

www.defra.gov.uk

Educational Attainment in Rural Areas

A Report Prepared for the Department for Environment, Food and Rural Affairs by the National Centre for Social Research (NatCen).

31 December 2009



department for
children, schools and families



© Crown copyright 2009

Copyright in the typographical arrangement and design rests with the Crown.

This publication (excluding the Royal Arms and departmental logos) may be re-used free of charge in any format or medium for research for non-commercial purposes, private study or for internal circulation within an organisation. This is subject to it being re-used accurately and not used in a misleading context. The material must be acknowledged as Crown copyright and the title of the publication specified.

For any other use of this material please apply for a Click-Use Licence for Public Sector Information (PSI) or core material at:

<http://www.opsi.gov.uk/click-use/psi-licence-information/index.htm>

write to:

Office of Public Sector Information
Information Policy Team
Kew
Richmond
Surrey
TW9 4DU

or email: licensing@opsi.gsi.gov.uk

Further information about this publication is available from:

rural.research@defra.gsi.gov.uk

Published by the Department for Environment, Food and Rural Affairs
Nobel House
17 Smith Square
London SW1P 3JR

Tel: 020 7238 6000

Website: www.defra.gov.uk

This report was prepared for the Department for Environment, Food and Rural Affairs by the National Centre for Social Research (NatCen).

Contents

1 Introduction

Introduction.....	5
Aims of the Study	5
Dataset and Methods	6
Key Findings.....	7
Conclusions and Recommendations	8
1.1 Background.....	10
1.2 Aims of the Project.....	11

2 Literature Review 12

2.1 Definitions of Rural and Urban Areas.....	12
Historical definitions.....	12
Summary of This Section	12
The 2004 definition	14
2.2 Studies Using Pre-2004 Definitions of Urban and Rural	15
2.3 Studies Using the 2004 Classification.....	17
2.4 Potential Factors Explaining Differences in Attainment	18
2.5 Points Arising from Literature Review	20
2.6 Conclusions	21

3 The LSYPE Dataset 22

3.1 Background of the Dataset	22
3.2 Purpose of the LSYPE Study.....	22
Summary of This Section	22
3.3 Information Available from the Study	23
Data Linkage	23
3.4 Sampling and Response Rates	24
3.5 Conclusions	25

4 Analysis Strategy 26

4.1 Variables Included in Analyses	26
Geographical Variables	26
Individual Pupil Characteristics.....	26
Summary of This Section	26
School-level Characteristics	27
LA-level Characteristics.....	27
Educational Outcomes	28
4.2 Stages of Analysis	28
Stage 1: Describing Differences Between Rural and Urban Areas.....	28
Stage 2: Exploring Attainment in Particular Groups.....	29
Stage 3: School and Local Authority Level Variation	29
Stage 4: Regional Variation.....	29

5 Results	30
5.1 Do Pupils Living in Rural Areas Really Have Higher Attainment?.....	30
Where do the LSYPE Respondents Live?	30
Summary of This Section	30
How Does Attainment Differ by Settlement Type?	31
5.2 What Are the Factors Behind These Differences?	36
Binary Definition of Rural and Urban Areas.....	36
Summary of This Section	36
5.3 Which Settlement Types Have the Highest Attainment?.....	40
2004 Definition of Rural and Urban Areas	40
Summary of This Section	40
5.4 Rural Pupils or Rural Schools?	43
Summary of This Section	43
5.5 Do Particular Groups of Pupils in Rural Areas Have Lower Attainment?	45
Summary of This Section	45
5.6 How Important is the School or Local Authority?	50
Variation in Attainment Between Schools and Local Authorities	50
Different Relationships in Different Schools	50
Summary of This Section	50
5.7 Are There Regional Differences?	51
Summary of This Section	51
6 Conclusions and Recommendations	56
6.1 Differences Between Rural and Urban Areas	56
Recommendations:.....	56
6.2 Particular Groups of Pupils with Lower Attainment	56
Recommendations:.....	57
6.3 Regional Differences in Attainment.....	58
Recommendations:.....	58
Appendix A Glossary of Terms	62
Appendix B Tables of Results.....	64
Appendix C Factors Not Included in Final Models	67
Appendix D Value-Added Analyses.....	68
Appendix E Standard Deviations.....	72
Appendix F Derived Variables	73
Appendix G Sample Sizes	75
Rural Areas	75
Minority Ethnic Groups	76
Government Office Region	76

Appendix H Technical Methods	77
Missing Data and Multiple Imputation	77
Multilevel Models	77
Multiple Membership Models	78
Appendix J Sample Design and Weighting	80
Sampling from Maintained Schools	80
Sampling from Independent Schools and PRUs	80
Sample Exclusions	81
Longitudinal Sampling	81
Response Rates	81
Weighting	81

Tables

Table 2.1 Definitions of Rural and Urban Areas	13
Table 2.2 Strengths and Weaknesses of Main Types of Classification ...	Error!
Bookmark not defined.	
Table 2.3 Rural and Urban Area Classification 2004 (Output Area Level).....	14
Table 2.4 ... Studies of Educational Attainment Using Pre-2004 Definitions of Urban and Rural	16
Table 2.5 Studies of Educational Attainment Using the 2004 Classification of Urban and Rural	17
Table 2.6 Potential Advantages of Living in Rural Areas.....	18
Table 2.7 Potential Disadvantages of Living in Rural Areas	19
Table 4.1 Geographical variables from NPD	26
Table 4.2 Individual pupil characteristics from LSYPE data	27
Table 4.3 School-level characteristics from LSYPE/NPD data	27
Table 4.4 LA-level characteristics from LSYPE/NPD data.....	28
Table 4.5 Educational outcomes from NPD data.....	28
Table 5.1 Factors Influencing the Relationship Between Settlement Type and Attainment.....	39
Table 5.2 Relationships Between Settlement Type and Attainment by Government Office Region, Key Stage 3.....	51
Table 6.1 Variables Affecting the Relationship Between Settlement Type and Attainment at Key Stage 3	64
Table 6.2 Variables Affecting the Relationship Between Settlement Type and Attainment at Key Stage 4	65
Table 6.3 Relationship Between Settlement Type (8 Categories) and Attainment at Key Stage 3, Adjusted	65
Table 6.4 Relationship Between Settlement Type (8 Categories) and Attainment at Key Stage 4, Adjusted	66
Table 6.5 Relationship Between Settlement Type of School (8 Categories) and Attainment at Key Stage 3, Adjusted	66
Table 6.6 Relationship Between Settlement Type of School (8 Categories) and Attainment at Key Stage 4, Adjusted	66
Table 6.7 Derived Variables Used in Analyses.....	73

1 Summary

Introduction

The comparison of key indicators between rural and urban areas has been of increasing interest in recent years, particularly since the inception of Defra. Educational attainment among secondary school pupils forms one of these key indicators, and is particularly interesting given that it appears to identify a rural advantage. Data from the National Pupil Database show that pupils living in rural areas tend to have higher secondary school attainment than those living in urban areas, suggesting that there may be intrinsic benefits of living in rural areas for young people. However, the same data also show that there is widespread regional variation in rural attainment, and that pupils living in some affluent rural areas do not have the high attainment that would be expected.

We undertook a detailed literature review of studies looking at rural attainment in order to determine whether previous studies had successfully investigated these differences. A number of the results we found were conflicting, but in general the studies tended to find that pupils in rural areas do have higher attainment, particularly if they live in smaller and more dispersed settlements. Some regional differences were also found, particularly in the East of England, where pupils living in sparse areas actually had lower attainment than urban pupils, and in the West Midlands, where rural pupils were more likely than urban pupils to leave school with no qualifications. However, none of these studies had attempted to adjust for factors like higher social position in rural areas, and only one had used value-added analyses which control for previous attainment. This study therefore represents the first in-depth investigation of these relationships using methods that can account for the complexity of the issue.

Aims of the Study

The main aims of this project are set out below:

- To provide an assessment of whether there are “true” differences in educational attainment at Key Stages 3 and 4 between rural and urban areas, once factors such as social position have been controlled for.
- To explore factors that particularly affect attainment in rural areas, and how these may vary across different types of rural settlement.
- To quantify how much of the variation in attainment between rural and urban areas is due to differences between individuals, between schools and between Local Authorities.
- To describe differences in attainment between regions of England and the factors that may explain these differences.

Dataset and Methods

To explore the relationship between settlement type and secondary school attainment, we used pupils recruited into the Longitudinal Study of Young People in England (LSYPE). This is a study which began in 2004 by interviewing over 15,000 young people (aged 14 at the beginning of the study) sampled from schools in England, as well as their main and secondary parents if these were available. The same young people have been re-interviewed every year, and we were therefore able to follow their progress up to age 16.

The LSYPE dataset contains information on the young person's family characteristics, theirs and their parents' attitudes and aspirations and their experiences of school. The data have also been linked to the National Pupil Database (NPD), which not only allows us to record these pupils' attainment at Key Stages 2, 3 and 4, but also provides information on school-level and Local Authority-level factors such as ethnic composition and the proportion of pupils receiving free school meals. Finally, the NPD also contains information on geographical indicators, including an 8-category measure of settlement type (based on settlement size and sparsity) as well as measures of area deprivation.

We analysed the LSYPE dataset using multilevel models, because in educational data pupils are usually clustered within schools, and schools are in turn clustered within Local Authorities. This means that any two pupils who attend the same school are likely to have more similar attainment than if they attended different schools (because they will attend the same classes and have the same teachers), and likewise any two schools within the same Local Authority are likely to have more similar average attainment (because they will have similar education policies). Clustering can cause problems for analyses because conventional models assume that all the pupils are independent of one another, which with clustering is not the case. Our multilevel models took account of this, and also allowed us to calculate the amount of variation in attainment attributable to differences between pupils, between schools and between Local Authorities. Finally, we also took account of pupils moving between different secondary schools from Key Stages 2-4, by weighting our analyses according to the amount of time each pupil had spent in each school.

First, we used the LSYPE data to explore whether there were differences in attainment at Key Stages 3 and 4 between rural and urban areas, as found in data from the NPD. Then, we added other factors measured by LSYPE and the NPD (including social position, previous attainment, individual pupil characteristics and factors measured at the school and Local Authority levels) into the models in order to see whether these differences were partly or wholly due to such other factors. Next, we tested for interactions to see whether particular groups of pupils might

have particularly high or low attainment if they lived in rural areas, and we also looked at the amount of variation in attainment that was due to differences between schools and Local Authorities. Finally, we used data on the Government Office Region of the pupils to explore whether there were any different relationships between settlement type and attainment in particular regions of England.

Key Findings

When we compared average attainment at Key Stages 3 and 4 between pupils living in rural and urban areas, we found that the rural pupils did have slightly higher attainment, although this difference was small and only statistically significant at Key Stage 3. However, when we adjusted for other factors that also affect attainment, we found that the higher rural attainment was largely due to the higher social position, lower area deprivation and higher prior attainment at Key Stage 2 found among rural pupils. Once these had been adjusted for, attainment among rural pupils was virtually no different from that of urban pupils. We also found that the pre-existing higher attainment at Key Stage 2 among rural pupils was likely to be due to their higher social position earlier in life¹. There was some evidence of slightly higher attainment among pupils living in very small and sparse areas, but because the sample sizes for these area types were so small we cannot confirm that these differences are genuine.

We therefore concluded that there are no characteristics intrinsic to rural areas that lead to higher attainment among rural pupils. Instead, we found that the higher attainment in rural areas is largely due to greater affluence and might not necessarily affect pupils who are not from affluent backgrounds. The next stage of our analyses followed on from these findings by examining particular groups of pupils that we felt might be particularly at risk of lower attainment if they lived in rural areas. We found that pupils from certain minority ethnic groups tended to make slightly less progress in attainment if they lived in rural compared to urban areas (specifically Black African pupils at Key Stage 3 and Indian pupils at both Key Stages 3 and 4). However, because there are only small numbers of minority ethnic pupils living in rural areas, these analyses are based on small sample sizes and require further confirmation using more detailed rural data. We also found that pupils whose mothers have low levels of qualifications make less progress in rural compared to urban areas, and this difference became larger and more significant from Key Stage 3 to Key Stage 4. These analyses are not prone to such problems of small sample sizes, and seem to indicate that certain types of disadvantaged pupils are additionally disadvantaged by living in rural areas.

Our models showed that around 20% of the variation in attainment among LSYPE respondents was attributable to differences between schools, whereas only a tiny percentage was attributable to differences between Local Authorities. When we investigated the between-school differences further, we found that different schools had different relationships between settlement type and attainment. The positive relationship between living in a rural area and attainment was strongest in schools with low average attainment and weakest in schools with high average attainment. This means that urban pupils in low-attaining schools have the lowest attainment of all, whereas rural pupils in these schools are more likely to be advantaged. However, rural pupils attending high-attaining schools are unlikely to have any advantage over urban pupils, and may even have lower attainment, since we found that in some cases the relationship was reversed.

Finally, our regional analyses showed that four regions had different relationships between settlement type and attainment when compared to the rest of England. The North West, Yorkshire and the Humber and the East Midlands all had a slightly stronger relationship between settlement type and attainment, which did seem to indicate some advantages for pupils living in rural areas. However, these relationships were often based on small sample sizes in the more sparse areas, and cannot therefore be confirmed without further investigation. We also found that pupils in the South East actually had *lower* attainment if they lived in rural areas, and this relationship was based on a much larger sample size. This seems to indicate a different pattern of attainment in rural and urban areas in the South East from that seen in the rest of England. This may be because of the relative affluence of this region, and may indicate that poorer rural pupils are being less well supported.

Conclusions and Recommendations

This study has shown that rural pupils do not tend to have substantially higher attainment once their higher social position is accounted for. Indeed, our results indicate that some groups of rural pupils may actually be disadvantaged compared to their urban counterparts. These groups may include certain minority ethnic pupils (although this relationship is yet to be confirmed), those with low qualified mothers and those living in the South East of England.

Research into inequalities has shown that disadvantaged people who live in affluent areas can actually be worse off than those who are surrounded by other disadvantaged people. Our results may indicate a similar pattern, in that some

¹ Because our subsequent analyses adjust for prior attainment at Key Stage 2, they are therefore value-added analyses and measure progress in attainment during secondary school rather than raw attainment.

types of disadvantaged pupils living in the generally more affluent rural areas of England appear to make less progress than disadvantaged pupils in the less affluent urban areas. These include rural pupils with low qualified mothers, those living in the affluent South East and those attending higher-attaining schools.

We would therefore recommend additional support for disadvantaged pupils in rural areas in order to close these gaps in attainment, as well as further dedicated research on education in rural areas in order to overcome the small rural sample sizes we encountered in LSYPE. We would also suggest that interventions could be organised at the school level as well as the individual level, since we found that a reasonably large amount of the variation in attainment was due to differences between schools.

2 Introduction

2.1 Background

This project aims to inform future policy by exploring differences in educational attainment between young people living in rural and urban areas of England. The delivery of education in rural areas has been an area of increasing interest in recent years, in particular since the inception of Defra in 2001 (Environment, Food and Rural Affairs Committee, 2003), and while there is still relatively sparse evidence for comparison of different types of settlement, differences between urban and rural areas have already been identified. For example, summary statistics comparing the educational attainment of rural pupils to those in urban areas clearly indicate that, on average, rural pupils tend to do better (Commission for Rural Communities, 2008). This difference is outlined in Defra's Departmental Strategic Objective for socially and economically sustainable rural communities (Defra, 2008). However, these average differences appear to mask a much more complicated relationship between area of residence and educational attainment that has not yet been thoroughly investigated. A recent State of the Countryside report showed that there is wide regional variation in educational attainment, and that this does not seem to correspond closely to regional economic performance, with some rural areas showing lower than expected attainment despite rural areas generally being more affluent and having better quality of life (Commission for Rural Communities, 2008). Clearly, the factors behind these statistics need to be investigated in more depth.

While there has been little formal research investigating possible factors behind differences in academic achievement in different types of settlement, previous reports have suggested that there may be a much wider range of achievement in rural areas, such that although average attainment is higher, certain groups of pupils might also have lower attainment (Commission for Rural Communities, 2008). In particular, there is evidence that low achievers in rural areas may have poorer results than low achievers in urban areas (Commission for Rural Communities, 2006). It is therefore important not only to determine whether average differences in attainment persist when adjusting for other factors that might explain them, but also to explore whether these average differences might conceal important variation within or between rural areas.

2.2 Aims of the Project

- To provide an assessment of whether there are “true” differences in educational attainment at Key Stages 3 and 4 between rural and urban areas, once factors such as socioeconomic position have been controlled for.
- To explore factors that particularly affect attainment in rural areas, and how these may vary across different types of rural settlement.
- To quantify how much of the variation in attainment between rural and urban areas is due to differences between individuals, between schools and between Local Authorities.
- To describe differences in attainment between regions of England and the factors that may explain these differences.

3 Literature Review

Summary of This Section

Previous studies have produced conflicting results about whether people living in rural areas tend to have higher or lower attainment than those living in urban areas. Adults living in rural areas tend to have slightly lower qualifications than those living in urban areas, but among young people currently in school, attainment appears to be slightly higher in rural areas. Most of these studies do not capture the differences between various types of rural area, and instead present a simple distinction between rural and urban. However, there is an 8-category classification of rural and urban areas that appears to be useful for distinguishing between different types of rural area. This will be used in our analyses.

previous research relating to educational attainment in rural areas. This is not a simple task, as most of the literature exploring educational attainment does not make the distinction between urban and rural areas, and most of the literature examining differences between rural and urban areas does not look at attainment (Commission for Rural Communities, 2006b). However, we were able to identify a number of sources that had explored this question, and a number of additional sources that had investigated factors that might be behind any differences in attainment between different settlement types.

3.1 Definitions of Rural and Urban Areas

Historical definitions

Historically, definitions of what constitutes urban and rural areas have been developed to serve a variety of purposes, and consequently they have been based on a number of different criteria. The need to distinguish between rural and urban areas of England had manifested itself at least as early as the nineteenth century (Office for National Statistics, 2001), when official statistics were reported separately for urban and rural areas due to the increasing separation of society from the land, a process which began in earnest with the Industrial Revolution. However, little is known about how these areas were defined until the 1950s, when various criteria began to be used for the definition of settlement types.

From this time until 2004, there was no universally agreed definition of what constituted an urban or rural area of England. The various definitions used relied on criteria such as land use, population density, social and economic characteristics and remoteness or deprivation – see Table 2.1 for a list of some

classifications that have been used by various government bodies in recent decades.

Table 2.1 Definitions of Rural and Urban Areas

Definition	Used by
Population density or settlement size	Department of the Environment (1974), DETR/RDC/CA (1995), RDC/Housing Corporation (1999)
Land use	Department of the Environment/OPCS (1981), Department of the Environment/OS/ONS (1991), ODPM (2001)
Social and economic characteristics	OPCS/ONS (1999), DETR/RDC/CA (1995)
Remoteness and deprivation	DETR/RDC/CA (1995)

Data taken from Office of National Statistics (2001).

Each type of definition has its associated strengths and weaknesses. The main categories of these associated with the different types of area classifications are summarised below in Table 2.2.

Definition	Main Strengths	Main Weaknesses
Population density or settlement size	Easy to calculate and understand, and is a continuous measure so the results can be ranked.	Difficult to characterise different area types that may have the same population size. Level of measurement can affect categorisation into different settlement types. Setting of arbitrary threshold can mean that areas similar in characteristics can fall on either side of the threshold.
Land use	Can take account of the character of non-residential as well as residential areas.	Land use is generally defined as urban, so rural areas are only defined in a residual capacity. This makes such definitions less useful for exploring characteristics of rural areas.
Socio-economic factors	Can characterise areas according to the types of people who live there – may be useful for measuring rural deprivation.	Difficulty in identifying what variables should be used to derive the classification. Social and economic structure of many areas is constantly and rapidly changing. In many cases urban and rural areas may no longer differ.
Remoteness and deprivation	Distance to nearest neighbour is a continuous measure so can be easily ranked.	Difficult concepts to define, and may rely on factors such as distance to local amenities which can rapidly change. Measures of crow-fly distance can be inaccurate and variable across different types of area.

The 2004 definition

As a consequence of this, in 2001 the then Department of Transport, Local Government and the Regions (DTLR) commissioned a review of the definitions in use with a view to suggesting recommendations for future definitions (Office of National Statistics, 2001). This led to the development of a new single classification of rural and urban areas in 2004, which is the classification now adopted by Defra (Bibby and Shepherd, 2004). The new classification uses the 2001 Urban Settlements definition to identify urban areas of more than 10,000 population, and classifies all other areas as rural. Further, areas are classified according to morphology (defined by looking at density profiles formed from densities at different scales) and sparsity (defined by looking at density profiles at much larger scales) based on hectare grid squares (Countryside Agency *et al*, 2004). The types of settlement distinguished by the classification are summarised in Table 2.3.

