset and for some uses in machine learning but is less likely to be useful for the types of modelling in this project and in models built by researchers to try to understand and explain how different factors affect labour market outcomes for individuals and local areas.

5.9 Models that failed to converge

A number of the models, particularly some of the more complex models that allowed for random coefficients at more than one geographic level, are reported in the results as having failed to converge and it has been suggested that these models are not used. The reasons for not using these models are that they may have failed to converge as they were overfitted. The documentation for the R lme4 package used for the modelling explains that overfitted models may have poor power, i.e. they have a high risk of producing Type 2 errors and a high risk of computational problems, Bates et al (2020). Where alternative,

simpler models are available that give a good fit with the data the simplest solution is to disregard the models that fail to converge and use one of the simpler models. Bates et al (2020) recommend using simpler alternative models rather than ones that fail to converge. There might be some circumstances where a researcher wishes to pursue the idea of allowing the chance for coefficients to vary despite the relevant model in this research failing to converge. In such cases it is suggested that they proceed with caution to see whether the model converges for their own dataset and perhaps experiment with data for different years to see how robust and consistent the model is for similar data that might be expected to exhibit similar relationships each year. If the results are consistent that would suggest there being more merit to the model where the coefficients are allowed to vary than if the results varied a lot from year to year.

5.10 Boundary (singular) fits

A relatively large number of the models in this research are reported as having a 'boundary (singular) fit'. Very often this is because the models in this research have large numbers of random effects. Singular fits are closely related to the concept of models failing to converge. Bates et al, 2020, say that the singular fit warning message means that the variance of one or more of the random effects in the model is either zero or close to zero and suggest that this may mean that the model includes more parameters than can be justified by the data. Random coefficient multilevel models may have a tendency to being overfitted due to the large number of coefficients they are required to estimate as a separate coefficient is required for each group at each level in a multilevel model if coefficients are allowed to vary at each level. Possible solutions to the issue of singular fit models include creating less complex models (Bates, et al, 2020).

6. Discussion of results

6.1 Introduction to discussion of results

This chapter sets out to show how the research aims for this project were addressed by the use of multilevel modelling of areal labour market statistics. It briefly sets out the research problem and describes the types of models built to investigate the problem. It then summarises the findings from the models. Next it compares the results from the VPC and model fit approaches used in this project. It then comments on the importance of null models before comparing the approaches and findings of this project to other researchers' approaches and solutions to the question of which geographic areas to use for publishing and analysing labour market statistics. Finally, the chapter discusses the wider implications of the findings including how VPC values can help in survey design by suggesting an appropriate level of geographic information to collect; the use of appropriate geographic scale for future research and practice; and speculates how the findings could be useful in the consideration of which geographic indicators could make microdata even more useful for those modelling local area labour market outcomes.

6.2 Research problem

The problem underlying research that involves the analysis of areal labour market and related statistics is that data are often only available for people aggregated by the areas that that live in rather than for individual people. Analysing data for the same people aggregated in different ways can give different results depending on the degree of aggregation and specific details of the aggregation. This project was focussed on how the degree of aggregation – the geographic scale - can affect the results of analyses. Specifically, it sought to investigate which geographic scales are helpful for researchers to include in their models of labour market statistics and which geographic scales might not be particularly helpful to include in their models or may even provide unhelpful or misleading information if they are included.

6.3 Modelling

The research was conducted by building multilevel statistical models of key labour market statistics and studying two aspects of the models. The first aspect was a within-model

comparison for each model of how much of the variation in each model occurred at each of the geographic scales used as a level in the model. This aspect was measured by calculating the variance partition coefficient (VPC) for each level of each model and comparing the values for each level within each model. The levels with the largest values relate to the levels with the most variation and as a result were deemed to be important levels for researchers to include in their models. The second aspect was a comparison between different models of how well different models of the same data fitted the data. This aspect was measured using the Akaike Information Criterion (AIC). As the AIC is essentially a measure of the 'lack of fit' of a model to the data, with an allowance made for the number of parameters to be estimated, those models with lower AIC values fit the data better than those with higher AIC values. In the case of negative AIC values, it is the most negative values that indicate the better fitting models (see section 4.2.6).

In this research AIC values were used to compare the fit of different multilevel models that used exactly the same variables but that were either two-level models with grouping of local authorities (the lowest level in each model) by a single geographic scale (NUTS 3 areas, or NUTS 2 areas, or NUTS 1 areas) or were four-level models with grouping of local authorities by all three of NUTS 3 areas, NUTS 2 areas and NUTS 1 areas. The model with the lowest AIC was the one that fitted the data best and the geographic scales that were used as levels in that model were deduced to be the most useful geographic scales for researchers to include in their models of the outcome variable.

Before addressing the question of which geographic scales to include as levels in multilevel modelling this research addressed the question of which geographic scales were the most helpful to use for the calculation of predictor or covariables to use in models of labour market statistics. This part of the research also made use of the AIC measure. Different four-level models were built of local authority level unemployment and employment outcome variables. All the models were built with grouping of local authorities by NUTS 3 areas level, NUTS 2 areas level and NUTS 1 areas level. The difference between each of the models was the geographic scale of the predictor variable. For each combination of local authority level outcome variable and predictor variable subject a separate four-level model was built which had either a single predictor variable (calculated at one of local authority scale, NUTS 3 areas scale, NUTS 2 areas scale or NUTS 1 areas scale) or had four predictor

variables (one calculated at each of these four geographic scales). By comparing the AIC values for models that used the same predictor variables calculated at different scales it was possible to see which model fitted the data best. This information could then be used to deduce which geographic scale predictor variable was the most useful to include in four-level models of local authority level unemployment and employment outcome variables. The pattern was very clearly that local authority level predictor variables led to better fitting models. As a result, only local authority level predictor variables were used for the rest of the project.

6.4 Overview of results

6.4.1 Output Areas

- Models of economic activity using the percentage of people who described their general health as good or very good in the 2011 Census as the predictor variable gave different regression lines depending on whether they were built at Output Area (OA) level or at NUTS 1 areas level.

- Separate OA-level models for each NUTS 1 area gave different regression lines for each NUTS 1 area.

- Analysis of VPC values showed there to non-negligible variation at county/unitary authority level and also at NUTS 1 areas level. This was true for null models and for random intercept models.

- VPC values for random intercept models showed that there was more variation at county/unitary authority level than at NUTS 1 areas level. This suggested that NUTS 1 areas do not tell the whole story of geographic variation and are not even the largest part of the story. This is an important finding as headline sub-UK statistics are often only provided at NUTS 1 areas level. Depending on the dataset in question there could be many reasons why NUTS 1 areas level is the lowest geographic scale used to provide 'local' statistics. It could be for historic reasons to provide statistics for the former government office regions which would have been important to those responsible for those regions; it could be to give simple summary information to see how broad areas of the UK compare; importantly it could be that when statistics for areas are based on survey data there might not be enough observations in the sample to give reliable estimates for areas smaller than NUTS 1 areas in some parts of the UK; finally there might be real or perceived data

protection issues that make it difficult to release statistics for smaller areas than NUTS 1 areas. This project highlights that important relationships between variables may be missed if economic activity statistics are not presented and analysed at a finer geographic scale than NUTS 1 areas suggesting that consideration might be given to whether more sub-UK statistics could be released at a finer geographic scale than NUTS 1 areas on a case by case basis rather than defaulting to releasing just NUTS 1 areas statistics.

6.4.2 Microdata

- Open access microdata only had NUTS 1 areas scale geographic identifiers so it was not possible to deduce which geographic scales to include in models by using open access microdata alone.

- Only a small proportion of the variance in unemployment status for individuals was shown to be at NUTS 1 areas scale suggesting that NUTS 1 areas on their own were not particularly useful in models of unemployment. This resonates with the finding from OA modelling that using NUTS 1 areas alone as the geographic scale for the analysis of statistics relating to economic activity can miss important relationships between variables.

6.4.3 Local authority level VPC

- For hours worked models it was more important to include NUTS 2 areas level than NUTS 3 areas level or NUTS 1 areas level.

- For residential earnings NUTS 1 areas level was more important to include in models than NUTS 2 areas level or NUTS 3 areas level.

- For workplace earnings models NUTS 1 areas were the most important to include. However they were less important in models of workplace earnings than they were in models of residential earnings.

- For job density models it was important to include NUTS 2 areas level but not important to include NUTS 1 areas level.

- Some variance in the proportion of working age adults who were unemployed was at NUTS 3 areas scale, some at NUTS 2 areas scale and some at NUTS 1 areas scale. Most of the variation was at local authority (or lower) level. This suggest that all of

NUTS 1 areas, NUTS 2 areas and NUTS 3 areas geographic scales should be
included as levels in models of unemployment.

- For unemployment the proportions of variance at NUTS 3 and NUTS 1 levels were twice those at NUTS 2 areas level.

- For employment NUTS 3 areas, NUTS 2 areas and NUTS 1 areas levels were all equally important.

- NUTS 2 areas level was more important to include in models of employment than in models of unemployment.

6.4.4 Predictors at different geographic scales

- Models with predictors calculated at four levels were not significantly better than models with predictors calculated just at local authority level or just at NUTS 3 areas level.

- Models of employment rates showed local authority level predictors to be the best for predictor variables related to people's characteristics and showed NUTS 3 areas predictors to be the best for certain predictor variables related to types and levels of jobs.

- Models of unemployment showed local authority level predictors were generally best for modelling local authority level unemployment.

- Although only employment and unemployment models were built to consider which geographic scales to use for predictor variables the results were clear enough to show the great importance of using local authority level variables. Only local authority level predictors were therefore used for the next stage of modelling for all outcome variables.

6.4.5 Overview of results from the comprehensive set of models for local authority level outcomes

- Earnings – four-level models were generally the best fitting. Exceptionally for some predictor variables two-level models with grouping by NUTS 2 areas produced better fitting models. Neither two-level models with grouping by NUTS 3 areas nor two-level models with grouping by NUTS 1 areas produced the best fitting model for any of the predictor variables used.

- Mean hours worked – generally two-level random intercept and/or random coefficient models with grouping of local authorities by NUTS 2 areas produced the best fitting models.

- Job density – generally two-level random intercept and/or random coefficient models with grouping by NUTS 2 areas produced the better fitting models.

- Employment – mixed results depending on which predictor variable was used suggested that there were influences on employment rates that operated at a variety of different geographic scales that were not necessarily predicable.

- Unemployment – four-level models fitted the data best for almost all predictor variables suggesting that there were influences on local authority level unemployment that operated at all of NUTS 1 areas, NUTS 2 areas and NUTS 3 areas. This suggested that all of these geographic scales should be included together as levels in models of local authority level unemployment.

6.4.6 Summary of findings

What people who are in work earn and whether economically active people are in work or are unemployed, are both related to factors at each of NUTS 3 areas level, NUTS 2 areas level and NUTS 1 areas level. This implies that each of these geographic scales should be included as a level in models of earnings and in models of unemployment rates.

Multilevel models of hours worked and of job density should include the NUTS 2 areas geographic scale as a level.

The results for employment rate models are less straightforward than those for other outcome variables. The geographic scale to include in models of employment depends on the predictor variable used in such models. This finding would be consistent with the idea that there may be positive reasons for a person not being in employment, such as planned, well-funded retirement, or negative reasons such as ill-health or lack of suitable job opportunities both of which will affect the numbers of people in employment. The idea of there being some positive reasons for working age people being economically inactive e.g. they are students investing in their education and skills, and some negative reasons e.g.

they have restricted access to the labour market due to work-limiting health conditions or disabilities or have responsibilities such as being carers or lone-parents is presented by the Department for Employment and Learning and Department of Enterprise, Trade and Investment (2015), page 3; see also section 5.3.3 of this thesis.

Researchers wanting to include both employment and unemployment in the same study would need to consider using different geographic scales to model each of these concepts. They would not necessarily need to collect predictor variable data for each geographic scale of interest, e.g. NUTS 3, NUTS 2 and NUTS 1. It may be more appropriate to collect or obtain predictor variable data at local authority level and then build multilevel models that include different geographic scales as the levels for the models for different output variables.

Table 31: Summary of the geographic levels that produce the lowest AIC values for different combinations of local authority level outcome and predictor variables

	Resident Earnings	Work-place Earnings	Mean Hours	Job Density	Employ-ment Rate	Unemploy-ment Rate
Null Models	NUTS 2	4-level	NUTS 2	NUTS 2	4-level / NUTS 2 or 1	4-level
NVQ level 4 or higher	4-level	4-level/ NUTS 2 *(random coeff.)*	NUTS 1 or 2 *random coeff/ 4-level)*	NUTS 1	4-level / NUTS 2 or 1	4-level
No qualifications	4-level	4-level	NUTS 2	NUTS 2	NUTS 1	4-level
Bad/Very Bad Health	4 -level/ NUTS 2 *rand.coeff*	4-level	NUTS 2	NUTS 2	NUTS 1	4-level
Working part-time	4-level	NUTS 2 *(random coeff.)*	NUTS 2	NUTS 2	4-level / NUTS 2	4-level
Age	4-level	4-level	NUTS 2	NUTS 2	4-level/ NUTS 2	4-level
Proportion female	4-level	4-level	NUTS 2	NUTS 2	NUTS 2	4-level
Proportion commuting 30km plus	4-level	4-level	NUTS 2	NUTS 2 *(random coeff.)*	NUTS2 / 4-level	4-level
Industrial Diversity Indicator	4-level	4-level	NUTS 2	NUTS 1 *(rand.coeff)/ NUTS 2*	4-level/ NUTS 2	4-level
Occupational Diversity Indicator	4-level	4-level	NUTS 2	NUTS 2	4-level / NUTS 2 / NUTS 1	4-level
Proportion managers/ professions	4-level	4-level	NUTS 2	NUTS 2	4-level/ NUTS 1/ NUTS2	4-level
Proportion plant/process workers	4-level	4-level	NUTS 2 *(random coeff.)*	NUTS 2 *(random coeff.)*	4-level/ NUTS 2/ NUTS 1	4-level
Proportion in rural areas	4-level	4-level	NUTS 2	NUTS 2 *(rand. coeff.)*	4-level / NUTS 2 or 1	4-level
Social housing variable	4-levels *+ NUTS 1 rand.coeff.*	4-level	NUTS 2	NUTS 2, *(random coeff.)*	NUTS 2 / 4-level	4-level
Proportion non-UK born	4-level	4-level	NUTS 1 / NUTS 2	NUTS 2 *(random coeff.)*	4-level	4-level
LA IMD variable	4-level	4-level	NUTS 2	NUTS 2	4-level /NUTS 2	4-level
Employment deprivation	4-level	4-level	NUTS 2	NUTS 3	NUTS3/ NUTS 2/ NUTS 1	4-level

Table 31 above gives a broad summary of the levels used in models which gave the lowest AIC value for each combination of outcome variable and predictor variable. Whilst this table indicates the models that gave the lowest AIC values it cannot give the details set out in section 5.7 in *Chapter 5 – Results for local authority areas* and in Annexes 2 to 6. Therefore, although this table indicates the levels that are likely to produce models that fit the data better than other models it does not show the degree to which these models are better fitting. For some combinations of variables there may be other models that fit the data almost as well. If such models are simpler, for example are two-level rather than four-level models, or random intercept rather than random coefficient, then depending on the purpose of model the simpler model may be more appropriate. If, however, the modelling is being carried out to study the strength of relationships at different geographic scales then a three-level or four-level model may be more helpful than a two-level model. Alternatively if the modelling is being carried out to see whether the relationships between variables are different in different parts of the country then random coefficient modelling may be more useful even if it does not produce a better fitting model. The geographic scales identified in table 31 should therefore be regarded as being recommended to the extent that they are likely to be among the best rather than that they are by far the best for all purposes. They can be used as a starting point for future modelling.

The models included in this research have all only included a single predictor variable rather than two or more predictor variables. For many of the outcome variables the same geographic scales have been found to be appropriate for many of the predictor variables. If more than one predictor variable is to be included in a model, then if the models for each of the predictors separately suggested including the same geographic scales these scales would be expected to produce a well-fitting model if two or more of the predictors are included in the model. Conversely, if the single predictor models for the outcome variable suggested different geographic scales for the two or more predictor variables to be included in the model, then it is harder to predict which geographic scales to include as levels in the model. A good place to start would be to include the scales suggested by the model for the predictor that led to the largest number of scales. The results from such a model for two or more predictor variables should then be studied carefully to see how the intercepts and coefficients differ at the different scales for each of the predictor variables. If different models are experimented with that include the same set of multiple predictor variables but different geographic scales as their levels then the model fit, measured by the AIC for example, could be compared to see which fit the data best and thus to indicate

which geographic scales to use as levels in multilevel models that include multiple predictor variables.

6.5 Wider implications of findings

6.5.1 Use of VPC findings in designing data collection

The aims of this research were focused on finding which geographic scales are the most appropriate and useful to include in statistical models and providing guidance to other researchers on the effects of using different geographic scales in their analysis of areal data. The motivation for this was that official and other business statistics are often published as areal statistics for different geographic scales presenting researchers with a choice to make about which geographic scale statistics to select for their models.

The findings and results from this research however have wider implications than which geographic scales should be used for secondary data that have been collected by others and are made available for different scales. For researchers setting out to collect their own data from individual people the results can show how much geographic information needs to be collected from each person in order to carry out the research successfully without collecting unnecessarily detailed personal geographic data.

For example, a researcher designing a survey concerning aspects of employment could look at the proportions of variance in employment rates that occur at different geographic scales and use that information to guide which geographic scale should be used to record peoples' locations. Section 5.3.3 shows that there was a higher proportion of variance in employment rates at NUTS 2 area than at NUTS 3 areas level. For example, in both the logit and probit link null models of local authority-level employment rates with a NUTS 3 areas level and a NUTS 2 areas level, the proportion of variance at NUTS 2 areas level was over twice that at NUTS 3 areas level. This explicitly suggests that it is more important to include a NUTS 2 areas level in a multilevel model of employment rates than a NUTS 3 areas level. It also suggests, more generally, that it is more important to have NUTS 2 areas level geographic information when studying employment rates than to have NUTS 3 areas

level information. This is important because it suggests that if a researcher is collecting original data from individuals, it can be sufficient, and more appropriate, to collect information on which NUTS 2 area they live in rather than which NUTS 3 area they live in.

Whilst it may be common and convenient to request a person's (unit) postcode in surveys there can be many advantages to collecting less detailed geographic information. If postcodes are collected and stored for individuals along with other information, such as age and gender, there is a much higher risk of an individual being identified than if a broader geographic area code is used. The broader the geographic area that can be used for data collection and storage the lower the chance of data protection issues arising. Individuals filling in a survey may be aware of this and may be less likely to provide a postcode, or even to take part in the survey at all, if precise geographic details are requested than if they are asked to tick a box indicating which NUTS 2 which area they live in (these would need to be listed as the NUTS nomenclature is not common knowledge). Collecting as little information as is required rather than as much as can be collected is good practice. It saves collecting, coding and storing unnecessary information and is in line with general data protection principles.

6.5.2 Relevance of results for future research and policy

Some of the most interesting results to come from this research are those where the geographic scale found to be useful is different to that which might be expected or that is commonly used for the publication and analysis of areal labour market statistics. In more detail, it is thought provoking that in some cases the geographic level that is particularly relevant to modelling is larger than might have been expected. The contribution to knowledge of such findings is that the geographic scale of influence on local authority-level outcomes is wider than the immediately surrounding local authorities. The implication of this is that future researchers should consider a wider 'hinterland', for example the county or NUTS 2 area that a local authority is part of, when studying what affects employment prospects, for example, for a town or local authority. This may enable them to both save resources by collecting only necessary geographic location information and to use more relevant larger geographic scales in their models (which as a by-product would run more quickly) whilst producing well fitting, informative models.

Another significant overall finding of the research is the inadequacy of NUTS 1 areas alone to provide very useful information on labour market outcomes in different parts of England. This makes a contribution to knowledge by providing evidence that there is a need for more detailed local statistics than NUTS 1 regions to be made available more widely than at present to enable better research, communications and policy making for local (sub-regional) areas. One implication of this is that regional authorities should have access to, and make use of, sub-regional statistics to help develop and implement policies that are as effective as possible for sub-regional areas (which would also be of benefit to the region as a whole).

6.6 Comparison of VPC and model building approaches to choosing geographic scales

For the two earnings outcome variables, the AIC values for various models showed that four-level random intercept models fitted the data better than two-level models. This is consistent with the VPC calculations which showed relatively large proportions of variance at all four levels under consideration. Thus, using VPC calculations alone to help choose which geographic scales to include in multilevel models of earnings data would have led to the same conclusion as the more resource consuming modelling process. From this it could be recommended that if VPC calculations show relatively large proportions of variance at each of the levels under consideration then a multilevel model using all of these levels should be built (see section 7.1.4).

For mean hours worked, the modelling exercise showed that two-level models with grouping by NUTS 2 areas fitted the data better than other models. For the same mean hours worked outcome variable, the VPC calculations showed a much higher proportion of variation at NUTS 2 areas level than at either NUTS 3 areas level (negligible proportion) or NUTS 1 areas level. This is consistent with the finding from the modelling exercise that the two-level module with grouping by NUTS 2 areas level fitted the data better than models that included other NUTS areas levels. Thus for mean hours worked, researchers just using VPC calculations to decide which geographic scales to include in their models are likely to have come to the same conclusion as if they had carried out the fuller modelling process to help the decide with scales to include. From this it could be concluded that if VPC calculations show one geographic scale to have a much larger proportion of variance than others, then a two-level model with grouping at that level is likely to lead to a relatively well-fitting model of the data (see section 7.1.4).

For job density excluding outliers, the modelling exercise showed that two-level models with grouping by NUTS 2 areas fitted the data better than other models. VPC calculations for job density suggested that it was important to include NUTS 2 areas as a level in models of job density but that it was not important to include NUTS 1 areas as a level in such models. These findings are consistent with each other as both show the importance of grouping by NUTS 2 areas level. Researchers using either the modelling exercise approach or the VPC calculations approach would have been drawn to the same idea of building a

multilevel model with grouping by NUTS 2 areas level. This provides more evidence that if VPC calculations show one scale to have a much larger proportion of variance than others then a two-level model with grouping at that level is likely to fit the data well (which supports the recommendations in section 7.1.4).

For unemployment rates, the modelling exercise showed that models of local authority unemployment rates would fit the data best if they were built as four-level models. The same suggestion would have been made if the VPC calculations alone had been used to choose which geographic scales to include in the models of unemployment rates. This finding again shows consistency between the modelling and VPC calculations approaches to deciding which scale geographic levels to include in multilevel models of labour market outcomes.

For employment rates, the VPC calculations would also have prompted a researcher to build four-level models. However, the AIC values from the modelling approach showed the situation to be more complex. It showed that the number of levels to include in models of employment rates, and which geographic scales to use as levels in two-level models of employment rates, depended on which predictor variables were used in the models. This shows that relying on VPC calculations alone, especially of null models, does not necessarily give sufficient information for researchers to decide which geographic scales to include in models that include various predictor variables.

