

JRF Programme Paper
Poverty and ethnicity**DATA SOURCES ON
ETHNICITY AND POVERTY
IN WALES****David Owen****October 2013**

This paper:

- describes statistical surveys undertaken in Wales;
- reviews official data sources containing information relevant to poverty and ethnicity;
- assesses the overall availability of quantitative data on poverty and disadvantage experienced by people from minority ethnic groups and new migrants in Wales.

The Joseph Rowntree Foundation (JRF) commissioned this paper as part of its programme on poverty and ethnicity, which aims to understand the underlying reasons for variations in low income and deprivation among different ethnic groups in the UK and the problems caused.

ISBN 9781859359723

© University of Warwick

Contents

	Page
Introduction	1
Section 1 Survey data sets yielding information on Wales	3
Section 2 Administrative and related data sets for Wales containing information relevant to ethnicity and poverty	42
Section 3 Assessment of the availability of data on minority ethnic and migrant groups and poverty in Wales	50
References	52
Appendix: Websites providing access to data	53

Introduction

The purpose of this report is to document the availability of quantitative data that can provide an insight into poverty and disadvantage experienced by people from minority ethnic groups and new migrant communities living in Wales. The report reviews the characteristics of official statistical sources and survey data sets that contain information relevant to the measurement of dimensions of poverty and identifies those that include information on minority ethnic groups and international migrants.

The measurement of poverty

The UK government defines a household as experiencing poverty if its income is less than 60 per cent of the median household income. The Joseph Rowntree Foundation has adopted a broader definition of 'poverty' for its programme on poverty and ethnicity. This encompasses not just low household income but also the impact of poverty and covers a number of life domains, specifically:

- lack of opportunities for sustained and rewarding employment;
- difficulty in accessing education and leisure opportunities;
- physical and/or mental ill health and disability;
- poor services, stigma and discrimination;
- social isolation, insecurity and anxiety;
- living in a deprived area with poor infrastructure and high levels of social problems such as crime and anti-social behaviour; and
- distribution of resources within households (between men and women, generations living together and adults and children).

Variables that can be used to create indicators of poverty under these domains are collected in a range of quantitative data sources. Survey data sets are usually collected using a household-level sampling frame and collect information on the household as well as for individuals in the household (in

most instances one individual is selected to be interviewed from those resident).

This report documents the major socio-economic data sources which exist in Wales. It focuses on data sources that have been available from 2001 onwards, which provide information on the current circumstances of minority ethnic and migrant groups. These can also be used to monitor how the incidence of poverty has changed over the recent past. The nature of each source and the information which can be derived from it is described and details provided about how it can be accessed.

Structure of the report

This report is structured as follows. Section 1 is concerned with describing the surveys currently undertaken in Wales. Section 2 describes the administrative data sets available for Wales containing information relevant to minority ethnic and migrant groups. Section 3 provides an overall assessment of the availability of data for the measurement of poverty among people from minority ethnic and migrant groups in Wales. An Appendix details the main websites providing access to the data discussed in the report.

Survey data sets yielding information on Wales

A number of statistical surveys containing information of relevance to the measurement of poverty (and the domains of interest to JRF) are conducted in Wales. Those which also collect information relevant to the study of ethnicity and migration are summarised in Table 1. This table lists the type of survey, which variables relevant to studying the characteristics of minority ethnic and migrant groups are included in the survey and whether the survey is currently being conducted.

The surveys being undertaken in Wales can be broadly classified as either:

- UK- or Great Britain-wide surveys conducted in Wales;
- Surveys that focus on Wales.

In the remainder of this section, the nature of each survey is described. The sample design, sample size, data on minority ethnic groups and methods of access are described in greater detail for each data set. The first part of the chapter is concerned with UK- or GB- wide surveys. The second part focuses on surveys specific to Wales.