Table 2.2 Rural and Urban Area Classification 2004 (Output Area Level)

Settlement Type	Features
Urban (less sparse)	Population \geq 10,000 and area is not in sparsest 5% of density at 10km, 20km and 30km radius.
Urban (sparse)	Population \geq 10,000 and area is in sparsest 5% of density at 10km, 20km and 30km radius.
Small town and fringe (less sparse)	Population < 10,000, defined according to density profile as small town or fringe area, and area is not in sparsest 5% of density at 10km, 20km and 30km radius.
Small town and fringe (sparse)	Population < 10,000, defined according to density profile as small town or fringe area, and area is in sparsest 5% of density at 10km, 20km and 30km radius.
Village (less sparse)	Population < 10,000, defined according to density profile as village, and area is not in sparsest 5% of density at 10km, 20km and 30km radius.
Village (sparse)	Population < 10,000, defined according to density profile as village, and area is in sparsest 5% of density at 10km, 20km and 30km radius.
Dispersed (less sparse)	Population < 10,000, defined according to density profile as dispersed, and area is not in sparsest 5% of density at 10km, 20km and 30km radius.
Dispersed (sparse)	Population < 10,000, defined according to density profile as dispersed, and area is in sparsest 5% of density at 10km, 20km and 30km radius.

Data taken from Bibby and Shepherd (2004).

As can be seen from Table 2.3, this definition is a settlement-based one (Countryside Agency *et al*, 2004), and relies on categorising the physical environment rather than the inhabitants of the area. The use of this definition solves a number of the weaknesses of previous definitions mentioned in Table 2.2. For example, the problem of sparsity at different levels of measurement that is

characteristic of classifications using settlement size or population density (Martin *et al*, 2000) has been overcome in the 2004 definition by using density profiles at different radii from each central cell.

However, this approach to a single definition of urban and rural areas is not without its problems, since no one measure of urban-rural can thoroughly capture such a multidimensional concept (Martin *et al*, 2000). The main potential difficulties with using such a classification are presented below:

- The classification does not consider land use beyond that used for residential purposes – this might be an important dimension of the character of an area e.g. in coastal or ex-mining communities.
- The classification cannot be used to consider the characteristics of *individual* local areas.
- It does not consider the residents' own conception of what their area is like.
- It does not consider socio-economic characteristics of areas – this may also be an advantage, as such characteristics can then be compared across different settlement types.
- The classification uses an arbitrary cut-off point to distinguish between urban and rural areas – e.g. some settlements may be on the urban side of the cut-off but may be more rural in character, such as some market towns.

These potential limitations will be discussed below in Section 2.3 with regard to recent studies that have used the 2004 classification to define urban and rural areas.

3.2 Studies Using Pre-2004 Definitions of Urban and Rural

Few studies could be found that had explicitly examined differences in educational attainment between rural and urban areas before the introduction of the 2004 definition. Three studies were found that had explored such differences using clearly defined definitions (one of which was actually published after the introduction of the 2004 definition), and the main features of these are summarised below:

Table 2.3 Studies of Educational Attainment Using Pre-2004 Definitions of Urban and Rural

Study	Year	Definition Used	Population	Results
Byner and Joshi	2002	Population density	NCDS and BCS70	Rural males from NCDS had significantly lower qualifications than urban males, but in BCS70 both males and females in rural areas had significantly higher qualifications. There was no interaction between rurality and social class.
Rural Evidence Research Centre	2004	1993 RDC definition (divided into accessible and remote rural areas)	2001 Census Data	People with no qualifications were more common in remote rural areas than in England as a whole, but there was no evidence of a north-south division. Good educational attainment (degree level) was slightly less common in rural areas, again particularly in remote rural areas, and particularly in the East and South West of England.
Gibbons and Silva	2008	School density, land use and population density	NPD/PLASC	In unadjusted analyses, pupils attending schools in rural areas tended to perform better, but in value-added analyses it was found that pupils who moved to secondary schools in areas with higher school density tended to improve their attainment more from primary to secondary school.

It can be seen from Table 2.4 that studies using varied definitions of urban and rural have tended to find differences in attainment between rural and urban areas, but there is no direct agreement as to the direction of these differences and little investigation into the factors behind them. The 2008 study, which is the closest in methodology to our current project, was able to investigate some possible factors behind the differences by using PLASC/NPD data, but did not have information on other individual-level factors which may also be very important in determining attainment and may differ between settlement types. The following section will examine research that has used the 2004 definition of rural and urban areas in order to see whether such studies are more comparable and can build up a clearer picture of differences between areas.

3.3 Studies Using the 2004 Classification

As described at the beginning of this review, very few studies exist which have explored differences in educational attainment specifically according to settlement type. In an extensive search, we found only three studies which had investigated such differences using the 2004 classification adopted by Defra, and two of these were government reports produced by the Commission for Rural Communities (CRC). The first was published in 2006, and explores the aspirations of young people in the rural West Midlands (Commission for Rural Communities, 2006a), and the second was published in 2008 and is part of a general State of the Countryside report (Commission for Rural Communities, 2008). The third was published in 2005 and compares the attainment of adults in rural and urban areas in the East of England. This report was produced in 2005 by the East of England Skills and Competitiveness Partnership, and uses 2001 census data in a similar way to the 2004 study identified in the previous section. The main results of these three reports are summarised below in Table 2.5.

Table 2.4 Studies of Educational Attainment Using the 2004 Classification of Urban and Rural

Study	Year	Area Covered	Population	Results
East of England Skills and Competitiveness Partnership	2005	East of England	Census (2001)	When Super Output Areas were divided into urban, sparse rural and less sparse rural, the less sparse rural areas were found to have the highest attainment but the sparse rural areas had the lowest attainment among adults.
Commission for Rural Communities	2006	West Midlands	NPD (2002/2003)	When regions within the West Midlands were ranked according to average A/AS Level point scores, the rural regions tended to perform better. However, in rural areas there were also slightly higher levels of young people with no qualifications.
Commission for Rural Communities	2008	England	NPD (2005/2006)	When percentages achieving 5+ A*-C grades and 5+ A*-G grades at GCSE were compared, rural areas (particularly villages and dispersed areas) had higher attainment, particularly in gaining 5+ A*-C grades.

It appears that thus far no studies have taken advantage of the 2004 classification to explore in detail the differences in attainment among different types of rural areas and the potential reasons behind this, a fact which is complicated by the

conflicting results produced by these studies. As a prelude to describing how we intend to build on the results of these previous studies and combine their various strengths, the next section will summarise various factors that are thought to affect attainment and that might also differ between urban and rural areas of England.

3.4 Potential Factors Explaining Differences in Attainment

The principal factors identified through a review of the literature as being potentially important for educational attainment in rural areas are summarised in tables 2.6 and 2.7 below, which comprise respectively a list of potential advantages and disadvantages of living in rural areas. Some factors may apply particularly to certain types of rural area, and this has been noted in the tables. Mention is also made of whether there is any evidence that these factors actually differ between urban and rural areas, and of whether there is any evidence that they relate to educational attainment.

Table 2.5 Potential Advantages of Living in Rural Areas

Factor	Differs between urban and rural areas?	Differences in educational attainment?
Less overcrowding (particularly in sparse areas)	Rural schools tend to be smaller and less oversubscribed (Environment, Food and Rural Affairs Committee, 2003)	Both small and large schools have been found to have lower attainment, so there may not be any particular advantage to rural schools (Bradley and Taylor, 1998), but no evidence could be found as to whether oversubscription in schools impacts on attainment
Better safety (particularly in sparse areas)	The risk of being a victim of crime is lower in rural areas for all age groups (Kershaw et al., 2008)	No UK-based data could be found that link safety to educational attainment.
Lower levels of population mobility (particularly in remote areas)	Mobility tends to be higher in urban schools (Lupton, 2004)	Pupil mobility has been found to impact negatively on school quality and is linked to deprivation (Dobson, 2008)
Less overcrowding in supporting services e.g. libraries (particularly in sparse areas)	Areas of higher population density have fewer libraries and shorter opening hours, potentially leading to more overcrowding (Hammond, 2002)	No UK-based data could be found that link overcrowding in supporting services to educational attainment
Higher social position (particularly in more accessible areas)	Rural England has lower average levels of deprivation and higher average gross income (Commission for Rural Communities, 2008)	Higher social position is strongly associated with educational attainment (Glaeser et al., 2000)

Of the potential advantages suggested, evidence was found linking all of these to settlement type and also to educational attainment, with the exception of safety and overcrowding in supporting services. It would therefore be desirable to examine school size (and other elements linked to possible overcrowding in and quality of schools), population mobility and social position in our analyses.

Table 2.6 Potential Disadvantages of Living in Rural Areas

Factor	Differs between urban and rural areas?	Differences in educational attainment?
Less competition between schools (particularly in sparse areas)	Urban pupils have the choice of more local schools and these are therefore likely to be more competitive (Gibbons and Silva, 2008)	Competition (between primary schools) has not been shown to be related to performance in England, although it has in the US (Gibbons and Silva, 2008)
Poor transport (particularly in remote areas)	Availability and cost of public transport in rural areas were thought to be disadvantages by contributors to an inquiry (Environment, Food and Rural Affairs Committee, 2003) – the average education journey is 5.1 miles in rural areas and 2.8 miles in urban areas (Fox, 2007)	Quality of school transport has been found to have an impact on school attendance, which in turn is linked to attainment (Murphy, 2007)
Smaller pool of teachers to draw from (particularly in sparse areas)	Urban schools tend to have larger pools of available teachers (Gibbons and Silva, 2008)	No UK-based evidence could be found linking the size of teacher pool to educational attainment
Smaller range of subjects taught (particularly in sparse areas)	Rural schools are often unable to provide a broad curriculum (Commission for Rural Communities, 2006b)	No UK-based evidence could be found linking the range of subjects taught to educational attainment
Poor career opportunities and lower aspirations (particularly in sparse and remote areas)	Low aspirations caused by poor availability of jobs may be a problem in particular rural areas (Commission for Rural Communities, 2006b)	Low parental aspirations are associated with educational attainment and may account for much of the variation associated with social position (Sacker et al., 2002)
More homogeneous ethnic makeup (particularly in remote areas)	Most people from ethnic minority groups in England live in urban areas (Commission for Rural Communities, 2006b)	Some minority groups tend to have higher attainment than the majority and some lower, so this may be a complex relationship (Heath and Brinbaum, 2007), although overall young people from ethnic minorities have higher attainment (Bradley and Taylor, 2004)
Poor support for pupils with SEN (particularly in sparse areas)	Rural schools are less likely to have specialist provision for those with learning difficulties (Commission for Rural Communities, 2006b)	Better provision for those with SEN has been found either to have no impact on attainment or to increase it (Kalambouka <i>et al</i> , 2007)

Factor	Differs between urban and rural areas?	Differences in educational attainment?
Fewer supplementary services (particularly in sparse areas)	Urban secondary schools are more likely to have family support workers and parent support facilities (Clemens et al., 2005)	No UK-based evidence could be found linking the provision of supplementary services to educational attainment

Of the potential disadvantages suggested, evidence was found linking all of these to settlement type and to educational attainment, apart from size of teacher pool, range of subjects taught and the provision of supplementary services. It would therefore be desirable to examine competition between schools, transport, parental and own aspirations, ethnic diversity and support for pupils with SEN in our analyses.

3.5 Points Arising from Literature Review

The above literature review has unearthed a number of important issues for the present research project, and these are summarised below. Most of these points have been specifically considered in the design of the analyses reported, and we hope that they address some of the limitations of previous analyses exploring differences in educational attainment between urban and rural areas.

- Simple dichotomies between urban and rural do not generally provide a detailed enough picture of differences – we hope to avoid this potential problem by using the 2004 classification, which can discriminate between different types of rural area, but there may still be a problem with distinguishing particularly remote areas, which one study has suggested may be the most disadvantaged.
- It may be important to identify regional differences in attainment, as rural areas in different parts of England appear to fare differently – this should be possible in our analyses, although sample sizes may be small for some regions.
- School density may be an important dimension of settlement type with regard to educational attainment – this is beyond the scope of the current research project but may be possible in future research.
- Cohort studies are a useful way of exploring attainment, because they can track changes in individuals over time – we can exploit this advantage by using the LSYPE dataset.

- When looking at educational attainment, it is important to consider prior attainment and to use value-added analyses (Gibbons and Silva, 2008) – we are able to do this using the NPD/PLASC database.
- It may be important to adjust for school-level and individual-level factors using multilevel models, as such models can properly represent the complexity of pupil data – this is possible in our analyses as we have school-level and individual-level data.
- The previous section showed that it would be desirable to investigate the relationship between inter-school competition and attainment - this is not possible using the data we have available, but we do have many other school-level measures that we can adjust for in the analyses.

3.6 Conclusions

The division of residences into urban and rural areas is both complex and problematic, as evidenced by the fact that until 2004 no single classification system had been adopted in England. A simple dichotomous classification separating rural and urban areas does not capture many of the nuances of different settlement types, and the use of thresholds of population size can be arbitrary in its division of settlements of different sizes. Nonetheless, it is helpful to have a single classification to refer to for the purposes of comparison, and we hope that the 2004 classification employed by Defra will make the task of reviewing urban/rural literature a more simple one in the future.

A number of factors were found to be associated in previous literature with both settlement type and educational attainment, and we will explore these factors in our own analyses, in addition to a number of other variables available within the LSYPE/NPD dataset. It is difficult to assess the results of previous studies that have explored the relationship between settlement type and attainment, as these are contradictory and are the result of very different statistical models. However, we have attempted to design analyses that use the positive features of these previous models while eliminating some of their design flaws, and a description of the present dataset along with clarification of the analyses used can be found in the following two sections.

4 The LSYPE Dataset

Summary of This Section

The dataset used in this project is taken from the first three waves of a longitudinal study of young people in England which began in 2004 when the respondents were aged 14. When matched to the National Pupil Database, this dataset provides information on whether the young people live in rural or urban areas and their attainment at Key Stages 3 and 4, as well as individual, school and Local Authority level characteristics.

4.1 Background of the Dataset

The Longitudinal Study of Young People in England (LSYPE) is a large, nationally representative survey designed to follow a single cohort of young people from the age of 14 to 25. The study began in 2004, when over 15,000 young people from all areas of England born between 1st September 1989 and 31st August 1990 were interviewed. These young people are tracked and re-interviewed every year, and the study is currently in its sixth wave of interviews, with the respondents now aged 19.

LSYPE is managed by the Department for Children, Schools and Families (DCSF), and fieldwork is carried out by a consortium led by the British Market Research Bureau (BMRB). It is a highly detailed and in-depth survey, and the data are publicly available from the UK Data Archive (Waves 1-3 are currently available). Because LSYPE is a longitudinal study, it is possible to link data between waves and explore young people's transitions and changing attitudes as they grow older.

4.2 Purpose of the LSYPE Study

The main objectives of the study are:

- To provide evidence on key factors affecting educational progress and attainment from the age of 14.
- To provide evidence about the transitions young people make from education or training to economic roles in early adulthood.
- To help monitor and evaluate the effects of existing policy and provide a strong evidence base for the development of future policy.

- To contextualise the implementation of new policies in terms of young people's current lives.

4.3 Information Available from the Study

As well as interviews with the sampled young people, LSYPE also includes interviews with parents or guardians (both main and secondary if available) in its first three waves. Only the main parent was interviewed at Wave 4, while at Wave 5 no parents or guardians were interviewed, as the young people are likely to be more independent at this stage. There is also a self-completion section used to record more sensitive information from the young person. The main types of information available from the core LSYPE dataset are listed below, divided into the categories in which the questions are asked:

- *Family background* – including household situation, languages spoken in the home, family activities, household responsibilities and resources, parental qualifications and education, parental occupations and employment history, parental health, household benefits and tax credits and estimates of household income.
- *Parental attitudes* – including attitudes to the young person's school and involvement in education, parental expectations and aspirations for the young person, school history, vocational courses and choice of current school.
- *Young person characteristics* – including demographics, health, Year 10 subject choices and reasons for these, rules and discipline at school, homework, ICT, study support, future plans and advice, household responsibilities, use of leisure time, subjects being studied and expected qualifications and knowledge of and intentions towards apprenticeships and related schemes.
- *Young person self-completion* – including relationships with parents, risk factors such as drinking and smoking and attitudes to school.
- *Household grid* – includes information about every household member (sex, marital status, employment status and ethnic group) and their relationship to other household members including the young person.

Data Linkage

The LSYPE data have been linked to administrative data held on the National Pupil Database (NPD), a pupil-level database which matches pupil and school characteristics to attainment. The data are also linked to school-level and Local Authority-level indicators such as school size, proportion of pupils gaining five or more GCSEs at grades A*-C and ethnic composition, and to geographical

indicators such as the Index of Multiple Deprivation (IMD) and classifications of urban and rural areas.

This data linkage enables researchers to draw links between the data collected at all waves of LSYPE and subsequent educational attainment in the same pupils. It also means that characteristics of particular schools or Local Authorities (e.g. ethnic composition or percentage of pupils receiving free school meals) can be investigated in conjunction with individual pupil characteristics. Linkage to the NPD database has enabled a range of other measures to be recorded, and these are listed below:

- *Individual-level data* – including attainment at Key Stages 2, 3 and 4, free school meal eligibility and special educational needs.
- *School-level data* – including OFSTED reports, numbers of pupils, percentage of pupils eligible for free school meals, percentage of pupils with special educational needs, ethnic composition, percentage for whom English is not a first language and school-level attainment at Key Stages 2, 3 and 4.
- *Local Authority-level data* – including percentage of pupils with special educational needs, ethnic composition and LA-level attainment at Key Stages 2, 3 and 4.
- *Geographical data* – including indicator of urban or rural residence, number of schools attended since Year 7, Index of Multiple Deprivation and Government Office Region.

4.4 Sampling and Response Rates

The original sample drawn for the first wave of the study was of over 33,000 young people in Year 9 attending maintained schools, independent schools and pupil referral units (PRUs) in England in February 2004 (Ward and D'Souza, 2008). The final issued sample was approximately 21,000 young people, all of whom were born between 1st September 1989 and 31st August 1990. The young people sampled for the study were aged 13-14 when the study began, and are now aged 17-18 as the study enters its fifth wave. Cleaned data are currently available for Waves 1-3.

The sample was taken from a school census database supplied by DCSF, and 892 schools were selected in total. Of these, 647 schools (73%) co-operated with the study. School-level non-response was a specific problem with LSYPE, especially in inner London, where only 56% of schools responded, and in the independent sector, where only 57% co-operated with the study. The final issued sample was therefore much smaller than the initial sample drawn from the census

database. Further information on the sample design and weighting can be found in Appendix J.

4.5 Conclusions

The LSYPE dataset is ideally suited for the purposes of this analysis, as it contains individual-level information on whether pupils are living in rural or urban areas, and through the NPD linkage also contains information on educational attainment at the ages of 14 and 16 (Key Stages 3 and 4). In addition, there are data on a number of other variables relating to the young person, their family, their school and their local area, which may help to clarify any relationships found between settlement type and attainment. The longitudinal nature of the study enables us to explore any differences that may occur in the relationship between settlement type and attainment at age 16 compared to age 14, and also enables us to take account of earlier attainment to create a value-added analysis. Finally, the sampling methods used in LSYPE ensure that minority ethnic groups and deprived areas are adequately represented, which may not always have been the case in previous analyses.

5 Analysis Strategy

Summary of This Section

The analyses of LSYPE data begin by describing differences between rural and urban areas both before and after adjustment for other important factors. We then investigate whether particular groups of young people might have higher or lower attainment depending on whether they live in rural areas. We next explore variation in attainment at the school and Local Authority Level, and finally look at variation between regions of England.

5.1 Variables Included in Analyses

The list of variables to be included in the analyses contains variables taken from Wave 1 and Wave 3 of LSYPE, corresponding to attainment at Key Stages 3 and 4 respectively. It also contains variables taken from the NPD at the pupil, school and Local Authority levels. The lists include variables that were selected for inclusion in the analytical models but which were not subsequently found to be significant predictors of attainment and which were therefore not included in the final models.

Geographical Variables

These include the main predictors to be used in the models (indicators of urban/rural residence), as well as a number of covariates including factors such as the distance from home to school and IMD. The Government Office Region variable is used to split the analyses by region in order to determine whether relationships with educational attainment differ among rural areas in different parts of England.

Table 4.1 Geographical variables from NPD

Variable	Source	Waves
Urban/rural indicator (England and Wales – also two other derived settlement type variables)	NPD	1, 3
Index of Multiple Deprivation	NPD	1, 3
Distance from home to school (derived using home and school postcodes)	NatGen	1, 3
Government Office Region	NPD	1, 3

Individual Pupil Characteristics

These are included in the models as individual-level covariates to determine the extent to which the covariates explain the variation in attainment between rural and urban areas and within different types of rural area.