6.7 The importance of null models

Calculating and comparing the proportion of variance at different geographic scales for the outcome variables using null models, i.e. models that have no predictors other than geographic identifiers is often recommended as an initial step to help decide how appropriate it is to carry out multilevel modelling. If almost all of the variation in the outcomes is at individual level then multilevel modelling is generally thought not to be necessary (see section 2.3.3 of the literature review and Nezlek (2008) for details of this

and the counter argument that multilevel modelling should be used whenever the data have a hierarchical structure). If there is appreciable variation at other levels then multilevel modelling is necessary and should include the levels at which the variation suggests most of the clustering of the data may occur. In other research building null models often comprises a relatively small proportion of the modelling activity. It is generally used just as an initial test to provide evidence that gives an indication of whether to carry out multilevel modelling or single level modelling.

The relatively large number of null models built and interpreted in this project provides a contribution to knowledge and, in particular, practice as it emphasises the importance of considering carefully and making good use of the information that may be gained from null models. It is recommended that future researchers should fully consider the results from null models as being of importance in their own right rather than just providing an indication of whether they should proceed to multilevel or single level modelling (see section recommendations in sections 7.1.3 and 7.3.1.7).

For example, researchers should use null models to experiment with using different geographic scales in multilevel models of labour market statistics. The first stage of this would be to ensure that they calculate the ICC/VPC 'by groups' to give not just the total proportion of variation in their data that is due to grouping but also to show separately the relative amount of the proportion of variation that is due to each of the different groups in their null models. By doing this, researchers will then gain not just a yes/no indicator about whether to use multilevel modelling but also gain more specific information about which of the geographic scales used in their null models are important to include in multilevel models. A second stage could be to build null models using a different set of hierarchical geographic areas as levels and calculate their ICC/VPC values. This would enable researchers to compare the ICC/VPC values from the two stages. The stage with the largest proportion of variance due to grouping would then indicate which set of geographic areas to use as levels in the multilevel models. A third stage could be to compare the AIC values of the null models built in stages one and two. The model(s) with the lowest AIC values could then be chosen as the best fitting model(s) to use. In addition, the geographic scales in the best fitting null model could be taken as a good set of geographic scales to use as levels in subsequent random intercept and random coefficient models (see section 7.1.3).

6.8 Comparison of findings with other researchers' solutions to the question of which geographic scales or areas to use for the modelling of labour market statistics

It is appropriate to compare the approach and findings of this project with existing research and current practice concerning the geographic scales used for the reporting and analysis of labour market statistics.

The specific aim of this project was to assess which geographic scales are the most appropriate and useful to include in the statistical modelling of selected UK labour market statistics and which geographic scales provide unhelpful or misleading information.

The main approach of this project was to use multilevel analysis of publicly available statistics for labour market outcomes, e.g. earnings, hours worked and unemployment rates, to assess which existing hierarchical statistical and administrative areas lead to better fitting models when used as levels in multilevel models. This approach generated findings which suggested that models should include either a NUTS 1 areas level, a NUTS 2 areas level and a NUTS 3 areas level, or just have a NUTS 2 areas level depending on the outcome variable. In essence this research has used published data and existing statistical and administrative boundary information to analyse the geographic scale of relationships between measures of the amount of work being undertaken by people and the covariates that can be used to model these measures.

The question of which geographic scales are useful to study labour market statistics, and in particular unemployment rates, has been addressed by other researchers in research to find practical geographic areas for which unemployment rates can be calculated that reflect functional labour markets. Such areas can then be used to analyse unemployment patterns and to implement policies designed to improve unemployment rates for example. Townsend (2019) notes that the first 'functional areas' were the areas covered by individual Labour Exchanges. Townsend (2019) explains that these were then grouped together by the then Department of Employment to create the first set of Travel to Work Areas (TTWAs) which were subsequently used by the government decide where to locate new large car factories for example. This demonstrates the use of TTWAs for policy implementation, i.e. resource allocation, to reduce unemployment.

In introducing the concept of TTWAs the ONS says 'For those involved in labour market analysis and planning, it is useful to be able to use data for labour market areas' (ONS, 2015b). The current approach taken to the creation of TTWAs by the ONS in collaboration with the University of Newcastle is to find a non-overlapping coverage of areas such that the majority of people both live and work within the same TTWA (ONS 2015b, 2015c, 2015d). The main reason for creating TTWAs is to find a set of geographic areas suitable to analyse unemployment statistics. This relates directly to the aim of the research presented in this thesis to find suitable geographic scales to model labour market outcomes such as unemployment statistics. However, whilst the TTWA approach has been to group together small areas (such as Lower Super Output Areas), this thesis has researched which existing administrative and statistical areas, for which data are already available for analysis, are the most useful to use for the study of unemployment rates and related labour market statistics.

In a report for Eurostat, Coombes et al (2012) explored the possibilities of producing a Europe-wide definition of local labour market areas, or functional areas, that could be used as an alternative to administrative areas for statistical and policy analyses. This demonstrates the need for functional labour market areas to be identified to help facilitate analysis and policy development and implementation. The identification of such areas could be informed by the results in this thesis which indicate the geographic scales that are useful in analysing and modelling labour market outcomes such as employment and unemployment. In the report Coombes et al (2012) noted that the results of statistical analysis for areas are sensitive to the choice of area scale and zoning boundaries (known as MAUP, see section 2.2.5 of this thesis) which raised the question of whether there was a set of areas that should be used and concluded that the most appropriate set of areas depended on the purpose of the analysis.

In the UK there are currently 228 TTWAs. Of these 149, are wholly within England, four cross the border between Wales and England, and two cross the border between England and Scotland. The remainder lie wholly within either Scotland, Wales or Northern Ireland. These TTWAs were defined using data from the 2011 Census and were published in 2015. As people tend to live and work in the same TTWA these areas would by definition appear to be highly appropriate areas to use to publish and analyse labour market statistics. As they were intended to be useful to 'those involved in labour market analysis and planning',

ONS (2015b), the TTWA coverage of England suggests itself as a good candidate for the geographic scale for researchers to use in statistical models of labour market statistics. A comparison of English and cross-border TTWAs and the geographic scales found by this project to be useful in modelling labour market statistics would therefore be a good way of comparing the findings of this project with existing research and practice.

The findings of this research emphasise the importance of the NUTS 2 areas geographic scale, often in combination with the NUTS 1 geographic scale. Rarely, if at all, were the NUTS 3 or NUTS 1 geographic scales found to be important or appropriate to use alone in statistical models of labour market statistics. A brief comparison of the 2011 TTWAs for England and the NUTS 1 areas and importantly the NUTS 2, areas for England was therefore undertaken. By comparing NUTS 2 areas and TTWAs it was found that 40% of the 155 English and cross-border TTWAs include area from more than one NUTS 2 area. Conversely, a single NUTS 2 area can be part of many TTWAs. The largest TTWA is that which covers London. This includes parts of three NUTS 1 areas (London, East of England and the South-East) and all or parts of nine NUTS 2 areas. It does not however include the area around Heathrow which is part of the large Slough and Heathrow TTWA.

This comparison of the areas covered by TTWAs and NUTS 1 and NUTS 2 areas provides evidence that labour markets within England can, and do, operate independently of administrative and statistical boundaries. This helps to explain and confirm the results from this project that more than one geographic scale should often be used in models of labour market outcomes. It does this by showing that labour markets in England (as identified by TTWAs) do not fit neatly into the NUTS 1 or NUTS 2 hierarchical statistical areas.

It is important to note that TTWAs are not constrained to fit within administrative boundaries by the algorithms used to create them. A singe TTWA can be part of more than one NUTS 1 area. TTWAs were created by grouping together Lower Super Output Areas (LSOAs) for England and Wales, and equivalent areas for Scotland and Northern Ireland, based on commuting flows between LSOAs to create a coverage of internally contiguous areas such that, ideally, 75% of those who work in a TTWA live in that TTWA and 75% of workers resident in each TTWA work in that TTWA (ONS, 2015d). The 75% rule is eased for larger areas (ONS, 2015c). As a result, six of the UK TTWAs cross either the England-Wales or England-Scotland border.

An insight into the complexity of labour market areas and hence the need identified by this project for more than one geographic scale to be included in statistical models of labour market outcomes may be obtained from summary statistics for 2011 TTWA (ONS, 2015e). These demonstrate a large range and an uneven distribution of English labour market sizes, both in terms of sizes of populations of workers and sizes of geographic areas. For example, the London TTWA has the largest workplace population, around 4.4 million people. This is approximately three to four times greater than that for the Manchester TTWA whilst London covers only a slightly larger land area. The smallest TTWA in terms of worker population is Whitby with a workplace population of around 10,000 people. These differences illustrate that one geographic scale is unlikely to suit statistical models of labour market outcomes for all parts of England and they are consistent with the findings from this project that a) more than one geographic scale is often needed to give better fitting models and b) that the NUTS 1 geographic scale is too coarse to capture all the information about relationships between labour market and related socioeconomic statistics that could be discovered by using a finer geographic scale.

The approach adopted in this project of building multilevel models of labour market outcomes with the lowest level units being grouped by larger geographic areas could be applied directly to the analyses of labour market and other related socioeconomic statistics for outcomes at LSOA level by building two-level models with LSOAs as level one grouped by TTWAs as level two then comparing the fit of models by comparing their AIC values. One application of this could be to compare the effects on model fit of using alternative sets of TTWAs which are produced to study the different labour market geographies that can be found if TTWAs are created separately for different groups of workers. Classic examples of alternative TTWAs are those calculated separately for male and female workers and those calculated separately for full-time and part-time workers. More information on alternative TTWAs is provided by ONS (2016c). Ward and Dale (1992) used separate TTWA for female workers in their geographic analysis of female labour force participation.

The approach developed by this project could add to knowledge and practice by providing an additional way of comparing and demonstrating the usefulness and importance of alternative TTWAs (by modelling LSOA labour market outcomes grouped by main and alternative TTWA coverages). This could be particularly important at a time of changing commuting patterns including more working from home and changing job availability in different industrial sectors. It may also be timely in advance of the results of the 2021

Census as Census statistics on commuting flows are used to produce the main and alternative TTWA coverages for the UK.

6.9 Geographic identifiers for anonymised microdata

An outcome of this project is that the scale used for the geographic identifiers in anonymised microdata can affect the results of analyses of the microdata and can affect the usefulness of the microdata to researchers building models for small areas. This in turn could affect the efficacy of any labour market intervention policies informed by the models. The broad geographic scale of NUTS 1 areas used in the microdata teaching set can cause models to miss the smaller geographic scale differences in labour market outcomes (see section 6.4.2) and miss the differences that can exist in the relationships between labour market outcomes and independent variables in small areas contained within the same NUTS 1 area. Section 4.1.3 of this thesis provides a discussion of the limitations of using microdata with only NUTS 1 areas geographic identifiers.

Consideration of the effects of the choice of geographic scale of location identifiers on the results that can be obtained from analysis of the data and therefore the usefulness of the microdata should be given at the time when the data are anonymised. The overriding concern when choosing the geographic scale for the geographic identifiers must be that the data are securely anonymised such that no individuals can be identified from the resulting microdata. This concern favours the broadest possible geographic scale, e.g. NUTS 1 areas level. Making the microdata more useful for modelling local area labour market outcomes would favour choosing a geographic scale that would provide anonymised data that produce models that fit the data better than microdata data with simply a NUTS 1 areas geographic identifier. This does not simply mean that the finest possible geographic scale available needs to be chosen. For example, many of the models of local authority level labour market outcomes were found to fit the data better when they were built as two-level models with local authorities grouped by NUTS 2 areas. It was rarely the case that two-level models with NUTS 3 areas forming the higher level produced the best fitting models. Section 6.4.6 of this thesis, including in particular Table 31, and also sections 5.7.13 and 6.4.6 highlight the importance of the NUTS 2 areas level. These results from this project suggest that microdata with NUTS 2 area identifiers would produce models that fit the data better than microdata with just NUTS 1 area identifiers and that there is no particular need to have NUTS 3 areas geographic identifiers in microdata for labour market

statistics in order to produce models that fit the data well. The practical explanation of this would be that the geographic identifiers required to produce well-fitting models of labour market outcomes should approximate to local labour market areas. In terms of the NUTS areas classification this would be closer to the NUTS 2 areas than to the smaller NUTS 3 areas. The advantage of this is that providing NUTS 2 areas geographic identifiers in microdata would present fewer confidentiality issues than providing NUTS 3 areas geographic codes. These findings lead to the recommendation that when anonymising microdata consideration should be given to: 1) whether it was possible to provide NUTS 2 areas geographic identifiers without loss of confidentiality; and 2) the degree of data security needed to give researchers access to microdata with NUTS 2 areas geographic codes (see recommendations in sections 7.1.1 and 7.2.19).

A possible area of future research for those responsible for anonymising microdata would be to investigate the effects on data confidentiality of providing microdata with NUTS 2 areas codes. If the effects were too great to make microdata with NUTS 2 areas codes widely available then one solution might be to group some of the less populous NUTS 2 areas to provide a set of merged NUTS 2 areas. This is outside the scope of this project and could instead be considered alongside research into finding labour market areas or travel to work areas suitable for the release of data from the 2021 Census.

7. Recommendations, Limitations and Conclusions

7.1 Specific Recommendations

7.1.1 Geographic scale should be considered

Researchers need to consider which geographic scales to use when modelling labour market and related statistics as different relationships may exist between the same variables measured at different geographic scales (described as the modifiable area unit problem which is a manifestation of Simpson's paradox both of which are described in section 2.2.5 of this thesis). Those responsible for the anonymisation of microdata should be aware that the choice of scale for geographic identifiers can affect the results of analyses and the usefulness of microdata to researchers and consequently to policymakers (further details are given in sections 4.1.3, 6.4.2, 6.9 and 7.2.1.9). In general, the primary purpose of microdata is research rather than teaching. The microdata 'teaching set' used in this research is unusual in that it is intended both for teaching and for research including the development of techniques to be used later on other microdata which are intended just for research.

7.1.2 Importance of different scales can vary for different outcome variables

It should not be assumed that the importance of different geographic levels is the same for all apparently similar outcome variables (as the same geographic scale can have very different proportion of variation in models of different outcome variables). The prime example of this is that the geographic scales most useful in models of unemployment (shown in sections 5.3.1 and 5.7.9) are not necessarily the same as the scales that are most useful for models of employment (shown in sections 5.3.3 and 5.7.8).

7.1.3 Extensive use should be made of null models

It is recommended that future researchers should make extensive use of the results from null models (as they can be of importance in their own right rather than just serve to provide an indication of whether to use multilevel or single level modelling). For example this could include calculating ICC/VPC values for null models 'by groups' to give not just the total proportion of variations due to grouping but also the relative amount of the proportion of variation due to the different levels in their null models. This would show

which geographic scales are important to include. Secondly, researchers could build null models using a different set of hierarchical geographic areas as levels and compare the ICC/VPC values and AIC values from both sets of models. Fuller details are given in section 6.7 above.

7.1.4 Variance Partition Coefficients should be calculated for random intercept models

VPC values should be calculated for random intercept models as well as for null models (as these can indicate the geographic scale to be included in multilevel models with particular predictor variables).

Recommendations from section 6.6:

- If VPC calculations show relatively large proportions of variance at each of the levels under consideration then a multilevel model using all of these levels should be built as it is likely fit the data relatively well;
- If VPC calculations show one geographic scale to have a much larger proportion of variance than others, then a two-level model with grouping at that level should be built as that is likely to lead to a relatively well-fitting model of the data.

7.1.5 Lowest practical geographic scale should be used for predictor variables

The lowest practical geographic scale should be used for predictor variables in multilevel models (as this work has shown that there was little or no advantage in terms of model fit in including predictor variables for larger areas when the same predictor variables were available for smaller areas). Fuller details are given in sections 4.2.6, 5.4, 5.4.3 and 6.4.4).

7.1.6 Fine scale areal statistics should be used in preference to NUTS 1 scale microdata

Areal statistics with geographic indicators for small areas should be used in preference to individual level data with only a high-level geographic indicator, such a NUTS 1 area code (as the proportion of variation at high level geographic scales can be very small). Fuller details are given in sections 6.4.2 and 6.9).

7.1.7 Models of unemployment rates should include NUTS 3 areas, NUTS 2 areas and NUTS 1 areas as levels

Multilevel models of unemployment rates should contain several geographic scale levels (as there can be appreciable proportions of variation at many different geographic scales in null and random intercept models of unemployment rates, as shown in sections 5.3.1 – 5.3.2, 5.3.13 and 5.5.6 – 5.5.7).

Recommendations from section 5.7.9:

- Researchers of unemployment rates should initially build four-level random intercept models of local authority level unemployment rates.

- Researchers could then allow the coefficients at either NUTS 2 areas level or at NUTS 1 areas level in the four-level models to vary to see if this produced a better fitting model (as indicated by the AIC values of the different models).

- Researchers of employment rates and researchers of unemployment rates should not assume that the geographic scales appropriate for either one of these labour market outcomes will necessarily be the most suitable geographic scales to use for the other one of these key labour market outcomes.

7.1.8 Choose geographic scales for employment models according to predictor variables

Multilevel models of employment rates should also use several geographic scales as levels (as there can be appreciable proportions of variation at many different geographic scales in null and random intercept models of employment rates, as shown in sections 5.3.3 - 5.3.4 and 5.5.5).

Recommendations from section 5.7.8:

- Researchers of employment rates should focus on building random intercept models and not put resources into building random coefficient models.
- Researchers of employment rates should initially build four-level random intercept models.
- Researchers may then experiment with building two-level random intercept models with grouping by NUTS 2 areas level or by NUTS 1 areas level to see whether for the particular predictor variable(s) in their models either of these two-level models fit the data appreciably better than the four-level model did.
- Researchers should not use NUTS 3 areas level alone to build models of employment rates .

7.1.9 Multilevel models of hours worked should include the NUTS 2 geographic scale

Multilevel models of hours worked per week should include NUTS 2 areas level, they do not need to include NUTS 3 areas level (as there was found to be far more variation at NUTS 2 areas level than NUTS 3 areas level – see section 5.5.3).

Recommendation from section 5.7.5:

- All researchers building models of mean hours worked should first build a random intercept two-level model with grouping by NUTS 2 area levels. They should then build a two-level random coefficient model with grouping by NUTS 2 areas level and compare the AIC values of the two models to see whether the random coefficient model fits the means hours data any better than the random intercept model for the particular predictor variables that they are including in their models.

7.1.10 Multilevel models of median weekly earnings should include NUTS 3 areas, NUTS 2 areas and NUTS 1 areas as levels

Multilevel models of median weekly earnings, both for residents and for those working in areas, should include NUTS 3 areas level, NUTS 2 areas level and NUTS 1 areas level. This is

recommended because a) there was appreciable variance at all three of these levels –
shown by VPC values in sections 5.3.7 - 5.3.10 and 5.5.1 - 5.5.2; and b) as the AIC values
showed four-level models fitted residential earnings data better than two-level models for
all but one predictor variable as shown in section 5.7.3. Section 2.3.3 of the literature
review and Nezlek (2008) give explanations of why multilevel models should be built even
though the overall amount of the proportion of variation due to grouping may be quite
small).

Residents earnings recommendations (from section 5.7.2):

- All researchers of residents' earnings should build four-level random intercept
  models.
- Researchers wishing to investigate whether there are differing relationships
  between residential earnings and predictor variables in different part of England
  should build: four-level models with random intercepts at all levels and random
  coefficients at NUTS 1 areas level; and then build two-level random coefficient
  models with grouping by NUTS 2 areas; and then compare the AIC values of the
  two models to see which model fits the data best.

Workplace earnings recommendation (from section 5.7.4)

- All researchers of workplace earnings should initially build four-level random
  intercept models. They should then build four-level models with random
  intercepts at all levels and random coefficients at NUTS 1 areas level and also build
  two-level random coefficient models with grouping by NUTS 2 areas and compare
  the AIC values of the three models to see if either of the models containing random
  coefficients fit the workplace earnings data appreciably better than the initial four-
  level random intercept model.

General earnings recommendation (from section 5.7.2 – 5.7.4):

- All researchers of areal earnings of any sort should initially build four-level random
  intercept models. They should then build four-level models with random
  intercepts at all levels and random coefficients at NUTS 1 areas level and compare
  the AIC values of the two models to see whether allowing a model to have random
  coefficients at NUTS 1 areas level produces a model that fits the earnings data
  better than the four-level random intercept model.

7.1.11 Models of job density should include the NUTS 2 geographic scale

Multilevel models of job density should include NUTS 2 areas level. It is not important for them to include NUTS 3 areas level. They do not need to include NUTS 1 areas level. Section 5.3.12 shows the VPC values that support these statements.

Recommendation from section 5.7.6):

- All researchers building models of job density should build two-level models in which local authorities are grouped by NUTS 2 areas.   Initially they should build a random intercept two-level model with grouping by NUTS 2 areas and then they should build a random coefficient two-level model with grouping by NUTS 2 areas and compare the AIC values of the two models to see whether the random coefficient model fits the job density data appreciably better than the random intercept model.

7.1.12 Models of earnings should account for the effect of higher earnings in London

Future research might find a theoretical missing level at a geographic scale higher than NUTS 1 areas that should be included in models of earnings (as the large proportion of variance at NUTS areas level might indicate a missing higher level whose variance is transferred to NUTS 1 areas level when it is missed from multilevel models of earnings – see section 5.3.7).

7.2 Limitations

7.2.1 Single models built for whole populations

All the models built in this research were for whole working age populations. Separate models were not built for people with different characteristics for whom different

geographic scales may have different influences. Whilst job opportunities may be 'open to all', access to public and private transport, health and long term conditions and caring responsibilities, for example, can affect a person's practical ability to commute or relocate for work. Further development of this research could include using the same methodology to build separate models for people with different characteristics. Starting points for this could be separate models: for males and females; for those with different levels of qualification, e.g. no qualifications, qualifications up to and including NVQ level 3, and those with NVQ level 4 or higher qualifications; for different age groups, e.g. 16-25 year olds, 26-50 year olds and 50-65 year olds; and for those with different transport opportunities e.g. drivers and non-drivers, or those in areas with good or poor public transport.

7.2.2 Reliance on the maximum likelihood estimation approach

Models are used to estimate parameters that describe the relationships between variables. The classical or frequentist view of statistics regards parameters as single values. These can be estimated using the maximum likelihood approach and this is the approach used throughout this research. In maximum likelihood estimation the value of the parameter most likely to have generated the sample data is deduced and used to estimate a single value for the parameter. An alternative to the frequentist view of statistics is the Bayesian view in which all parameters are regarded as random variables with a distribution of possible values rather than a single value. In order to estimate the distribution of a parameter a computer simulation can be run to generate a large number of values for the parameter and thereby generate a possible distribution for the parameter. To simulate parameter values Markov Chain Monte Carlo (MCMC) methods can be used. The 'chain' part of the MCMC name signifies that the output from one simulation run is used as input for the next run. The Monte Carlo part of the name indicates that there is a random or stochastic part to the process.