Table 1: Summary of statistical surveys relevant to ethnicity and poverty in Wales

Data set	Type of survey	Contains ethnicity information	Contains migrant information	Contains information on religion	Poverty-relevant	Currently collected?
Census of Population	Survey (100%)	Yes	Yes	Yes	Yes	Yes
The Labour Force Survey	Household survey	Yes	Yes	Yes	Yes	Yes
The Family Resources Survey	Household survey	No	No	No	Yes	Yes
Infant Feeding Survey	Survey of all new mothers	Yes	Yes	Yes	Yes	Yes
British Cohort Study 1970	Repeat survey of birth cohort	No	No	No	Yes	Yes
Millennium Cohort Study	Repeat survey of birth cohort	Yes	No	Yes	Yes	Yes
British Household Panel Survey	Longitudinal household survey	Yes	Yes	Yes	Yes	No
Living Costs and Food / Expenditure and Food Survey	Survey	Yes	No	No	Yes	Yes
Annual Population Survey	Household survey, 5 successive waves	Yes	Yes	Yes	Yes	Yes

Table 1: Summary of statistical surveys relevant to ethnicity and poverty in Wales (continued)

Data set	Type of survey	Contains ethnicity information	Contains migrant information	Contains information on religion	Poverty-relevant	Currently collected?
Local Area Labour Force Survey	Household survey, 5 successive waves	Yes	Yes	Yes	Yes	No
Integrated Household Survey	Household survey, 5 successive waves	Yes	Yes	Yes	Yes	Yes
Understanding Society	Longitudinal household survey	Yes	Yes	Yes	Yes	Yes
Living in Wales Survey / Welsh Household and Dwelling Survey	Household survey and physical housing survey	Yes	In 2004 and 2007, but not 2008	Yes	Yes	No
National Survey for Wales	Household survey	Yes	No	Yes	Yes	Yes
Welsh Health Survey	Household survey	Yes	No	No	Yes	Yes
Annual Survey of Hours and Earnings	Business survey	No	No	No	Yes	Yes

UK or Great Britain-wide surveys

These can be further grouped into two broad types: (i) regular cross-sectional surveys that are concerned with measuring the population as a whole; and (ii) longitudinal surveys. The latter either provide a picture of the population as a whole or focus on a particular topic.

Cross-sectional surveys of the population and labour market

The Census of Population

The Census of Population is the most comprehensive source of data on people and households available for Wales. In Wales, it is conducted by the Office for National Statistics (ONS). The census has been held every ten years (except 1941) since 1841, using a questionnaire completed by one person on behalf of the household. It asks a number of questions about each individual living in the household (e.g. age, gender, highest educational qualification, employment status, type of work and travel to work) and about the dwelling (e.g. number of rooms, availability of amenities such as exclusive use of a bathroom and central heating). Derived variables (such as socio-economic group, overcrowding and industry of employment) are created from the answers to census questions.

The census aims for complete enumeration of the population, but non-response (particularly by certain sections of the population) has been an increasing problem. Since 1991, the national statistical agencies have attempted to adjust for census under-enumeration. In 1991, adjustment factors were published for England and Wales and Scotland for certain

variables (age, sex and ethnic group) based on estimates of the undercount. In 2001, the census offices used a complex statistical modelling process to adjust the entire census data set for under-enumeration (based on a comparison of census results with those of a Census Coverage Survey undertaken six weeks later; Brown *et al* 2011.) before it was published. This approach has also been followed for the 2011 census data. The first data from the 2011 Census (rounded population totals by local authority) was published in July 2012 (the long delay is because of the time taken to input the data and then create estimates and quality assure them by comparison with other data sources) and increasingly detailed data sets will be published from November 2012 onwards. In Wales, an extra question is asked concerned with ability to use the Welsh language and additional tables are produced detailing the characteristics of Welsh speakers. The 2011 Census differs from previous censuses in having questions about the first language of all people, citizenship, national identity and more detailed questions on international migration.