Table 4.2 Individual pupil characteristics from LSYPE data

Variable Label	Source	Waves
Whether English is first or main language	LSYPE	1, 3
Young person's ethnic group (grouped)	LSYPE	1, 3
Gender	LSYPE	1, 3
Highest qualification held by mother (grouped)	LSYPE	1, 3
Parental NS-SEC class (derived from father or mother if father not present)	NatCen	1, 3
Parental attitudes to education (derived additive score)	NatCen	1, 3
Parental involvement with school (derived additive score)	NatCen	1, 3
Whether pupil is identified as having SEN (any type)	LSYPE	1, 3
Young person's attitude to school (derived additive score)	NatCen	1, 3
Family cohesion according to main parent (derived additive score)	NatCen	1, 3
Family cohesion according to young person (derived additive score)	NatCen	1, 3
Anti-social behaviour (derived additive score)	NatCen	1, 3
Whether young person is a carer	LSYPE	1, 3
Whether young person has a disability/long term illness or health problem	LSYPE	1, 3
Whether household is single parent household	LSYPE	1, 3

School-level Characteristics

These are included in the models as school-level covariates, and interactions with rural-urban residence are then tested for. Attainment points scores from the NPD have been converted to normal scores so that they are normally distributed.

Table 4.3 School-level characteristics from LSYPE/NPD data

Variable Label	Source	Waves
Pupil teacher ratio (for relevant year)	NPD	1, 3
Percentage of pupils known to be eligible for free school meals (for relevant year)	NPD	1, 3
Percentage of pupils with SEN (for relevant year)	NPD	1, 3
Percentage of white British pupils (for relevant year)	NPD	1, 3
Percentage of pupils whose first language is other than English (for relevant year)	NPD	1, 3
Attainment average score per pupil (continuous variable for Key Stage 3 or 4)	NatCen	1, 3

LA-level Characteristics

These are included in the models as LA-level covariates, and interactions with rural-urban residence are again tested for. Attainment points scores from the NPD have been converted to normal scores so that they are normally distributed.

Table 4.4 LA-level characteristics from LSYPE/NPD data

Variable Label	Source	Waves
Percentage of pupils with SEN	NPD	1, 3
Percentage of white British pupils (for relevant year)	NPD	1, 3
Percentage of pupils whose first language is other than English (for relevant year)	NPD	1, 3
Attainment average score per pupil (continuous variable for Key Stage 3 or 4)	NPD	1, 3

Educational Outcomes

Educational outcomes at Key Stages 3 and 4 are used in the models, corresponding to Waves 1 and 3 of the LSYPE data.

Table 4.5 Educational outcomes from NPD data

Variable Name	Source	Waves
Achieved Level 5 or above in English, Maths and Science at KS2	NPD	1, 3
Attainment score for Key Stage 2	NPD	1, 3
Achieved level 5 or above in English, Maths and Science at KS3	NPD	1, 3
Attainment score for Key Stage 3	NPD	1, 3
Achieved Level 2 at GCSE	NPD	1, 3
Attainment score for Key Stage 4	NPD	1, 3

5.2 Stages of Analysis

Stage 1: Describing Differences Between Rural and Urban Areas

First, we produced graphs summarising how the LSYPE respondents vary according to the settlement types they live in. We also examined the distribution of attainment scores at Key Stages 3 and 4 and how these vary by settlement type. Then we constructed regression models in order to replicate the results of the graphs showing differences in attainment between rural and urban areas, but controlling for other factors found in the literature to also influence attainment. These models are an accepted means of analysing educational data, as they can estimate the effect of a particular factor of interest on attainment while holding constant the effects of other factors that might distort the relationship, such as family background or gender.

For the initial models, we decided to include settlement type as a binary variable only distinguishing between rural and urban residence, as this would indicate whether there appeared to be any basic differences in attainment between rural and urban areas. In later analyses, the 8-category classification developed by Defra was used in order to investigate these relationships in more detail. These models were all employed in a multilevel framework in order to take account of the fact that pupils attending the same school and schools located in the same Local Authority might be expected to have more similar attainment levels.

Covariates were added to the models in sequence: first individual-level variables such as family background, gender and individual characteristics, then school-level variables and finally LA-level variables. This strategy was designed to enable us to distinguish the most important predictors of educational attainment, and which of these explain the most variation by type of settlement. We theorised that it was likely that variables such as social position and Index of Multiple Deprivation might obscure the true relationship between settlement type and attainment. We therefore added variables sequentially in order to assess the effects of each factor in turn on the primary relationship between settlement type and attainment.

Stage 2: Exploring Attainment in Particular Groups

Having explored the differences in attainment between urban and rural areas and factors that were behind these differences, we then added interactions to the models in order to investigate whether particular groups of young people might have higher or lower attainment if they were resident in certain settlement types (e.g. pupils from particular ethnic groups in rural compared to urban areas). We also investigated whether attending a rural school might have a similar effect on attainment to living in a rural area.

Stage 3: School and Local Authority Level Variation

We used our multilevel models to explore the amount of unexplained variation in attainment that was due to differences between schools and differences between Local Authorities. We also allowed the relationship between settlement type and attainment to vary between schools in order to determine whether pupils in different schools might have higher or lower attainment depending on whether they live in rural or urban areas. In these analyses we took into account all the secondary schools the pupils had attended from Year 7 to Year 11 by weighting the models according to the amount of time spent in each school.

Stage 4: Regional Variation

The analyses from Stage 1 were split according to Government Office Region in order to explore regional differences between rural and urban pupils. Regions of particular interest were followed up in more detail in order to determine how their patterns of attainment differed from that of England as a whole.

6 Results

6.1 Do Pupils Living in Rural Areas Really Have Higher Attainment?

Summary of This Section

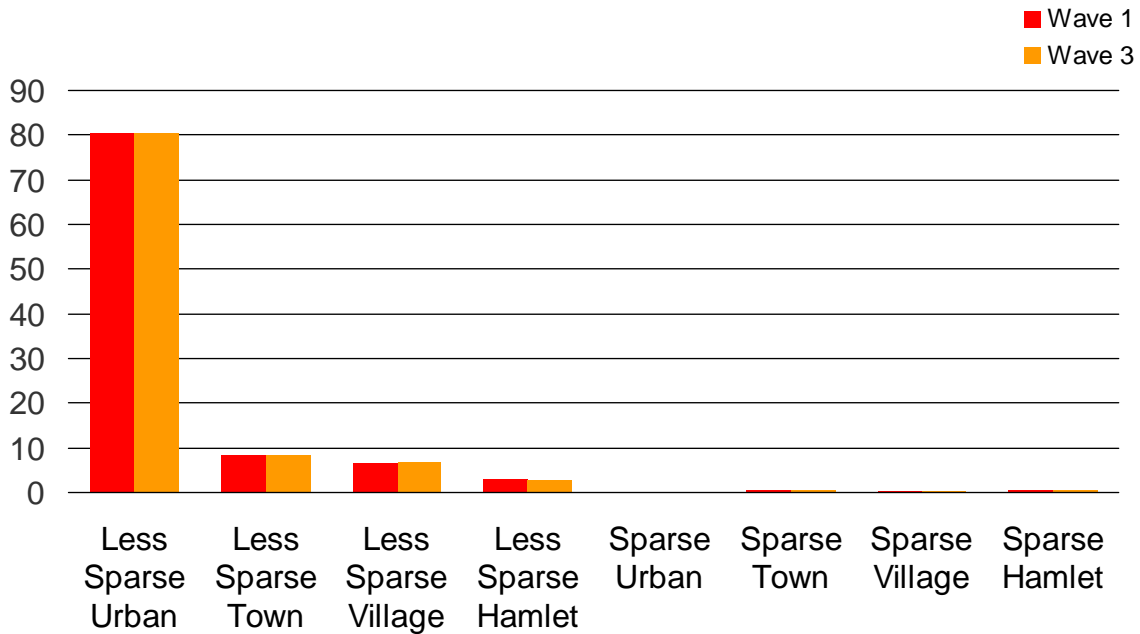
The data from the LSYPE show that pupils living in rural areas have higher attainment than pupils living in urban areas at Key Stages 3 and 4, particularly those living in the smaller settlement types such as villages, hamlets and scattered dwellings. However, the differences in attainment are small, and only significant at Key Stage 3.

Where do the LSYPE Respondents Live?

We first looked at the distribution of rural and urban pupils at Key Stages 3 and 4 (corresponding to Waves 1 and 3 of the study). As can be seen from Figure 5.1, the vast majority of respondents were resident in less sparse urban areas at both waves, with only a tiny proportion of young people living in the more sparse areas (some of these do not show up as percentages at all on the graph). These percentages are very similar to those found in Census data and therefore indicate that LSYPE is representative of the population in this regard. Interestingly, the pattern of movement between Wave 1 and Wave 3 does not appear to reflect recent trends in population movement whereby families with teenagers appear to migrate away from rural areas (Commission for Rural Communities, 2008), since the proportion of respondents resident in urban areas is almost identical at Wave 1 and Wave 3. This may reflect the relatively short period of time between the two waves of the study.

When we examined movement between different area types from Wave 1 to Wave 3, only 50 respondents (0.4% of the total at Wave 3) had moved from a rural to an urban area between the two waves, and only 31 respondents (0.3%) had moved from an urban to a rural area. A further 40 respondents (0.3%) had moved between different types of rural area, and the remaining 99% of respondents had stayed in the same type of area. These figures show that population movement between different settlement types was indeed very small over the two years between waves.

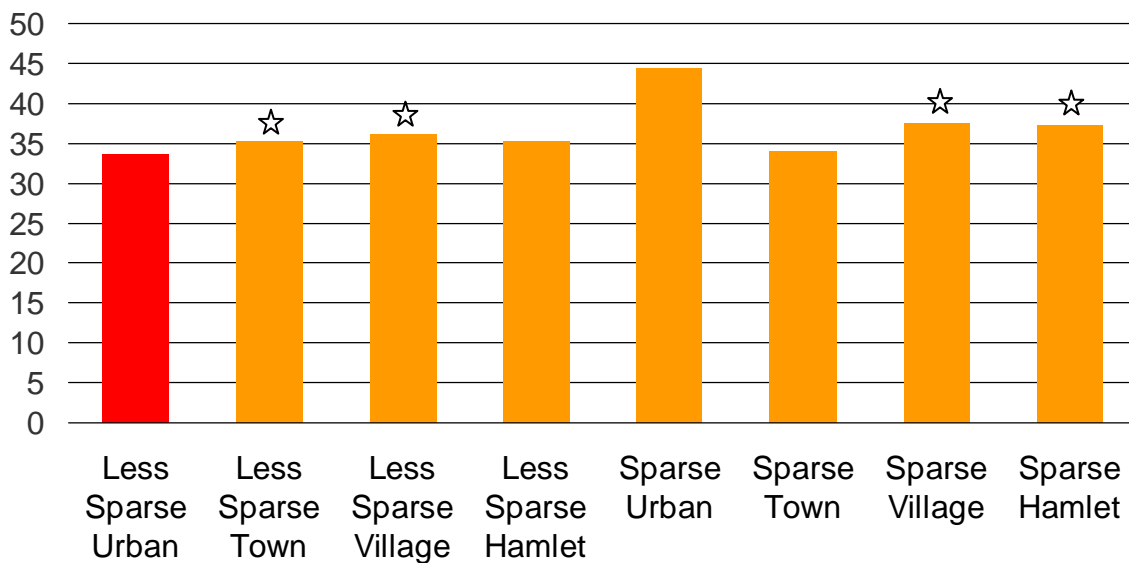
Figure 5.1 Settlement Type at Waves 1 and 3 (8 Categories)



How Does Attainment Differ by Settlement Type?

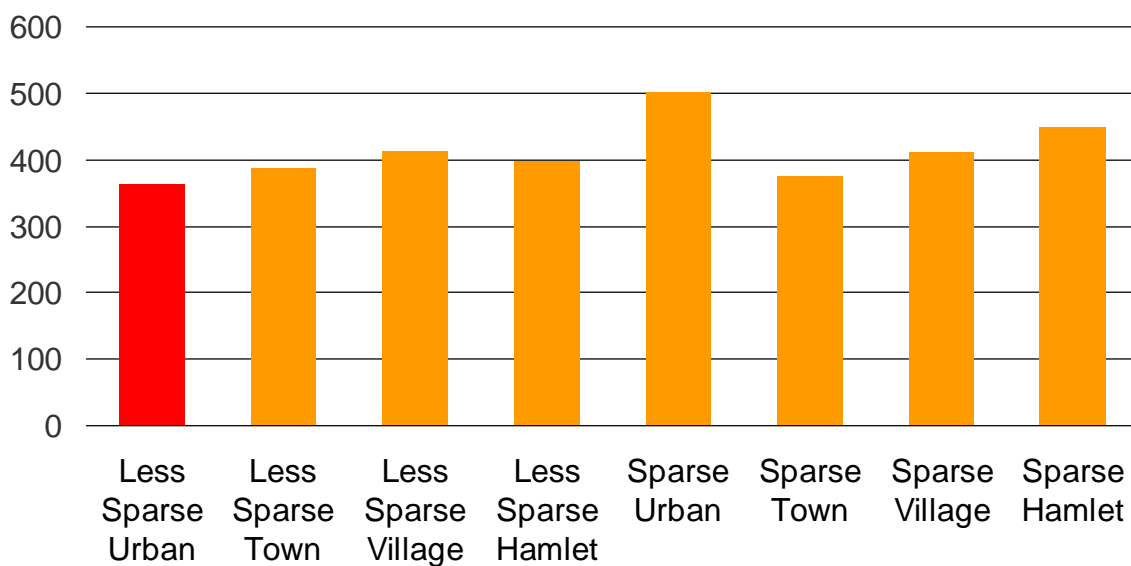
Next, we examined raw attainment scores at Waves 1 and 3, comparing the results across the 8 settlement types identified above (see Figure 5.2). First, examining the mean points score achieved at Key Stage 3, we found that in general, average attainment was higher in rural areas, as stated in the 2008 State of the Countryside report (Commission for Rural Communities, 2008). However, Figure 5.2 also shows that average attainment was considerably higher in the sparse urban areas. This is likely to be a chance result generated by the fact that there were very few respondents living in sparse urban areas, and is not significant. Significant differences in points scores when compared to less sparse urban areas are denoted by the stars on the graph.

Figure 5.2 Mean Points Score at Key Stage 3 (Wave 1) by Settlement Type (8 Categories)



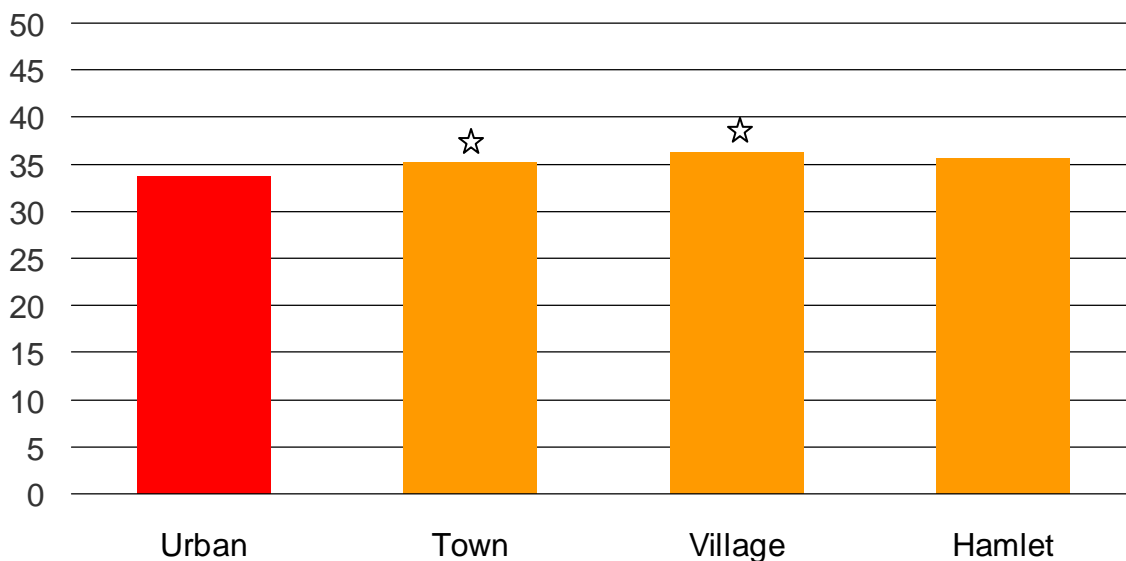
A similar pattern of results was found in the Wave 3 data looking at average Key Stage 4 points scores (see Figure 5.3). Again, the highest attainment occurred in sparse urban areas, with the lowest attainment occurring in the less sparse urban areas. Generally, rural areas were again found to have slightly higher average attainment than urban areas, but as with the Key Stage 3 data, some settlement types had very small numbers of respondents. It should also be noted here that the differences in attainment at Key Stage 4 may appear larger than those found at Key Stage 3, but this is partly due to the different scales of measurement, with Key Stage 4 scores having much wider variability in general. It can also be seen from this graph that none of the settlement types had significantly different attainment from the less sparse urban areas at Key Stage 4.

Figure 5.3 Mean Points Score at GCSE (Wave 3) by Settlement Type (8 Categories)



When the sparse and less sparse areas were combined, some of the variation in attainment disappeared and the differences became smaller (see Figure 5.4). Urban areas (with sparse and less sparse urban areas being combined in this graph) were seen to have lower attainment on average than all three types of rural area, with respondents from villages having the highest average attainment at Key Stage 3, and pupils living in both villages and towns having significantly higher attainment than those in urban areas. There was approximately a 2-point difference found in average attainment between the urban and rural areas, as opposed to the 11-point difference found between the lowest and highest scoring areas in Figure 5.2.

Figure 5.4 Mean Points Score at Key Stage 3 (Wave 1) by Settlement Type (4 Categories)



A similar picture was observed at Key Stage 4, with respondents living in villages again having the highest average attainment at GCSE level, and respondents living in urban areas having the lowest attainment (see Figure 5.5). Here there was approximately a 40-point difference found in average attainment between the urban and rural areas, but none of these differences was significant.

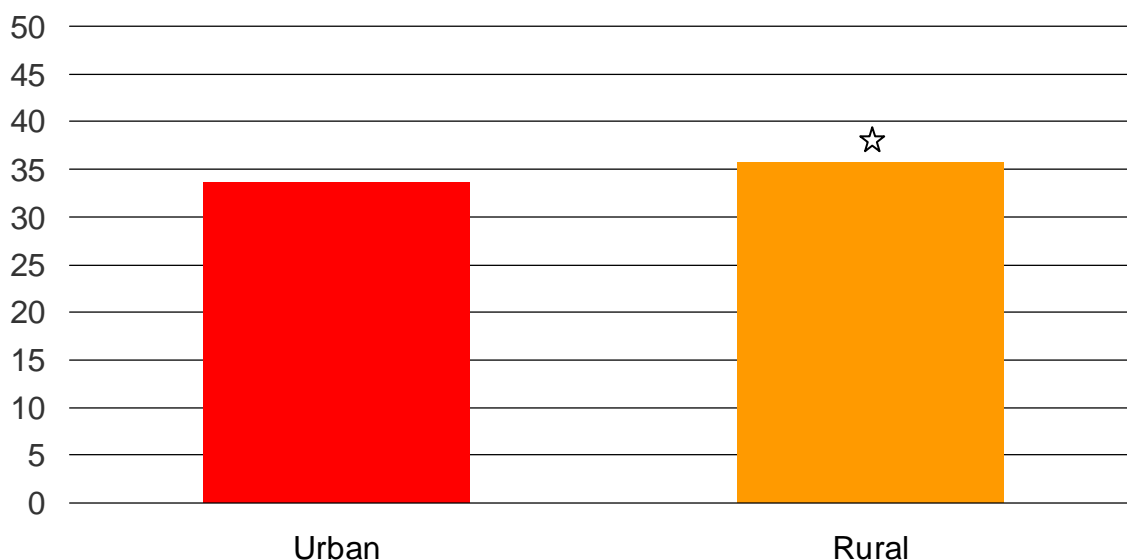
Figure 5.5 Mean Points Score at GCSE (Wave 3) by Settlement Type (4 Categories)



Finally, we compared attainment at each Key Stage to settlement type as a binary variable, to determine whether overall attainment tended to be higher in urban (sparse and less sparse again being combined) or rural areas. The Key Stage 3

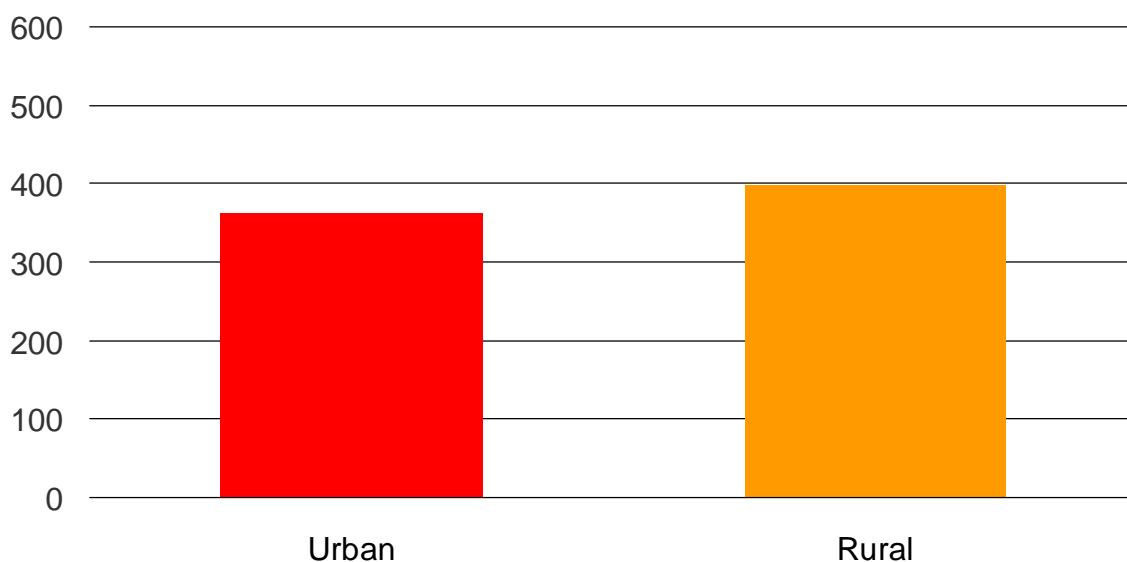
results clearly confirm that average attainment is significantly higher (by about 2 points) in rural areas than in urban areas, as indicated by the results in Figure 5.6.