In the R environment the package MCMCglmm can be used to model binomial distributions by setting 'family = multinomial2'. In line with the Bayesian approach a prior distribution needs to be set before the MCMC process is run. The prior distribution can either be

informative, e.g. when previous research has provided a likely distribution for the model or, if no useful previous estimates of the parameter exist, then a diffuse prior can be used which has a much smaller effect on the model than an informative prior would have.

The advantages of the MCMC approach include that it can be used for outcomes that do not have a standard distribution e.g. they are skewed. A big disadvantage is that MCMC models can take a lot longer to run than maximum likelihood models. This is especially true for more complex models. This is particularly relevant for this project where the number of individual models built was very large and a number of them are very complex, for example the models for proportions that include a pseudo level to account for binomial overdispersion. By itself the extra time that would be needed to run MCMC models would be sufficient reason to use just use the maximum likelihood approach for the large number of models that form the core of this research. In addition, for each MCMC model the researcher needs to set a prior distribution before running the model. If these were to be informed by previous research a lot of consideration would need to go into these for each of the models. This could be addressed by using a very diffuse prior (as is often used). Another step needed for each MCMC model is for the output (often in the form of charts) to be studied by the researcher after each model is run to gauge whether the model has converged. This would again be time consuming for the very large number of models run in the research.

On balance, for reasons of time given the large number of models and complexity of the random coefficient models, especially those which include a pseudo level to model binomial overdispersion and thus estimate a very large number of parameters, the MCMC approach has not been used for this project. This is a limitation as the findings from the maximum likelihood approach have not been compared with those that might have been obtained from the MCMC simulation approach to estimating model parameters. Future research, especially if it were focused on a smaller number of variables and models, could experiment with using both the MCMC approach and the maximum likelihood approach to see how the parameter values and findings about which geographic scales are the most helpful and appropriate to use compare for the two approaches. The important point to note would be whether decisions about which geographic scale to use in multilevel models were dependent on the choice of parameter estimation method.

7.2.3 Separate models rather than cross-classified models used for income outcomes

In this research separate multilevel models were built for the two separate variables of median income for people living in each local authority and median income for people working in each local authority. If data for individual workers were available with geographic indicators to show both home and workplace local authority, then there would be the potential to build cross-classified multilevel models of individual incomes. Such models could take account of the fact that workers can be nested in both their home local authority and in their workplace local authority rather than just being part of a strict hierarchy of local authorities.

7.2.4 Use of R scripts rather than building packages

The models used in this research were written and saved in the form of R scripts and the datasets used were stored separately as csv files. If the work were to be developed further to create ready to use models for other researchers, then the writing of packages containing code and data would be one way of making the models more accessible for other researchers.

7.2.5 Separate models for each predictor variable rather than models containing two or more predictor variables

All the models in this research were either null models (i.e. had no predictor variables other than group identifiers) or were models with just one predictor variable. This approach was used partly in order to study the geographic scale effects for each independent variable separately. It was also a pragmatic approach to use given the very large number of independent variables used in the research. Adding models with two or more predictor variables could have increased the number of models considerably. If the research were to be developed further in a way that focused on a smaller number of outcome and predictor variables it would be appropriate to complement the single predictor variable models with models that use a combination of different predictor variables to see how this affects the proportions of variance at each geographic scale.

7.2.6 Reliance on unweighted models for the interval outcome variables

Whilst some account was taken of the size of each local authority in the models for binomial outcome variables, i.e. employment and unemployment rates, as each model contained the number of 'positive' binary outcomes and the number of 'negative' binary outcomes, no account was taken of the size of the local authorities in the models of the interval outcomes of hours worked and median earnings for example. By not adding weights to the models for the interval outcomes each local authority contributed the same amount of information to the models regardless of its overall population or number of workers etc. This could be seen as treating each local authority fairly regardless of its size or as treating the individuals in each local authority unequally as those in more populous areas could be said to be under-represented in the models whilst those in less populous areas could be said to be over-represented. This would not be a problem if the local authority sizes were all approximately equal or in future research based on this research where the areas used at level one in models all had similar population sizes. If this research were to be developed further, or future research were to follow the methodology set out in this research, then it would be possible to build population or workforce weighted models as well as unweighted models and compare the results to see if weighting the models leads to different conclusions about which geographic scales are the most useful and appropriate to use in statistical models of labour market and related socio-economic statistics.

7.3 Broader Conclusions

7.3.1 General conclusions arising from specific findings/recommendations

7.3.1.1 The relationships between areal measures of economic activity including employment and unemployment vary when measured at different geographic scales (see section 7.1.1).

7.3.1.2 The relations between measures of economic activity and covariables can be different in different parts of a study area (see section 4.1.1).

7.3.1.3 Large geographic scales, such as NUTS 1 areas, are not generally sufficient to create models that are as good as those that could be built using finer geographic scales, such as NUTS 2 areas (see sections 6.4.1 – 6.4.2).

7.3.1.4 Covariates to include in models of labour market outcomes should normally be measured or calculated at the same geographic scale as the outcomes (see section 5.4). For some covariates related directly to types and levels of jobs it may be helpful for them to be calculated or measured at the next geographic scale up. This would be consistent with the fact that people often work slightly outside the local area in which they live.

7.3.1.5 Models of local earnings and unemployment rates should be built as multilevel models with several hierarchical geographic scales. This will help to take account of factors that affect whether or not economically active people are in work and how much they are paid that operate at a number of different geographic scales. See sections 7.1.7 – 7.1.8.

7.3.1.6 Models of hours worked and job density, which could both be seen as measures of the availability of work, may be built as two-level models with grouping of the local levels by an intermediate geographic scale, e.g. NUTS 2 areas, that is related to the availability of work in a slightly wider geographic areas than the local or immediate surrounding area whilst being finer than regional, e.g. NUTS 1 areas, scale.

7.3.1.7 Future research should fully consider the results from null models as being of importance in their own right rather than just providing an indication of whether the research should be carried out using multilevel as opposed to single level modelling. (See 6.6 and 7.1.3)

7.3.1.8 As is true for single level regression modelling, histograms of outcome variables being considered for multilevel modelling should be examined to check for normality and outlying observations before carrying out multilevel modelling. This is particularly important for multilevel models as irregularly distributed outcome variables can lead to multilevel models that either fail to converge or result in models with a singular fit (see section 3.3.3.3). It is therefore important that outliers are removed before building multilevel models (see section 4.2.4.2). It has also been found in this project that mean areal outcome variables lead to better fitting models than median outcome variables (see section 4.2.4.2).

For the above reasons it is recommended that researchers intending to build multilevel models should pay particular attention to histograms of potential outcome variables as they may indicate which will lead to successful multilevel models and which will lead to multilevel models that fail to converge or have a singular fit (see section 3.3.3.3 above).

7.3.1.9 Often in anonymised datasets the only geographic information provided is the NUTS 1 area code for each record. As this thesis has shown the NUTS 1 areas geographic scale on its own is often not particularly useful to have in a multilevel model of labour market outcomes. Moreover, for some labour market outcomes two-level models that include grouping by NUTS 2 areas fit the data better than those with grouping at other geographic scales. One implication of this is that for those statistics where anonymised microdata could safely be made available with NUTS 2 areas geographic information, without risking the security of personal information, the availability of such data could greatly enhance the fit of models that researchers could build for such statistics. This could potentially increase the usefulness of such models to inform policy for areas smaller than NUTS 1 areas and to allow the targeting of resources to those NUTS 2 areas where the resources may have the greatest influence. Further details of the issues concerning the geographic scale of microdata are discussed above in sections 6.4.2, 6.9 and 7.1.6 of this thesis.

7.3.2 Hierarchy and proximity

The aims of the project included a statement that the modelling would take account of both the hierarchical and proximity effects present in the statistics. The models built have all been multilevel models where the levels have all been either tiers of the hierarchy of English local government administration or levels of the NUTS hierarchical areas that, for England, are groupings of local authority areas. Thus, the modelling very clearly has taken account of the hierarchical effects in the statistics. That multilevel modelling using hierarchical administrative areas also takes account of the proximity of different areas is supported by research by Moellering and Tobler (1972) into the geographic scales to use to study processes. They argued that as higher level administrative areas were physically larger than the lower level areas that they contained then analysing the amount of variance at different hierarchical areas was equivalent to analysing variance at different geographic

scales. It can be concluded therefore that the project's aim to use modelling that takes account of both hierarchical and proximity effects has been met by the use of multilevel modelling of hierarchical administrative areas and hierarchical statistical areas based upon them (see section 2.3.10).

7.3.3 Types of complexity and purpose of model

There is often a choice to be made between complex, i.e. random coefficient, models with just two-levels and simpler, i.e. random intercept, models with four levels. A logical reason for this is that both a) using random coefficient models rather than random intercept models, and b) using models with a larger number of levels, increase the number of model parameters that need to be estimated. Generally in linear modelling, in order to choose a parsimonious model (one that fits the data well with as few parameters as necessary rather than one which fits the data very well but which requires a much larger number of parameters) it is often appropriate to choose a model with a lower number of parameters. The results of this project suggest that in multilevel modelling, in order to choose a parsimonious model, it is often necessary to choose between a random coefficient model with a lower number of levels and a random intercept model with a higher number of levels or to use a model with several levels but only with random coefficients at one of the levels. Given that this choice may have to be made, consideration should be given to what sort of information or enlightenment is sought from the model. In order to learn about influences coming from different geographic scales, for example to determine what geographic scales to use to formulate or implement policy, a random intercept model with many different levels would be appropriate. However, in order to learn about different strengths of effects in different parts of a study area a random coefficient multilevel model with just two or possibly three levels would be more useful. Such a model could show, for example, whether having a more highly qualified workforce had more or less effect on average local authority income in different parts of a country.

In summary researchers may have to choose between the complexity of random coefficient models that have to be constricted to two-level models in order to converge successfully without giving warnings messages and the theoretically simpler null or random intercept models which can have a more complicated hierarchical structure with a larger number of geographic scales. Depending on the purpose of the model a choice could be made between a two-level random coefficient model which might fit the data better or a four-

level random intercept model that may fit the data less well but provide useful information about the relationships between the outcome and predictor variable(s) at more geographic scales.

7.3.4 Summary

This project has addressed the aims of assessing which geographic scales are the most appropriate for researchers to include in their models of labour market and related socioeconomic statistics by means of comparing proportions of variance at different geographic scales and by comparing the fit of multilevel models that used different geographic scales as the levels. The project has shown that the geographic scales to include in models vary depending on the labour market outcomes being modelled but tend to be finer than the NUTS 1 regions scale. A comparison of the findings from the project with the creation and use of travel to work areas by other researchers and organisations is included in this thesis. Guidance to other researchers on steps to use to help choose which geographic scales to include in their multilevel models of labour market outcomes is provided in the toolkit in section 7.3.5 below.

7.3.5 Toolkit for other researchers: How to choose which geographic scales to use in multilevel modelling of labour market and related socioeconomic statistics

Step 1. Examine potential areal outcome variables carefully and consider removing outliers as they can affect whether or not multilevel models converge successfully. If a potential outcome variable consists of median values for each area then consider whether mean values have a more regular distribution suggesting that they will lead to better fitting models.

Step 2. For each outcome variable, create a null multilevel model with all geographic scales under consideration. For example, the model may contain several hierarchical levels. Calculate VPC values for each level to see which levels have <u>some</u>, i.e. non-negligible, variance. Calculate AIC values for null models with different geographic scales and use the models with the lowest AIC values to indicate with geographic scales to use as levels in subsequent random intercept and random coefficient models.

Step 3. Consider the intended purpose of the model. If the purpose is to study the geographic scales/distances for which predictor variables have effects then build random intercept models that includes all the levels that have <u>some</u> variance. If the purpose is to see whether influences have the same strength in different regions of the study area then build two-level models with random coefficients for those levels with the most variation, e.g. NUTS 1 areas or NUTS 2 areas or NUTS 3 areas in the context of this project, or add random coefficients at just one level at a time to multilevel models with several levels.

Step 4. Compare the AIC values for different models. Those models with lower AIC values fit the data better than those with higher AIC values. Hence the geographic scales included in the models with the lowest AIC values are the scales which create the better fitting models and are therefore the geographic scales which researchers should include in their models of labour market statistics.

Step 5 (optional). Consider extracting the coefficients from the model and comparing them for different regions of the study area (which could include mapping the coefficients and observing patterns across the study area).

Bibliography

Anselin, L., Cohen, J., Cook, D., Gorr, W. and Tita, G. (2000) 'Spatial Analyses of Crime', *Criminal Justice,* 4(Measurement and Analysis of Crime and Justice), pp. 213-262.

Arrow, J., O (1996) 'Estimating the influence of health as a risk factor on unemployment: A survival analysis of employment durations for workers surveyed in the German Socio-Economic Panel (1984-1990)', *Social Science & Medicine,* 42, pp. 1651-1659.

Ashworth, M. and Armstrong, D. (2003) 'Partnership effects in general practice: identification of clustering using intra-class correlation coefficients', *British Journal of General Practice,* November 2003.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B. and Scheipl, F. (2020) *Package lme4, Linear Mixed-Effects Models using 'Eigen' and S4*: R Cran. Available at: https://cran.r-project.org/web/packages/lme4/lme4.pdf.

Belkin, M., Hsu, D., Ma, S., Mondal, S. (2019) *Reconciling modern machine learning practice and the bias-variance trade off.* Proceedings of the National Academy of Sciences of the United States of America. Available at: https://www.pnas.org/content/116/32/15849Bell, D. and Blanchflower, D. (2010) 'UK Unemployment in the Great Recession', *National Institute Economic Review,* 214.

Bell, D. and Blanchflower, D. (2011) 'Young People and the Great Recession', *Oxford Review of Economic Policy,* 27, pp. 241-267.

Borg, I. and Brandén, M. (2018) 'Do high levels of home-ownership create unemployment? Introducing the missing link between housing tenure and unemployment', *Housing Studies,* 33, pp. 501-524.

Browne, W., J,, Subramanian, S., V,, Jones, K. and Goldstein, H. (2005) 'Variance partitioning in multilevel logistic models that exhibit overdispersion', *Journal of the Royal Statistical Society, A,* 168, pp. pp. 599-613.

Browne, W. J., Goldstein, H. and Rasbash, J. (2001) 'Multiple membership multiple classification (MMMC) models', *Statistical Modelling,* 1, pp. 21.

Burns, R. and Burns, R. (2008) *Business research methods and statistics using SPSS.* Los Angeles ; London: SAGE.

CACI (2019) *What is Acorn?* Available at: https://acorn.caci.co.uk/.

Cartmel, F., Furlong, A. (2000) *Youth Unemployment in Rural Areas*, York: Joseph Rowntree Foundation.

Chen, D.-R. and Truong, K. (2012) 'Using multilevel modelling and geographically weighted regression to identify spatial variations in the relationship between place-level disadvantages and obesity in Taiwan', *Applied Geography,* 32, pp. 8.

Clark, K. and Summers, L. (1982) 'The Dynamics of Youth Unemployment', in Freeman, R. and Wise, D. (eds.) *The Youth Labor Market Problem: Its Nature, Causes, and Consequences*. Chicago: University of Chicago Press.

CMM (2019a) *Centre for Multilevel Modelling*. Available at: http://www.bristol.ac.uk/cmm/.

CMM (2019b) *MLwiN software*. Available at: http://www.bristol.ac.uk/cmm/software/mlwin/.

Commission for Rural Communities (CRC). (2012) *Barriers to Education, Employment and Training for Young People in Rural Areas*, Gloucester, CRC.

Coombes, M., Wymer, C., Casado-Diaz, J., M, Martinez-Bernabeu, L. and Carausa, F. (2012) *Study on comparable Labour Market Areas*: University of Newcastle.

Cribb, J., Emmerson, C., Sibieta, L. (2015) *Public sector pay in the UK*, IFS Report R97, Available from: http://www.ifs.org.uk.

Culliney, M. (2016) *Escaping the rural pay penalty: location, migration and the labour market*, Work, Employment and Society. Vol 31, Issue 3, pp. 429-446.

de Smith, Goodchild, M., F and Longley (2015) *Geospatial Analysis - 5th Edition* Available at: http://www.spatialanalysisonline.com/HTML/?global_spatial_autocorrelation.htm 2018).

DEFRA (2005) *DEFRA Classification of Local Authority Districts and Unitary Authorities in England – A Technical Guide*. Available at: Microsoft Word - LAClassifications_technicalguide.doc (nationalarchives.gov.uk)

DEFRA (2013a) *Urban and Rural Area Definitions for Policy Purposes in England and Wales: Methodology (v1.0)*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/239477/RUC11methodologypaperaug_28_Aug.pdf.

DEFRA (2013b) *The 2011 Rural-Urban Classification For Small Area Geographies: A User Guide and Frequently Asked Questions (v1.0)*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/239478/RUC11user_guide_28_Aug.pdf.

Deller, S. and Watson, P. (2015) 'Spatial Variations in the relationship between economic diversity and stability', *Applied Economics Letters*.

Department for Communities and Local Government (2015a) *English Indices of Deprivation 2015, File 10 Local Authority District Summaries*. Available at: http://www.gov.uk/government/statistics/english-indices-of-deprivation-2015.

Department for Communities and Local Government (2015b) *The English Index of Multiple Deprivation (IMD) 2015 – Guidance*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/464430/English_Index_of_Multiple_Deprivation_2015_-_Guidance.pdf.

Department for Employment and Learning and Department of Enterprise Trade and Investment (2015) *Enabling Success' Supporting the succession from economic inactivity to employment: A strategy to tackle economic inactivity in Northern Ireland, version 1.0*. Available at: https://www.economy-ni.gov.uk/publications/enabling-success-strategy-tackle-economic-inactivity-northern-ireland#:~:text=Enabling%20Success%20is%20the%20new%20strategy,for%20reducing%20economic%20inactivity%20in%20Northern%20Ireland.

Diggle, P. (2013) *Statistical analysis of spatial and spatio-temporal point patterns.* 3rd ed. edn. Boca Raton, FL: CRC Press, 2013.

Dong, G. and Harris, R. (2015) 'Spatially Autoregressive Models for Geographically Hierarchical Data Structures', *Geographical Analysis,* 47(2), pp. 173-191.

Dong, G., Harris, R. and Mimis, A. (2016) 'HSAR: An R Package for Integrated Spatial Econometric and Multilevel Modelling', *GISRUK 2016*.

Edward, C., N and Fisher, L., M (2009) 'Housing tenure and labor market impacts: The search goes on', *Journal of Urban Economics,* 65.3, pp. 252-264.

Eurostat (2018) *NUTS, background*. Available at: http://ec.europa.eu/eurostat/web/nuts/background 2018).

Experian (2018) *Mosaic: The consumer classification solution for consistent cross-channel marketing*. Available at: https://www.experian.co.uk/assets/marketing-services/brochures/mosaic_uk_brochure.pdf.

Fabozzi, F., J, Focardi, S., M, Rachev, S., T and Arshanapalli, B., G (2014) 'Appendix E, Model Selection Criterion: AIC and BIC', *The Basics of Financial Econometrics: Tools, Concepts, and Asset Management Applications: Vol. John Wiley and Sons, inc*.

Felix, A. (2012) 'Industrial Diversity, Growth and Volatility in the Seven States of the Tenth District', *Economic Review,* Fourth Quarter.

Felstead, A., Gallie, D., Green, F., Inanc, H. (2013) *Skills at Work in Britain: First Findings from the Skills and Employment Survey 2012*. London: Centre for Life Chances in Knowledge Economies and Societies, Institute of Education.

Feser, E., Mix, T., White, M., Poole, K., Markley, D. and Pages, E. (2014) *Economic Diversity in Appalachia*: Appalachian Regional Commission. Available at: http://economicdiversityinappalachia.creconline.org/ 2018).

Fielding, A. and Goldstein, H. (2006) *Cross-classified and Multiple Membership Structures in Multilevel Models: An Introduction and Review, Department for Education and Skills Research Report RR791*.

Fisher, R. A. (1925) *Statistical Methods for Research Workers.*

Flowerdew, R., Geddes, A. and Green, M. (2001) 'Behaviour of Regression Models under Random Aggregation', in Tate N, J. and Atkinson, P., M (eds.) *Modelling Scale in Geographic Information Science*: Wiley.

Fotheringham, A., S (2016) *Local Modelling: An Introduction, presentation at the Glasgow Urban Big Data Centre*. Available at: http://www. ubdc.ac.uk/media/1434/glasgow-lecture-1.pdf 2018).

Fotheringham, A., S, Brunsdon, C. and Charlton, M. (2000) *Quantitative Geography.* SAGE.

Freeman, D., G (2012) 'On (Not) Closing
the Gaps: The Evolution of National and Regional Unemployment Rates by Race and Ethnicity', *Review of Black Political Economy,* 39, pp. 267-284.

Gale, C. and Longley, P. (2013) 'The 2011 Output Area Classification for the UK'.

Glo.be (2019) *ILO Centenary: how the organisation was founded.* Available at: https://www.glo-be.be/en/articles/ilo-centenary-how-organisation-was-founded.

Goldstein, H. (1995) *Multilevel statistical models. Kendall's library of statistics* 2nd edn. London: Edward Arnold.

Goldstein, H., Browne, W., and Rasbash, J. (2002) Partitioning Variance in Multilevel Models. *Understanding Statistics*, 1:4, p.p. 223-231.

Goldstein, H., Huiqi, P. and Rath, T. (2000) *The use of value added information in judging school performance.* Institute of Education.

Gomez-Rubio, V., Best, N., Richardson, S. and Clark, P. (2008) *Bayesian Statistics for Small Area Estimation, presentation at GSS Methodology Conference 2008*. Available at: http://www.bias-project.org.uk/papers/GSS08.pdf. Available at: http://www.bias-project.org.uk/papers/BayesianSAE.pdf.

Gov.uk (2020) *Ethnicity Facts and Figures: Employment*. Available at: https://www.ethnicity-facts-figures.service.gov.uk/work-pay-and-benefits/employment/employment/latest.

Greig-Smith, P. (1952) 'The use of random and contiguous quadrats in the study of the structure of plant communities', *Annals of Botany,* 16(2).

Hedges, L., V, and Hedburg, E., C (2007) 'Intraclass correlation values for planning group-randomized trials in education', *Educational Evaluation and Policy Analysis,* 29, pp. 60-87.

Hirschman, A. O. (1964) *The Paternity of an Index*, The American Economic Review, 54(5).

House of Commons Library (2019) *Labour market statistics: UK regions and countries* (CBP-7950).

Hox, J. (2002) *Multilevel Analysis : Techniques and Applications.* Lawrence Erlbaum Associates.

Huang, F., L (2018) 'Multilevel Modeling Myths', *School Psychology Quarterly,* 33(3), pp. 429-499.