The results of the census are published in the form of pre-designed cross-tabulations of each variable against one or two key dimensions (e.g. age of person or household type). Before 1991, the bulk of the census output took the form of printed reports published for local authority districts and counties. From 1961 onwards, an increasing amount of census tables, and all the data for small geographical areas, has been produced in electronic format. From 2001 onwards, electronic media became the primary means for distributing census data. In 2001, much data was distributed on CD/DVD and census data was also made available via the Casweb service run by the UK academic

community (www.census.ac.uk/casweb), via subscription to the SASPAC software package managed by the Greater London Authority (www.saspac.org) and via NOMIS (www.nomisweb.co.uk). Tables of data derived from the census were also made available on the StatsWales web site (<https://statswales.wales.gov.uk/Catalogue>) under particular themes (e.g. equality).

For 2011, new web-based routes to census data will be available. The ONS Neighbourhood Statistics web site (<http://www.neighbourhood.statistics.gov.uk>) and the NOMIS web site (<http://www.nomisweb.co.uk>) will also host census tables as they are released. In addition, both ONS and the providers of access to census data (primarily SASPAC, NOMIS and the ESRC) will provide API interfaces to the data, enabling users greater flexibility in accessing data. The ONS will also make more detailed data available than previously on sub-sets of the population, such as minority ethnic groups, and may publish some data in the form of 'hypercubes' enabling users to tabulate data in more flexible ways than before.

Even though it does not include questions on income, the census is of great importance for the measurement of poverty and socio-economic differentials related to poverty. Derived variables on occupation and socio-economic group are created from the answers to the census questionnaire, which provide a proxy for income. The questions asked in the census cover demography, health differentials, education, the labour market, migration and identity and education.

Figure 1: The 2001 and 2011 Census of Population ethnic group questions

a) 2001	b) 2011
<p>8 What is your ethnic group? Choose ONE section from A to E, then tick the appropriate box to indicate your cultural background.</p> <p>A White</p> <p><input type="checkbox"/> British <input type="checkbox"/> Irish</p> <p><input type="checkbox"/> Any other White background, please write in</p> <p>_____</p> <p>_____</p> <p>B Mixed</p> <p><input type="checkbox"/> White and Black Caribbean</p> <p><input type="checkbox"/> White and Black African</p> <p><input type="checkbox"/> White and Asian</p> <p><input type="checkbox"/> Any other Mixed background please write in</p> <p>_____</p> <p>_____</p> <p>C Asian or Asian British</p> <p><input type="checkbox"/> Indian <input type="checkbox"/> Pakistani</p> <p><input type="checkbox"/> Bangladeshi</p> <p><input type="checkbox"/> Any other Asian background please write in</p> <p>_____</p> <p>_____</p> <p>D Black or Black British</p> <p><input type="checkbox"/> Caribbean <input type="checkbox"/> African</p> <p><input type="checkbox"/> Any other Black background please write in</p> <p>_____</p> <p>_____</p> <p>E Chinese or other ethnic group</p> <p><input type="checkbox"/> Chinese</p> <p><input type="checkbox"/> Any other, please write in</p> <p>_____</p> <p>_____</p>	<p>16 What is your ethnic group? Choose one section from A to E, then tick one box to best describe your ethnic group or background</p> <p>A White</p> <p><input type="checkbox"/> Welsh / English / Scottish / Northern Irish / British</p> <p><input type="checkbox"/> Irish</p> <p><input type="checkbox"/> Gypsy or Irish Traveller</p> <p><input type="checkbox"/> Any other White background, write in</p> <p>_____</p> <p>B Mixed / multiple ethnic groups</p> <p><input type="checkbox"/> White and Black Caribbean</p> <p><input type="checkbox"/> White and Black African</p> <p><input type="checkbox"/> White and Asian</p> <p><input type="checkbox"/> Any other Mixed/multiple ethnic background, write in</p> <p>_____</p> <p>C Asian / Asian British</p> <p><input type="checkbox"/> Indian</p> <p><input type="checkbox"/> Pakistani</p> <p><input type="checkbox"/> Bangladeshi</p> <p><input type="checkbox"/> Chinese</p> <p><input type="checkbox"/> Any other Asian background, write in</p> <p>_____</p> <p>D Black / African / Caribbean / Black British</p> <p><input type="checkbox"/> African</p> <p><input type="checkbox"/> Caribbean</p> <p><input type="checkbox"/> Any other Black / African / Caribbean background, write in</p> <p>_____</p> <p>E Other ethnic group</p> <p><input type="checkbox"/> Arab</p> <p><input type="checkbox"/> Any other ethnic group, write in</p> <p>_____</p>