Figure 5.6 Mean Points Score at Key Stage 3 (Wave 1) by Settlement Type (2 Categories)



Again, a similar picture was observed at GCSE level, with average achievement at Wave 3 being approximately 35 points higher in rural areas than in urban areas (see Figure 5.7). However, like the other analyses of Key Stage 4, this relationship was not significant.

Figure 5.7 Mean Points Score at GCSE (Wave 3) by Settlement Type (2 Categories)



6.2 What Are the Factors Behind These Differences?

Summary of This Section

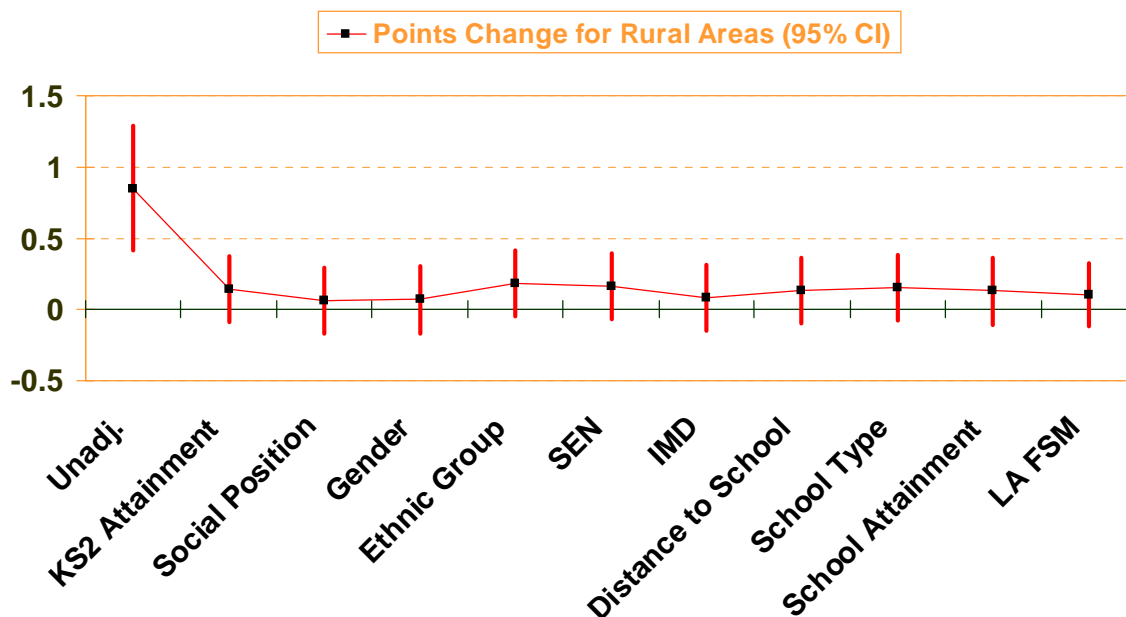
The differences in attainment between rural and urban areas are much reduced after adjustment for social position, area deprivation and previous attainment, and are no longer statistically significant even at Key Stage 3. This indicates that the observed unadjusted differences are likely to be due to social factors rather than any intrinsic quality of rural areas.

Binary Definition of Rural and Urban Areas

Having established that there are unadjusted differences in attainment between rural and urban areas at both Key Stage 3 and Key Stage 4, we then constructed multilevel models in order to look at factors that might partially or completely explain these differences. For each model, we began by including only settlement type and attainment, and then progressively added other factors in order to assess how much of the relationship between settlement type and attainment these factors explained. We began by including settlement type as a binary variable (distinguishing only between rural and urban areas rather than also including the different sub-categories) in order to maximise the number of individuals within each settlement type.

In all the graphs below, the 0 line represents attainment for pupils in urban areas, and the black points on the graph show the change in attainment associated with living in a rural area. The red error bars either side of the points indicate 95% confidence intervals, so that if these bars overlap the 0 line, the relationship is not significantly different from 0. Each of the factors listed along the bottom of the graph was added into the model in turn, to provide cumulative adjustment. This adjustment for other factors at Key Stage 3 is shown in Figure 5.8 below (the equivalent figures are also shown in Table 6.1 in Appendix B).

Figure 5.8 Points Change in Key Stage 3 Score Associated with Living in a Rural Area, with Cumulative Adjustment for Other Factors



From the above models, it can be seen that there is just under a 1-point increase in Key Stage 3 attainment (or 0.13 SD) associated with living in a rural area, before adjustment for any other factors. This is slightly less than the difference observed in Figure 5.6 above, which is likely to be due to the multilevel models used for this analysis. When other factors are adjusted for, this increase is no longer significant and the difference in points is much smaller (around 0.2 points, or 0.03 SD)², indicating that the other factors in the model are largely explaining the differences found between rural and urban areas.

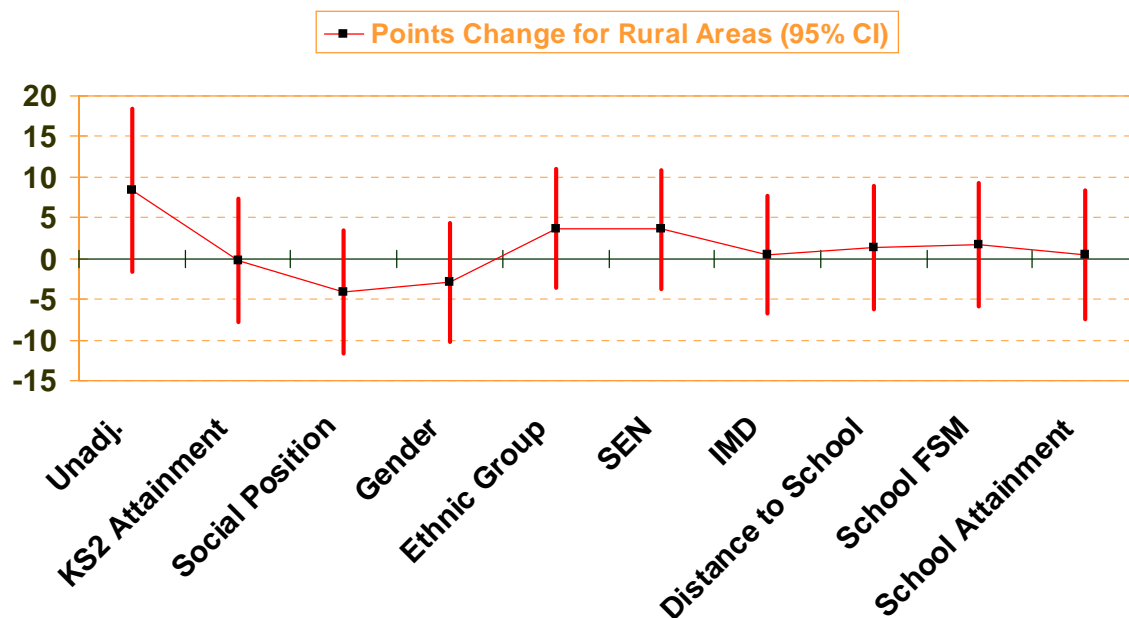
The overall effect of adjustment for all of the above factors is that there seem to be no qualities intrinsic to the character of rural areas that improve attainment at Key Stage 3 once the effect of other factors has been removed. Instead, the difference in attainment seems to be due to a mixture of factors including higher average social position, lack of local deprivation and pre-existing higher attainment among rural pupils at Key Stage 2. We experimented with adding all the other factors into the model in different sequences, but always found that these three main factors explained most of the difference in attainment. For a list of factors tested in the models but removed because they had no effect, see Appendix C.

We also ran the same set of analyses using the Wave 3 (Key Stage 4) data from LSYPE, in order to check that the same variables tended to affect the relationship

² For the LSYPE respondents, mean progress in attainment from Key Stages 2-3 is 7.06 points (SD 3.41 points).

between settlement type and attainment. The results can be found in Figure 5.9 below, and are also shown in Appendix B, Table 6.2.

Figure 5.9 Points Change in Key Stage 4 Score Associated with Living in a Rural Area, with Cumulative Adjustment for Other Factors



As can be seen from Figure 5.9, the results for Key Stage 4 are not especially different from those for Key Stage 3, but there are a small number of factors that are different in predicting attainment at Key Stage 4. School type was not found to be a significant predictor of attainment at Key Stage 4 once social position was accounted for, so this was removed from the model. Instead, the percentage of pupils in the school who were eligible for free school meals was included and was found to slightly increase the size of the relationship between settlement type and attainment. Also, none of the Local Authority-level factors were found to be important in predicting attainment at Key Stage 4, so we therefore restricted the models to the individual and school levels.

The overall unadjusted difference in attainment between pupils living in urban and rural areas at Key Stage 4 is around 8 points (or 0.05 SD). This difference is not significant and once other factors are adjusted for it is reduced virtually to zero (0.46 points)³. At Key Stage 4, previous attainment at Key Stage 2 has a smaller effect on the relationship with settlement type than at Key Stage 3 (as might be expected as there have been an additional 2 years of secondary schooling compared to Key Stage 3), but social position, ethnic group and IMD all have a

³ For the LSYPE respondents, mean progress in attainment from Key Stages 2-4 is 350 points (SD 153 points).

larger relative effect on the relationship, indicating that social and cultural factors have become increasingly important predictors of attainment over the additional 2 years.

Table 5.1 below presents a summary of the other factors included in the models at Key Stage 3 and Key Stage 4, including whether these tend to operate as rural advantages or disadvantages and whether they have a large or significant effect on the relationship with settlement type.

Table 5.1 Factors Influencing the Relationship Between Settlement Type and Attainment

Factor	Rural Advantage or Disadvantage?	Size of Effect	Significant?
Key Stage 2 Attainment	Advantage at Key Stage	Large	Yes
Social Position – Parental Social Class	3 and 4	Medium	Yes
Social Position – Mother’s Highest	Advantage at Key Stage	Medium	Yes
Qualification	3 and 4	Medium	Yes
Social Position – Whether Single Parent	Advantage at Key Stage	None	No
Family	3 and 4	Large	Yes
Gender	Advantage at Key Stage	None	No
Ethnic Group	3 and 4	Large	Yes
Special Educational Needs	No effect	Medium	Yes
Index of Multiple Deprivation	Disadvantage at Key	Small	Yes
Distance Travelled to School	Stage 3 and 4	Medium	Yes
School Type (City Technology Colleges Vs. Others)	No effect	Small	Yes
School Average Attainment	Advantage at Key Stage	Small	No
Proportion of Pupils in School Receiving FSM	3 and 4		
Proportion of Pupils in LA Receiving FSM	Disadvantage at Key Stage 3 and 4		
	Disadvantage at Key Stage 3 only		
	Advantage at Key Stage 3 and 4		
	Disadvantage at Key Stage 4 only		
	No effect		

This section has established that pupils in rural areas do tend to have higher attainment than pupils in urban areas at Key Stage 3 and to a lesser extent at Key Stage 4, but that these differences are explained by other factors. In particular, differences in social position, area deprivation and pre-existing differences between pupils and schools at Key Stage 2 make up the large part of the observed differences.

Because we found that previous attainment at Key Stage 2 had a large effect on the relationship between settlement type and attainment, the analyses that follow

in subsequent sections are analyses of *progress* in attainment from Key Stage 2 rather than analyses of raw attainment scores. Descriptions of the differences between these two measures of attainment and examples of analyses unadjusted for Key Stage 2 attainment can be found in Appendix D.

6.3 Which Settlement Types Have the Highest Attainment?

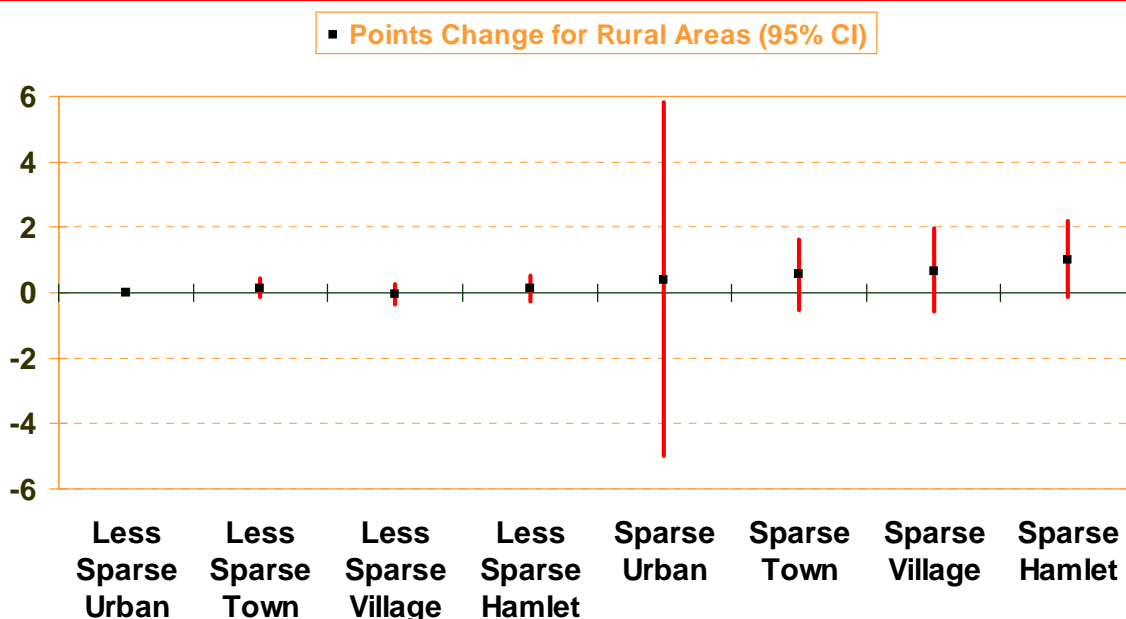
Summary of This Section

Breaking down the settlement types into smaller categories shows that there is some evidence of greater progress in attainment among pupils living in smaller and sparser areas even after adjustment for other factors. However, these differences are small and do not quite reach statistical significance.

2004 Definition of Rural and Urban Areas

Using the 8-category classification of rural and urban areas, we repeated the fully-adjusted analyses described above in Section 5.2. Using this variable, we were able to obtain an estimate of the increase or decrease in attainment progress associated with living in each type of settlement as compared to less-sparse urban settlements (the reference category). The results of this model for Wave 1 (Key Stage 3) can be found below in Figure 5.10, and are also shown in table form in Appendix B, Table 6.3.

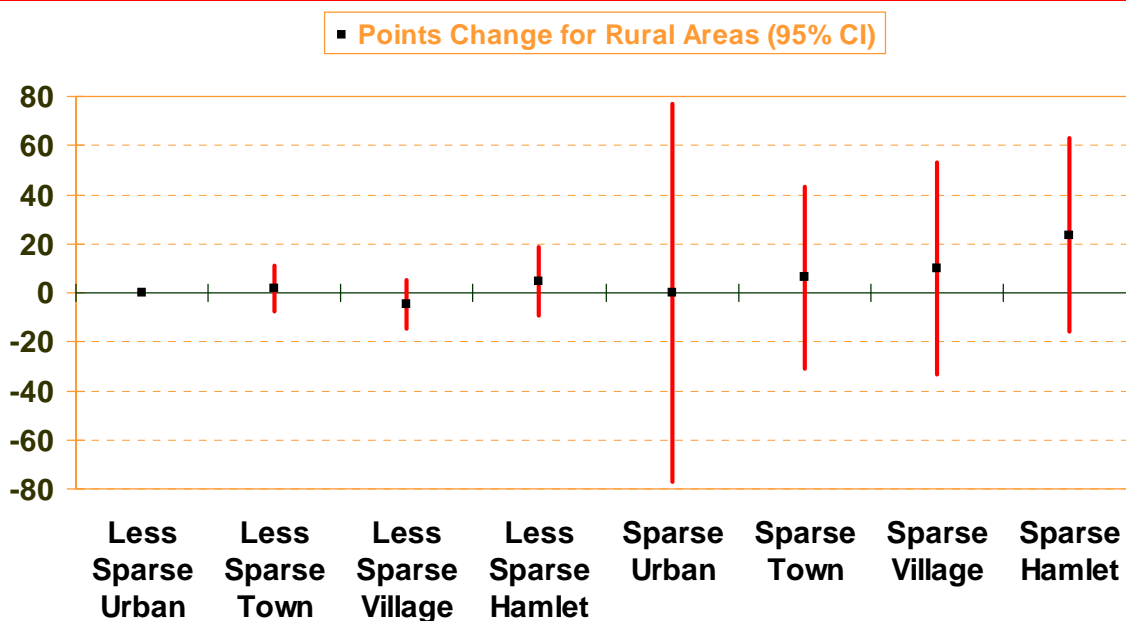
Figure 5.10 Points Change in Key Stage 3 Score Associated with Different Settlement Types (with Less Sparse Urban Areas as Reference Category)



It can be seen from the above results that there is generally little difference in Key Stage 2-3 progress between the different settlement types after we adjust for other factors. However, there does seem to be some evidence of slightly greater progress among pupils living in sparse areas, with the points change for sparse villages and hamlets almost reaching significance. For sparse hamlets, progress is just over 1 point greater than in less sparse urban areas. This corresponds to an increase in progress of 0.33 SD (see Appendix E for an explanation of standard deviations).

These analyses were again repeated for the Wave 3 (Key Stage 4) data, using the same model adjusting for all the factors previously found to affect attainment. The results of the analysis using the 8-category settlement type variable are documented in Figure 5.11 below, and the table of figures can be found in Appendix B, Table 6.4.

Figure 5.11 Points Change in Key Stage 4 Score Associated with Different Settlement Types (with Less Sparse Urban Areas as Reference Category)



It can be seen from this figure that again there are no settlement types with significantly greater or lesser progress from Key Stage 2-4 than less sparse urban areas. There is some evidence that pupils living in hamlets have higher attainment (by about 25 points or 0.15 SD for sparse hamlets, although this difference is smaller for less sparse hamlets), but for all the other settlement types there was very little evidence of any difference, as at Key Stage 3.

There therefore seems to be some evidence indicating that pupils living in smaller and more remote areas may have slightly higher attainment even after adjusting for their social position and prior attainment, but these differences do not reach statistical significance and appear larger (in SD terms) at Key Stage 3 than at Key Stage 4. The lack of significance may be at least partly due to the small numbers of pupils in some of the settlement categories (as described in Appendix G), particularly in the sparse areas which often appear to show the biggest differences in attainment when compared to less sparse urban areas.