ILO (2015) 1919 *ILO between the two world wars. Available at:* https://www.ilo.org/legacy/english/lib/century/

ILO (2019) *About the ILO*. Available at: https://www.ilo.org/global/about-the-ilo/lang--en/index.htm.

Improvement and Development Agency (2010) *Tackling worklessness: How to help people into employment in rural areas*. Available at: http://www.employabilityinscotland.com/media/121038/how_to_help_people_into_employment_in_rural_areas.pdf.

Jackson, R. W. (1984) 'An Evaluation of Alternative Measures of Regional Industrial Diversification. Regional Studies', *Regional Studies,* 18, pp. 10.

Kingston, G., McGinnity, F., O'Connell and P (2015) 'Discrimination in the labour market: Nationality, ethnicity and the
recession', *Work, Employment & Society,* 29, pp. 213-232.

Koo, T. K. and Li, M. Y. (2016) 'A guideline of selecting and reporting intraclass correlation coefficients for reliability research', *Journal of Chiropractic Medicine,* 15, pp. 155-163.

Kreft, I. and de Leeuw, J. (1998) *Introducing multilevel modeling (Introducing statistical methods series)* London: SAGE.

Lee, V., E, (2000) 'Using Hierarchical Linear Modelling to Study Social Contexts: The Case of School Effects', *Educational Psychologist,* 35(2), pp. 125-141.

Lindstrom, M., Moghaddassi, M. and Merlo, J. (2003) 'Social capital and leisure time physical activity: a population based multilevel analysis in Malmo, Sweden', *Journal of Epidemiology and Community Health,* 57, pp. 23-28.

London Data Store (2020) *London Living Wage (LLW)*. Available at: https://data.london.gov.uk/economic-fairness/labour-market/london-living-wage-llw/ (Accessed: 17 June 2020.

Longley, P., A, Goodchild, M., F., Maguire, D., J and Rhind, D., W (2005) *Geographic Information Systems and Science, Second Edition.* Wiley.

Ma, A., Phimister, E. and Battu, H. (2008) 'Housing Tenure, Job Mobility and Unemployment in the UK', *Economic Journal,* 118, pp. 311-328.

Ma, Z. (2015) 'Simpson's paradox in GDP and per capita GDP growths', *Empirical Economics,* 49.

Martin, D., Nolan, A. and Tranmer, M. (2001) 'T*he Application of Zone-Design Methodology in the 2001 UK Census*', *Environment and Planning A: Economy and Space,* 33(11), pp. 1949-1962.

Mason, S. and Howard, D. 2010. Regional Industrial Diversity and its Impact of Regional Unemployment. *Postprint of Proceeding of 12th Path to Full Employment/17th National Unemployment conference, Newcastle, New South Wales.*

Mayor of London/London Assembly (2020) *London Living Wage*. Available at: https://www.london.gov.uk/what-we-do/business-and-economy/london-living-wage (Accessed: 17 June 2020.

McLaughlin, G. E. (1930) 'Industrial Diversification in American Cities', *Quarterly Journal of Economics,* 45, pp. 19.

Moellering, H. and Tobler, W. (1972) 'Geographical Variances', *Geographic Analysis,* 4, pp. 17.

Moerbeek, M. (2004) 'The consequences of ignoring a level of nesting in multilevel analysis', *Multivariate Behavioural Research,* 39(1).

Moran (1948) 'The Interpretation of Statistical Maps', *Journal of the Royal Statistical Society, Series B,* 10, pp. 243-251.

Murray, E., Zaninotto, P., Fleischmann, M., Stafford, M., Carr, E., Shelton, N., Stansfeld, S, Kuh, D. and Head, J. (2019) 'L*inking local labour market conditions across the life course to retirement age: Pathways of health, employment status, occupational class and educational achievement, using 60 years of the 1946 British Birth Cohort*', *Social Science & Medicine,* 226, pp. 113-122.

Nalborczyk, L. (2017) 'Three methods for computing the intra-class correlation in multilevel logistic regression'. Available at: https://www.barelysignificant.com/post/icc/ 2020].

National Records of Scotland (2018) *Scotland's Census: Microdata Teaching File*. Available at: https://www.scotlandscensus.gov.uk/microdata-teaching-file.

Nezlek, J., B (2008) 'An Introduction to Multilevel Modeling for Social and Personality Psychology', *Social and Personality Psychology Compass,* 2, pp. 842-860.

Nomis (2019) *Official labour market statistics service provided by the ONS*. Available at: https://www.nomisweb.co.uk/.

Northern Ireland Statistics and Research Agency (2018) *Specialist products: Microdata teaching file*. Available at: https://www.nisra.gov.uk/statistics/2011-census/results/specialist-products.

O'Donoghue, D. (1999) 'The Relationship Between Diversification and Growth: Some Evidence from the British System 1978 to 1991', *International Journal of Urban and Regional Research,* 1999.

ONS (2003) *Development of improved estimation measures for local area unemployment levels and rates, Labour Market Trends, 37*. Available at: http://www.nomisweb.co.uk/articles/news/files/lmt0301_rates.pdf.

ONS (2006) *Model Based Estimates of ILO Unemployment for LAD/UAs in Great Britain, Guide for Users*. Available at:
https://www.ons.gov.uk/file?uri=/employmentandlabourmarket/peopleinwork/employment andemployeetypes/methodologies/subnationallabourmarket/modelbasedestimatesu nemploymenttcm77252969.pdf.

ONS (2012) *Annual Population Survey (APS) QMI*. Available at:
https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentande mployeetypes/methodologies/annualpopulationsurveyapsqmi.

ONS (2014) *2011 rural/urban classification*. Available at:
https://www.ons.gov.uk/methodology/geography/geographicalproducts/ruralurbanclassi fications/2011ruralurbanclassification. Available at:
http://geoportal.statistics.gov.uk/datasets/rural-urban-classification-2011-of-local-authority-districts-in-england.

ONS (2015a) *Opinions and Lifestyle Survey: 2013*. Available at:
https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthan dlifeexpectancies/compendium/opinionsandlifestylesurvey/2015-03-19.

ONS (2015b) *Travel to Work Areas*. Available at:
https://webarchive.nationalarchives.gov.uk/20160106004211/http://www.ons.gov.uk/on s/guide-method/geography/beginner-s-guide/other/travel-to-work-areas/index.html.

ONS (2015c) *Overview of 2011 Travel to Work Areas*. Available at:
https://webarchive.nationalarchives.gov.uk/20160106004211/http://www.ons.gov.uk/on s/guide-method/geography/beginner-s-guide/other/travel-to-work-areas/index.html 2020).

ONS (2015d) *Methodology note on 2011 Travel to Work Areas*. Available at:
https://webarchive.nationalarchives.gov.uk/20160106004211/http://www.ons.gov.uk/on s/guide-method/geography/beginner-s-guide/other/travel-to-work-areas/index.html.

ONS (2015e) *2011 Travel to Work Areas Summary Statistics*. Available at:
https://webarchive.nationalarchives.gov.uk/20160106004211/http://www.ons.gov.uk/on s/guide-method/geography/beginner-s-guide/other/travel-to-work-areas/index.html.

ONS (2016) *Output Area production: An overview of the system used to create Output Areas*. Available at:
https://www.ons.gov.uk/census/2001censusandearlier/dataandproducts/outputgeograp hy/outputareaproduction (Accessed: 15 June 2020.

ONS (2016a) *Annual Survey of Hours and Earnings (ASHE) methodology and guidance*. Available at:
https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworki nghours/methodologies/annualsurveyofhoursandearningsashemethodologyandguidance.

ONS (2016aa) *Census Geography: An overview of the various geographies used in the production of statistics collected via the UK Census*. Available at:
https://www.ons.gov.uk/methodology/geography/ukgeographies/censusgeography#out put-area-oa.

ONS (2016b) *Business Register and Employment Survey (BRES)*. Available at: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/methodologies/businessregisterandemploymentsurveybres.

ONS (2016c) *Travel to work area analysis in Great Britain: 2016*. Available at: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/articles/traveltoworkareaanalysisingreatbritain/2016.

ONS (2017) *2011 Area Classification for Local Authorities, revised 15 September 2017*. Available at: https://www.ons.gov.uk/methodology/geography/geographicalproducts/areaclassifications/2011areaclassifications/datasets.

ONS (2017b) *The Spatial Distribution of Industries in Great Britain: 2015*.

ONS (2017d) *Exploring labour productivity in rural and urban areas in Great Britain: 2014*. Available at: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/labourproductivity/articles/exploringlabourproductivityinruralandurbanareasingreatbritain/2014.

ONS (2018) *UNEM02: Unemployment by previous occupation*. Available at: https://www.ons.gov.uk/employmentandlabourmarket/peoplenotinwork/unemployment/datasets/unemploymentbypreviousoccupationunem02.

ONS (2018c) *2011 Census microdata teaching set*. Available at: https://www.ons.gov.uk/census/2011census/2011censusdata/censusmicrodata/microdatateachingfile 2018).

ONS (2019a) *A Guide to Labour Market Statistics*. Available at: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/methodologies/aguidetolabourmarketstatistics.

ONS (2019b) *Despite record employment, nearly 10% of adults have never done paid work*. Available at: https://www.ons.gov.uk/employmentandlabourmarket/peoplenotinwork/economicinactivity/articles/despiterecordemploymentnearly10ofadultshaveneverdonepaidwork/2019-02-28.

ONS (2019c) *LI01 Regional labour market: Local indicators for counties, local and unitary authorities*. Available at: https://tinyurl.com/Previous-Local-Indicators. Available at: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/datasets/locallabourmarketindicatorsforcountieslocalandunitaryauthoritiesli01/current.

ONS (2019d) *All data related to Regional labour market statistics in the UK: April 2019*. Available at: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/bulletins/regionallabourmarket/april2019/relateddata?:uri=employmentandlabourmarket/peopleinwork/employmentandemployeetypes/bulletins/regionallabourmarket/april2019/relateddata&:uri=employmentandlabourmarket/peopleinwork/employmentandemployeetypes/bulletins/regionallabourmarket/april2019/relateddata&:uri=em

ploymentandlabourmarket/peopleinwork/employmentandemployeetypes/bulletins/regionallabourmarket/april2019/relateddata&:uri=employmentandlabourmarket/peopleinwork/employmentandemployeetypes/bulletins/regionallabourmarket/april2019/relateddata&page=3.

ONS (2019e) *M01 Regional labour market: Modelled unemploy*ment for local and unitary authorities.

ONS (2019f) Median weekly earnings of full-time workers by workplace, *LI01 Local labour market indicators by unitary and local authority (data from Annual Survey* of *Hours and Earnings (ASHE))*.

ONS (2019g) *EARN06: Gross Weekly Earnings by Occupation*. Available at: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/datasets/grossweeklyearningsbyoccupationearn06.

ONS (2019h) *2011 Area Classifications*. Available at: https://www.ons.gov.uk/methodology/geography/geographicalproducts/areaclassifications/2011areaclassifications.

ONS (2019i) *Health State Life Expectancies by national deprivation indices deciles, England and Wales, 2015 to 2017*. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthinequalities/bulletins/healthstatelifeexpectanciesbyindexofmultipledeprivationimd/2015to2017.

ONS (2019j) *Socioeconomic inequalities in avoidable mortality, England and Wales: 2001 to 2017*. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/causesofdeath/articles/measuringsocioeconomicinequalitiesinavoidablemortalityinenglandandwales/2001to2017.

ONS (2019k) *Labour Market Overview, UK: March 2019*. Available at: http://Labour market overview, UK - Office for National Statistics (ons.gov.uk).

ONS (June 2018) *UK mid-year 2017 population estimates*. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland.

ONS (October 2018) *Earnings and hours worked, place of work by local authority: ASHE Table 7*
Earnings and hours worked, place of residence by local authority: ASHE Table 8. Available at:
https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/datasets/placeofworkbylocalauthorityashetable7. Available at:
https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/datasets/placeofresidencebylocalauthorityashetable8.

ONS Open Geography Portal (2019) *Search the Open Geography portal*. Available at: http://geoportal.statistics.gov.uk/.

ONS UK data service (2017) *InFuse*. Available at: http://infuse.ukdataservice.ac.uk/ 2017).

ONS UK Geographies (2019) *Detailed information on the administrative structure within England*. Available at: https://www.ons.gov.uk/methodology/geography/ukgeographies/administrativegeography/england (Accessed: July 2019).

Opdenakker, M. C. and Van Damme, J. (2000) 'The importance of identifying levels in multilevel analysis', *School Effectiveness and School Improvement,* 11(1).

Openshaw, S. (1984) *The modifiable areal unit problem. Catmog.* Norwich: Geo.

Oxford Consultants for Social Inclusion (2010) *Why IMD are still useful in era of open data*. Available at: https://ocsi.uk/2011/03/24/why-the-imd-is-still-important-in-the-open-data-age/.

Oxford Consultants for Social Inclusion (2018) *Indices of Deprivation*. Available at: https://ocsi.uk/indices-of-deprivation/.

Park, Y., M and Kim, Y. (2014) 'A spatially filtered multilevel model to account for spatial dependency: application to self-rated health status in South Korea', *International Journal of Health Geographics,* 13(6).

Phimister, E., Theodossion, I., Upward, R., (2006) *Is it easier to escape from the low pay in urban areas? Evidence from the United Kingdom*. Environment and Planning A, 38, pp. 693-710.

R Cran Project (2019) *Package 'lme4' March 5 2019*. Available at: https://cran.r-project.org/web/packages/lme4/lme4.pdf (Accessed: 13/3/2020.

Robson, K. and Pevalin, D., J (2016) *Multilevel modeling in plain language.*

Robson, M., T (2003) 'Housing Markets and Regional Unemployment Flows in Great Britain', *Manchester School,* 71.

Rossbacher, L., A (1986) 'Nearest-Neighbour Analysis: A Technique for Quantitative Evaluation of Polygonal Ground Patterns', *Geografiska Annaler. Series A, Physical Geography,* 68, pp. 101-105.

Shahidi, F. V., Siddiqi, A. and Muntaner, C. (2016) 'Does social policy moderate the impact of unemployment on health? A multilevel analysis of 23 welfare states', *The European Journal of Public Health,* 26, pp. 1017-1022.

Shrout, P. E. and Fleiss, J. L. (1979) 'Correlations: Uses in Assessing Rater Reliability', *Psychological Bulletin,* 86, pp. 420-428.

Simpson, E., H (1951) 'The Interpretation of Interaction in Contingency Tables', *Journal of the Royal Statistical Society, Series B,* 13, pp. 238–241.

Skrondal, A. and Rabe-Hesketh, S. (2007) 'Redundant overdispersion parameters in multilevel models for categorical responses', *Journal of Educational and Behavioural Statistics,* 32, pp. pp. 419-430.

Snijders, T., A, B and Bosker, R., J (2012) *Multilevel Analysis An Introduction to Basic and Advanced Multilevel Modeling.* Second edn.: Sage.

Somerset Intelligence (2019) *Unemployment and Economic Inactivity: Rural-Urban Analysis*. Available at: http://www.somersetintelligence.org.uk/unemployment-and-economic-inactivity-rural-urban-analysis/ 2019).

Sommet, N. and Morselli, D. (2017) 'Keep Calm and Learn Multilevel Logistic Modeling: A Simplified Three-Step Procedure Using Stata, R, Mplus, and SPSS', *International Review of Social Psychology,* 30, pp. 203-218.

SpringerOpen (2020) *Journal for Labour Market Research*. Available at: https://labourmarketresearch.springeropen.com/about.

Support, U. D. S. C. (2020a) *Census aggregate data guide*. Available at: https://census.ukdataservice.ac.uk/use-data/guides/aggregate-data.aspx 2020).

Support, U. D. S. C. (2020b) *Census microdata guide*. Available at: https://census.ukdataservice.ac.uk/use-data/guides/microdata.aspx 2020).

Titheridge, H., Christie, N., Mackett, R., Hernan, D. O. and Ye, R. (2014) *Transport and Poverty, A review of the evidence*: UCL. Available at: https://www.ucl.ac.uk/transport/sites/transport/files/transport-poverty.pdf.

Tobler, W., R 'A Computer Movie Simulating Urban Growth in the Detroit Region', *Economic Geography,* 46, Supplement: Proceedings. International Geographical Union. Commission on Quantitative Methods (June 1970), pp. 234-240.

Townsend, A. (2019) *Combined Authorities for more sub-regions? – Learning the adverse lesson from England beyond the metropolitan conurbations.* Local Economy Vol 34 (2), pp. 123-138

Tran, H. P. D. (2011) *Industrial Diversity and Economic Performance: A Spatial Analysis.* PhD, University of Nebraska-Lincoln [Online] Available at: https://digitalcommons.unl.edu/cgi/viewcontent.cgi?referer=https://www.bing.com/&httpsredir=1&article=1022&context=businessdiss (Accessed.

Tranmer, M. and Steel, D. G. (2001a) 'Using local Census data to investigate scale effects', in Tate, N. and Alkinson, P. (eds.) *Modelling Scale in Geographic Information Science*: Wiley.

Tranmer, M. and Steel, D. G. (2001b) 'Ignoring a level in a multilevel model: evidence from UK Census data', *Environment & Planning A,* 33(5), pp. 941-948.

Trendle, B. and Shorney, G. (2004) *The Effect of Industrial Diversification on Regional Economic Performance, Working Paper No 20*: Queensland Government.

Tzavidis, N., Zhang, L.-C., Luna, A., Schmid, T. and Rojas-Perilla, N. (2018) 'From start to finish: a framework for the production of small area official statistics', *J. R. Statist. Soc. A,* 181, pp. 927–979.

UK Data Service (2019) *InFuse, 2011 Geographies*. Available at: http://infuse.mimas.ac.uk/help/definitions/2011geographies/ (Accessed: 12 July 2019).

UK Data Service Census Support (2020a) *Census aggregate data guide*. Available at: https://census.ukdataservice.ac.uk/use-data/guides/aggregate-data.aspx 2020).

UK Data Service Census Support (2020b) *Census microdata guide*. Available at: https://census.ukdataservice.ac.uk/use-data/guides/microdata.aspx 2020).

UN (2019) *E-handbook on SDG Indicators*. Available at: https://unstats.un.org/wiki/display/SDGeHandbook/Indicator+8.5.2 (Accessed: 7 July 2019).

United States Census Bureau (2011) *What are Census blocks?* Available at: https://www.census.gov/newsroom/blogs/random-samplings/2011/07/what-are-census-blocks.html 2020).

United States Census Bureau (2020) *Glossary*. Available at: https://www.census.gov/programs-surveys/geography/about/glossary.html#par_textimage_13.

Van den Noortgate, W., Opdenakker, M.-C. and Onghena, P. (2005) 'The effects of ignoring a level in a multilevel analysis', *School Effectiveness and School Improvement,* 16(3).

Van Landeghem, G., Fraine, D. and Van Damme, J. (2005) 'The consequences of ignoring a level of nesting in multilevel analysis: a comment', *Multivariate Behavioral Research,* 40(4).

Virtanen, P., Janlert, U. and Hammarstrom, A. (2013) 'Health Status and health behaviour as predictors of the occurrence of unemployment and prolonged unemployment', *Public Health,* 127.

Ward, C. and Dale, A. (1992) 'Geographical Variation in Female Labour Force Participation: An Application of Multilevel Modelling', *Regional Studies,* 26, pp. 243-255.

Wilson, R. (2013) 'Changing Geographic Units and the Analytical Consequences: An Example of Simpson's Paradox', *Cityscape,* 15(2).

Wu, S., Crespi, C. M. and Wong, W. K. (2012) 'Comparison of methods for estimating the intraclass correlation coefficient for binary responses in cancer prevention cluster randomized trials', *Contemporary Clinical Trials,* 33(5), pp. 869-880.

Yu, W. and Sun, S. (2019) 'Race-ethnicity, class, and unemployment dynamics: Do macroeconomic shifts alter existing disadvantages?', *Research in Social Stratification and Mobility,* 63.

Annex 1: Extracts of R code

1. Example of R code for multilevel models of Output Area data with pseudo level

> Over_Null_Logistic_01<-

glmer(**cbind(EconActive_16_74,EconInactive_16_74)**~(**1|GEO_CODE/OA.Data_01_GEO_LABEL**), data=data4regression, family = binomial)

Important points to note in this code are:-

cbind(EconActive_16_74,EconInactive_16_74) – the dependent/outcome variable which is made up by combining the number of people aged 16 to 74 who are economically active and the number who are economically inactive to make a variable suitable to model the proportions of people who are economically active.

(1|GEO_CODE/OA.Data_01_GEO_LABEL)

-the **1|** indicates a random (variable) intercept for each group of OAs in the model

-**GEO_CODE** is the identifier for the NUTS 1 areas

-**OA.Data_01_GEO_LABEL** is the OA level label which is used here to add a pseudo level to the model to try to take account of any overdispersion.

2. Null models for OA data without pseudo level

```
Null_Logistic_01<-glmer(cbind(EconActive_16_74,EconInactive_16_74)~
(1|GEO_CODE), data=data4regression, family = binomial )

Null_Logistic_02<-glmer(cbind(EconActive_16_74,EconInactive_16_74)~
(1|County.Code), data=data4regression, family = binomial )

Null_Logistic_03<-glmer(cbind(EconActive_16_74,EconInactive_16_74)~
(1|GEO_CODE/County.Code), data=data4regression, family = binomial )
```

In each of these models the cbind(EconActive_16_74,EconInactive_16_74) element of the code again provides the dependent variable.

In the first model (1|GEO_CODE) indicates that the OAs are to be grouped by the NUTS 1 areas

codes in the GEO_CODE column of the input data.

In the second model the (1|County.Code) indicates that the OAs are to be grouped by the county/unitary authority codes in the County.Code column of the input data.

In the third model the (1|GEO_CODE/County.Code) indicates that the OAs are to be grouped by the county/unitary authority codes and by the NUTS 1 areas codes.