The ethnic group classification for the Census of Population was changed slightly between the 2001 and 2011 Censuses. Under 'White', 'British' was replaced by 'Welsh/English/Scottish/Northern Irish/ British'. The order in which some ethnic groups appear has also changed. For example, the category 'Chinese' appeared under the 'Chinese or other' heading in 2001, but was under 'Asian/Asian British' in 2011. The 2011 classification added 'Gypsy or Irish Traveller' to the 'White' group and 'Arab' to the 'Other' heading.

The population base for the 2011 Census will change to represent the usually resident population i.e. those resident in the country for at least 12 months (to be consistent with international census practice). It will also collect information on short-term migrants more accurately.

Microdata (i.e. data for an anonymous sample of individual census forms) is also available from the census. This enables all variables and recoded variables from the census to be cross-tabulated in a completely flexible fashion and enables regression analyses to be undertaken. In 1991 and 2001, 'Samples of Anonymised Records' (SAR) were created for 1 per cent of households, and 3 per cent of individuals. The Small Area Microdata Sample is based on a 5 per cent sample and provides data on a more limited range of variables for local authority districts. In Wales, the sample size is typically around five thousand per local authority district, with a maximum sample size of fifteen thousand in Cardiff. The minority ethnic group sample for Wales is just under three thousand. These data sets include a range of variables merged from other data sources (e.g. small area deprivation indicators). Because of the extreme confidentiality risk involved, much geographical and variable detail is removed and the most detailed data sets can only be

accessed via the Secure Data Service. In 2001, the Household SAR did not contain any geographical information.

It is anticipated that microdata will be produced from the 2011 Census. With a total population of 3 million in Wales, a 3 per cent Controlled Access Microdata Sample would contain at least 90 thousand records. This is much larger than other survey data sets for Wales, but would still not permit detailed cross-tabulations for unitary authorities. A data set based on a larger sample but with a less detailed breakdown for the individual variables will be made available for public access.

Labour Force Survey

The ONS Labour Force Survey (LFS) represents the most comprehensive source of information on economic participation in the UK. It is intended to be representative of the total UK population. A quarterly sample survey of households living at private addresses, it also includes those living in NHS accommodation and young people (aged 16–24) in student halls of residence or similar institutions. Its purpose is to provide information on the labour market that can then be used to develop, manage, evaluate and report on labour market policies.

The survey was established when the UK joined the European Community in 1973. It was conducted biennially from 1973 to 1983 and annually from 1984 to 1992 (results referred to the spring quarter of each year). The survey has been conducted quarterly in Wales since 1992. The survey changed from yielding data for four seasonal quarters to four calendar quarters in 2006.

The sample of addresses for the LFS is drawn from the Post Office's Postcode Address File. The sample for each quarter is made up of five 'waves', each containing approximately 60 thousand households across the UK, with a sample for Wales of around three thousand households. Every sampled address is interviewed in five successive quarters, such that any one quarter will include households in one of the five waves. At each address, information is collected on the economic status and activity of all residents aged 16 and over during a specified week in the quarter (termed the reference week).

The LFS includes ethnic group, religion and national identity questions. It also asks about country of birth, nationality and year of entry to the UK. The most recent LFS surveys include questions about planned duration of migration. It is therefore possible to generate data on white minority groups with migrant origins (e.g. East Europeans).

Results from the LFS are presented on the StatsWales website (<http://statswales.wales.gov.uk/index.htm>), which presents a number of tables (in downloadable Excel format), focused upon labour market participation, employment and unemployment. The standard breakdowns provided include gender, but no information is provided by ethnic group or migrant origin.