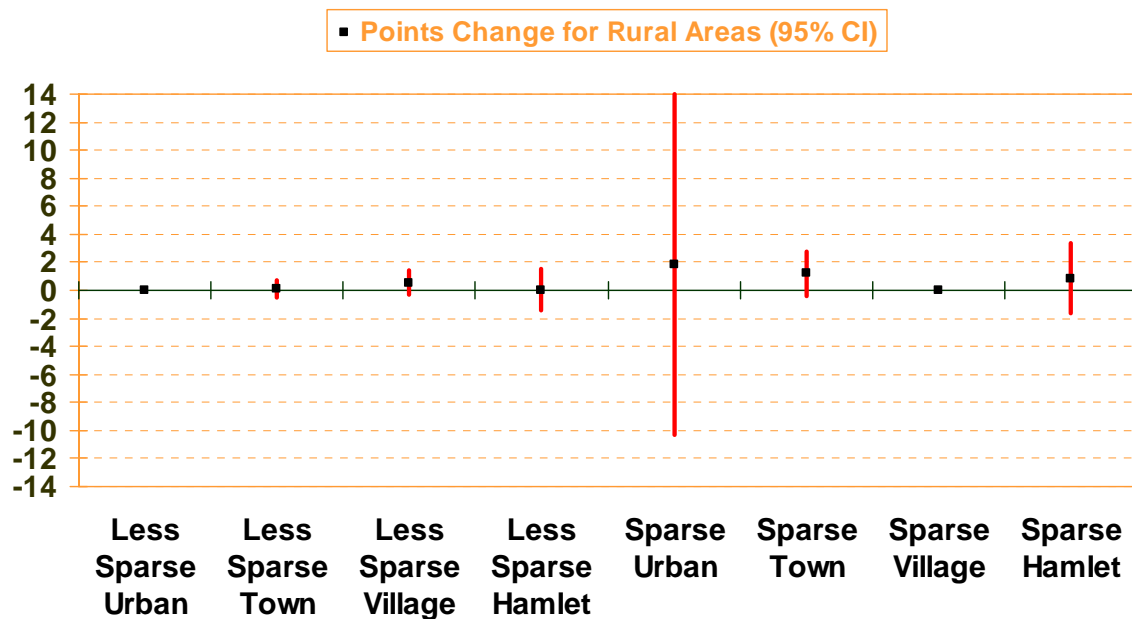
6.4 Rural Pupils or Rural Schools?

Summary of This Section

There is some evidence that attending rural schools leads to greater progress in attainment at Key Stages 3 and 4, particularly for pupils who live in urban areas at Key Stage 4. These results are based on small numbers of rural schools, but seem to be consistent across different area types and time periods.

In addition to information on the settlement type in which each individual pupil in LSYPE lives, we were also able to calculate whether the schools they attend are in rural or urban areas, and the results of this analysis for Key Stage 3 can be found in Figure 5.12 below (results are also shown as a table in Appendix B, Table 6.5).

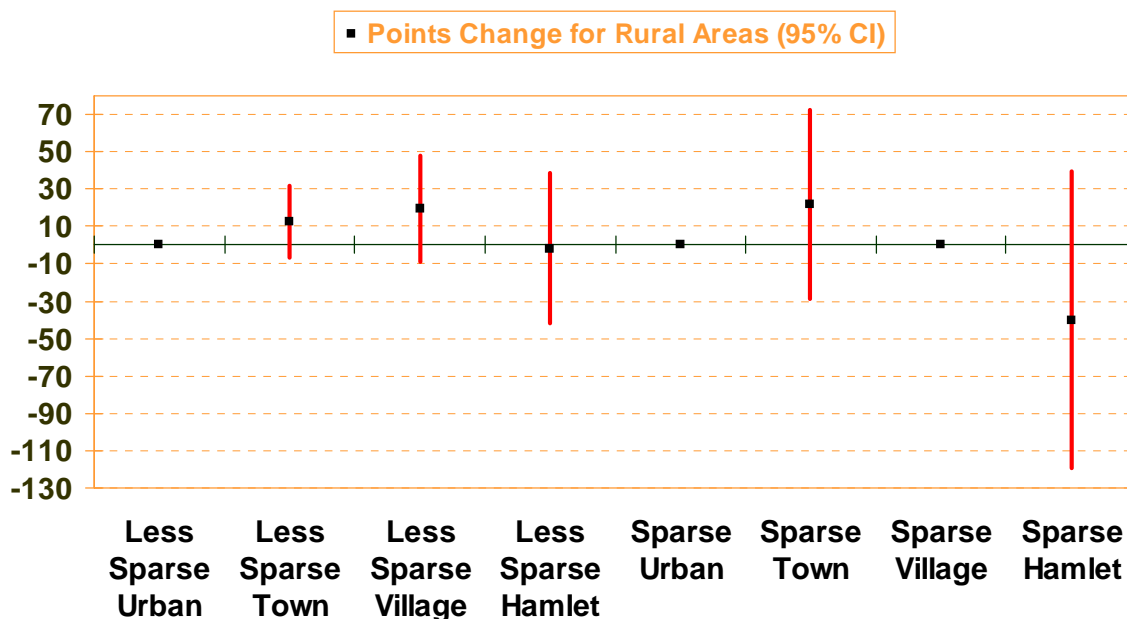
Figure 5.12 Points Change in Key Stage 3 Score Associated with Attending Schools in Different Settlement Types (with Less Sparse Urban Areas as Reference Category)



These results show that the differences between school settlement types tend to be of a similar size but are less close to being significant than those for individual settlement types (see Figure 5.10 above for comparison). There is very little evidence of pupils in rural schools having greater progress in attainment, although for pupils attending schools in sparse towns progress is approximately 1.2 points (or 0.35 SD) greater than for those attending less sparse urban schools. This difference is the closest to being significant, but is still a very small difference between settlement types.

The results for pupils attending schools in different settlement types at Key Stage 4 are shown below in Figure 5.13 below, and are also reproduced as a table in Appendix B, Table 6.6.

Figure 5.13 Points Change in Key Stage 4 Score Associated with Attending Schools in Different Settlement Types (with Less Sparse Urban Areas as Reference Category)



These results show that at Key Stage 4, while pupils attending schools in towns and less sparse villages appeared to have slightly greater progress (as they did at Key Stage 3), pupils attending schools in hamlets actually appeared to make slightly less progress than those attending schools in urban areas. For pupils attending schools in sparse hamlets, this progress was approximately 40 points (or 0.26 SD) less, but none of the differences reached significance due to small sample sizes and the relatively small differences observed.

We also explored whether there was any interaction between the areas pupils live in and the areas where they go to school. At Key Stage 3, we found that pupils living in urban areas tended to make significantly greater progress in attainment if they attended a rural compared to an urban school, but that there was little difference for pupils living in rural areas. At Key Stage 4 we found a similar relationship, and there was a statistically significant positive relationship between attending a rural school and progress among pupils living in urban areas (these pupils on average made 30 points or 0.2 SD more progress at Key Stages 2-4 compared with those who attended urban schools). Again, there was no relationship at all between school settlement type and progress among those living in rural areas (results not shown). These results seem to indicate that attending rural schools provides an advantage for pupils who live in urban areas, but it is

difficult to confirm these conclusions given the relatively small numbers of rural schools available.

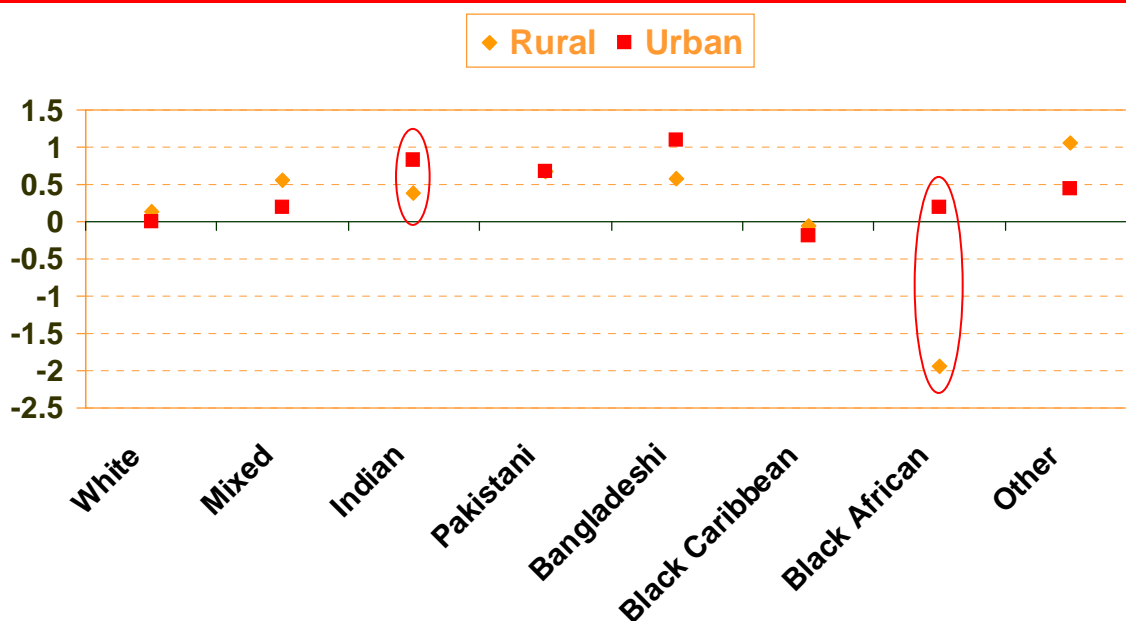
6.5 Do Particular Groups of Pupils in Rural Areas Have Lower Attainment?

Summary of This Section

Pupils from certain minority ethnic groups (although sample sizes are small for these groups) and those whose mothers have lower qualifications tend to make less progress in attainment if they live in rural areas compared to those in urban areas. This indicates that pupils with particular characteristics may be worse off living in rural areas even though rural pupils tend to do slightly better overall.

We next explored the possible effect of interactions between settlement type and various different factors, to determine whether particular groups of pupils might be particularly advantaged or disadvantaged if they live in rural areas. At Key Stage 3, the only significant interaction found when we looked at where the respondents lived was with ethnic group. The results of this analysis are summarised in Figure 5.14, and show that different ethnic groups have patterns of progress in attainment depending on whether they live in rural or urban areas.

Figure 5.14 Points Change in Key Stage 3 Score Associated with Different Ethnic Groups in Rural and Urban Areas (White Pupils in Urban Areas are Reference Category)



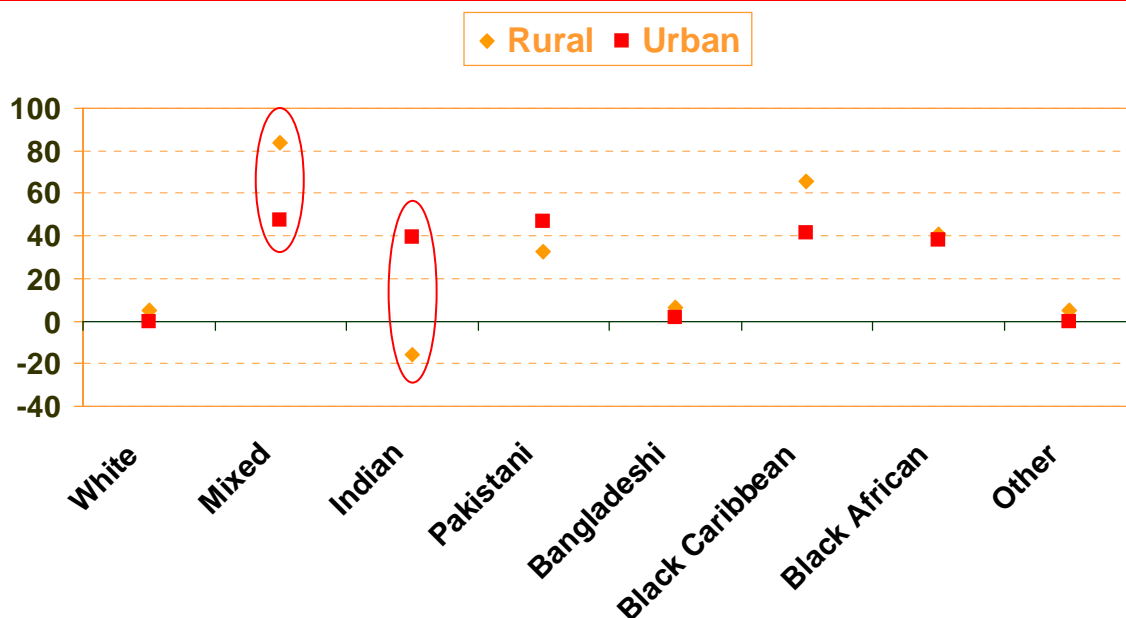
The figure above shows that in urban areas, almost all minority ethnic pupils make more progress than White pupils, with the exception of Black Caribbean pupils whose progress is slightly less on average. Previous studies have also shown similar patterns in different ethnic groups. For most ethnic groups, progress is similar in rural and urban areas, with average differences of 0.5-1 points (or 0.15-0.3 SD) between rural and urban pupils in the same ethnic group. However, it can be seen from Figure 5.14 that both Indian and Black African pupils tend to make significantly less progress if they live in rural areas compared to urban areas (circles show significant differences). A similar pattern was also observed for Bangladeshi pupils, although this difference was not significant. Sample sizes are very small for ethnic minority pupils living in rural areas (see Appendix G), and more research is therefore needed to determine whether these differences are genuine.

Further analyses of the characteristics of different pupils show that ethnic minority groups (including Black African and Indian pupils) tend to have slightly higher social position in rural areas than those in the same ethnic groups living in urban areas. It therefore seems unlikely that socioeconomic factors would be behind the differences we found between urban and rural pupils in the Indian and Black African groups. Despite this, we found that among Indian pupils, social position is a less important determinant of progress than it is for White pupils, and that the proportion of non-White pupils in the school is much more important for progress among Indian pupils than for White pupils (progress is greater among Indian pupils

attending more ethnically-mixed schools). As pupils living in rural areas are more likely to attend ethnically homogeneous schools, this may partly explain the lesser progress among Indian pupils living in rural areas. However, we also found that these differences do not apply to Black African pupils, and it is therefore possible that the lesser progress among Black African pupils in rural areas might be a chance finding due to small sample size.

We also found a significant interaction between ethnic group and settlement type at Key Stage 4, and the results are shown below in Figure 5.15, indicating the differences in points scores associated with the different ethnic groups in both rural and urban areas.

Figure 5.15 Points Change in Key Stage 4 Score Associated with Different Ethnic Groups in Rural and Urban Areas (White Pupils in Urban Areas are Reference Category)

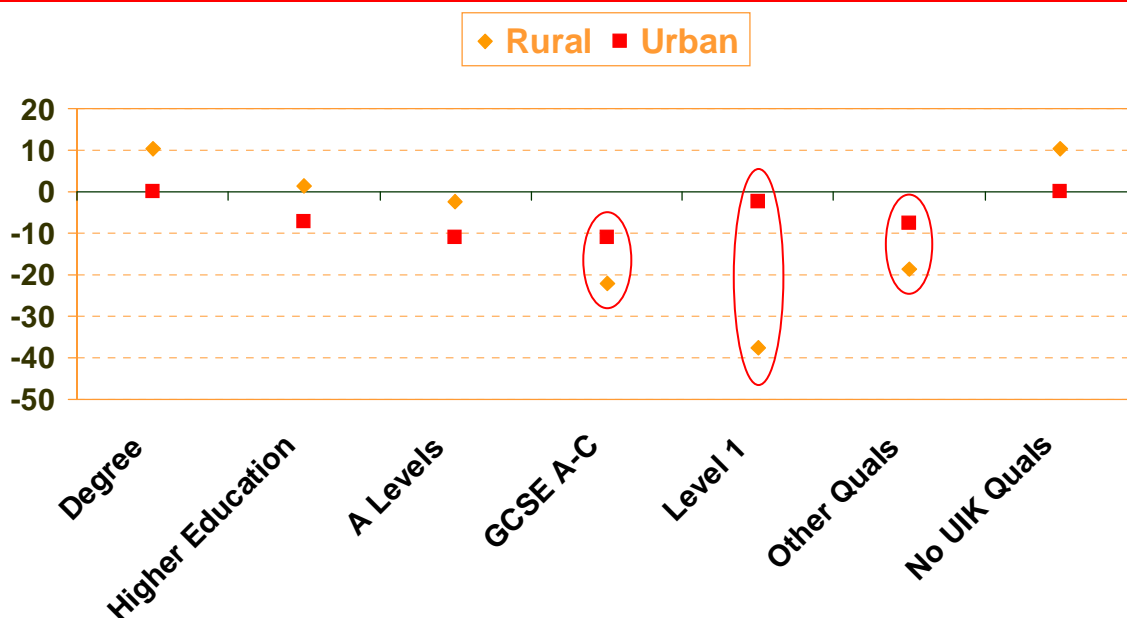


These results show that in urban areas, White pupils again tend to make the least progress, and that for most ethnic groups progress is again similar between rural and urban areas. As with the results for Key Stage 3, Indian pupils tend to make significantly less progress in rural than in urban areas (by approximately 55 points, or 0.36 SD). However, Black African pupils no longer make less progress if they live in rural areas, and indeed the progress of all Black African pupils is greater in relation to that of White pupils when compared to the results for Key Stage 3 (see Figure 5.14). As with Key Stage 3, the White, Mixed, Black Caribbean and Other pupils all make greater progress if they live in rural areas, and for the Mixed pupils this is a significant difference at Key Stage 4.

Again, sample sizes are very small (see Appendix G) and so these relationships cannot be confirmed even though they are significant. As with the results at Key Stage 3, further descriptive analyses show that among Indian pupils, social position is a less important predictor of progress, so despite the higher average social position of minority ethnic pupils in rural areas, rural Indian pupils still make less progress. At Key Stage 4 however, unlike Key Stage 3, Indian pupils do not tend to be disadvantaged by schools with a more homogeneous ethnic makeup, and so the higher proportions of White pupils in rural areas are unlikely to contribute to the lesser progress of Indian pupils at this Key Stage. Further investigation of these results is required in order to confirm whether they are genuine and if so what mechanisms are behind the relationships.

In addition to the interaction between settlement type and ethnic group, at Key Stage 4 we also found a significant interaction between settlement type and mother's highest qualification (no significant interaction was found at Key Stage 3, although there was a tendency towards a similar relationship). The results of this analysis are shown below in Figure 5.16.

Figure 5.16 Points Change in Key Stage 4 Score Associated with Different Levels of Mother's Highest Qualification in Rural and Urban Areas (Pupils Whose Mothers Have Degrees in Urban Areas are Reference Category)



This Figure shows that pupils with highly qualified mothers tend to make greater progress at Key Stages 2-3 if they live in rural compared to urban areas (although these differences are not significant). However, pupils with lower qualified mothers (i.e. 5 GCSEs at grades A-C and below) make significantly worse progress if they live in rural areas. The exceptions to this trend are pupils whose mothers have no

qualifications at all, who do slightly better in rural areas, but this difference is not significant. For pupils whose mothers have Level 1 qualifications (5 GCSEs at grades D-G), there is a 35-point difference (or 0.23 SD) between pupils living in rural and urban areas. Sample sizes for rural pupils whose mothers have all levels of qualifications are reasonable, and we therefore conclude that this relationship is robust.

Further investigations of the relationship show that, as expected, pupils in LSYPE whose mothers have lower qualifications are more likely to have SEN, more likely to be carers and more likely to come from single parent families. They are also more likely to have a disability. Pupils whose mothers have no qualifications at all appear to have slightly different characteristics – they are much more likely to have English as a second language and to belong to a minority ethnic group. Such pupils tend to have higher attainment than those whose mothers have low qualifications, and this may be because they tend to come from more affluent families who have recently arrived in England (the ‘no qualifications’ category also includes mothers with qualifications not part of the recognised list for LSYPE, so a number of these mothers might have unrecognised overseas qualifications). When we examined only those pupils living in rural areas, we found that the relationship between having a low qualified mother and both SEN and being a carer appeared to be stronger in rural areas than in LSYPE as a whole. Since such pupils are less likely to gain the support they need if they live in rural compared to urban areas, this may partly explain the relationship found between settlement type and mother’s highest qualification. It may be the case that disadvantaged pupils are further disadvantaged if they live in rural areas because of the lower levels of support available.

It seems that, while in general pupils appear to be slightly better off living in rural areas in terms of progress in attainment, there are particular groups that may be disadvantaged by living in rural areas and who may need additional support in order to increase their secondary school progress in attainment. In particular, pupils whose mothers have low educational qualifications may need additional support if they live in rural areas where supporting services may be less widespread.

6.6 How Important is the School or Local Authority?

Summary of This Section

Around 15% of the variation in attainment was found to be due to differences between schools, while a further 2% at Key Stage 3 was due to differences between Local Authorities. Different schools not only have very different average attainment, they also have different relationships between settlement type and attainment, with lower-achieving schools having the strongest relationship between living in a rural area and making better progress at school.