To obtain the ICC proportions for each of the higher levels in the third model the icc automatic
command was adjusted by add 'by_group = TRUE' to split the total proportion of variance due to
the model grouping structure to the proportions at each higher level as shown below.

icc(Null_Logistic_03, by_group = TRUE)

3. Random intercept models for OA data

Logistic_04<-glmer(cbind(EconActive_16_74,EconInactive_16_74)~**OA_PropHealthy** + (1|GEO_CODE/County.Code), data=data4regression, family = binomial)

Logistic_05<-glmer(cbind(EconActive_16_74,EconInactive_16_74)~**OA_PropDegreePlus** + (1|GEO_CODE/County.Code), data=data4regression, family = binomial)

4. Random coefficient models for OA data

Null_Logistic_06<-glmer(cbind(EconActive_16_74,EconInactive_16_74)~**OA_PropHealthy** + (1+**OA_PropHealthy**|GEO_CODE/County.Code), data=data4regression, family = binomial)

Logistic_07<-glmer(cbind(EconActive_16_74,EconInactive_16_74)~**OA_PropDegreePlus** + (1+**OA_PropDegreePlus**|GEO_CODE/County.Code), data=data4regression, family = binomial)

5. Microdata models

Null microdata model code and output

```
MLM_00<-glmer(Unemployed~(1|Region), family= binomial("logit"), data
=adultUsualResidents)
> summary(MLM_00)
Generalized linear mixed model fit by maximum likelihood (Laplace Ap
proximation) ['glmerMod']
 Family: binomial  ( logit )
Formula: Unemployed ~ (1 | Region)
   Data: adultUsualResidents

     AIC      BIC   logLik deviance df.resid
147707.1 147729.0 -73851.6 147703.1   411624

Scaled residuals:
    Min      1Q  Median      3Q     Max
-0.2399 -0.2310 -0.2169 -0.1918  5.3221

Random effects:
 Groups Name         Variance Std.Dev.
 Region (Intercept) 0.0257   0.1603
Number of obs: 411626, groups:  Region, 10

Fixed effects:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -3.07977    0.04795  -64.23   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> icc(MLM_00)
# Intraclass Correlation Coefficient

     Adjusted ICC: 0.008
   Conditional ICC: 0.008
```

Random intercept microdata model code and output

```
> MLM_01<-glmer(Unemployed~Healthy + (1|Region), family =
binomial("logit"), data=adultUsualResidents)
> summary(MLM_01)
Generalized linear mixed model fit by maximum likelihood (Laplace
Approximation) ['glmerMod']
 Family: binomial  ( logit )
Formula: Unemployed ~ Healthy + (1 | Region)
   Data: adultUsualResidents

     AIC      BIC   logLik deviance df.resid
147450.8 147483.5 -73722.4 147444.8   411623

Scaled residuals:
    Min      1Q  Median      3Q     Max
-0.2441 -0.2345 -0.2209 -0.1937  7.2479

Random effects:
 Groups Name         Variance Std.Dev.
 Region (Intercept) 0.02691  0.164
Number of obs: 411626, groups:  Region, 10
```

241

```
Fixed effects:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -3.69246    0.05975   -61.8   <2e-16 ***
Healthy      0.64041    0.04079    15.7   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
        (Intr)
Healthy -0.568
> icc(MLM_01)
# Intraclass Correlation Coefficient

      Adjusted ICC: 0.008
   Conditional ICC: 0.008
```

Random coefficient model code and output

> MLM_02<-glmer(Unemployed~**Healthy** + (1+**Healthy**|Region), family = binomial("logit"),
data=adultUsualResidents)

boundary (singular) fit: see ?isSingular

```
> summary(MLM_02)
Generalized linear mixed model fit by maximum likelihood (Laplace
Approximation) ['glmerMod']
 Family: binomial  ( logit )
Formula: Unemployed ~ Healthy + (1 + Healthy | Region)
   Data: adultUsualResidents

     AIC       BIC    logLik deviance df.resid
147447.8 147502.4 -73718.9 147437.8    411621
Scaled residuals:
    Min       1Q  Median      3Q      Max
-0.2452 -0.2336 -0.2222 -0.1927   6.2864

Random effects:
 Groups Name          Variance Std.Dev. Corr
 Region (Intercept) 0.00000  0.0000
        Healthy      0.02846  0.1687    NaN
Number of obs: 411626, groups:  Region, 10


Fixed effects:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -3.67677    0.03982  -92.34   <2e-16 ***
Healthy      0.62630    0.06042   10.37   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
        (Intr)
Healthy -0.577
convergence code: 0
```
boundary (singular) fit: see ?isSingular

> icc(MLM_02)

[1] NA

Warning message:

Can't compute random effect variances. Some variance components equal zero.

  Solution: Respecify random structure!

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
      (Intr)
Healthy -0.568
> icc(MLM_01)

6. Example code for local authority null models for binomially distributed outcomes

NullModelaa<-glmer(**Output_variable**

~(**1|NUT115CD/NUTS215CD/NUTS315CD/LAD13CD**), data=England_LA_data, family =

binomial)

Important elements in the code above are:-

1|NUT115CD/NUTS215CD/NUTS315CD – which indicates a random (variable) intercept

with the local authority level dependent variable observations grouped by NUTS 1 areas,

NUTS 2 areas and NUTS 3 areas

/LAD13CD - the pseudo level at local authority level used to account of binomial

overdispersion

Output_variable – this was set earlier in the R script using the following lines of code as

appropriate for the unemployment and unemployment dependent variables:-

#For Employment model
Output_variable<-cbind(England_LA_data$Employed_16_64,

England_LA_data$Not_Employed_16_64)

#For Unemployment model
Output_variable<-cbind(England_LA_data$Unemp_16plus,

England_LA_data$Employed_16plus)

Summary information for these models was obtained using R code such as:-
summary(NullModelaa)

The ICC/VPC values giving the proportions of variance at each level in the different models

was extracted from the models using R code such as:

icc(NullModelaa, by_group = TRUE)

Code to set either the employment rate or unemployment rate as the dependent variable

Example of R code used to model employment and unemployment with the independent variable of the proportion of people with no qualifications used as the independent variable at four different geographic scales.

```
Output_variable<-cbind(England_LA_data$Employed_16_64,
England_LA_data$Not_Employed_16_64)
Output_variable<-cbind(England_LA_data$Unemp_16plus,
England_LA_data$Employed_16plus)
```


Code to run separate models with independent variables at different geographic scales

```
#Null four-level model for reference
NullModel<-glmer(Output_variable ~(1|NUT115CD/NUTS215CD/NUTS315CD/LAD13CD),
data=England_LA_data, family = binomial )
summary(NullModel)
AIC(NullModel)
#No qualifications as a predictor at different geographic scales
NoQuals_LA<-glmer(Output_variable ~PropNoQuals_LA +
(1|NUT115CD/NUTS215CD/NUTS315CD/LAD13CD), data=England_LA_data, family =
binomial)
AIC(NoQuals_LA)

NoQuals_NUTS3<-glmer(Output_variable ~PropNoQuals_NUTS3 +
(1|NUT115CD/NUTS215CD/NUTS315CD/LAD13CD), data=England_LA_data, family =
binomial)
AIC(NoQuals_NUTS3)

NoQuals_NUTS2<-glmer(Output_variable ~PropNoQuals_NUTS2 +
(1|NUT115CD/NUTS215CD/NUTS315CD/LAD13CD), data=England_LA_data, family =
binomial)
AIC(NoQuals_NUTS2)
```

```
NoQuals_NUTS1<-glmer(Output_variable ~PropNoQuals_NUTS1 +
(1|NUT115CD/NUTS215CD/NUTS315CD/LAD13CD), data=England_LA_data, family =
binomial)
AIC(NoQuals_NUTS1)


NoQuals_4Levels<-glmer(Output_variable ~PropNoQuals_LA + PropNoQuals_NUTS3 +
PropNoQuals_NUTS2 + PropNoQuals_NUTS1 +
(1|NUT115CD/NUTS215CD/NUTS315CD/LAD13CD), data=England_LA_data, family =
binomial)
AIC(NoQuals_4Levels)
```

7. Example code for local authority null models for Normally distributed outcomes

```
NullModela<-lmer(Output_variable ~(1|NUT115CD/NUTS215CD/NUTS315CD),
data=England_LA_data )
```

Elements of the code are:

lmer – a linear model (suitable for Normally distributed dependent variables)

(1|NUT115CD/NUTS215CD/NUTS315CD) – grouping of the level one units (local authorities) is by NUTS 1 areas, NUTS 2 areas and NUTS 3 areas

Output_variable – this is set to each appropriate Normally distributed dependent variable is turn to run the null model for that variable using one of the following lines of codes.

```
Output_variable<-England_LA_data$MeanHoursWorked
Output_variable<-England_LA_data$ResidentEarnings
Output_variable<-England_LA_data$WorkPlaceEarnings
Output_variable<-England_LA_data$JobDensity.without3outliers
```

As was the case for the case for models of binomially distributed dependent variables, summary information and ICC/VPC proportions were extracted from the models using code such as that shown below.

summary(NullModela)

icc(NullModela, by_group = TRUE)

8. Code to calculate an industrial diversity index measure at local authority level

##Herfindahl Index = chance of two randomly selected workers in an area being in the same industry

##Calculating Herfindahl_Index Industrial Diversity Indicator for Local Authorities

```
a<-100
b<-England_LA_data$LA_Percent_01_AgFish
c<-England_LA_data$LA_Percent_02_Mining
d<-England_LA_data$LA_Percent_03_Manu
e<-England_LA_data$LA_Percent_04_Construction
f<-England_LA_data$LA_Percent05_MotorTrades
g<-England_LA_data$LA_Percent06_Wholesale
h<-England_LA_data$LA_Percent_07_Retail
i<-England_LA_data$LA_Percent08_Transport
j<-England_LA_data$LA_Percent09_Accom
k<-England_LA_data$LA_Percent10_Information
l<-England_LA_data$LA_Percent11_Financial
m<-England_LA_data$LA_Percent12_Property
n<-England_LA_data$LA_Percent13_Professional
o<-England_LA_data$LA_Percent14_Business
p<-England_LA_data$LA_Percent15_PublicAdmin
q<-England_LA_data$LA_Percent16_Education
r<-England_LA_data$LA_Percent17_Health
s<-England_LA_data$LA_Percent18_Arts

Frequency_b<-b/a
Frequency_c<-c/a
Frequency_d<-d/a
Frequency_e<-e/a
Frequency_f<-f/a
Frequency_g<-g/a
Frequency_h<-h/a
Frequency_i<-i/a
Frequency_j<-j/a
Frequency_k<-k/a
```

```
Frequency_l<-l/a
Frequency_m<-m/a
Frequency_n<-n/a
Frequency_o<-o/a
Frequency_p<-p/a
Frequency_q<-q/a
Frequency_r<-r/a
Frequency_s<-s/a
Herfindahl_Index<-((Frequency_b)^2 + (Frequency_c)^2 + (Frequency_d)^2 +
(Frequency_e)^2 + (Frequency_f)^2 + (Frequency_g)^2 + (Frequency_h)^2 +
(Frequency_i)^2 + (Frequency_j)^2  + (Frequency_k)^2 + (Frequency_l)^2 +
(Frequency_m)^2 +
        (Frequency_n^2 + (Frequency_o)^2 + (Frequency_p)^2 + (Frequency_q)^2 +
(Frequency_r)^2 + (Frequency_s)^2 ))
Herfindahl_Index
England_LA_data$LA_IndustryDiversity<-Herfindahl_Index
```

9. Random intercept models for local authority outcomes

```
Predictor_LAa<-lmer(Output_variable ~Predictor_variable +
(1|NUT115CD/NUTS215CD/NUTS315CD), data=England_LA_data)
summary(Predictor_LAa)
install.packages('sjstats')
#library(sjstats)
icc(Predictor_LAa, by_group = TRUE)
```

For binomially distributed dependent variables the following R code was used.

```
#LOGISTIC MODELS
NullModela<-glmer(Output_variable ~(1|NUT115CD/NUTS215CD/NUTS315CD),
data=England_LA_data, family = binomial )
summary(NullModela)
AIC(NullModela)
icc(NullModela, by_group = TRUE)
```

10. Models for possible higher level

#fifth level model intended for earnings outcomes variables

England_LA_data$London<-0

England_LA_data$London<-as.numeric(England_LA_data$NUT115CD=="UKI")

sum(England_LA_data$London)

Null_Five_Levels<-lmer(Output_variable

~(1|London/NUT115CD/NUTS215CD/NUTS315CD), data=England_LA_data)

summary(Null_Five_Levels)

AIC(Null_Five_Levels)

icc(Null_Five_Levels, by_group = TRUE)

icc(Null_Five_Levels)

Null_London_or_Not<-lmer(Output_variable ~(1|London), data=England_LA_data)

summary(Null_London_or_Not)

AIC(Null_London_or_Not)

icc(Null_London_or_Not, by_group = TRUE)

icc(Null_London_or_Not)

#London and SouthEast combined v. rest of England

England_LA_data$London_SouthEast<-0

England_LA_data$London_SouthEast<-
as.numeric(England_LA_data$NUT115CD=="UKI"|England_LA_data$NUT115CD=="UKJ")

sum(England_LA_data$London_SouthEast)

Null_London_SouthEast<-lmer(Output_variable

~(1|London_SouthEast/NUT115CD/NUTS215CD/NUTS315CD), data=England_LA_data)

summary(Null_London_SouthEast)

AIC(Null_London_SouthEast)

icc(Null_London_SouthEast, by_group = TRUE)

icc(Null_London_SouthEast)

Null_LonSE_2level<-lmer(Output_variable ~(1|London_SouthEast), data=England_LA_data
)

summary(Null_LonSE_2level)

AIC(Null_LonSE_2level)

icc(Null_LonSE_2level, by_group = TRUE)

11. Code for random coefficient models

Models for Normal Variables

```
NullModela<-lmer(Output_variable ~(1|NUT115CD/NUTS215CD/NUTS315CD),

data=England_LA_data )

summary(NullModela)

AIC(NullModela)

icc(NullModela, by_group = TRUE)

NullModelb<-lmer(Output_variable ~ (1|NUTS315CD), data=England_LA_data)

AIC(NullModelb)

NullModelc<-lmer(Output_variable ~ (1|NUTS215CD), data=England_LA_data)

AIC(NullModelc)

NullModeld<-lmer(Output_variable ~ (1|NUT115CD), data=England_LA_data)

AIC(NullModeld)

icc(NullModeld, by_group=TRUE)


Predictor_LAa<-lmer(Output_variable ~Predictor_variable +

(1|NUT115CD/NUTS215CD/NUTS315CD), data=England_LA_data)

summary(Predictor_LAa)

install.packages('sjstats')

#library(sjstats)

icc(Predictor_LAa, by_group = TRUE)

icc(Predictor_LAa)

AIC(Predictor_LAa)

Predictor_LAb<-lmer(Output_variable ~Predictor_variable + (1|NUTS315CD),

data=England_LA_data)

AIC(Predictor_LAb)

Predictor_LAc<-lmer(Output_variable ~Predictor_variable + (1|NUTS215CD),

data=England_LA_data)

AIC(Predictor_LAc)

Predictor_LAd<-lmer(Output_variable ~Predictor_variable + (1|NUT115CD),

data=England_LA_data)

AIC(Predictor_LAd)
```

```
Random_Slopea<-lmer(Output_variable ~ Predictor_variable +

(1+Predictor_variable|NUT115CD/NUTS215CD/NUTS315CD), data=England_LA_data)

summary(Random_Slopea)

AIC(Random_Slopea)

plot(fitted(Random_Slopea), resid(Random_Slopea))


Random_Slopeb<-lmer(Output_variable ~ Predictor_variable +

(1+Predictor_variable|NUTS315CD), data=England_LA_data)

#summary(Random_Slopeb)

AIC(Random_Slopeb)

#plot(fitted(Random_Slopeb), resid(Random_Slopeb))


Random_Slopec<-lmer(Output_variable ~ Predictor_variable +

(1+Predictor_variable|NUTS215CD), data=England_LA_data)

#summary(Random_Slopec)

AIC(Random_Slopec)

#plot(fitted(Random_Slopec), resid(Random_Slopec))

#plot(observed(Random_Slopec), fitted(Random_Slopec))

#plot(Output_variable, fitted(Random_Slopec))


Random_Sloped<-lmer(Output_variable ~ Predictor_variable +

(1+Predictor_variable|NUT115CD), data=England_LA_data)

summary(Random_Sloped)

AIC(Random_Sloped)

#plot(fitted(Random_Sloped), resid(Random_Sloped))
```

Models binomial variables

```
## LOGISTIC MODELS WITH PSEUDO LEVEL TO USE ADDITIVE APPROACH TO

OVERDISPERSION##

NullModelaa<-glmer(Output_variable ~(1|NUT115CD/NUTS215CD/NUTS315CD/LAD13CD),

data=England_LA_data, family = binomial )

summary(NullModelaa)

AIC(NullModelaa)
```

```
icc(NullModelaa, by_group = TRUE)

NullModelaaa<-glmer(Output_variable ~(1|LAD13CD), data=England_LA_data, family =
binomial )

summary(NullModelaaa)

AIC(NullModelaaa)

icc(NullModelaaa, by_group = TRUE)

NullModelbb<-glmer(Output_variable ~ (1|NUTS315CD/LAD13CD), data=England_LA_data,
family = binomial )

summary(NullModelbb)

AIC(NullModelbb)

icc(NullModelbb, by_group = TRUE)

NullModelcc<-glmer(Output_variable ~ (1|NUTS215CD/LAD13CD), data=England_LA_data,
family = binomial )

AIC(NullModelcc)

NullModeldd<-glmer(Output_variable ~ (1|NUT115CD/LAD13CD), data=England_LA_data,
family = binomial )

AIC(NullModeldd)


Predictor_LAaa<-glmer(Output_variable ~Predictor_variable +
(1|NUT115CD/NUTS215CD/NUTS315CD/LAD13CD), data=England_LA_data, family =
binomial )

coef(Predictor_LAaa)

summary(Predictor_LAaa)

AIC(Predictor_LAaa)

icc(Predictor_LAaa, by_group = TRUE)

Predictor_LAbb<-glmer(Output_variable ~Predictor_variable + (1|NUTS315CD/LAD13CD),
data=England_LA_data, family = binomial )

AIC(Predictor_LAbb)

Predictor_LAcc<-glmer(Output_variable ~Predictor_variable + (1|NUTS215CD/LAD13CD),
data=England_LA_data, family = binomial )

AIC(Predictor_LAcc)

Predictor_LAdd<-glmer(Output_variable ~Predictor_variable + (1|NUT115CD/LAD13CD),
data=England_LA_data, family = binomial )

AIC(Predictor_LAdd)
```

```
#summary(Predictor_LAd)


Random_Slopeaa<-glmer(Output_variable ~ Predictor_variable +

(1+Predictor_variable|NUT115CD/NUTS215CD/NUTS315CD/LAD13CD),

data=England_LA_data, family = binomial )

#summary(Random_Slopeaa)

AIC(Random_Slopeaa)

Random_Slopebb<-glmer(Output_variable ~ Predictor_variable +

(1+Predictor_variable|NUTS315CD/LAD13CD), data=England_LA_data, family = binomial )

#summary(Random_Slopebb)

AIC(Random_Slopebb)

Random_Slopecc<-glmer(Output_variable ~ Predictor_variable +

(1+Predictor_variable|NUTS215CD/LAD13CD), data=England_LA_data, family = binomial )

#summary(Random_Slopecc)

AIC(Random_Slopecc)

Random_Slopedd<-glmer(Output_variable ~ Predictor_variable +

(1+Predictor_variable|NUT115CD/LAD13CD), data=England_LA_data, family = binomial )

#summary(Random_Slopedd)

AIC(Random_Slopedd)
```

Annex 2: Workplace earnings

Annex 2 - Table 1: Workplace Earnings at LA level: AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level NVQ4+ and No qualifications predictors

	Workplace Earnings – Null Models	**Workplace Earnings** – Random Intercept models with LA level **NVQ4+** predictor variable	**Workplace Earnings** – Random coefficient models with LA level **NVQ4+**predictor variable	**Workplace Earnings** – Random Intercept models with LA level **No qualifications** predictor variable	**Workplace Earnings** – Random Coefficient models with LA level 'No qualifications' predictor variable
Four Level model	**3,599**	**3,549**	*3,525 Model failed to converge*	**3,426**	*3,437 boundary (singular) fit*
Grouping by NUTS 3 areas	3,649	3,563	*3,533 Model failed to converge*	3,472	*3,475 boundary (singular) fit*
Grouping by NUTS 2 areas	3,629	3,562	**3,545**	3,451	3,455
Grouping by NUTS 1 areas	3,652	3,572	*3,560 boundary (singular) fit*	3,468	3,471
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level		*3,540 boundary (singular) fit*		*3,428 boundary (singular) fit*	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level		*3,537 boundary (singular) fit*		**3,428**	

The workplace earnings four level null model gave a lower AIC value than the other null models of workplace earnings. In contrast to the four-level null model of resident earnings, the four-level model of workplace earnings converged successfully without any warnings of a singular fit.

Of the models that used the NVQ level 4 or higher qualification predictor to model workplace earnings the two-level, random coefficient model with grouping by NUTS 2 areas had the lowest AIC indicating that it fitted the data best. The other five models that included random coefficients either failed to converge or produced a singular fit (see section 3.3.3.3 above).

Considering the models using the 'No qualifications' predictor to model workplace earnings the four-level random intercept model and the four level model with random intercepts at all levels and random coefficients at NUTS 1 areas level had similar AIC values which were much lower than the AIC values of any of the other models indicating that these two models fitted the data better than any of the other models. Three of the models with included random coefficients produced singular fits (see section 3.3.3.3 above). The remaining two models which included random coefficients did converge but had slightly higher AIC values than the corresponding random intercept values showing that random coefficient models do not always give better fitting models than the corresponding random intercept models.

Annex 2 - Table 2: Workplace Earnings at LA level: AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level Proportion with Bad/Very Bad Health and Proportion working Part-time predictors

	Workplace Earnings – Null Models	Workplace Earnings – Random Intercept models with LA level **Bad/Very Bad Health** predictor variable	Workplace Earnings – Random coefficient models with LA level **Bad/Very Bad Health** 'predictor variable	Workplace Earnings – Random Intercept models with LA level **working Part-time** predictor variable	Workplace Earnings – Random Coefficient models with LA level **working Part-time** predictor variable
Four Level model	**3,599**	**3,559**	3,564 *Model failed to converge*	3,561	3,530 *Model failed to converge*
Grouping by NUTS 3 areas	3,649	3,601	3,603 *Model failed to converge*	3,605	3,565
Grouping by NUTS 2 areas	3,629	3,578	3,574	3,579	**3,547**
Grouping by NUTS 1 areas	3,652	3,605	3,599	3,594	3,576
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level		3,561 *Model failed to converge*		3,563	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level		3,561 *Model failed to converge*		3,563	

Of the models using the proportion of people with bad or very bad health predictor to model workplace earnings the four-level, random intercept model fitted the data best.

Observing the AIC values of the models using the proportion of people working part-time predictor to model workplace earnings, it was the two-level random coefficient model with grouping by NUTS 2 areas that fitted the data best. Only the four-level random coefficient model failed to converge for this predictor variable.

Annex 2 - Table 3: Workplace Earnings at LA level: AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level Proportion with Age and proportion Female predictors

	Workplace Earnings – Null Models	**Workplace Earnings** – Random Intercept models with LA level **Age** predictor variable	**Workplace Earnings** – Random coefficient models with LA level **Age** 'predictor variable	**Workplace Earnings** – Random Intercept models with LA level **proportion female** predictor variable	**Workplace Earnings** – Random Coefficient models with LA level **proportion female** predictor variable
Four Level model	**3,599**	**3,577**	*3,587 Model failed to converge*	**3,562**	*3,532 singular (boundary) fit*
Grouping by NUTS 3 areas	3,649	3,615	3,616	3,609	3,568
Grouping by NUTS 2 areas	3,629	3,600	*3,602 Model failed to converge*	3,586	*3,571 singular (boundary) fit*
Grouping by NUTS 1 areas	3,652	3,624	3,622	3,605	*3,592 Model failed to converge*
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level	**3,579**			*3,564 Model failed to converge*	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level	**3,579**			*3,564*	

Of the models using the local authority level Age predictor to model workplace earnings the four-level random intercept model and the four level models with random intercepts at all levels and random coefficients at either NUTS 2 areas level or NUTS 1 areas level had similar AIC values that were lower than those for any of the other models indicating that these three models fitted the data better than any of the other models with this predictor variable.