Microdata (survey responses for anonymised individuals) from the LFS is available from the UK Data Archive (ESDS Government). There are three variants of the LFS corresponding to increasing level of detail on a number of sensitive variables (particularly age, occupation, geography and ethnicity). The most detailed information (including 4-digit occupation and ward) can only

be accessed by Approved Researchers via the Secure Data Service and all outputs must be checked by ONS statisticians. The data set can be browsed using the NESSTAR catalogue (<http://nesstar.esds.ac.uk/webview/index.jsp>). Household data sets are produced twice a year. The geographical breakdown available in these data files is the four nations of the UK and Government Office Regions within England.

LFS data is generally weighted to provide estimates of the population. ONS advises that weighted numbers of less than ten thousand for a given quarter should not be published, because the associated standard error exceeds 20 per cent. When combining quarters, this threshold is reduced because a weighted figure is based on a larger numbers of observations. Thus, for annual averages, the threshold for publication is reduced to a weighted value of six thousand.

Local Area Labour Force Survey (LALFS)

The Local Area Labour Force Survey ran from 1999 to 2006 and comprised a boosted version of the Labour Force Survey, intended to yield data which was statistically robust at the local scale. Aggregate estimates were made available for Local Authority Districts (and the now defunct Training and Enterprise Council areas) for the following variables:

- employment (by age group);
- employees;
- self-employed;

- ILO unemployed (by age group);
- economically active (by age group);
- economically inactive (by age group);
- industry sectors (by Standard Industrial Classification (SIC) code);
- occupations (by Standard Occupational Classification (SOC) code);
- ethnic minorities;
- full-time education;
- qualifications;
- job related training.

Data for local authorities can be accessed via the NOMIS website for a set of pre-designed tables and selected variables. This includes the percentage of the population aged 16 and above and of working age who are from ethnic minorities. Data is provided as rolling averages of four quarters. Data is also available on StatsWales alongside the APS as the LALFS was the predecessor to the APS which started in 2004.

Annual Population Survey (APS)

Despite being a relatively large survey, the LFS is not designed to provide robust information on detailed population sub-groups living within particular areas of the UK. The Annual Population Survey (APS) goes some way to satisfying this need. Data is combined from four successive quarters of the LFS and includes a local booster sample intended to yield a sample of at least 875 economically active individuals in most local authority districts. The Welsh Government has funded a booster sample for Wales since 2001. The April

2009 version of the APS data file includes information on some 36 thousand individuals living in Wales (and 333 thousand people living in the UK as a whole). Even a sample this large is too small to yield robust information on the circumstances of people from small population sub-groups (e.g. some individual ethnic or religious groups) and hence it is necessary to combine data for a number of years to undertake detailed analysis. The consequence is that it is difficult to both analyse the population in detail and investigate trends over time.

The data can be accessed for local authority districts via NOMIS in the form of a series of pre-specified tables. Those relevant to ethnicity and country of birth are:

- Table T04 Economic activity by ethnic group and country of birth;
- Table T05 Economic activity by ethnic group and nationality;
- Table T14a Employment by industry (SIC 2007) and ethnic group;
- Table T15a Employment by occupation (SOC 2010) and ethnic group;
- Table T18 Ethnicity by age;
- Table T21 Adult learning by age/ethnicity/disability;
- Table T24 National identity.

NOMIS suppresses data items for which the relative standard error is greater than 20 per cent. Because of the small sample size and the small percentage of the population from minority ethnic groups, many items of data for minority ethnic groups will be suppressed, even for Wales as a whole. Aggregating data for several years may minimise this problem. In most NOMIS APS tables, the minority population is not broken down to individual ethnic groups.

Microdata from the APS can be ordered from ESDS Government. There are three levels of access, with an 'End User Licence' version containing a restricted range of variables and the coding of variables which could disclose information on individuals is less detailed. Access to more detailed information requires the user to apply for access to the 'Special Licence' version of the data set. This involves becoming an 'Approved Researcher' and adhering to conditions, mainly concerned with preserving the confidentiality of the data. Sensitive variables are suppressed. Access to the most detailed version of the data for small geographical areas can only be obtained via the Secure Data Service, which places greater confidentiality demands upon Approved Researchers.