Variation in Attainment Between Schools and Local Authorities

Because we used multilevel models in our analyses, we were able to look at the amount of variation in attainment at each Key Stage that is due to differences between individuals, between schools and between Local Authorities, and this helped us to investigate which levels are most important in determining attainment.

At Key Stage 3, we found that a 3-level model explained the variation in attainment the best, with average attainment varying not just between individuals but also between schools and between Local Authorities. 80% of the variation in this model (once we included weighting for movement between schools – see Appendix H) was explained by differences between individuals, showing that individual (and no doubt family) characteristics are most important for determining attainment. Of the remaining variation in attainment, 18% was accounted for by the school and only 2% by the LA, showing that the school is by far the more important of these two levels.

The picture was slightly different at Key Stage 4, where a 2-level model (allowing attainment to vary only between individuals and between schools) seemed to fit the data better than a 3-level one. This is likely to mean that differences in attainment between Local Authorities decrease between Key Stage 3 and Key Stage 4. However, in other respects the amount of variation was similar to that at Key Stage 3, with 79% at the individual level and 21% at the school level.

Different Relationships in Different Schools

Because we found that a relatively large amount of the variation in attainment was due to differences between schools, we then allowed the relationship between settlement type and attainment to vary between different schools, so that some schools might have stronger relationships than others.

At Key Stage 3, this significantly improved the model compared to a model assuming that the relationship was the same in all schools. This model shows that not only is there significant variation in average attainment between schools, but the relationship between settlement type and attainment also varies between schools. We found that the relationship was strongest in those schools with lower average attainment. This means that rural pupils in schools with lower average attainment will benefit more than those in schools with higher average attainment. In the highest-achieving schools, rural and urban pupils will have similar attainment, whereas in the lowest-achieving schools there will be a much larger difference between the urban and rural pupils. We found exactly the same results at Key Stage 4, indicating that these findings are likely to be robust.

6.7 Are There Regional Differences?

Summary of This Section

At Key Stage 3, there is a significant interaction between region of England and settlement type, with respondents living in the North West, Yorkshire and the Humber and the East Midlands tending to make greater progress in attainment if they lived in rural areas. Respondents living in the South East, however, made less progress if they lived in rural areas. At Key Stage 4, there was no significant interaction between region and settlement type, suggesting that these differences do not persist over the added two years.

In order to examine the relationship between settlement type and progress in attainment across different regions of England, we added an additional variable to the analyses to represent the Government Office Region of England pupils were living in. The results for Wave 3 showed that certain regions appear to have significantly different relationships between settlement type and progress when compared to the rest of England, and these are summarised in Table 5.2 below.

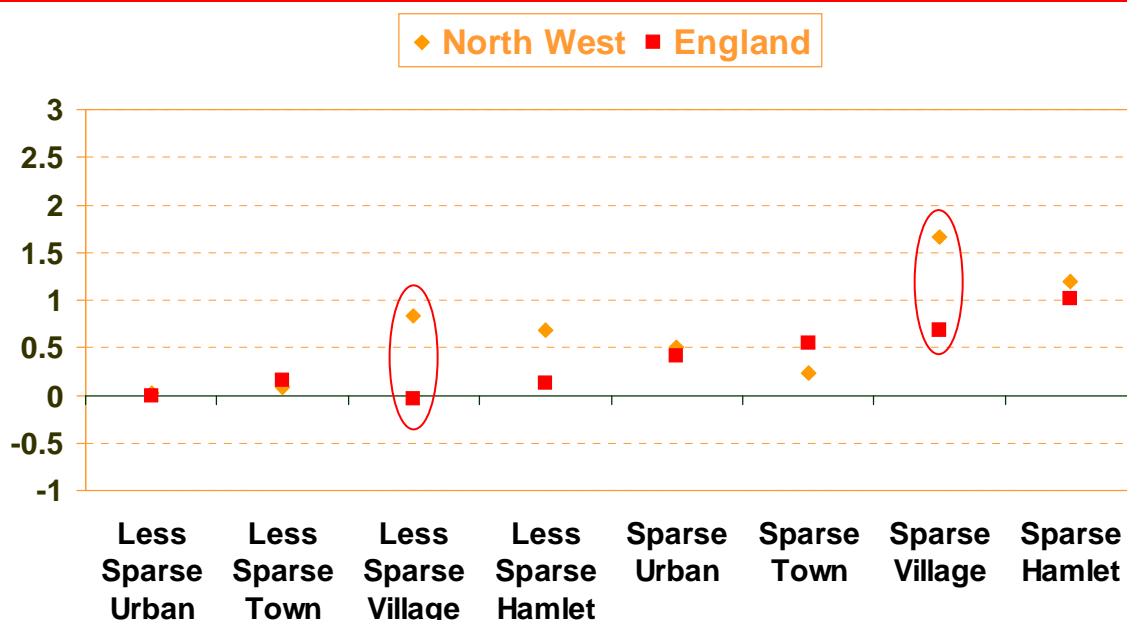
Table 5.2 Relationships Between Settlement Type and Attainment by Government Office Region, Key Stage 3

Government Office Region	Difference in Attainment in Rural Compared to Urban Areas
North East	-0.43 points
North West	+0.53 points (significant difference)
Yorkshire and the Humber	+0.64 points (significant difference)
East Midlands	+0.68 points (significant difference)
West Midlands	+0.13 points
East of England	+0.13 points
London	+0.09 points

South East	-0.55 points (significant difference)
South West	-0.17 points

It can be seen from these results that four regions had significant differences in progress between rural and urban areas. For three of these regions (the North West, Yorkshire and the Humber and the East Midlands), pupils living in rural areas made significantly greater progress, but in the South East the rural pupils made significantly less progress, indicating that different factors may be operating in this region. We explored these relationships in more detail in the regions found to have significant differences, and the results of these analyses can be seen below in Figures 5.17-5.20.

Figure 5.17 Relationships Between Settlement Type and Key Stage 3 Attainment in the North West Compared to the Rest of England



Figures 5.17-5.19 show that in the North West, Yorkshire and the Humber and the East Midlands, rural pupils tend to make more progress in attainment than urban pupils. However, there is only a significant difference from the rest of England in a small minority of rural areas in these regions (see circled relationships), and sample sizes are again small, particularly in the sparse areas. It is therefore difficult to conclude that the relationship between settlement type and progress is much stronger than in England as a whole.

Figure 5.18 Relationships Between Settlement Type and Key Stage 3 Attainment in Yorkshire and the Humber Compared to the Rest of England

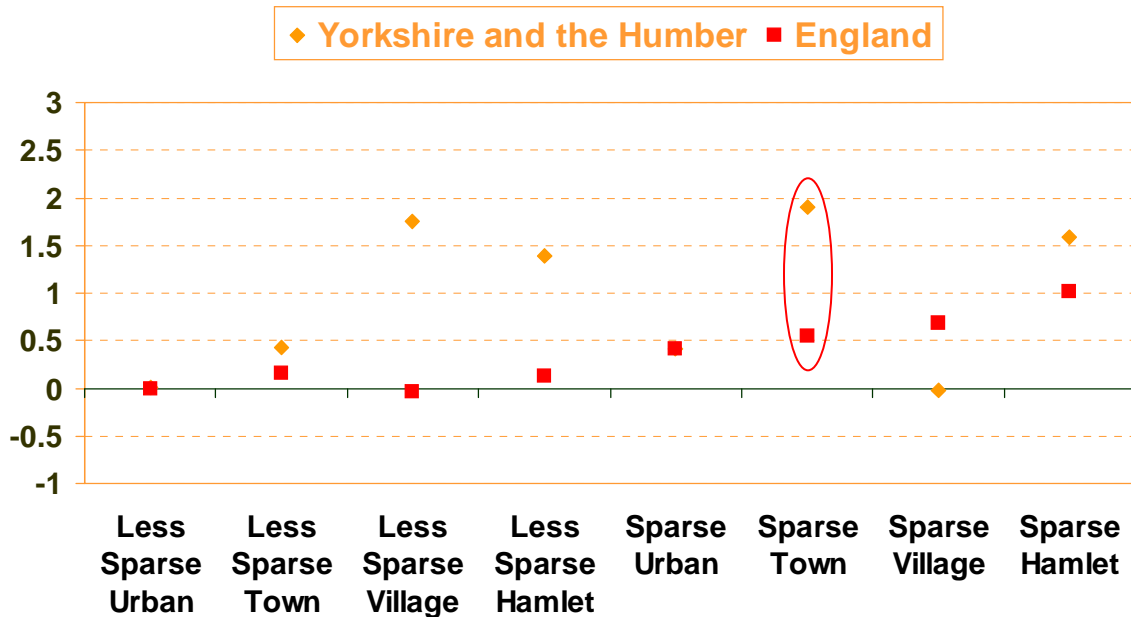
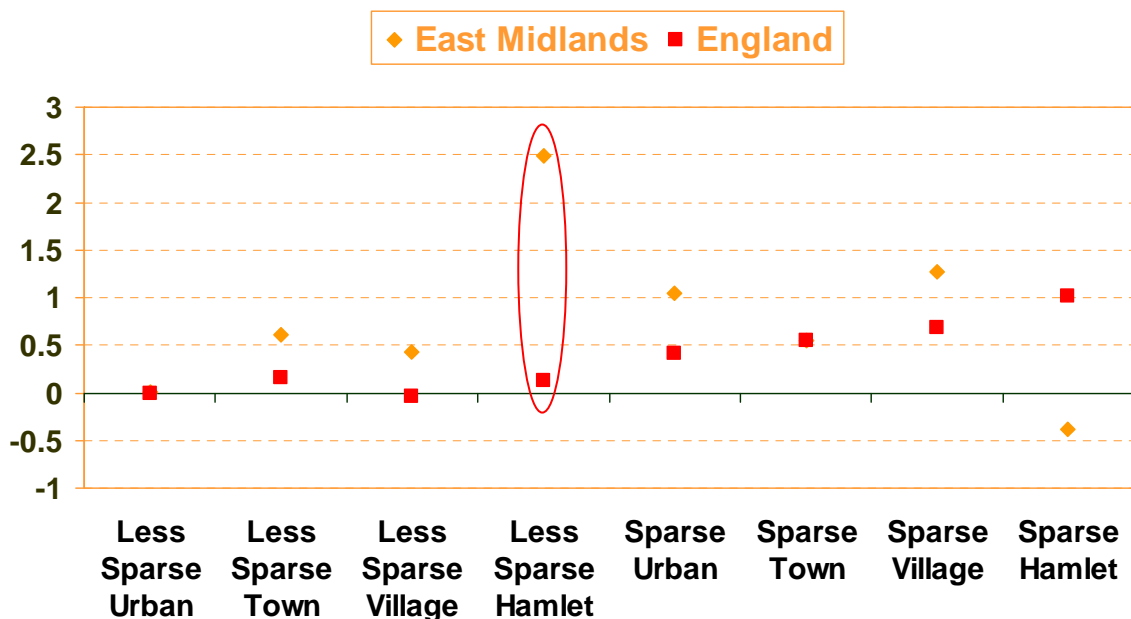
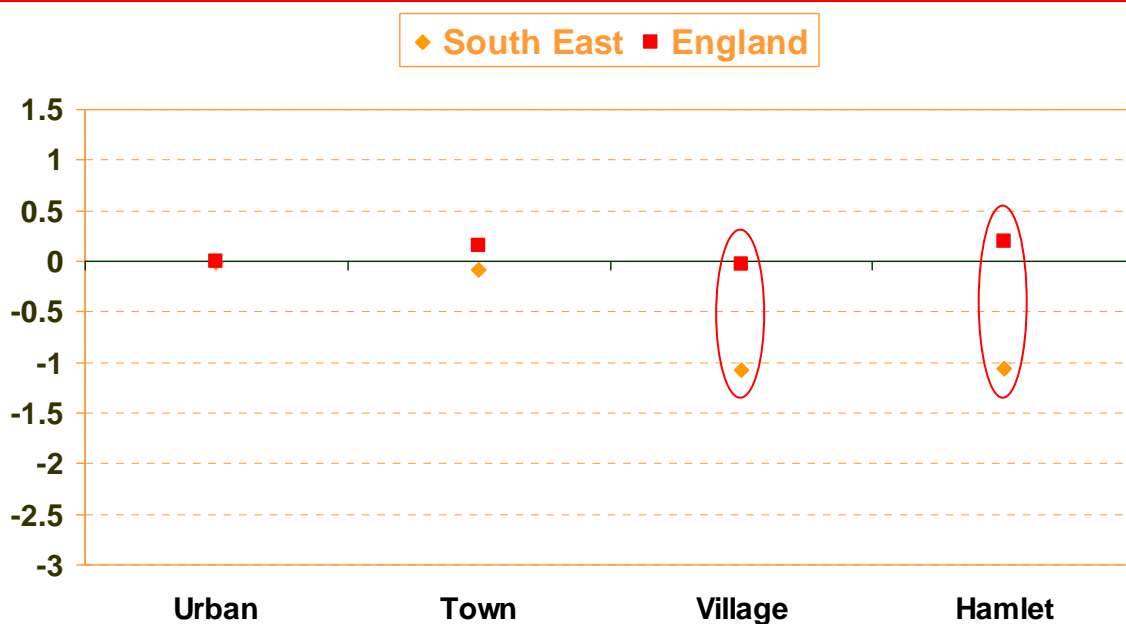


Figure 5.19 Relationships Between Settlement Type and Key Stage 3 Attainment in the East Midlands Compared to the Rest of England



The final region with a significant difference from the rest of England in the relationship between settlement type and attainment progress was the South East. In this region the relationship was inverse, showing that pupils in rural areas actually made less progress than those in urban areas. When we examined this relationship in more detail, we found that there were no pupils living in sparse areas in the South East, so we used the 4-category settlement size variable to distinguish between different types of rural area instead. The results are shown below in Figure 5.20.

Figure 5.20 Relationships Between Settlement Type and Key Stage 3 Attainment in the South East Compared to the Rest of England



These results show that pupils living in all sizes of rural settlement, but particularly villages and hamlets (the relationship is significant for these two settlement types), make less progress than those living in urban areas in the South East. Pupils living in hamlets were found to make around 1.5 points (or 0.44 SD) less progress if they lived in rural areas compared to those living in urban areas, which is one of the larger relationships found by this study. This suggests that the South East has a very different relationship between settlement type and attainment at Key Stage 3 from other regions of England.

We carried out some further analyses to investigate possible reasons behind the regional differences found, particularly with regard to the South East of England. The results of these showed that, as expected, average social class and levels of mothers' qualifications are both higher in the South East than in the rest of England. We also found a slightly stronger relationship between mother's highest qualification and attainment in the South East than in the rest of England, which may indicate that – although the region is more affluent overall – pupils from more deprived or less well-educated families may be worse off in terms of attainment than in other regions. This may be particularly true in rural areas where less support might be available for these pupils. We also found that pupils in the South East were more likely to have SEN than those in other regions (30.1% compared to an average of 20%). Since rural pupils with SEN may be less well supported at school than urban pupils, this may also contribute to the different relationship seen in the South East compared to the rest of England.

We performed the same analyses at Key Stage 4, but found no significant interaction between region and settlement type at this Key Stage, although there was a tendency for similar relationships to those seen at Key Stage 3. This may indicate that any inequalities in attainment across different regions at Key Stage 3 have been lessened or removed entirely by Key Stage 4.

7 Conclusions and Recommendations

7.1 Differences Between Rural and Urban Areas

Our analyses have shown that pupils living in rural areas do have higher average attainment than those living in urban areas, as found in Census data. However, this difference is almost entirely due to differences in social position, area deprivation and pre-existing attainment on entry to secondary school. It therefore seems that there are no intrinsic benefits to attainment that come from living in a rural area, at least not during secondary school.

Despite this, the analyses show that pupils living or attending schools in sparse areas may have slightly higher attainment than those in less sparse areas. Due to the small numbers of respondents in sparser areas this relationship is not statistically significant and therefore may be due to chance, but a larger sample may yield significant results. There is also some evidence that urban pupils attending rural schools may have slightly higher attainment than those attending urban schools.

Recommendations:

The perceived benefits for attainment of living in a rural area seem to be largely due to differences in social position, and therefore it is important to recognise that pupils with less advantageous circumstances who live in rural areas may need as much support as those living in urban areas. We would therefore discourage the assumption that all rural pupils are educationally advantaged.

We would recommend further research looking at differences in primary school attainment between rural and urban areas, in case there are as-yet unidentified benefits of living in rural areas at earlier ages. The present research does not rule out this possibility, although our results unadjusted for Key Stage 2 attainment show little difference from the adjusted results (see Appendix D). We would also recommend consideration of a rural boost to be added to future cohorts of LSYPE, particularly for sparse rural areas, in order to increase sample sizes in these areas (see Appendix G for details) so that the relationships can be further clarified.

7.2 Particular Groups of Pupils with Lower Attainment

We found that pupils in some ethnic minority groups appear to make less progress in attainment if they live in rural areas compared to the same ethnic group living in urban areas. However, sample sizes for these analyses were small and we therefore cannot confirm this finding without further analysis on larger groups of pupils. We also found that respondents with lower qualified mothers tended to

make less progress if they were living in rural areas, while those with higher qualified mothers tended to do particularly well if they lived in rural areas. Finally, we found that there is a larger gap in attainment between rural and urban pupils in lower achieving schools. This means that urban pupils in low-achieving schools have the lowest attainment of all and need the most support, whereas in high-achieving schools there is little difference in attainment between rural and urban pupils.

Taken together, these results show that particular groups of pupils may be especially disadvantaged if they live in rural areas, and may need additional support. In particular, the lesser progress of rural pupils with low qualified mothers is of concern, as this may indicate a possible pathway by which low attainment is passed down through generations in rural areas. The size of this relationship also increased from Key Stage 3 to Key Stage 4, where most of the other relationships we found were diminishing over time, and this may indicate an increasing problem for some rural pupils.

Recommendations:

Small sample sizes proved to be a problem in identifying particular groups of rural pupils who make less progress during secondary school. This is due to the small numbers of pupils living in some types of rural area and also the small numbers of particular groups of pupils (particularly minority ethnic groups) living in rural areas (see Appendix G). It is difficult to surmount this problem with currently available national data, so we would recommend that future large studies include rural sample boosts in order to raise sample sizes. In the meantime, dedicated smaller studies of rural pupils could be undertaken in order to investigate those pupils particularly at risk of low attainment.

Despite these problems of sample size, our analyses did find that certain types of pupils living in rural areas tended to make less progress. We would therefore suggest that strategies designed to benefit disadvantaged pupils living in rural areas could be implemented, since the groups we found to have lower attainment in rural areas tend to be disadvantaged pupils (i.e. they have mothers with low educational qualifications or they are from minority ethnic groups). Additional support could be made available at the school or the individual level, as our analyses have established that both of these levels are important for attainment. However, targeting low-achieving schools is likely to benefit urban pupils more than rural pupils, as urban pupils tend to be more disadvantaged in such schools.

Narrowing attainment gaps between advantaged and disadvantaged pupils is a key priority for DCSF, and it appears that these gaps may be wider in rural than in urban areas for certain types of pupils. Rural areas may therefore require

particular strategies to target particularly disadvantaged young people in order to narrow such gaps in attainment.

7.3 Regional Differences in Attainment

Our regional analyses showed that, while in most regions of England living in a rural area is beneficial, in the South East those living in rural areas actually made less progress in attainment from Key Stages 2-3 than pupils living in urban areas. This difference did not persist to Key Stage 4, although there was a tendency towards a similar relationship. However, the relationship we found at Key Stage 3 appeared to be robust, and was one of the larger relationships found by the study, which does appear to indicate a different experience for pupils living in the more affluent South East of England.

Recommendations:

We would recommend further investigation of the lower attainment found among rural pupils living in the South East and possible factors behind this, which may be due to the relative affluence of the South East compared to the rest of England. If these differences are found to be systematic across the South East, they are again likely to be linked to disadvantaged pupils and we would again suggest that strategies designed to support these pupils at secondary school would be beneficial. Since DCSF's 14-19 policies focus on delivery at the regional, as well as the local and national level, we feel that these regional results are important and should be further investigated.

References

Bibby P and Shepherd J (2004). *Developing a new classification of urban and rural areas for policy purposes – the methodology*. Rural Evidence Research Centre, http://www.nerc.ac.uk/findings/documents_rural/RD2Definition_Methodology.pdf.

Bradley S and Taylor J (2004). Ethnicity, educational attainment and the transition from school. *The Manchester School* 72: 317.

Bradley S and Taylor J (1998). The effect of school size on exam performance in secondary schools. *Oxford Bulletin of Economics and Statistics* 60: 291.

Bynner J and Joshi H (2002). Equality and opportunity in education: evidence from the 1958 and 1970 birth cohort studies. *Oxford Review of Education* 28: 405.