Observing the AIC values of the models using the proportion of people who were female as the predictor to model resident earnings, it was the four-level random intercept model and

the four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level that fitted the data best. Four of the models that included random coefficients for this predictor variable either failed to converge or had a singular fit (see section 3.3.3.3 above).

	Workplace Earnings – Null Models	Workplace Earnings – Random Intercept models with LA level **Proportion travelling 30km** predictor variable	Workplace Earnings – Random coefficient models with LA level **Proportion travelling 30km** 'predictor variable	Workplace Earnings – Random Intercept models with LA level **Industrial Diversity Indicator** predictor variable	Workplace Earnings – Random Coefficient models with LA level **Industrial Diversity Indicator** predictor variable
Four Level model	**3,599**	**3,588**	*3,592 Model failed to converge*	**3,561**	*3,494 Model failed to converge*
Grouping by NUTS 3 areas	3,649	3,633	3,622	3,609	*3,576 boundary (singular) fit*
Grouping by NUTS 2 areas	3,629	3,617	3,616	3,596	*3,573 boundary (singular) fit*
Grouping by NUTS 1 areas	3,652	3,642	3,644	3,613	*3,565 Model failed to converge*
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level		*3,590 Model failed to converge*		*3,540 Model failed to converge*	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level		*3,590 Model failed to converge*		*3,532 Model failed to converge*	

Of the models using the local authority level proportion of people travelling 30 km or more to work to model workplace earnings, it was the four-level random intercept model that fitted the data best. In contrast, the four-level random coefficient model failed to converge.

Observing the AIC values of the models using the local authority level industrial diversity indicator as the predictor to model workplace earnings, as was the case for the equivalent set of resident earnings models, it was the four-level random intercept model that fitted

the data best. All six of the models that included random coefficients for this predictor

variable either failed to converge or gave rise to a singular fit (see section 3.3.3.3 above).

Annex 2 - Table 5: Workplace Earnings at LA level: AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level Proportion with Occupational Diversity Indicator and proportion managers and professionals predictors

	Workplace Earnings – Null Models	**Workplace Earnings** – Random Intercept models with LA level **Occupational Diversity Indicator** predictor variable	**Workplace Earnings** – Random coefficient models with LA level **Occupational Diversity Indicator** 'predictor variable	**Workplace Earnings** – Random Intercept models with LA level **proportion managers and professionals** predictor variable	**Workplace Earnings** – Random Coefficient models with LA level **proportion managers and professionals** predictor variable
Four Level model	**3,599**	**3,529**	3,529 *boundary (singular) fit*	**3,575**	3,584 *boundary (singular) fit*
Grouping by NUTS 3 areas	3,649	3,544	3,542 *boundary (singular) fit*	3,609	3,605 *boundary (singular) fit*
Grouping by NUTS 2 areas	3,629	3,540	3,539 *boundary (singular) fit*	3,595	3,599 *Model failed to converge*
Grouping by NUTS 1 areas	3,652	3,553	3,547 *boundary (singular) fit*	3,608	3,607 *Model failed to converge*
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level	3,526 *Model failed to converge*		**3,577**		
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level	3,525 *boundary (singular) fit*		**3,577**		

Of the models using the local authority level occupational diversity indicator as the predictor to model workplace earnings, it was yet again the four-level random intercept model that appeared to fit the data best. For this predictor all six of the models that included random coefficients either produced a singular fit or failed to converge (see section 3.3.3.3 above).

Observing the AIC values of the models using the proportion of managers and professional as the predictor to model workplace earnings, it was the four-level random intercept model and the four-level models with random intercepts at all levels and random coefficients as

either NUTS 2 areas level or NUTS 1 areas level that fitted the data best. For this predictor the other four models that included random coefficients either failed to converge or had singular fits.

Annex 2 - Table 6: Workplace Earnings at LA level: AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level Proportion plant/process workers and proportion of people in rural areas inc. hub towns predictors

	Workplace Earnings – Null Models	Workplace Earnings – Random Intercept models with LA level **Proportion plant/process workers** predictor variable	Workplace Earnings – Random coefficient models with LA level **Proportion plant/process workers** 'predictor variable	Workplace Earnings – Random Intercept models with LA level **proportion of people in rural areas inc. hub towns** predictor variable	Workplace Earnings – Random Coefficient models with LA level **proportion of people in rural areas inc. hub towns** predictor variable
Four Level model	**3,599**	**3,578**	*3,565 Model failed to converge*	**3,588**	*3,628 Model failed to converge*
Grouping by NUTS 3 areas	3,649	3,619	*3,587 Model failed to converge*	3,633	3,628
Grouping by NUTS 2 areas	3,629	3,599	3,586	3,615	*3,611 boundary (singular) fit*
Grouping by NUTS 1 areas	3,652	3,616	3,600	3,640	3,634
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level	*3,580 boundary (singular) fit*			**3,590**	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level	**3,580**			**3,588**	

Of the models using the proportion of plant and process workers as the predictor to model workplace earnings, it was the four-level random intercept model and the four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level fitted the data best.

Observing the AIC values of the models using the proportion of people who live in rural areas including hub towns as the predictor to model workplace earnings, it was again the four-level random intercept model and the four-level models with random intercepts at all

levels and random coefficients at either NUTS 2 areas or NUTS 1 areas level that fitted the

data best.

Annex 2 - Table 7: Workplace Earnings at LA level: AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level social housing and proportion managers and proportion non-UK born

	Workplace Earnings – Null Models	**Workplace Earnings** – Random Intercept models with LA level **social housing** predictor variable	**Workplace Earnings** – Random coefficient models with LA level **social housing** predictor variable	**Workplace Earnings** – Random Intercept models with LA level **proportion non-UK born** predictor variable	**Workplace Earnings** – Random Coefficient models with LA level **proportion non-UK born** predictor variable
Four Level model	**3,599**	3,584	*3,587 Model failed to converge*	3,530	*3,523 Model failed to converge*
Grouping by NUTS 3 areas	3,649	3,625	3,619	3,539	3,529
Grouping by NUTS 2 areas	3,629	3,613	*3,612 boundary (singular) fit*	3,549	*3,547 Model failed to converge*
Grouping by NUTS 1 areas	3,652	3,631	*3,625 boundary (singular) fit*	3,562	3,554
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level		3,585		**3,524**	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level		**3,580**		*3,527 Model failed to converge*	

The models using the proportion of social housing in each local authority as the predictor to model workplace earnings produced a slightly unusual pattern in that the four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level fitted the data better than any of the other models with its AIC value being four units lower than that for the four-level model with random intercepts at all levels.

The AIC values of the models using the proportion of people who born outside the UK as the predictor to model workplace earnings also showed the four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level to fit to fit the data

better than any of the other models using this predictor variable. Three out of the six models that included random coefficient models failed to converge.

Annex 2 - Table 8: Workplace Earnings at LA level: AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level IMD 2015 and employment deprivation predictors

	Workplace Earnings – Null Models	Workplace Earnings – Random Intercept models with LA level **IMD 2015** predictor variable	Workplace Earnings – Random coefficient models with LA level **IMD 2015** predictor variable	Workplace Earnings – Random Intercept models with LA level **employment deprivation 2015** predictor variable	Workplace Earnings – Random Coefficient models with LA level **employment deprivation 2015** predictor variable
Four Level model	**3,599**	**3,581**	*3,588 Model failed to converge*	**3,569**	*3,573 boundary (singular) fit*
Grouping by NUTS 3 areas	3,649	3,637	*3,636 Model failed to converge*	3,619	*3,620 Model failed to converge*
Grouping by NUTS 2 areas	3,629	*3,605 Model failed to converge*	3,604	3,590	3,587
Grouping by NUTS 1 areas	3,652	3,630	3,626	3,615	3,617
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level		*3,582 Model failed to converge*		*3,571 boundary (singular) fit*	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level		**3,581**		**3,571**	

Of the models using the average IMD 2015 rank for each local authority as the predictor to model workplace earnings it was, as is often the case for other predictor variables, the four-level random intercept model and the four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level fitted the data best. Three out of the six models that included random coefficients failed to converge.

Of the models using the average 2015 employment deprivation rank for each local authority as the predictor to model workplace earnings, it was again the four-level random intercept model and the four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level that fitted the data best. For this predictor also, three out of the six models that included random coefficients either produced a singular fit or failed to converge (see section 3.3.3.3 above).

Annex 3: Mean hours

Annex 3 - Table 1: Mean Hours at LA level: AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level NVQ4+ and No qualifications predictors

	Mean Hours – Null Models	**Mean Hours** – Random Intercept models with LA level **NVQ4+** predictor variable	**Mean Hours** – Random coefficient models with LA level **NVQ4+**'predictor variable	**Mean Hours** – Random Intercept models with LA level **No qualifications** predictor variable	**Mean Hours** – Random Coefficient models with LA level **No qualifications** predictor variable
Four Level model	1,199	**1,186**	*1,186 Model failed to converge*	*1,144 Model failed to converge*	*1,157 boundary (singular) fit*
Grouping by NUTS 3 areas	1,204	1,194	*1,189 Model failed to converge*	1,147	1,150
Grouping by NUTS 2 areas	**1,195**	**1,185**	**1,183**	**1,141**	1,145
Grouping by NUTS 1 areas	1,202	1,188	**1,182**	1,144	*1,147 boundary (singular) fit*
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level	*1,186 Model failed to converge*			1,146	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level	*1,188 Model failed to converge*			*1,157 boundary (singular) fit*	

Of the null models of mean hours worked, the model with grouping by NUTS 2 areas had the lowest AIC value indicating that it fitted the data better than the other null models of mean hours worked. However, the AIC value for the four-level null model was only slightly higher indicating that it fitted the model almost as well. Indeed, there is only a difference of nine between the highest and lowest AIC values for the null models showing that there was not very much difference between the fits of each of the different models.

Of the models that used the NVQ level 4 or higher qualification predictor to model mean hours worked the two-level, random coefficient model with grouping by NUTS 1 areas and

the random coefficient models with grouping by NUTS 2 areas had the lowest AIC values indicating that they fitted the data better than the other models with this predictor variable. Again, the differences between the highest and lowest AIC values for all the models using this predictor variable were not very large.

The AIC values of the models of hours worked that used the 'No qualifications' predictor mean that the random intercept, two-level model with grouping by NUTS 2 areas fitted the data better than the other models that used this predictor variable. However, that there were not great differences between any of the AIC values for this predictor variable.

Annex 3 - Table 2: Mean Hours at LA level: AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level Proportion with Bad/Very Bad Health and Proportion working Part-time predictors

	Mean Hours – Null Models	Mean Hours – Random Intercept models with LA level **Bad/Very Bad Health** predictor variable	Mean Hours – Random coefficient models with LA level **Bad/Very Bad Health** 'predictor variable	Mean Hours – Random Intercept models with LA level **working Part-time** predictor variable	Mean Hours – Random Coefficient models with LA level **working Part-time** predictor variable
Four Level model	1,199	*1,195 Model failed to converge*	*1,204 boundary (singular) fit*	1,162	*1,173 boundary (singular) fit*
Grouping by NUTS 3 areas	1,204	1,201	*1,204 boundary (singular) fit*	1,164	*1,168 Model failed to converge*
Grouping by NUTS 2 areas	**1,195**	**1,192**	*1,195 Model failed to converge*	**1,159**	1,162
Grouping by NUTS 1 areas	1,202	1,198	*1,198 boundary (singular) fit*	1,162	*1,166 Model failed to converge*
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level		*1,197 boundary (singular) fit*		*1,164 boundary (singular) fit*	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level		*1,194 boundary (singular) fit*		*1,164 boundary (singular) fit*	

Of the models using the proportion of people with bad or very bad health predictor to model mean hours worked the two-level, random intercept model with grouping by NUTS 2 areas fitted the data best. All the models that included random coefficients for this predictor variable either gave a singular fit or failed to converge (see section 3.3.3.3 above).

Observing the AIC values of the models using the proportion of people working part-time predictor to model workplace earnings, it was also the two-level random intercept model with grouping by NUTS 2 areas that fitted the data best although all of the models that converged without a singular fit had very similar AIC values indicted that that fitted the data about equally as well as each other. For this predictor variable five out of the six models that included random coefficients either failed to converge or had singular fits.

Annex 3 Table 2 above reiterates the value of using grouping by NUTS 2 areas in multilevel models of mean hours worked.

Annex 3 - Table 3: Mean Hours at LA level: AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level Age and proportion Female predictors

	Mean Hours – Null Models	**Mean Hours** – Random Intercept models with LA level **Age** predictor variable	**Mean Hours** – Random coefficient models with LA level **Age** 'predictor variable	**Mean Hours** – Random Intercept models with LA level **proportion female** predictor variable	**Mean Hours** – Random Coefficient models with LA level **proportion female** predictor variable
Four Level model	1,199	*1,196 boundary (singular) fit*	*1,202 boundary (singular) fit*	1,191	*1,200 boundary (singular) fit*
Grouping by NUTS 3 areas	1,204	1,201	*1,198 boundary (singular) fit*	1,194	*1,196 Model failed to converge*
Grouping by NUTS 2 areas	**1,195**	**1,193**	*1,194 boundary (singular) fit*	**1,187**	*1,191 Model failed to converge*
Grouping by NUTS 1 areas	1,202	1,199	*1,202 boundary (singular) fit*	1,191	*1,195 boundary (singular) fit*
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level	*1,196 boundary (singular) fit*			*1,164 boundary (singular) fit*	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level	*1,198 boundary (singular) fit*			*1,164 boundary (singular) fit*	

Of the models using the local authority level Age predictor to model mean hours worked the two-level random intercept model with grouping by NUTS 2 areas had the lowest AIC indicating that it fitted the data better than other models with this predictor variable. All of the models that included random coefficients for this predictor variable had a singular fit indicating that they were over complex (see section 3.3.3.3 above).

The findings for the models using the proportion of people who were female as the predictor to model resident earnings were very similar to those for the Age predictor

models. It was again the two-level random intercept model with grouping by NUTS 2 areas that fitted the data best. All of the models that included random coefficients for this predictor variable either failed to converge or had a singular fit indicates that random coefficient models were over complex for this combination of output and predictor variables.

Annex 3 - Table 4: Mean Hours at LA level: AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level Proportion travelling 30km + and Industrial Diversity Indicator predictors

	Mean Hours – Null Models	Mean Hours – Random Intercept models with LA level **Proportion travelling 30km** predictor variable	Mean Hours – Random coefficient models with LA level **Proportion travelling 30km** 'predictor variable	Mean Hours – Random Intercept models with LA level **Industrial Diversity Indicator** predictor variable	Mean Hours – Random Coefficient models with LA level **Industrial Diversity Indicator** predictor variable
Four Level model	1,199	1,194	*1,207 boundary (singular) fit*	1,194	*1,192 boundary (singular) fit*
Grouping by NUTS 3 areas	1,204	1,198	1,200	1,198	*1,197 boundary (singular) fit*
Grouping by NUTS 2 areas	**1,195**	**1,191**	*1,194 Model failed to converge*	**1,191**	*1,191 boundary (singular) fit*
Grouping by NUTS 1 areas	1,202	1,198	1,202	1,197	*1,200 boundary (singular) fit*
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level	*1,196 boundary (singular) fit*		*1,195 boundary (singular) fit*		
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level	*1,196 Model failed to converge*		*1,193 boundary (singular) fit*		

Of the models using the local authority level proportion of people travelling 30 km or more to work to model mean hours worked it was the two-level random intercept model with grouping by NUTS 2 areas that fitted the data best.

Observing the AIC values of the models using the local authority level industrial diversity indicator as the predictor to model mean hours worked, it was also the two-level random intercept model with grouping by NUTS 2 areas that appeared to fit the data best. For this predictor all of the models that included random coefficients gave rise to a singular fit (see section 3.3.3.3 above).

The importance of grouping by NUTS 2 areas is clearly supported by the AIC values in Annex 3 Table 4.

Annex 3 - Table 5: Mean Hours at LA level: AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level Proportion with Occupational Diversity Indicator and proportion managers and professionals predictors

	Mean Hours – Null Models	**Mean Hours** – Random Intercept models with LA level **Occupational Diversity Indicator** predictor variable	**Mean Hours** – Random coefficient models with LA level **Occupational Diversity Indicator** 'predictor variable	**Mean Hours** – Random Intercept models with LA level **proportion managers and professionals** predictor variable	**Mean Hours** – Random Coefficient models with LA level **proportion managers and professionals** predictor variable
Four Level model	1,199	1,194	*1,193 boundary (singular) fit*	*1,191 Model failed to converge*	*1,196 boundary (singular) fit*
Grouping by NUTS 3 areas	1,204	1,201	*1,201 boundary (singular) fit*	1,198	*1,199 Model failed to converge*
Grouping by NUTS 2 areas	**1,195**	**1,191**	*1,189 boundary (singular) fit*	**1,188**	**1,188**
Grouping by NUTS 1 areas	1,202	1,199	*1,192 Model failed to converge*	1,197	*1,197 boundary (singular) fit*
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level		*1,191 boundary (singular) fit*		*1,190 boundary (singular) fit*	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level		*1,194 boundary (singular) fit*		*1,191 boundary (singular) fit*	

Of the models using the local authority level occupational diversity indicator as the predictor to model mean hours worked, it was again the two-level random intercept model with grouping by NUTS 2 areas that appeared to fit the data best. For this predictor all of the models that included random coefficients either gave rise to a singular fit or failed to converge (see section 3.3.3.3 above).

The AIC values of the models using the proportion of managers and professional as the predictor to mean hours worked show that for this predictor the two-level random intercept model with grouping by NUTS 2 areas and the random coefficient model with grouping by NUTS 2 both had almost identical AIC values that were lower than those for the other models with this predictor. This supports the importance of using two-level models with grouping by NUTS 2 areas for modelling means hours worked whether by random intercept models or by models that include random coefficients. For this predictor variable the other models that included random coefficients either failed to converge or had singular fits.

Annex 3 - Table 6: Mean Hours at LA level: AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level Proportion plant/process workers and proportion of people in rural areas including hub towns predictors

	Mean Hours – Null Models	**Mean Hours** – Random Intercept models with LA level **Proportion plant/process workers** predictor variable	**Mean Hours** – Random coefficient models with LA level **Proportion plant/process workers** 'predictor variable	**Mean Hours** – Random Intercept models with LA level **proportion of people in rural areas inc. hub towns** predictor variable	**Mean Hours** – Random Coefficient models with LA level **proportion of people in rural areas inc. hub towns** predictor variable
Four Level model	1,199	*1,197 boundary (singular) fit*	*1,199 boundary (singular) fit*	*1,201 boundary (singular) fit*	*1,212 boundary (singular) fit*
Grouping by NUTS 3 areas	1,204	1,202	1,203	1,207	1,209
Grouping by NUTS 2 areas	**1,195**	**1,193**	**1,191**	**1,198**	**1,200**
Grouping by NUTS 1 areas	1,202	1,201	*1,204 Model failed to converge*	1,204	*1,207 boundary (singular) fit*
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level	*1,199 Model failed to converge*			*1,202 boundary (singular) fit*	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level	*1,198 boundary (singular) fit*			*1,203 boundary (singular) fit*	

Of the models using the proportion of plant and process workers as the predictor to model mean hours worked, it was the two-level random coefficient model with grouping by NUTS 2 areas and the two-level random intercept model with grouping by NUTS 2 areas that fitted the data best, both having similar AIC values.

For the models using the proportion of people who live in rural areas including hub towns as the predictor to model mean hours worked, it was again the two-level random intercept and random coefficient models with grouping by NUTS 2 areas that fitted the data best,

although they fitted the data less well than the null model with grouping by NUTS 2 areas level. Indeed all the models with this predictor variable fitted the data less well than the corresponding null model showing that adding this predictor variable led to worse fitting models than just using geographic area indicators when used to model the mean hours worked outcome variable. This means that the proportion of people living in rural areas is not, at least on its own, a useful variable for modelling local authority level hours worked per week. This is consistent with the total conditional variance for hours worked modelled using the rural predictor variable being higher than the total unconditional variance in the null model for hours worked (see section 5.5.3).

Annex 3 - Table 7: Mean Hours at LA level: AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level social housing and proportion managers and proportion non-UK born

	Mean Hours – Null Models	**Mean Hours** – Random Intercept models with LA level **social housing** predictor variable	**Mean Hours** – Random coefficient models with LA level **social housing** predictor variable	**Mean Hours** – Random Intercept models with LA level **proportion non-UK born** predictor variable	**Mean Hours** – Random Coefficient models with LA level **proportion non-UK born** predictor variable
Four Level model	1,199	1,185	*1,190 Model failed to converge*	*1,176 boundary (singular) fit*	*1,185 Model failed to converge*
Grouping by NUTS 3 areas	1,204	1,187	1,188	1,183	*1,184 boundary (singular) fit*
Grouping by NUTS 2 areas	**1,195**	**1,181**	*1,178 Model failed to converge*	**1,176**	1,180
Grouping by NUTS 1 areas	1,202	1,188	*1,189 boundary (singular) fit*	**1,175**	*1,178 boundary (singular) fit*
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level		*1,187 boundary (singular) fit*		*1,177 boundary (singular) fit*	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level		*1,197 boundary (singular) fit*		*1,178 boundary (singular) fit*	

The models using the proportion of social housing in each local authority as the predictor to model mean hours worked were typical of those for other predictor variables in the that is was the two-level random intercept model with grouping by NUTS 2 areas that had the lowest AIC value, indicating that it fitted the data better than other models. Five of the models that included random coefficients either failed to converge or had a singular fit implying that they were over complex for the data (see section 3.3.3.3 above).

The AIC values of the random intercept models using the proportion of people who born outside the UK as the predictor to model mean hours worked showed a slightly different pattern. They showed the two-level random intercept model with grouping by NUTS 1 areas and the two-level random intercept model with grouping by NUTS 2 areas to have almost equally low AIC values showing that they fitted the data equally well. More typically, five out of six of the models that included random coefficients either failed to converge or produced a singular fit.