Integrated Household Survey (IHS)

The rationale behind the Integrated Household Survey (IHS) was to take advantage of the efforts of the ONS in conducting a range of social surveys in order to effectively increase the sample size for a number of key measures of the population and thereby improve the accuracy with which they are measured. Another aim was to provide greater geographical detail than existing surveys could achieve. The IHS is therefore a composite survey combining questions asked on a number of ONS social surveys to produce a dataset of around 100 'core' variables. From the data set for April 2009 to March 2010, the IHS has encompassed approximately 450,000 individuals from interviews undertaken in Great Britain and Northern Ireland.

The IHS includes the following ONS Social Surveys:

- General Lifestyle Survey (GLF);
- Living Cost and Food Survey (LCF);
- English Housing Survey (EHS);
- Labour Force / Annual Population Survey (LFS/APS);
- Life Opportunities Survey (LOS).

The IHS includes two sections: a suite of 'core' IHS questions and individual survey modules 'bolted' onto the core. The core questions have been developed to provide estimates for a number of themes, including:

- economic activity;
- education;
- health and disability;
- identity;
- income.

Details of the survey and the methodology used are provided in the User Guides

(<http://www.ons.gov.uk/ons/guide-method/method-quality/specific/social-and-welfare-methodology/integrated-household-survey/index.html>). It should be noted that, since the IHS started, many of the

surveys that formed the IHS are no longer run, and the achieved sample size for the UK is now much smaller (although for Wales it is broadly the same)

The survey includes questions on ethnic group, religion, country of origin, and date of migration. A series of questions about individual and household income are asked. The survey is the first to include a question on sexual identity. The sample size is similar to the APS for Wales (see below), and thus

there is potential for some disaggregated analysis from the IHS, at least for the larger minority ethnic groups.

Reports based on the survey are published by the ONS. For example, experimental statistics from the survey for 2010–11 can be found at: <http://www.ons.gov.uk/ons/rel/integrated-household-survey/integrated-household-survey/april-2010-to-march-2011/stb---integrated-household-survey-april-2010-to-march-2011.html#tab-Summary-about-the-IHS>.

Data sets for rolling four-quarter periods containing anonymised individual records are available from the UK Data Archive. The IHS does not include all the variables collected in the constituent surveys. There are also two versions of the IHS study made available: one is available under the standard End User Licence agreement, and the other is a Special Licence version. The Special Licence version contains more detailed variables on a range of topics including country of birth, ethnicity and country of residence history. It is not clear whether income data is available in the Special Licence version of the data set.

Sample Size: The IHS for January to December 2010 contains data on 425 thousand individuals in the UK, 39,425 of whom were in Wales. Of this total, 1,041 were from minority ethnic groups. Nearly half of this total was from Asian or Asian British ethnic groups.

The Annual Survey of Hours and Earnings (ASHE)

This has been the prime source of data on earnings patterns by industry and occupation in the UK since it replaced the New Earnings Survey in 2004. It is based on a 1 per cent sample of employee jobs taken from HM Revenue and Customs (HMRC) PAYE records. Information on earnings and hours is obtained from employers and treated confidentially. It does not cover the self-employed or employees not paid during the reference period.

The survey collects details of the earnings, hours worked, job details, home and workplace location of all workers in the sampled workplace. The survey collects information on basic pay, overtime and pensions. The only demographic variables collected about workers are age and gender.

Access to data: Data from the survey is available from the ONS website, NOMIS and the StatsWales website. The data on NOMIS comprises the number of jobs and mean and median wage rates for full- and part-time male and female workers for each local authority district.