Centre for Multilevel Modelling (2009). 'Realcom' – Developing multilevel models for REAListically COMplex social science data. Accessed from: <http://www.cmm.bristol.ac.uk/research/Realcom/index.shtml>.

Clemens S, Kinnaird R, Mackey T, Deakin G and Ulman A (2005). *Extended Services in Schools: Baseline Survey of Maintained Schools in 2005*. Department for Education and Skills Research Report RR681.

Commission for Rural Communities (2008). *The State of the Countryside 2008*. Cheltenham: Commission for Rural Communities.

Commission for Rural Communities (2006a). *Research into the aspirations of young people in the rural West Midlands*. Birmingham: Ecotech.

Commission for Rural Communities (2006b). *Rural Disadvantage: Reviewing the evidence*. London: Countryside Agency.

Countryside Agency, Department for Environment, Food and Rural Affairs, Office of the Deputy Prime Minister, Office for National Statistics and Welsh Assembly Government (2004). *Rural and Urban Area Classification 2004: An Introductory Guide*, [http://www.defra.gov.uk/rural/ruralstats/rural-defn/Rural Urban Introductory Guide.pdf](http://www.defra.gov.uk/rural/ruralstats/rural-defn/Rural_Urban_Introductory_Guide.pdf).

Defra (2008). *Departmental Strategic Objective: Socially and economically sustainable rural communities*. Department for the Environment, Food and Rural Affairs. <http://www.defra.gov.uk/rural/dso/index.htm>.

- Dobson J (2008). Pupil mobility, choice and the secondary school market: assumptions and realities. *Educational Review* 60: 299.
- Environment, Food and Rural Affairs Committee (2003). *Ninth Report of Session 2002-03. The Delivery of Education in Rural Areas*. London: The Stationery Office Report HC467.
- Fox R (2007). *Rural and Coastal Study Literature Review*. Aim higher in the East of England.
- Gibbons S and Silva O (2008). Urban density and pupil attainment. *Journal of Urban Economics* 63: 631.
- Glaeser E, Kahn M and Rappaport J (2000). *Why do the poor live in cities?* Working Paper 7636. NBER.
- Hammond C (2002). Efficiency in the provision of public services: a data envelopment analysis of UK public library systems. *Applied Economics* 34: 649.
- Heath A and Brinbaum Y (2007). Explaining ethnic inequalities in educational attainment. *Ethnicities* 7: 291.
- Kalambouka A, Farrell P, Dyson A and Kaplan I (2007). The impact of placing pupils with special educational needs in mainstream schools on the achievement of their peers. *Educational Research* 49: 365.
- Kershaw C *et al.* (eds.) (2008). *Home Office Statistical Bulletin: Crime in England and Wales 2007/08: Findings from the British Crime Survey and police recorded crime*. London: Office for National Statistics.
- Lupton R (2004). *Schools in disadvantaged areas: Recognizing context and raising quality*. Discussion paper 76. CASE.
- Martin D, Brigham P, Roderick P, Barnett S and Diamond I (2000). The (mis)representation of rural deprivation. *Environment and Planning* 32: 735.
- Murphy H (2007). School transport, inclusion and pupil attendance. *Proceedings of the Institution of Civil Engineers – Municipal Engineer* 160: 197.
- Office for National Statistics (2001). *A Review of Urban and Rural Definitions*. Project Report, The Stationery Office, London.

Piesse A and Kalton G (2009). *A Strategy for Handling Missing Data in the Longitudinal Study of Young People in England (LSYPE)*. Research Report No RW086, London: Department for Children, Schools and Families.

Rural Evidence Research Centre (2004). *Social and Economic Change and Diversity in Rural England: A report by the Rural Evidence Research Centre, Birkbeck College, University of London*. Department for Environment Food and Rural Affairs, London.

Sacker A, Schoon I and Bartley M (2002). Social inequality in educational achievement and psychosocial adjustment throughout childhood: magnitude and mechanisms. *Social Science and Medicine* 55: 863.

Ward K and D'Souza J (2008). *LSYPE User guide to the datasets: Wave One to Wave Three*. UK Data Archive Study Number 5545 – Longitudinal Study of Young People in England.

Appendix A Glossary of Terms

BCS70 – the 1970 Birth Cohort Study

CA – Countryside Agency

CRC – Commission for Rural Communities

DETR – Department for the Environment, Transport and the Regions

FSM – Free school meals

HC – Housing Commission

LSYPE – Longitudinal Study of Young People in England

NCDS – the National Child Development Study

NPD – National Pupil Database

ODPM – Office of the Deputy Prime Minister

ONS – Office for National Statistics

OPCS – Office for Population Censuses and Surveys

OS – Ordnance Survey

PLASC – Pupil Level Annual School Census

PRU – Pupil Referral Unit

RDC – Rural Development Commission

RERC – Rural Evidence Research Centre

SD – Standard Deviation – see Appendix E

Wave 1 – the first data collection phase of LSYPE, conducted in 2004 when the respondents were in Year 9. This corresponds to Key Stage 3 in their education

Wave 3 – the third data collection phase of LSYPE, conducted in 2006 when the respondents were in Year 11. This corresponds to Key Stage 4 in their education.

Appendix B Tables of Results

Table 6.1 Variables Affecting the Relationship Between Settlement Type and Attainment at Key Stage 3

Variables in Model	Points Change	Significant?
Settlement Type	+0.85	Yes
Settlement Type + Key Stage 2 Points Score Compared to School Average	+0.14	No
Above Variables + Mother's Highest Qualification Parental Socio-Economic Class Single Parent Family	+0.06	No
Above Variables + Gender	+0.07	No
Above Variables + Ethnic Group	+0.18	No
Above Variables + Special Educational Needs	+0.16	No
Above Variables + Index of Multiple Deprivation	+0.08	No
Above Variables + Distance from Home to School	+0.13	No
Above Variables + Type of School	+0.15	No
Above Variables + School-Level Average Attainment at Key Stage 2 Interaction Between Own Attainment and School-Level Average at KS2	+0.13	No
Above Variables + Percentage of LA Pupils Receiving Free School Meals	+0.10	No

Table 6.2 Variables Affecting the Relationship Between Settlement Type and Attainment at Key Stage 4

Variables in Model (all models based on 8,495 cases)	Points Change	Significant?
Settlement Type	+8.34	No
Above + Key Stage 2 Points Score	-0.15	No
Above + Mother's Highest Qualification		
Above + Parental Socio-Economic Class		
Above + Single Parent Family	-4.02	No
Above + Gender	-2.93	No
Above + Ethnic Group	+3.71	No
Above + Special Educational Needs	+3.55	No
Above + Index of Multiple Deprivation	+0.46	No
Above + Distance from Home to School	+1.39	No
Above + Percentage of Pupils in School Eligible for Free School Meals	+1.70	No
Above + School Level Average Attainment at Key Stage 2		
Above + Interaction Between Own Attainment and School-Level Average at KS2	+0.31	No
Final 2-Level Model Without LA Variance	+0.46	No

Table 6.3 Relationship Between Settlement Type (8 Categories) and Attainment at Key Stage 3, Adjusted

Settlement Type (less-sparse urban areas are the reference category)	Points Change	Significant?
Town and fringe less sparse	+0.16	No
Village less sparse	-0.04	No
Hamlet and scattered dwellings less sparse	+0.13	No
Urban sparse	+0.41	No
Town and fringe sparse	+0.55	No
Village sparse	+0.69	No
Hamlet and scattered dwellings sparse	+1.02	No

Table 6.4 Relationship Between Settlement Type (8 Categories) and Attainment at Key Stage 4, Adjusted

Settlement Type (less-sparse urban areas are the reference category)	Points Change	Significant?
Town and fringe less sparse	+1.70	No
Village less sparse	-4.79	No
Hamlet and scattered dwellings less sparse	+4.63	No
Urban sparse	+0.15	No
Town and fringe sparse	+6.18	No
Village sparse	+10.04	No
Hamlet and scattered dwellings sparse	+23.63	No

Table 6.5 Relationship Between Settlement Type of School (8 Categories) and Attainment at Key Stage 3, Adjusted

Settlement Type (urban less sparse areas are the reference category)	Points Change	Significant?
Town and fringe less sparse	+0.11	No
Village less sparse	+0.54	No
Hamlet and scattered dwellings less sparse	+0.04	No
Urban sparse	+1.82	No
Town and fringe sparse	+1.18	No
Village sparse	N/A	No
Hamlet and scattered dwellings sparse	+0.83	No

Table 6.6 Relationship Between Settlement Type of School (8 Categories) and Attainment at Key Stage 4, Adjusted

Settlement Type (urban less sparse areas are the reference category)	Points Change	Significant?
Town and fringe less sparse	+12.66	No
Village less sparse	+19.30	No
Hamlet and scattered dwellings less sparse	-1.70	No
Urban sparse	N/A	No
Town and fringe sparse	+22.08	No
Village sparse	N/A	No
Hamlet and scattered dwellings sparse	-40.31	No

Appendix C Factors Not Included in Final Models

Variables not found to be strongly associated with attainment or not found to improve the overall model were left out of the analyses. These were:

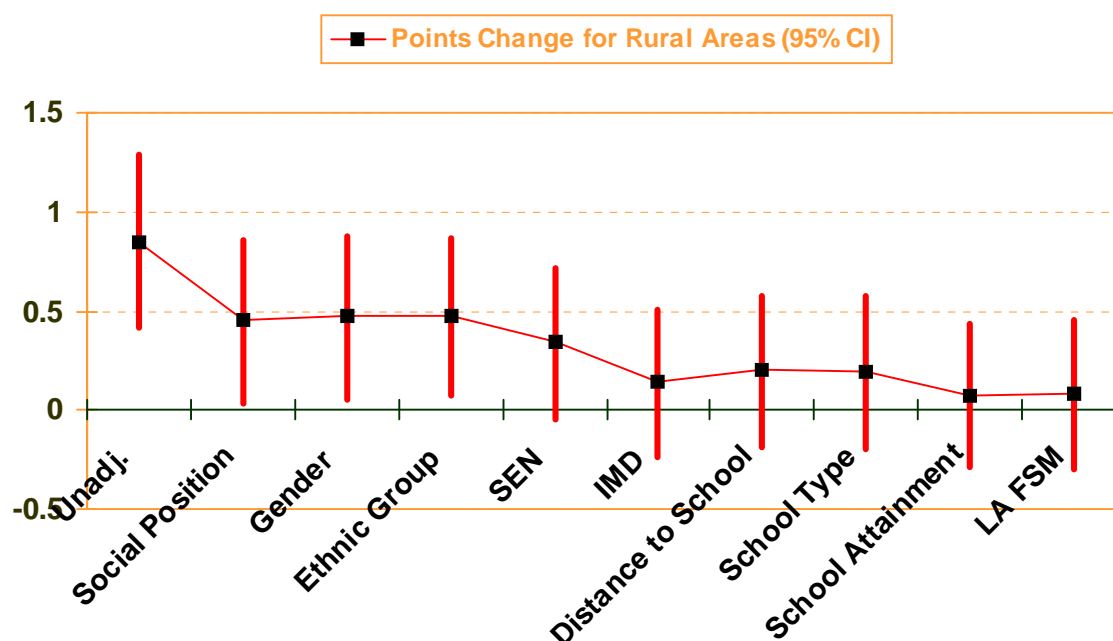
- Receipt of free school meals
- Language spoken at home
- Whether the young person is a carer
- Whether the young person has a disability
- Parental attitudes to school
- Parental aspirations for the young person
- Young person's attitudes to school
- Young person's aspirations
- Family cohesion
- Anti-social behaviour
- Percentage of pupils in a school with SEN
- Percentage of pupils in a school receiving free school meals
- Percentage of pupils in a school without English as a first language
- Pupil to teacher ratio
- Percentage of non-White pupils in a school
- School admissions policy
- Gender of school
- Percentage of non-White pupils in a Local Authority
- Local Authority level average attainment
- Percentage of pupils in a Local Authority without English as a first language.

Appendix D Value-Added Analyses

The adjusted analyses in this report include adjustment for prior attainment at Key Stage 2, and are therefore measuring not raw attainment scores at Key Stages 3 and 4, but the progress in attainment (or otherwise) between Key Stages 2 and 3 (at Wave 1) or between Key Stages 2 and 4 (at Wave 3). A 1-point adjusted difference therefore refers to a 1-point increase in progress between these Key Stages compared to the average pupil.

These analyses are known as ‘value-added’ analyses, because they provide an estimate of the effect that a given period of schooling has had on a pupil’s attainment – in this case, the period of secondary school up to Key Stage 4. In this analysis, we were unable to adjust for factors that may have affected the pupils’ attainment at primary school, so any models not adjusted for Key Stage 2 attainment would have to assume that all pupils living in rural areas at Wave 1 of LSYPE had also been living in rural areas all the way through primary school. This is unlikely to be the case. However, the value-added analyses do not allow us to speculate on overall differences in attainment between rural and urban areas across the entirety of schooling, so we also constructed models without this adjustment for comparison. Figure 6.1 below shows the same results as Figure 5.8 in the main part of the report, but without adjustment for Key Stage 2 attainment.

Figure 6.1 Points Change in Key Stage 3 Score Associated with Living in a Rural Area, with Cumulative Adjustment for Other Factors

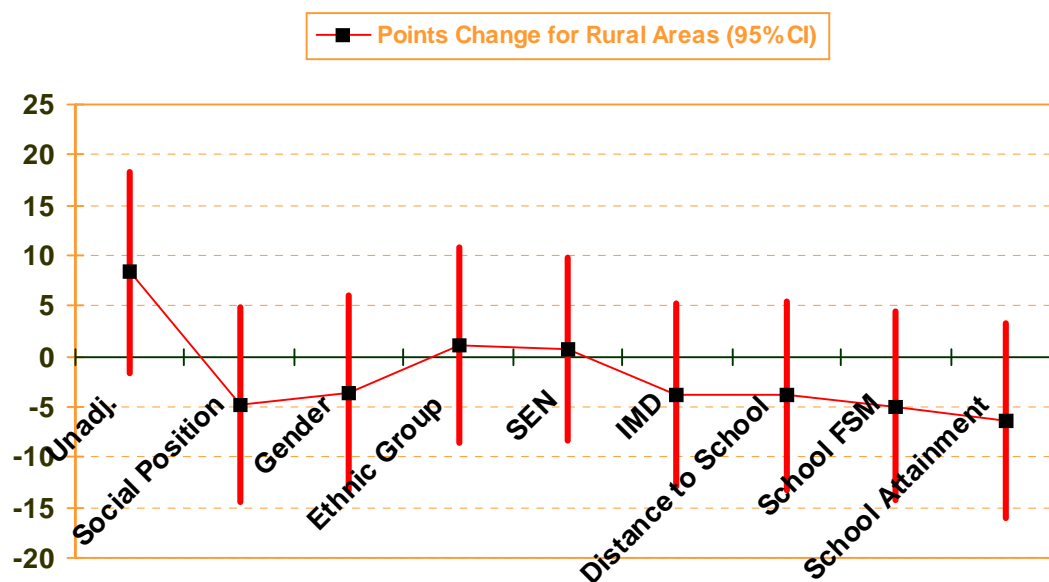


It can be seen from this figure that the lack of adjustment for Key Stage 2 attainment makes virtually no difference to the fully-adjusted relationship between binary settlement type and Key Stage 3 attainment. The main difference is that

when we do not take previous attainment into account, more of the relationship is explained by differences in social position and IMD. This indicates that these factors are also likely to be associated with higher attainment at primary school, and that again there appears to be no intrinsic benefit to living in a rural area once they are adjusted for.

We also completed the same analyses for Key Stage 4. Figure 6.2 below shows the same results as Figure 5.9 in the main report, but again without any adjustment for previous attainment at Key Stage 2.

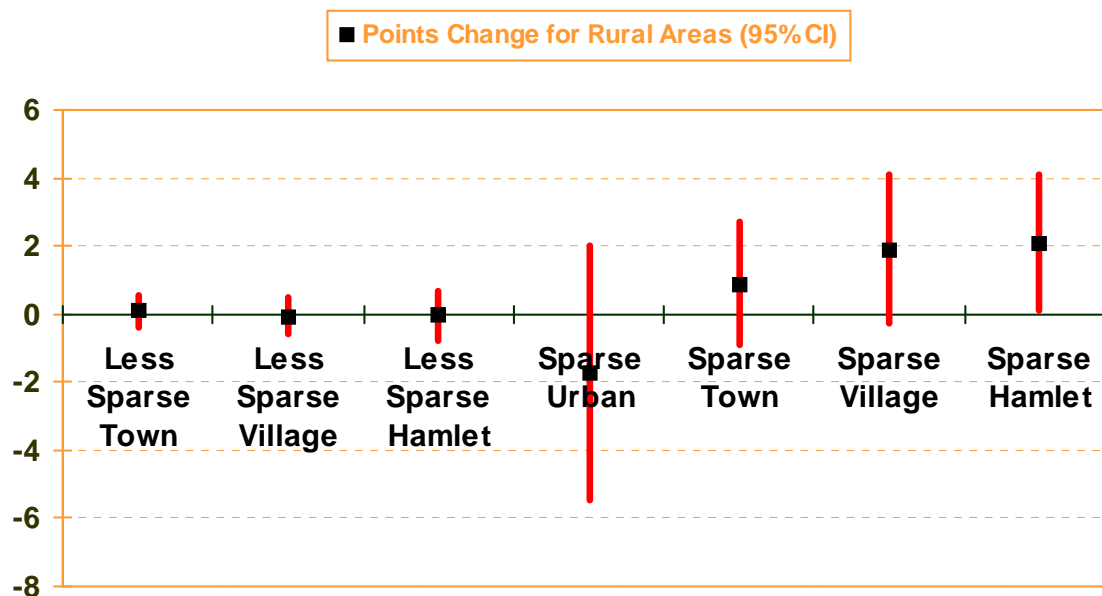
Figure 6.2 Points Change in Key Stage 4 Score Associated with Living in a Rural Area, with Cumulative Adjustment for Other Factors



It can be seen from this Figure that again there is still no significant fully-adjusted relationship between binary settlement type and attainment at Key Stage 4, although in this case the rural pupils appear to have slightly lower attainment without adjustment for Key Stage 2 attainment. This difference is small and not significant, however. Again, social position and area deprivation exert a larger influence when we do not adjust for previous attainment, indicating that they are likely to be important during primary school as well as secondary school.

We also looked at our fully-adjusted analyses for the 8-category settlement type variable without adjustment for Key Stage 2 attainment, in order to see whether this adjustment made a large difference to the relationships. The results of this analysis for Key Stage 3 can be found in Figure 6.3 below, which corresponds to Figure 5.10 in the main report but without adjustment for Key Stage 2 attainment.

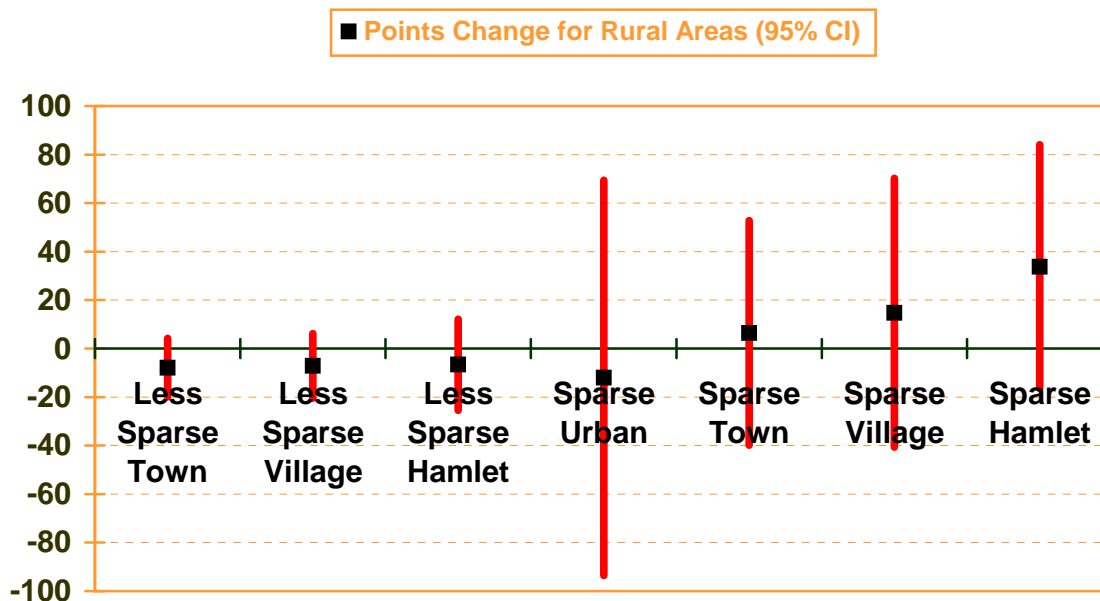
Figure 6.3 Points Change in Key Stage 3 Score Associated with Different Settlement Types (with Less Sparse Urban Areas as Reference Category)



These results show that there is again very little difference between raw attainment scores and value-added scores in the relationship with settlement type. The relationship appears slightly stronger in the sparse area types when we do not adjust for Key Stage 2 attainment, and becomes borderline significant for pupils living in sparse hamlets, but the differences in attainment are small.