Annex 3 - Table 8: Mean Hours at LA level: AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level IMD 2015 and employment deprivation predictors

	Mean Hours – Null Models	**Mean Hours** – Random Intercept models with LA level **IMD 2015** predictor variable	**Mean Hours** – Random coefficient models with LA level **IMD 2015** predictor variable	**Mean Hours** – Random Intercept models with LA level **employment deprivation 2015** predictor variable	**Mean Hours** – Random Coefficient models with LA level **employment deprivation 2015** predictor variable
Four Level model	1,199	1,200	*1,210 Model failed to converge*	*1,200 Model failed to converge*	*1,205 Model failed to converge*
Grouping by NUTS 3 areas	1,204	1,204	1,208	1,206	1,210
Grouping by NUTS 2 areas	**1,195**	**1,196**	*1,200 boundary (singular) fit*	**1,197**	*1,201 Model failed to converge*
Grouping by NUTS 1 areas	1,202	1,204	1,205	1,203	*1,202 boundary (singular) fit*
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level	*1,202 boundary (singular) fit*			1,202	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level	*1,201 boundary (singular) fit*			*1,198 boundary (singular) fit*	

Of the models using the average IMD 2015 rank for each local authority as the predictor to model mean hours worked, in common with models using many other predictor variables, it was the two-level random intercept model with grouping by NUTS 2 areas fitted the data best. However its fit, as measured by AIC values, was very similar to that of the corresponding null model showing that adding the IMD predictor to the two-level model with grouping by NUTS 2 areas had not improved the model.

Similarly for the models using the average 2015 employment deprivation rank for each local authority as the predictor to model mean hours worked, it was the two-level random intercept model with grouping by NUTS 2 areas that fitted the data best for this predictor

variable although again the fit was no better than that for the corresponding null model. For this predictor four out of six of the models that included random coefficients either failed to converge or had a singular fit (see section 3.3.3.3 above).

Annex 4: Job density – models excluding outliers (City of London, Westminster and Camden)

Annex 4 - Table 1: Job Density at LA level (excluding three outliers): AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level NVQ4+ and No qualifications predictors

	Job Density – Null Models	**Job Density** – Random Intercept models with LA level **NVQ4+**predictor variable	**Job Density** – Random coefficient models with LA level **NVQ4+** predictor variable	**Job Density** – Random Intercept models with LA level **No qualifications** predictor variable	**Job Density** – Random Coefficient models with LA level **No qualifications** predictor variable
Four Level model	*-159 boundary (singular) fit*	-196	*-199 boundary (singular) fit*	*-153 boundary (singular) fit*	*-144 boundary (singular) fit*
Grouping by NUTS 3 areas	-156	-189	*-202 Model failed to converge*	-153	*-151 boundary (singular) fit*
Grouping by NUTS 2 areas	**-162**	-196	*-200 boundary (singular) fit*	**-157**	-155 boundary (singular) fit
Grouping by NUTS 1 areas	-147	**-199**	**-199**	-148	*-145 boundary (singular) fit*
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level	*-196 boundary (singular) fit*			*-152 boundary (singular) fit*	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level	*-196 boundary (singular) fit*			*-152 boundary (singular) fit*	

Note: With negative AIC values it is the most negative, i.e. the lowest, that indicates the 'best' model (see section 4.2.6).

During the exploratory analysis of the proposed local authority level outcome variables it was observed that three London authorities, The City of London, Westminster and Camden had job densities that were far higher than those in any other local authorities. The high densities are due to large numbers of people coming in to work in the financial and

administrative centres based in the UK's capital city. As these job densities are far outside the normal range of job densities including them in models of local authority job density would have an influential effect on the model, i.e. they are influential outliers in the dataset.

The null models for local authorities excluding the three outliers show the lowest AIC value, i.e. the most negative (see section 4.2.6), was for the null model with grouping by NUTS 2 areas showing that this model fitted the data better than any of the other null models.

For the NVQ level 4 or higher qualification predictor variable models, the random intercept and random coefficient two-level models with grouping by NUTS 1 areas had the equal lowest AIC values (for models that converged without producing a warning message). This showed that these models fitted the data better than other models with this predictor variable.

For the 'No qualifications' predictor variable it was the random intercept model with grouping by NUTS 2 areas that had the lowest AIC value indicating that this model fitted the data better than the other models that used this predictor variable. All of the models that included random intercepts for this predictor variable produced a singular fit (see section 3.3.3.3 above).

Annex 4 - Table 2: Job Density at LA level (excluding three outliers): AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level Proportion with Bad/Very Bad Health and Proportion working Part-time predictors

	Job Density – Null Models	Job Density – Random Intercept models with LA level **Bad/Very Bad Health** predictor variable	Job Density – Random coefficient models with LA level **Bad/Very Bad Health** 'predictor variable	Job Density – Random Intercept models with LA level **working Part-time** predictor variable	Job Density – Random Coefficient models with LA level **working Part-time** predictor variable
Four Level model	*-159 boundary (singular) fit*	-188	*-177 boundary (singular) fit*	*-160 boundary (singular) fit*	*-157 boundary (singular) fit*
Grouping by NUTS 3 areas	-156	-188	-184	-155	*-159 boundary (singular) fit*
Grouping by NUTS 2 areas	**-162**	**-192**	*-189 Model failed to converge*	**-164**	*-166 Model failed to converge*
Grouping by NUTS 1 areas	-147	-188	-185	-148	*-146 Model failed to converge*
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level	*-186 boundary (singular) fit*		*-158 boundary (singular) fit*	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level	*-186 boundary (singular) fit*		*-158 boundary (singular) fit*	

Of the models using the proportion of people with bad or very bad health predictor to model job density, the two-level random intercept model with grouping by NUTS 2 areas had the lowest AIC value indicating that it fitted the data better than the other models using this predictor variable.

Of the models using the proportion of people working part-time predictor it was the two-level random intercept model with grouping by NUTS 2 areas that had the lowest AIC value. This indicated that it fitted the data better than the other models built using this predictor variable. All the models that included random coefficients for this predictor variable either

failed to converge or had a singular fit suggesting that they may be over complex for the data (see section 3.3.3.3 above). For each of the models using this predictor variable the AIC value was only very slightly better than that for the corresponding null model indicating that adding this predictor variable to the model did not have an appreciable effect on the fit of the model. This suggests that the geographic scales used in the model had more effect on the model than did this particular predictor variable. This shows the importance of the choice of geographic scale used to model job density.

Annex 4 - Table 3: Job Density at LA level (excluding three outliers): AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level Proportion with Age and proportion Female predictors

	Job Density – Null Models	**Job Density – Random Intercept models with LA level Age predictor variable**	**Job Density – Random coefficient models with LA level Age 'predictor variable**	**Job Density – Random Intercept models with LA level proportion female predictor variable**	**Job Density – Random Coefficient models with LA level proportion female predictor variable**
Four Level model	*-159 boundary (singular) fit*	*-169 boundary (singular) fit*	*-169 boundary (singular) fit*	*-169 boundary (singular) fit*	*-166 Model failed to converge*
Grouping by NUTS 3 areas	-156	-160	-168	-166	*-168 Model failed to converge*
Grouping by NUTS 2 areas	**-162**	**-172**	**-170**	**-173**	*-169 Model failed to converge*
Grouping by NUTS 1 areas	-147	-152	-155	-158	*-154 Model failed to converge*
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level		*-167 Model failed to converge*		*-167 boundary (singular) fit*	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level		-167		*-167 boundary (singular) fit*	

Of the models using the local authority level age predictor to model job density the two-level random intercept and random coefficient models with grouping by NUTS 2 areas had

similar low AIC values indicating that they fitted the data better than other models with this predictor variable.

For the proportion of people who were female predictor variable models only the two-level random intercept models converged without having a singular fit (see section 3.3.3.3 above). Of the two-level random intercept models the one with grouping by NUTS 2 areas had the lowest (most negative – see section 4.2.6) AIC value indicating that it fitted the data better than the other models built using this predictor variable. All of the models that included random coefficients for this predictor variable either failed to converge or produced a singular fit.

	Job Density – Null Models	Job Density – Random Intercept models with LA level **Proportion travelling 30km** predictor variable	Job Density – Random coefficient models with LA level **Proportion travelling 30km** 'predictor variable	Job Density – Random Intercept models with LA level **Industrial Diversity Indicator** predictor variable	Job Density – Random Coefficient models with LA level **Industrial Diversity Indicator** predictor variable
Four Level model	*-159 boundary (singular) fit*	*-157 boundary (singular) fit*	*-153 Model failed to converge*	*-173 boundary (singular) fit*	*-195 Model failed to converge*
Grouping by NUTS 3 areas	-156	-153	-158	-169	*-175 boundary (singular) fit*
Grouping by NUTS 2 areas	**-162**	**-160**	**-162**	-174	*-174 boundary (singular) fit*
Grouping by NUTS 1 areas	-147	-146	*-145 boundary (singular) fit*	-158	**-178**
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level		*-155 boundary (singular) fit*		-174	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level		*-155 boundary (singular) fit*		*-171 boundary (singular) fit*	

Of the models using the local authority level proportion of people travelling 30 km or more to work predictor variable, the random intercept and random coefficient models with grouping by NUTS 2 areas had the lowest AIC values showing that they fitted the data better than the other models with this predictor variable, although they did not fit the data any better than the null model with grouping by NUTS 2 areas.

Looking at the AIC values of the models using the local authority level industrial diversity indicator as the predictor to model job density shows that the two-level random coefficient model with grouping by NUTS 1 areas had the lowest AIC value showing that it fitted that

data better than the other models with the industrial diversity predictor variable. It is worth noting that it fitted the data considerably better than the two-level random intercept model with grouping by NUTs 1 areas that used this predictor variable.

Annex 4 - Table 5: Job Density at LA level (excluding three outliers): AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level Proportion with Occupational Diversity Indicator and proportion managers and professionals predictors

	Job Density – Null Models	**Job Density** – Random Intercept models with LA level **Occupational Diversity Indicator** predictor variable	**Job Density** – Random coefficient models with LA level **Occupational Diversity Indicator** 'predictor variable	**Job Density** – Random Intercept models with LA level **proportion managers and professionals** predictor variable	**Job Density** – Random Coefficient models with LA level **proportion managers and professionals** predictor variable
Four Level model	*-159 boundary (singular) fit*	*-181 Model failed to converge*	*-178 boundary (singular) fit*	*-172 boundary (singular) fit*	*-164 Model failed to converge*
Grouping by NUTS 3 areas	-156	-179	*-177 boundary (singular) fit*	-172	-172
Grouping by NUTS 2 areas	**-162**	**-184**	*-181 Model failed to converge*	**-176**	-173
Grouping by NUTS 1 areas	-147	-178	*-177 Model failed to converge*	-170	*-167 Model failed to converge*
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level		*-179 Model failed to converge*		*-171 boundary (singular) fit*	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level		*-179 boundary (singular) fit*		*-170 boundary (singular) fit*	

Of the random intercept models using the local authority level occupational diversity to model job density the two-level random intercept model with grouping by NUTS 2 areas gave the lowest AIC value indicating that it fitted the data better than the other models using this predictor variable. All six of the models that included random coefficients for the

occupational diversity predictor failed to converge or gave singular fits (see section 3.3.3.3 above).

For the models using the proportion of managers and professionals as their predictor it was also the two-level random intercept model with grouping by NUTS 2 areas that had the lowest AIC value indicating that it fitted the data better than other models with this predictor variable.

Annex 4 - Table 6: Job Density at LA level (excluding three outliers): AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level Proportion plant/process workers and proportion of people in rural areas inc. hub towns predictors

	Job Density – Null Models	**Job Density** – Random Intercept models with LA level **Proportion plant/process workers** predictor variable	**Job Density** – Random coefficient models with LA level **Proportion plant/process workers** 'predictor variable	**Job Density** – Random Intercept models with LA level **proportion of people in rural areas inc. hub towns** predictor variable	**Job Density** – Random Coefficient models with LA level **proportion of people in rural areas inc. hub towns** predictor variable
Four Level model	*-159 boundary (singular) fit*	*-161 boundary (singular) fit*	*-156 boundary (singular) fit*	*-152 boundary (singular) fit*	*-151 boundary (singular) fit*
Grouping by NUTS 3 areas	-156	-159	*-159 Model failed to converge*	-149	*-155 Model failed to converge*
Grouping by NUTS 2 areas	**-162**	**-165**	**-166**	-155	**-158**
Grouping by NUTS 1 areas	-147	-154	-156	-141	-140
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level		*-159 boundary (singular) fit*		*-150 boundary (singular) fit*	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level		*-159 boundary (singular) fit*		*-150 boundary (singular) fit*	

Of the random intercept models using the local authority level proportion of plant or process workers the two-level random coefficient and random intercept models with grouping by NUTS 2 areas had the lowest AIC values indicating that they fitted the data better than the other models built using this predictor variable.

For models of job density using the proportion of people living in rural areas including hub towns as their predictor the two-level random coefficient model with grouping by NUTS 2 areas had the lowest AIC value indicating that it fitted the data better than other models using this predictor variable. The corresponding random intercept model had an AIC that was only slightly higher indicating that it fitted the data almost as well as the random coefficient model. For this predictor variable all the models had higher AIC values than the corresponding null models.

Annex 4 - Table 7: Job Density at LA level (excluding three outliers): AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level social housing and proportion managers and proportion non-UK born

	Job Density – Null Models	Job Density – Random Intercept models with LA level **social housing** predictor variable	Job Density – Random coefficient models with LA level **social housing** predictor variable	Job Density – Random Intercept models with LA level **proportion non-UK born** predictor variable	Job Density – Random Coefficient models with LA level **proportion non-UK born** predictor variable
Four Level model	*-159 boundary (singular) fit*	-169	*-171 Model failed to converge*	-166	*-181 Model failed to converge*
Grouping by NUTS 3 areas	-156	-159	*-165 boundary (singular) fit*	-157	*-183 Model failed to converge*
Grouping by NUTS 2 areas	**-162**	**-173**	**-174**	-168	**-179**
Grouping by NUTS 1 areas	-147	-152	*-156 Model failed to converge*	-160	*-161 Model failed to converge*
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level		*-172 Model failed to converge*		*-175 boundary (singular) fit*	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level		*-178 boundary (singular) fit*		*-169 boundary (singular) fit*	

Of the random intercept models using the local authority level proportion of social housing to model job density the two-level random coefficient and random intercept models with grouping by NUTS 2 areas had the lowest AIC values indicating that they fitted the data better than the other models using this predictor variable.

For the models of job density using the proportion of people who were born outside the UK as their predictor variable the two-level random coefficient model with grouping by NUTS 2 areas had by far the lowest AIC value indicating that it fitted the data better than the other models built using this predictor variable. The other models that included random

coefficients for this predictor variable either failed to converge or produced singular fits (see section 3.3.3.3 above).

Annex 4 - Table 8: Job Density at LA level (excluding three outliers): AIC values for Null, Random Intercept and Random Coefficient models using Local Authority level IMD 2015 and employment deprivation 2015 predictors

	Job Density – Null Models	**Job Density – Random Intercept models with LA level IMD 2015 predictor variable**	**Job Density – Random coefficient models with LA level IMD 2015 predictor variable**	**Job Density –** Random Intercept models with LA level **employment deprivation 2015** predictor variable	**Job Density –** Random Coefficient models with LA level **employment deprivation 2015** predictor variable
Four Level model	*-159 boundary (singular) fit*	*-162 boundary (singular) fit*	*-153 boundary (singular) fit*	-183 boundary (singular) fit	*-171 boundary (singular) fit*
Grouping by NUTS 3 areas	-156	-161	*-159 Model failed to converge*	**-184**	*-180 boundary (singular) fit*
Grouping by NUTS 2 areas	**-162**	**-166**	*-162 Model failed to converge*	-181	*-183 boundary (singular) fit*
Grouping by NUTS 1 areas	-147	*-158 boundary (singular) fit*	*-157 boundary (singular) fit*	*-183 boundary (singular) fit*	*-180 boundary (singular) fit*
Four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level	*-160 boundary (singular) fit*			*-181 boundary (singular) fit*	
Four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level	*-160 boundary (singular) fit*			*-181 boundary (singular) fit*	

All of the models that included random coefficients for the average IMD 2015 rank for each local authority as the predictor to model job density either failed to converge or had singular fits (see section 3.3.3.3 above). The two-level random intercept model with

grouping by NUTS 2 areas had the lowest AIC value indicating that it fitted the data better than the other random intercept models.

Of the models of job density using the average 2015 employment deprivation rank to model job density it was the two-level random intercept model with grouping by NUTS 3 areas that had the lowest AIC value indicating that it fitted the data better than other models using this predictor variable. All of the models that included random coefficients for this predictor had a singular fit.

Annex 5: Employment

Annex 5 - Table 1: Employment Rate at LA level: AIC values for Null and Random Intercept models using Local Authority level NVQ4+ and No qualifications predictors

	Employment Rate – Null Models	Employment Rate – LA level NVQ4+ predictor variable	Employment Rate – LA level No qualifications predictor variable
Four Level model with random intercepts only	**6,395**	**6,374**	*6,121 boundary (singular) fit*
Two-level model with grouping by NUTS 3 areas, random intercepts only	6,406	6,393	6,129
Two-level model with grouping by NUTS 2 areas, random intercepts only	**6,396**	**6,377**	6,127
Two-level model with grouping by NUTS 1 areas, random intercepts only	**6,398**	**6,373**	**6,118**
Four level with random intercepts at all levels and random coefficients at NUTS 3, NUTS 2 and NUTS 1 levels but not at pseudo level		*6,379 boundary (singular) fit*	*6,127 boundary (singular) fit*
Four level with random intercepts at all levels and random coefficients at NUTS 2 areas level only		*6,376 boundary (singular) fit*	*6,123 boundary (singular) fit*
Four level with random intercepts at all levels and random coefficients at NUTS 1 areas level only		**6,375**	*6,123 boundary (singular) fit*

Although the four-level null model gave a slightly lower AIC value than the null two-level model with grouping by NUTS 2 areas and the null two-level model with grouping by NUTS 1 areas, the values for these three models were all very similar showing they all fitted the data to the same degree.

Of the random intercept models that used the NVQ level 4 or higher qualification predictor to model employment rates the four-level random intercept model, and the two-level random intercept models with grouping by NUTS 2 areas and with grouping by NUTS 1 areas, and the four level model with random intercepts at all levels and random

coefficients just at NUTS 1 areas level had similar AIC values to each other. They were all noticeably lower than the AIC value for the two-level model with grouping by NUTS 3 areas levels. This indicated that these four models fitted the data better than the two-level model with grouping by NUTS 3 areas level.

The AIC values for the models using the 'No qualifications' predictor to model employment rates showed that the two-level random intercept model with grouping by NUTS 1 areas fitted the data better than the other models built using this predictor variable. The next best fitting model was the four-level model which fitted the data almost as well but as it required more parameters to be estimated it would be a less parsimonious and therefore less desirable model. This means that the two-level random intercept model with grouping by NUTS 1 areas is recommended as a better choice than the four-level model for the 'No qualifications' predictor variable.

Annex 5 - Table 2: Employment Rate at LA level: AIC values for Null and Random Intercept models using Local Authority level Proportion with Bad/Very Bad Health and Proportion working Part-time predictors

	Employment Rate – Null Models	**Employment Rate** – LA level **Bad/Very Bad Health** predictor variable	**Employment Rate** – LA level **working Part-time** predictor variable
Four Level model with random intercepts only	**6,395**	6,302 *boundary (singular) fit*	**6,396**
Two-level model with grouping by NUTS 3 areas, random intercepts only	6,406	6,306	6,408
Two-level model with grouping by NUTS 2 areas, random intercepts only	**6,396**	6,304	**6,398**
Two-level model with grouping by NUTS 1 areas, random intercepts only	6,398	**6,298**	6,400
Four level with random intercepts at all levels and random coefficients at NUTS 3, NUTS 2 and NUTS 1 levels but not at pseudo level		6,308 *boundary (singular) fit*	6,401 *boundary (singular) fit*
Four level with random intercepts at all levels and random coefficients at NUTS 2 areas level only		6,304 *boundary (singular) fit*	6,398 *boundary (singular) fit*
Four level with random intercepts at all levels and random coefficients at NUTS 1 areas level only		6,304 *boundary (singular) fit*	6,398

The four-level random intercept model using the proportion of people with bad or very bad health predictor to model employment rates generated a warning that the model had a singular fit (see section 3.3.3.3 above). This is because this model had zero variance at both NUTS 3 areas level and NUTS 2 areas level. The singular fit warning and that the lowest AIC value was for the two-level model with grouping at NUTS 1 areas level both indicate that models of employment using poor health as the only predictor variable should include the NUTS 1 areas geographic scale as a level in preference to the NUTS 3 areas scale or the NUTS 2 areas scale. They also suggest that it is not helpful and could even be detrimental to include NUTS 3 areas and NUTS 2 areas geographic scales as levels in models of local authority level employment rates that use the proportion of people with bad health as a predictor variable.

In contrast the AIC values of the models using the proportion of people working part-time predictor variable to model employment rate showed the four-level random intercept model, the two-level random intercept model with grouping at NUTS 2 areas level and the four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level only to be the three best fitting models.

Annex 5 - Table 3: Employment Rate at LA level: AIC values for Null and Random Intercept models using Local Authority level Age and proportion Female predictors

	Employment Rate – Null Models	Employment Rate – LA level Age predictor variable	Employment Rate – LA level proportion female predictor variable
Four Level model with random intercepts only	**6,395**	**6,371**	*6,393 Model failed to converge*
Two-level model with grouping by NUTS 3 areas, random intercepts only	6,406	6,380	6,404
Two-level model with grouping by NUTS 2 areas, random intercepts only	**6,396**	**6,372**	**6,394**
Two-level model with grouping by NUTS 1 areas, random intercepts only	6,398	6,375	6,398
Four level with random intercepts at all levels and random coefficients at NUTS 3, NUTS 2 and NUTS 1 levels but not at pseudo level		*6,377 boundary (singular) fit*	*6,399 boundary (singular) fit*
Four level with random intercepts at all levels and random coefficients at NUTS 2 areas level only		*6,373 boundary (singular) fit*	*6,395 boundary (singular) fit*
Four level with random intercepts at all levels and random coefficients at NUTS 1 areas level only		*6,373 boundary (singular) fit*	*6,395 boundary (singular) fit*

Of the random intercept models using the local authority level Age predictor to model employment rates the four-level model and the two-level random intercept model with grouping by NUTS 2 areas had the lowest AIC values showing that they fitted the data better than other models with this predictor variable.

The findings for the random intercept models using the proportion of people who were female as the predictor to model resident earnings were similar to those for the Age predictor models. It was again the four-level model and the two-level random intercept model with grouping by NUTS 2 areas that fitted the data best.