Again, we constructed the same analyses for Key Stage 4 raw attainment using the 8-category settlement type variable. The results can be found below in Figure 6.4, which corresponds to Figure 5.11 in the main report.

Figure 6.4 Points Change in Key Stage 4 Score Associated with Different Settlement Types (with Less Sparse Urban Areas as Reference Category)



This figure again shows little difference from the value-added analyses, although again the effect sizes are slightly larger when we look at raw attainment rather than progress. The relationships with different area types appear to be similar, and again there is no significant relationship between living in any particular type of rural area and having higher attainment.

In summary, the additional analyses without adjustment for prior attainment at Key Stage 2 are little different from the value-added analyses in the main part of the report. However, the unadjusted analyses are more difficult to interpret because we do not have any detailed information on the pupils or where they were living when they attended primary school. It is therefore encouraging that we found little difference between the two sets of results, since using the value-added analyses allows us to better explain secondary school attainment but cannot account for any factors that might lead to differences in attainment between rural and urban areas at primary school. The available evidence from the unadjusted analyses indicates that the pre-existing differences in attainment at Key Stage 2 may be due largely to socioeconomic factors (as is the case in our fully-adjusted analyses in Section 5), and that there appears to be no intrinsic benefit to living in a rural area either at primary school or at secondary school.

Appendix E Standard Deviations

This Appendix will present a brief description of standard deviations (SD), as these are referred to throughout the results in Section 5 of the report. The SD of a given measure (e.g. attainment) represents the level of dispersion of the data, so a small SD indicates that the values tend to be clustered closely around the mean, whereas a large SD indicates that they are more widely spread out. For most continuous measures such as scores (if they are normally distributed), 95% of people will have a score within 2 SD of the mean, so that if a mean test score is 100 and the SD is 10, 95% of people will score between 80 and 120 on the test.

Differences between groups (e.g. pupils living in rural and urban areas) can also be measured in terms of standard deviations as an alternative to using percentages where these are not appropriate. SD changes are therefore presented to supplement the changes quoted in points scores throughout Section 5 of the report. For the scores looking at Key Stage 2-3 progress, the mean progress made by pupils in LSYPE was 7.06 points, with an SD of 3.41 points. A 1-point difference in progress between rural and urban areas therefore corresponds to a 0.29 SD change. For the scores looking at Key Stage 2-4 progress, the mean progress made by pupils in LSYPE was 350 points, with an SD of 153 points. A 10-point difference in progress between rural and urban areas therefore corresponds to a 0.07 SD change.

Appendix F Derived Variables

A number of derived variables were used in the analyses for this study, although not all were included in the final models as they did not necessarily reach significance or improve the model as a whole. A list of the variables derived by NatCen and how they were created can be found below in Table 6.7.

Table 6.7 Derived Variables Used in Analyses

Variable Description	Method of Derivation	Waves
Binary indicator of urban/rural	8-category settlement type variable collapsed to 2 categories: urban (less sparse and sparse urban) and rural (less sparse and sparse towns, villages and hamlets)	1, 3
4-category indicator of urban/rural	8-category settlement type variable collapsed to 4 categories: urban (less sparse and sparse urban), town (less sparse and sparse towns), village (less sparse and sparse villages) and hamlet (less sparse and sparse hamlets)	1, 3
Sparsity	8-category settlement type variable collapsed to 2 categories: sparse (sparse urban, towns, villages and hamlets) and less sparse (less sparse urban, towns, villages and hamlets)	1, 3
Number of schools attended in Years 7-11	Different school IDs for each pupil from NPD (Years 7-9 for Wave 1 and Years 7-11 for Wave 3) combined into new variable indicating number of schools attended	1, 3
Distance from home to school	Home and school postcodes obtained from DCSF and direct distances in miles calculated using batch tool	1, 3
Young person's ethnic group	Information on ethnic group taken from young person interview, and coded into one of 8 groups (White, Mixed, Indian, Pakistani, Bangladeshi, Black African, Black Caribbean and Other) – if no young person interview this information was taken from the household grid	1, 3
Mother's highest qualification	List of 50 possible qualifications for main and second parent coded into 7 groups (degree or equivalent, higher education below degree level, GCE A-level or equivalent, GCSE grades A-C or equivalent, qualifications at Level 1 and below, other qualifications, and no qualification), with only highest qualification of mother recorded	1, 3
Main parent's NS-SEC class	Main parent's occupational category calculated from ONS lookup table and grouped into 8 classes (higher managerial and professional, lower managerial and professional, intermediate, small employers and own account workers, lower supervisory and technical, semi-routine, routine and never worked/unemployed)	1, 3
Parental involvement with school	Scale comprised of main parent's answers to questions (whether they attend parents' evenings, how often they speak to teachers, how involved they feel in the young person's school life, if and how often they talk about school reports with the young person, and activities they get involved in at school), recoded so that higher scores indicate greater involvement	1, 3
Parental attitudes to education		1, 3

Young person's attitude to school	Scale comprised of main parent's answers to questions (agreement with statements that young people need qualifications to get a good job, that leaving school at 16 limits opportunities and that they want the young person to have a better education than they had, plus details of what the parent would do to help the young person to stay in education), recoded so that higher scores indicate more positive attitudes	1, 3
Anti-social behaviour score	Scale comprised of whether the young person is happy at school, whether they feel schoolwork is a waste of time, whether they are bored at school and whether they are engaged with schoolwork, recoded so that higher scores indicate a more positive attitude to school	1, 3
Family cohesion score	Scale comprised of smoking, alcohol use, cannabis use, graffiti, vandalism, shoplifting and fighting, recoded so that higher scores indicate more anti-social behaviour	1, 3
Whether young person has a disability/long term illness or health problem	Scale comprised of how well the young person gets on with their parent(s), how often they talk to their parent(s) about things that matter, how often they have a family meal, how often parent(s) know where they are going in the evening	1, 3
Whether household is a single parent household	and how often they talk to their parent(s) about their day at school	1, 3
Percentage of pupils in school / Local Authority with special educational needs	Calculated from two variables present in dataset which code whether the young person has a disability and, if so, whether this makes it hard for them to attend school regularly	
Percentage of non-White British pupils in school / Local Authority	Uses household grid relationships to identify whether none, one or two parents of the young person are present in the household	
School establishment type	Percentages of pupils with and without Statements of Needs (taken from NPD) were combined into a single variable for relevant year (2004 for Wave 1 or 2006 for Wave 3)	
Z-scores for attainment at Key Stages 2, 3 and 4	Percentage of White pupils was removed from total (taken from NPD) for relevant year (2004 for Wave 1 or 2006 for Wave 3) Collapsed into 7 categories: community schools, City Technology Colleges, foundation/voluntary controlled/voluntary aided schools, independent schools and special schools/pupil referral units Calculated using raw points scores minus the population mean score, divided by the population standard deviation. These were included in models and then back-transformed to raw scores for report	

Appendix G Sample Sizes

LSYPE is a large dataset, with over 15,000 respondents at Wave 1 and over 12,000 at Wave 3, and as such it is possible to detect small differences between different groups within the dataset with a relatively small margin of error. However, there are a number of exceptions to this which have created sample size problems with some sections of the project, and these will be explained below.

Rural Areas

As discussed in Section 5 of the report, approximately 80% of the LSYPE respondents live in urban areas, leaving a population of 2,391 young people living in rural areas at Key Stage 3, and 2,009 at Key Stage 4. For this project's analyses, this is an adequate sample size for Key Stage 3 – power calculations completed in Stata indicate that we can detect a (fully-adjusted) 1-point difference in Key Stage 3 progress between urban and rural areas at 80% power and at a 5% significance level with a sample of only 400 pupils living in rural areas. This would correspond to one term of progress between Key Stage 2 and Key Stage 3.

However, at Key Stage 4 the same number of pupils living in rural areas only enables us to detect a (fully-adjusted) 10-point difference in progress, and the actual differences between rural and urban areas observed at Key Stage 4 tended to be much smaller than this (less than 1 point). Since 10 points is still a small difference at Key Stage 4, we conclude that the sample size for these analyses is adequate, but that the differences found at Key Stage 4 are so small that they are unlikely to be significant even with a large sample.

In the analyses of Section 4, we also explore differences between the 8 different area types classified by Defra. Sample sizes are very small for a number of these types, and at Key Stage 3 we calculated that approximately 350 pupils would be required in each rural settlement type in order for us to be able to detect a 1-point difference in progress with 80% power and at a 5% significance level. This means that the sample size is adequate for those pupils living in less sparse rural areas, but too small for those living in the sparse area types (which only contain between 9 and 67 pupils at Key Stage 3).

At Key Stage 4, we calculated that approximately 2,000 pupils would be required in each rural settlement type for us to detect a 10-point difference in progress, and that approximately 500 pupils in each settlement type would be required even to detect a 20-point difference. The sample sizes in all rural areas therefore appear to be too small to detect differences that might be meaningful, although in less sparse towns (n=888) and less sparse villages (n=686) the sample size may be adequate. In any case, the differences we actually observed were again so small

that finding a significant relationship would be very unlikely even if the sample size had been much larger.

In future cohorts of LSYPE, we would suggest that a boost sample or sample redistribution would be beneficial in order to provide additional power for analyses focusing on rural areas.

Minority Ethnic Groups

Waves 4 and 5 of LSYPE contain an ethnic boost which recruits extra pupils from minority ethnic groups in the dataset. However, Waves 1-3 of the study do not contain this boost and therefore have often small numbers of pupils from minority ethnic groups despite the oversampling mentioned in Appendix J). This is a particular problem when studying pupils living in rural areas, as the numbers of minority ethnic pupils in rural areas are particularly small.

At Key Stage 3, the numbers of minority ethnic pupils are sufficient to detect a 1-point difference in attainment with 80% and at a 5% level of significance. However, the numbers living in rural areas are very small and it is therefore difficult to make inferences from any differences found. At Key Stage 4, the numbers of minority ethnic pupils are smaller and it is therefore more difficult to detect small differences in attainment. The sample boost added from Wave 4 of LSYPE onwards is likely to improve these difficulties, but they do make the interactions with ethnic group found in this project more difficult to interpret.

Government Office Region

Finally, the analyses broken down by region suffer from small sample size in the North East region at both Key Stage 3 and Key Stage 4. All other regions have adequate sample sizes, but in the North East there are only 708 pupils at Key Stage 3 and 562 pupils at Key Stage 4. This means that the analyses focusing on the North East are likely to be underpowered.

Appendix H Technical Methods

The analyses conducted for this project were necessarily complex in order to take account of the complexity of the data and properly explore the relationships between different groups. These will be described in brief in this Appendix, with reference to further literature where necessary.

Missing Data and Multiple Imputation

Not all the variables selected for inclusion in the analyses (see Section 3.1) had complete data, and this therefore reduced the number of cases present in the analyses. Some variables (particularly the derived scales describing family cohesion, parental attitudes and risky behaviours) had large numbers of missing values which reduced the number of cases included in the analyses considerably. We decided that, rather than undertaking a complete case analysis which would restrict the analyses to only those respondents who had answered all the questions to be included in the models (which would have included approx. 70% of respondents at each wave) we would employ multiple imputation to produce a complete dataset.

This was done by running multiple imputation models in MLwiN, which are designed to include plausible values for variables that have missing data. Multiple imputation is thought to be a better way of dealing with missing data than complete case analysis, because it can make use of all the available data and is relatively simple to compute. It has also been particularly recommended for cases of item non-response in LSYPE (Piesse and Kalton, 2009). Using this method, the values of missing variables are predicted using the values of other variables, and can therefore reduce bias in the resulting estimates.

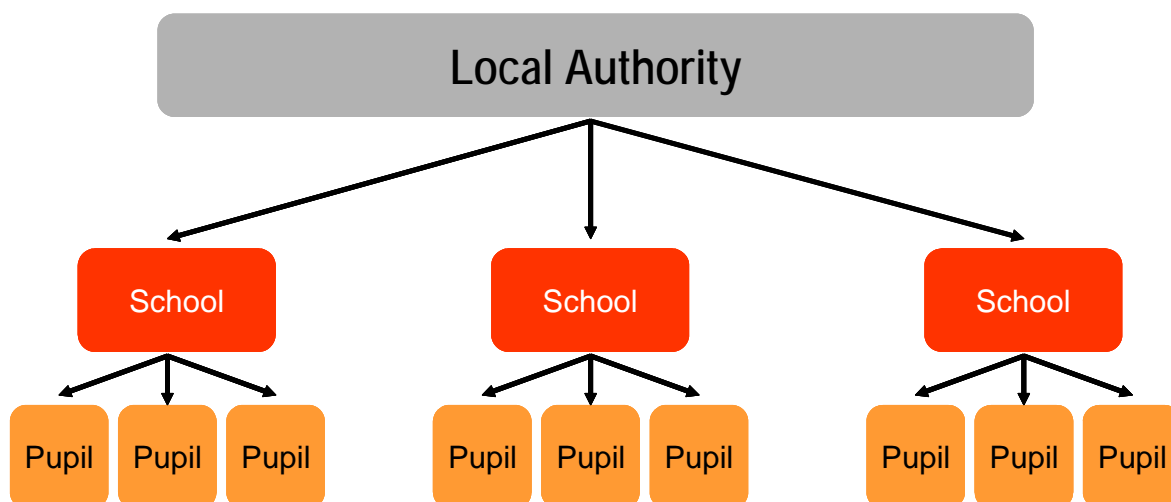
For these analyses, we imputed 10 datasets (including all the variables to be used in the statistical models) using REALCOM-impute (Centre for Multilevel Modelling, 2009) for each wave of the study used, and combined the estimates produced by these imputations in all the analyses. The imputation models described were not designed to account for attrition in the study after Wave 1, so the multiple imputations were also combined with weights that are designed to combat such attrition. The use of multiple imputation in this study should provide more accurate estimates of relationships than a complete case analysis, and should also increase the power of the analyses so that small differences between settlement types can be more easily distinguished.

Multilevel Models

We used a technique known as multilevel modelling to explore the relationship between settlement type and attainment, as these models are known to be useful

for analysing educational data. In particular, in the LSYPE dataset pupils are clustered within schools and schools are clustered within Local Authorities, because of shared factors such as teaching methods, pupil demographics and education policies. This means that any two pupils who attend the same school will tend to be more similar in attainment than if they attended different schools, and that any two schools within the same Local Authority will also tend to be more similar. Using multilevel models helps us to take account of this clustering, and also allows us to estimate the amount of variation in attainment that is due to differences between pupils, between schools and between Local Authorities. The structure of the models used for this project's analyses is summarised below in Figure 6.1.

Figure 6.5 Multilevel Model Structure for Analyses



In the models shown above, attainment is allowed to vary between schools and also between Local Authorities, so that each school and each Local Authority has its own baseline level of attainment that individual pupils and schools can vary around. These models reflect the real structure of the data much better than ordinary regression models, and can also help to prevent us from overestimating the differences in attainment between individual pupils.

Multiple Membership Models

Because this project examines changes in attainment across secondary school, a number of the pupils in the LSYPE dataset had moved school at least once during the time period we were studying (from Year 7 to Year 11). This movement between schools can affect some of the relationships found, because factors that operate at the school level will change when the pupil changes schools. Ordinary

models do not capture this movement and may therefore underestimate school-level variation in attainment. We therefore decided to use multiple membership models in order to take movement between schools in LSYPE into account. This involved using the NPD to calculate the amount of time pupils had spent in each secondary school they had attended in Years 7-9 (for Key Stage 3) and Years 7-11 (for Key Stage 4). We then weighted our models according to the amount of time pupils had spent in each school so that the school-level factors reflected all the different secondary schools each pupil had attended.

At Key Stage 3, we found that 1,553 pupils (10% of the total) had moved school at least once between Year 7 and Year 9, and 59 pupils (0.4%) had moved school more than once. At Key Stage 4 this percentage had increased to 1,413 pupils (or 12% of the total), of whom 72 (or 0.6%) had moved school more than once. It can be seen from Section 5.6 that at both Key Stages taking account of movement between schools increased the amount of variation in attainment between schools. Our analyses therefore show that it is important to take this movement between schools into account when looking at changes in attainment.

The present analyses focus only on progression in secondary school attainment, and because the NPD only began in 2001 it is therefore not possible to include primary school factors. However, in future analyses it would also be desirable to examine the primary school attended and take primary school factors into account, since the transition between primary and secondary school may also be an important time point to study in terms of attainment.

Appendix J Sample Design and Weighting

Sampling from Maintained Schools

In the maintained sector, the sample was drawn using the Pupil Level Annual Schools Census (PLASC), and there was a two-stage probability proportional to size (PPS) sampling design, with disproportionate stratification. The primary sampling unit (PSU) was the school, and maintained schools were stratified into deprived/non-deprived, with deprived schools (defined by schools in the top quintile according to the proportion of pupils receiving free school meals) being over-sampled by a factor of 1.5. Within each deprivation stratum, school selection probabilities were calculated based on the number of pupils in Year 9 from major minority ethnic groups (Indian, Pakistani, Bangladeshi, Black African, Black Caribbean and Mixed). Within each stratum, maintained schools were ordered and thus implicitly stratified by region then by school admissions policy before selection. 838 schools were selected in the maintained sector.

The second stage sampled the pupils within schools. Pupils from the six major minority ethnic groups identified above were over-sampled at pupil level in order to achieve target sample numbers of 1000 in each group. The school sampling stage took into account the number of pupils from each of these minority groups. Taken together, the school selection probabilities and the pupil selection probabilities ensured that, within each stratum of deprivation, all pupils had an equal chance of selection. The average number of pupils sampled per school was 33.25, although this varied according to the ethnic group composition of the school.

Sampling from Independent Schools and PRUs

A two-stage sampling design was also used for independent schools and PRUs, but these were sampled using the School Level Annual Schools Census (SLASC). Independent schools were stratified by percentage of pupils achieving five or more A*-C GCSE grades in 2003 within boarding status (i.e. whether or not they had any boarding pupils), within gender of pupils (i.e. boys, girls and mixed). PRUs formed a stratum of their own. Both independent schools and PRUs were sampled with probability proportional to the number of pupils aged 13 at that institution. 52 independent schools and 2 PRUs were sampled in this way.

Pupils in independent schools and PRUs were sampled directly from school rolls by LSYPE interviewers using a sampling program. An average of 33.25 pupils was randomly selected at each school/PRU containing 34 or more Year 9 pupils. All Year 9 pupils were selected in schools/PRUs containing fewer than 34 but more than five Year 9 pupils.

Sample Exclusions

Excluded from the original sample were young people educated solely at home (and therefore not present on a school roll), pupils in schools with fewer than ten (maintained sector) or six (independent sector) Year 9 pupils, boarders (including weekly boarders) and young people residing in the UK solely for educational purposes.

Longitudinal Sampling

At each subsequent wave, the survey attempted to follow all the households who took part in the previous wave where the young person was still alive and living in the UK. Movers were traced using the stable contact address collected at Wave 1, and where this failed, DCSF sent a letter to the head teacher of the school from which the young person was sampled to locate up-to-date address details for them.

Response Rates

Of the 21,000 young people sampled at Wave 1, the survey reached 15,770 households (74%) in England. This comprises 13,914 full interviews (66%) and 1,856 partial interviews (9%), most of which were cases where the second adult in the household was not interviewed. At Wave 2, the survey reached 86% of the total households, and at Wave 3 it reached 92% of the total households.

Weighting

The LSYPE data were weighted to account for the survey design for each wave of the study, and pupils from maintained and non-maintained schools were weighted separately at Wave 1. For pupils from independent schools and PRUs, responses were found to vary according to the sex of the pupil and the size of the school, so these pupils were weighted accordingly and the weights combined with design weights which were taken from the reciprocal of the pupil's selection probability. Calibration weights were also applied, so that the achieved sample size matched the population breakdown by type of school and by region. Pupils from maintained schools were first weighted according to school non-response (found to be linked to the school's deprivation status and its region), and then according to pupil non-response (found to be linked to region, ethnicity and qualifications). These were again combined with the design weights, and the two sets of weights for maintained and non-maintained schools were then combined and weighted so that the maintained/non-maintained split matched the population proportions.

For subsequent waves of the study, statistical models were used to model the differences between those who responded at each wave and those who did not. These non-response weights were again calculated separately for pupils from maintained and non-maintained schools and then combined.