Annex 5 - Table 4: Employment Rate at LA level: AIC values for Null and Random Intercept models using Local Authority level Proportion travelling 30km + and Industrial Diversity Indicator predictors

	Employment Rate – Null Models	**Employment Rate** – LA level **Proportion travelling 30km** predictor variable	**Employment Rate** – LA level **Industrial Diversity Indicator** predictor variable
Four Level model with random intercepts only	**6,395**	**6,383**	**6,390**
Two-level model with grouping by NUTS 3 areas, random intercepts only	6,406	6,385	6,400
Two-level model with grouping by NUTS 2 areas, random intercepts only	**6,396**	**6,381**	**6,390**
Two-level model with grouping by NUTS 1 areas, random intercepts only	6,398	6,385	6,394
Four level with random intercepts at all levels and random coefficients at NUTS 3, NUTS 2 and NUTS 1 levels but not at pseudo level		*6,388 boundary (singular) fit*	*6,394 boundary (singular) fit*
Four level with random intercepts at all levels and random coefficients at NUTS 2 areas level only		*6,384 boundary (singular) fit*	*6,392 boundary (singular) fit*
Four level with random intercepts at all levels and random coefficients at NUTS 1 areas level only		*6,385 boundary (singular) fit*	**6,692**

Of the random intercept models using the local authority level proportion of people travelling 30 km or more to work to model employment rates it was the two-level random intercept model with grouping by NUTS 2 areas that fitted the data best. The AIC values were very similar for all four models showing the fit to be similar for all four models.

Looking at the AIC values of the random intercept models using the local authority level industrial diversity indicator as the predictor to model employment rates shows that it was the four-level random intercept model, the two-level random intercept model with grouping by NUTS 2 areas and the four level model with random intercepts at all levels and random coefficients at NUTS 1 areas level only that fitted the data best.

Annex 5 - Table 5: Employment Rate at LA level: AIC values for Null and Random Intercept models using Local Authority level Proportion with Occupational Diversity Indicator and proportion managers and professionals predictors

	Employment Rate – Null Models	Employment Rate – LA level Occupational Diversity Indicator predictor variable	Employment Rate – LA level proportion managers and professionals predictor variable
Four Level model with random intercepts only	**6,395**	**6,393**	**6,388**
Two-level model with grouping by NUTS 3 areas, random intercepts only	6,406	6,406	6,400
Two-level model with grouping by NUTS 2 areas, random intercepts only	**6,396**	**6,394**	**6,389**
Two-level model with grouping by NUTS 1 areas, random intercepts only	6,398	**6,395**	**6,388**
Four level with random intercepts at all levels and random coefficients at NUTS 3, NUTS 2 and NUTS 1 levels but not at pseudo level		*6,399 boundary (singular) fit*	*6,394 boundary (singular) fit*
Four level with random intercepts at all levels and random coefficients at NUTS 2 areas level only		*6,395 boundary (singular) fit*	*6,390 boundary (singular) fit*
Four level with random intercepts at all levels and random coefficients at NUTS 1 areas level only		**6,395**	*6,390 boundary (singular) fit*

Of the random intercept models using the local authority level occupational diversity indicator as the predictor to model employment rates, it was the four-level random intercept model that had the lowest AIC value indicating that it fitted the data best. However, three of the other models all had similar AIC values to that model (6,394 and 6,395 as shown in Annex 5 Table 5 above) showing they fitted the data almost equally as

well. Furthermore, none of the models had AIC values that were significantly lower than the AIC values of the corresponding null models indicating that the occupational diversity predictor variable was not particularly helpful to include in multilevel models of the local authority level employment rate outcome variable

The AIC values of the models using the proportion of managers and professional as the predictor to model employment rates show that the four-level random intercept model, the two-level random intercept model with grouping by NUTS 1 areas and the two-level random intercept model with grouping by NUTS 2 are all very similar and are all lower than that for the two-level random intercept model with grouping by NUTS 3 areas level which is appreciably higher.

Annex 5 - Table 6: Employment Rate at LA level: AIC values for Null and Random Intercept models using Local Authority level Proportion plant/process workers and proportion of people in rural areas including hub towns predictors

	Employment Rate – Null Models	**Employment Rate** – LA level **Proportion plant/process workers** predictor variable	**Employment Rate** – LA level **proportion of people in rural areas inc. hub towns** predictor variable
Four Level model with random intercepts only	**6,395**	**6,382**	**6,380**
Two-level model with grouping by NUTS 3 areas, random intercepts only	6,406	6,395	6,388
Two-level model with grouping by NUTS 2 areas, random intercepts only	**6,396**	**6,382**	**6,381**
Two-level model with grouping by NUTS 1 areas, random intercepts only	6,398	**6,383**	**6,381**
Four level with random intercepts at all levels and random coefficients at NUTS 3, NUTS 2 and NUTS 1 levels but not at pseudo level		*6,386 boundary (singular) fit*	*6,386 boundary (singular) fit*
Four level with random intercepts at all levels and random coefficients at NUTS 2 areas level only		**6,383**	**6,382**
Four level with random intercepts at all levels and random coefficients at NUTS 1 areas level only		**6,385**	**6,382**

Of the random intercept models using the proportion of plant and process workers as the predictor to model employment rates, all the models that converged without producing warnings except for the two-level random intercept model with grouping by NUTS 3 areas fitted the data equally well. The two-level random intercept model with grouping by NUTS 3 areas fitted the data noticeably less well and the four-level model with random intercepts at all levels and random coefficients at NUTS 3, NUTS 2 and NUTS 1 levels but not at the pseudo level produced a singular fit (see section 3.3.3.3 above).

Observing the AIC values of the models using the proportion of people who live in rural areas including hub towns as the predictor all of the models fitted the data equally well except for the two-level random intercept model with grouping by NUTS 3 areas level which had a higher AIC value and the four-level model with random intercepts at all levels and random coefficients at NUTS 3, NUTS 2 and NUTS 1 areas levels but not at the pseudo level produced a singular fit.

Annex 5 - Table 7: Employment Rate at LA level: AIC values for Null and Random Intercept models using Local Authority level social housing and proportion managers and proportion non-UK born

	Employment Rate – Null Models	**Employment Rate** – LA level **social housing** predictor variable	**Employment Rate** – LA level **proportion non-UK born** predictor variable
Four Level model with random intercepts only	**6,395**	**6,360**	**6,361**
Two-level model with grouping by NUTS 3 areas, random intercepts only	6,406	6,363	6,380
Two-level model with grouping by NUTS 2 areas, random intercepts only	**6,396**	**6,357**	6,365
Two-level model with grouping by NUTS 1 areas, random intercepts only	**6,398**	6,362	6,368
Four level with random intercepts at all levels and random coefficients at NUTS 3, NUTS 2 and NUTS 1 levels but not at pseudo level		*6,364 boundary (singular) fit*	*6,367 boundary (singular) fit*
Four level with random intercepts at all levels and random coefficients at NUTS 2 areas level only		**6,360**	**6,363**
Four level with random intercepts at all levels and random coefficients at NUTS 1 areas level only		*6,361 boundary (singular) fit*	*6,363 boundary (singular) fit*

Of the models using the proportion of social housing in each local authority as the predictor to model employment rates it was the two-level random intercept model with grouping by NUTS 2 areas that had the lowest AIC value indicating that it fitted the data better than other models (although a number of the other models fitted the data almost as well).

The AIC values of the random intercept models using the proportion of people born outside the UK as the predictor to model employment rates showed the four-level random intercept model had the lowest AIC value indicating that it fitted the data better than other models. However, and the four-level model with random intercepts at all levels and random coefficients at NUTS 1 areas level only fitted the data almost equally as well.

Annex 5 - Table 8: Employment Rate at LA level: AIC values for Null and Random Intercept models using Local Authority level IMD 2015 and employment deprivation 2015 predictors

	Employment Rate – Null Models	**Employment Rate** – LA level **IMD 2105** predictor variable	**Employment Rate** – LA level **employment deprivation 2015** predictor variable
Four Level model with random intercepts only	**6,395**	**6,384**	*6,290 boundary (singular) fit*
Two-level model with grouping by NUTS 3 areas, random intercepts only	6,406	6,406	**6,287**
Two-level model with grouping by NUTS 2 areas, random intercepts only	**6,396**	**6,396**	**6,287**
Two-level model with grouping by NUTS 1 areas, random intercepts only	**6,398**	6,398	**6,286**
Four level with random intercepts at all levels and random coefficients at NUTS 3, NUTS 2 and NUTS 1 levels but not at pseudo level		*6,389 boundary (singular) fit*	*6,296 boundary (singular) fit*
Four level with random intercepts at all levels and random coefficients at NUTS 2 areas level only		*6,385 boundary (singular) fit*	*6,292 boundary (singular) fit*
Four level with random intercepts at all levels and random coefficients at NUTS 1 areas level only		*6,385 boundary (singular) fit*	*6,292 boundary (singular) fit*

Of the random intercept models using the average IMD 2015 rank for each local authority as the predictor to model employment rates it was the four-level random intercept model that fitted the data best.

For the random intercept models using the average 2015 employment deprivation rank for each local authority as the predictor to model employment rates the three two-level models with random intercepts at all levels and random coefficients at one of NUTS 1 areas level, NUTS 2 areas level and NUTS 3 areas level all fitted the data equally well.

Annex 6: Unemployment

Annex 6 - Table 1: Unemployment Rate at LA level: AIC values for Null and Random Intercept models using Local Authority level NVQ4+ and No qualifications predictors

	Unemployment Rate – Null Models	Unemployment Rate – LA level NVQ4+ predictor variable	Unemployment Rate - LA level No qualifications predictor variable
Four Level model with random intercepts only	**5,097**	**5,042**	**4,907**
Two-level model with grouping by NUTS 3 areas, random intercepts only	5,133	5,103	4,948
Two-level model with grouping by NUTS 2 areas, random intercepts only	5,138	5,079	4,939
Two-level model with grouping by NUTS 1 areas, random intercepts only	5,158	5,094	4,929
Four level with random intercepts at all levels and random coefficients at NUTS 3, NUTS 2 and NUTS 1 levels but not at pseudo level		**5,042**	*4,911 boundary (singular) fit*
Four level with random intercepts at all levels and random coefficients at NUTS 2 areas level only		**5,043**	*4,907 boundary (singular) fit*
Four level with random intercepts at all levels and random coefficients at NUTS 1 areas level only		**5,042**	*4,909 boundary (singular) fit*

The four-level null model gave a lower AIC value than the other null models of local authority unemployment rates.

Of the models that used the NVQ level 4 or higher qualification predictor to model unemployment rates the four-level random intercept model and the three two-level models with random intercepts at all levels and random coefficients at one or more of NUTS 1 areas level, NUTS 2 areas level or NUTS 3 areas level had the lowest, almost identical, AIC values indicating that they fitted the data best. This clearly shows the importance of four-level models.

The AIC values for the models using the 'No qualifications' predictor to model unemployment rates clearly showed that the four-level random intercept model fitted the data better than any of the other models that used this predictor variable.

These findings are different to those for models of employment rates shown in Annex 5 Table 1 in that for employment the two-level models with grouping by NUTS 1 areas levels and with grouping at NUTS 2 areas level, gave models that fitted the data as well as the four-level models whereas for unemployment models the two-level models with grouping by NUTS 1 areas levels and with grouping at NUTS 2 areas level did not fit the data as well as the four-level model.

Annex 6 - Table 2: Unemployment Rate at LA level: AIC values for Null and Random Intercept models using Local Authority level Proportion with Bad/Very Bad Health and Proportion working Part-time predictors

	Unemployment Rate – Null Models	**Unemployment Rate** – LA level **Bad/Very Bad Health** predictor variable	**Unemployment Rate** – LA level **working Part-time** predictor variable
Four Level model with random intercepts only	**5,097**	**4,961**	**5,092**
Two-level model with grouping by NUTS 3 areas, random intercepts only	5,133	5,029	5,122
Two-level model with grouping by NUTS 2 areas, random intercepts only	5,138	4,986	5,134
Two-level model with grouping by NUTS 1 areas, random intercepts only	5,158	4,998	5,151
Four level with random intercepts at all levels and random coefficients at NUTS 3, NUTS 2 and NUTS 1 levels but not at pseudo level		*4,961 boundary (singular) fit*	*5,094 boundary (singular) fit*
Four level with random intercepts at all levels and random coefficients at NUTS 2 areas level only		*4,960 boundary (singular) fit*	**5,093**
Four level with random intercepts at all levels and random coefficients at NUTS 1 areas level only		*4,963 boundary (singular) fit*	**5,094**

Of the random intercept models using the proportion of people with bad or very bad health predictor to model unemployment rates the four-level model fitted the data best. This is different to the finding for equivalent models of employment where the two-level model with grouping by NUTS 1 areas had the best fit with the data.

The AIC values of the random intercept models using the proportion of people working part-time predictor to model unemployment rates showed the four-level random intercept model and the two four-level model with random intercepts at all levels and random coefficients at one of either NUTS 2 areas level or NUTS 1 areas level fitted the data better than the other models that used this predictor variable.

Annex 6 - Table 3: Unemployment Rate at LA level: AIC values for Null and Random Intercept models using Local Authority level Proportion with Age and proportion Female predictors

	Unemployment Rate – Null Models	Unemployment Rate – LA level **Age** predictor variable	Unemployment Rate – LA level **proportion female** predictor variable
Four Level model with random intercepts only	**5,097**	**5,002**	**5,078**
Two-level model with grouping by NUTS 3 areas, random intercepts only	5,133	5,029	5,113 *Failed to converge*
Two-level model with grouping by NUTS 2 areas, random intercepts only	5,138	5,052	5,121 *Failed to converge*
Two-level model with grouping by NUTS 1 areas, random intercepts only	5,158	5,063	5,143
Four level with random intercepts at all levels and random coefficients at NUTS 3, NUTS 2 and NUTS 1 levels but not at pseudo level		*5,005 boundary (singular) fit*	*5,083 boundary (singular) fit*
Four level with random intercepts at all levels and random coefficients at NUTS 2 areas level only		**5,004**	*5,080 boundary (singular) fit*
Four level with random intercepts at all levels and random coefficients at NUTS 1 areas level only		*5,004 boundary (singular) fit*	*5,080 Failed to converge*

Of the models using the local authority level age predictor to model unemployment rates the four-level random intercept model and the four-level model with random intercepts at all levels and random coefficients at NUTS areas level only had the lowest AIC values indicating that they fitted the data better than other models with this predictor variable. This is a different finding to that for the equivalent models of employment where the model with grouping by NUTS 2 areas had an equally low AIC showing it fitted the data as well as the four-level model.

The findings for the models using the proportion of people who were female as the predictor to model unemployment rates also showed the four-level model to fit the data better than the other models. Five out of six of the other models either failed to converge or produced a singular fit for this predictor variable (see section 3.3.3.3 above).

Annex 6 - Table 4: Unemployment Rate at LA level: AIC values for Null and Random Intercept models using Local Authority level Proportion travelling 30km + and Industrial Diversity Indicator predictors

	Unemployment Rate – Null Models	Unemployment Rate – LA level **Proportion travelling 30km** predictor variable	Unemployment Rate – LA level **Industrial Diversity Indicator** predictor variable
Four Level model with random intercepts only	**5,097**	**5,067**	**5,067**
Two-level model with grouping by NUTS 3 areas, random intercepts only	5,133	5,090	5,103
Two-level model with grouping by NUTS 2 areas, random intercepts only	5,138	5,102	5,101
Two-level model with grouping by NUTS 1 areas, random intercepts only	5,158	5,127	5,134
Four level with random intercepts at all levels and random coefficients at NUTS 3, NUTS 2 and NUTS 1 levels but not at pseudo level		*5,069 Model failed to converge*	*5,069 boundary (singular) fit*
Four level with random intercepts at all levels and random coefficients at NUTS 2 areas level only		**5,065**	*5,069 boundary (singular) fit*
Four level with random intercepts at all levels and random coefficients at NUTS 1 areas level only		**5,068**	**5,069**

Of the random intercept models using the local authority level proportion of people travelling 30 km or more to work to model unemployment rates it was the four-level model random intercept model and the two four-level models with random intercepts at all levels and random coefficients at either NUTS 1 areas only or NUTS 2 areas only that fitted the data best.

Similarly, looking at the AIC values of the random intercept models using the local authority level industrial diversity indicator as the predictor to model unemployment rates, , it was the four-level model random intercept model and the four-level model with random intercepts at all levels and rando coefficients at NUTS 1 area level only that had the lowest AIC values indicating that they fitted the data better than any of the other models.

Interestingly the four-level random intercept model had the same AIC value as the four-level random intercept model for the proportion of people travelling 30 km or more to work. For both predictor variables in Annex 6 Table 4 it was the two-level model with grouping by NUTS 1 areas level that fitted the data the least well.

Annex 6 - Table 5: Unemployment Rate at LA level: AIC values for Null and Random Intercept models using Local Authority level Proportion with Occupational Diversity Indicator and proportion managers and professionals predictors

	Unemployment Rate – Null Models	Unemployment Rate –LA level **Occupational Diversity Indicator** predictor variable	Unemployment Rate –LA level **proportion managers and professionals** predictor variable
Four Level model with random intercepts only	**5,097**	**5,083**	**5,047**
Two-level model with grouping by NUTS 3 areas, random intercepts only	5,133	5,128	5,100
Two-level model with grouping by NUTS 2 areas, random intercepts only	5,138	5,118	5,083
Two-level model with grouping by NUTS 1 areas, random intercepts only	5,158	5,136	5,099
Four level with random intercepts at all levels and random coefficients at NUTS 3, NUTS 2 and NUTS 1 levels but not at pseudo level		*5,085 boundary (singular) fit*	*5,052 boundary (singular) fit*
Four level with random intercepts at all levels and random coefficients at NUTS 2 areas level only		**5,085**	**5,049**
Four level with random intercepts at all levels and random coefficients at NUTS 1 areas level only		*5,081 boundary (singular) fit*	**5,048**

Of the models using the local authority level occupational diversity indicator as the predictor to model unemployment rates, it was the four-level random intercept model and the four-level model with random intercepts at all levels and random coefficients just at NUTS 2 areas level that fitted the data best.

The AIC values of the models using the proportion of managers and professional as the predictor to model unemployment rates show the four-level random intercept model and

the two four-level models with random coefficients at all levels and random intercepts at one of NUTS 2 areas level or NUTS 1 areas level that fitted the data best.

Annex 6 - Table 6: Unemployment Rate at LA level: AIC values for Null and Random Intercept models using Local Authority level Proportion plant/process workers and proportion of people in rural areas including hub towns predictors

	Unemployment Rate – Null Models	**Unemployment Rate** – LA level **Proportion plant/process workers** predictor variable	**Unemployment Rate** – LA level **proportion of people in rural areas inc. hub towns** predictor variable
Four Level model with random intercepts only	**5,097**	**5,035**	**5,017**
Two-level model with grouping by NUTS 3 areas, random intercepts only	5,133	5,087	5,049
Two-level model with grouping by NUTS 2 areas, random intercepts only	5,138	5,062	5,062
Two-level model with grouping by NUTS 1 areas, random intercepts only	5,158	5,086	5,071
Four level with random intercepts at all levels and random coefficients at NUTS 3, NUTS 2 and NUTS 1 levels but not at pseudo level		*5,035 boundary (singular) fit*	*5,021 boundary (singular) fit*
Four level with random intercepts at all levels and random coefficients at NUTS 2 areas level only		**5,034**	**5,019**
Four level with random intercepts at all levels and random coefficients at NUTS 1 areas level only		*5,037 boundary (singular) fit*	**5,019**

Of the random intercept models using the proportion of plant and process workers as the predictor to model unemployment rates, it was the four-level random intercept model and the four-level model with random intercepts at all levels and random coefficients at just NUTS 2 areas level that fitted the data best.

Observing the AIC values of the models using the proportion of people who live in rural areas including hub towns as the predictor to model employment rates, it was the four-level model random intercept model and the two four-level models with random intercepts

at all levels and random coefficients at either NUTS 2 areas level or NUTS 1 areas level that fitted the data best.

	Unemployment Rate – Null Models	**Unemployment Rate** – LA level **social housing** predictor variable	**Unemployment Rate** – LA level **proportion non-UK born** predictor variable
Four Level model with random intercepts only	**5,097**	**4,998**	**5,046**
Two-level model with grouping by NUTS 3 areas, random intercepts only	5,133	5,013	5,079
Two-level model with grouping by NUTS 2 areas, random intercepts only	5,138	5,030	5,089
Two-level model with grouping by NUTS 1 areas, random intercepts only	5,158	5,051	5,118
Four level with random intercepts at all levels and random coefficients at NUTS 3, NUTS 2 and NUTS 1 levels but not at pseudo level		*5,001 boundary (singular) fit*	*5,052 boundary (singular) fit*
Four level with random intercepts at all levels and random coefficients at NUTS 2 areas level only		**4,998**	**5,048**
Four level with random intercepts at all levels and random coefficients at NUTS 1 areas level only		**4,997**	5,048 *boundary (singular) fit*

The models using the proportion of social housing in each local authority as the predictor to model unemployment rates produced a typical pattern. That is to say, for this predictor for unemployment rates it was the four-level random intercept model and the two four-level models with random intercepts at all levels and random coefficients at either NUTS 2 areas level or at NUTS 1 areas level that had the lowest AIC values indicating that they fitted the data better than the other models.

The AIC values of the models using the proportion of people who born outside the UK as the predictor to model unemployment rates also showed a fairly typical pattern with the four-level random intercept model and the four-level model with random intercepts at all

levels and random coefficients at NUTS 2 areas level only having the lowest AIC values indicating that they fitted the data better than all the other models.

Annex 6 - Table 8: Unemployment Rate at LA level: AIC values for Null and Random Intercept models using Local Authority level IMD 2015 and employment deprivation predictors

	Unemployment Rate – Null Models	Unemployment Rate – LA level IMD 2015 predictor variable	Unemployment Rate – LA level employment deprivation 2015 predictor variable
Four Level model with random intercepts only	5,097	**4,836**	**4,862**
Two-level model with grouping by NUTS 3 areas, random intercepts only	5,133	4,886	4,929
Two-level model with grouping by NUTS 2 areas, random intercepts only	5,138	4,858	4,886
Two-level model with grouping by NUTS 1 areas, random intercepts only	5,158	4,883	4,895
Four level with random intercepts at all levels and random coefficients at NUTS 3, NUTS 2 and NUTS 1 levels but not at pseudo level		*4,842 boundary (singular) fit*	*4,868 boundary (singular) fit*
Four level with random intercepts at all levels and random coefficients at NUTS 2 areas level only		**4,838**	**4,864**
Four level with random intercepts at all levels and random coefficients at NUTS 1 areas level only		*4,838 boundary (singular) fit*	*4,864 boundary (singular) fit*

Of the models using the average IMD 2015 rank for each local authority as the predictor to model unemployment rates, in common with models using many other predictor variables, it was the four-level random intercept model and one of the four-level models with random intercepts at all levels and random coefficients at just one level, in this case just at NUTS 2 areas level, that fitted the data best.

Similarly, for the models using the average 2015 employment deprivation rank for each local authority as the predictor to model unemployment rates it was again the four-level random intercept model and the four-level model with random intercepts at all levels and random coefficients at NUTS 2 areas level only that fitted the data best.

These findings are different to those for the equivalent models of employment rates as in those the AIC values for all of the different models were much more similar to each other.