Microdata is available via the Secure Data Service run by the UK Data Archive. Since the data is extremely sensitive, users have to become Approved Researchers in order to access ASHE data. A research proposal describing why and how the data will be analysed must be submitted for approval by the ONS. All tables generated are checked by ONS to ensure that confidentiality is preserved. Tables based on small numbers are likely to be suppressed. The main use of the secure data service is to undertake multiple regression analysis, which can involve merging information from other data

sets held within the secure data service (e.g. census data for small areas such as wards) to investigate relationships with earnings variables.

Specialised cross-sectional surveys

The Family Resources Survey (FRS)

The Family Resources Survey (FRS) is the most comprehensive survey of the income, benefits and expenditure of households in the UK. It was introduced in Great Britain in October 1992 to provide the data required by (what is now) the Department for Work and Pensions, because the sample size of the Family Expenditure Survey and General Household Survey was too small to yield the detailed information required. The survey collects information on housing type and tenure and household finances, employment status and details of income, savings and benefits from all sources. It also asks about difficulty in paying utility bills and eligibility for free school meals. It is a continuous household survey with an achieved UK sample size of around 24 thousand households.

The FRS questionnaire contains a group of questions on national identity and ethnicity. These include questions on country of birth, the year and month in which a person born outside the UK first arrived, whether they have lived in the UK continuously since then, the year of most recent arrival, nationality, national identity, ethnic group and religion.

The questionnaire asks for the ethnic group to which respondents consider that they belong (or for proxy responses, the view of the person answering the

questionnaire about the person for whom responses are being provided). The FRS questionnaire uses the standard 2001 Census of Population ethnic group categories.

The FRS has the potential to yield extremely detailed information on the factors underlying poverty and financial exclusion by ethnic group and country of birth. The small size of the FRS sample means that data for several years has to be aggregated for analysis of income within minority ethnic groups in Wales to be undertaken (in fact, the published national report for 2009/10 presented results aggregated over three surveys: 2007/8, 2008/9 and 2009/10 in order to increase the sample size for some tables). Inevitably, this means that the rapid growth of the minority population is obscured and differences between longer established and new migrant communities is obscured.

Access to data: Microdata and documentation for each FRS is deposited with the UK Data Archive (ESDS Government). However, access is only granted to Approved Researchers, who can access Special Licence and Secure Data Service versions of the data set.

Living Costs and Food Survey (LCFS)

This is the current household expenditure survey conducted annually in the United Kingdom. It replaced the Expenditure and Food Survey (EFS) in January 2008 and is collected as a module of the Integrated Household Survey. The EFS had itself replaced the Family Expenditure Survey (FES) and National Food Survey (NFS) in April 2001. The primary use of the LCFS

is to provide information for the Retail Prices Index, National Accounts estimates of household expenditure, analysis of the effect of taxes and benefits, and trends in nutrition. However, the EFS questionnaire includes a core set of questions common to all of the separate modules of the IHS and hence the LCF survey also provides a wider range of economic and social indicators. Fieldwork is conducted by the ONS in Great Britain.

The first part of the LCFS/EFS questionnaire collects information about households. Most questions are asked at a household-level, with the household reference person typically responding on behalf of the household as a whole. The household questionnaire includes questions on a range of subjects including family relationships, ethnicity, employment details and the ownership of household durables. It is also the source of all expenditure information not recorded in the diary element of the survey; principally concerning regular payments typically made by all households and large, infrequently purchased items such as vehicles, package holidays and home improvements.

The income questionnaire follows on immediately from the household questionnaire and collects the key person-level variables used on the survey. The LCF income questionnaire includes sections covering income from employment, benefits and assets. These together form an overview of the total income received by each household, as well as each household member individually. In addition to the household and income questionnaires, each individual aged 16 years and over in the household is asked to keep a diary of daily expenditure for two weeks.

The sample size for the survey was 12,178 for the UK in 2010 and the average number of households surveyed was 5,640 during 2008 to 2010. The LCFS includes a question on ethnicity, which uses the ONS ethnic group classification (including a write-in option for those whose ethnicity is 'Other'). 10,783 of the 11,823 respondents to this question in 2010 were from the 'White' ethnic group.

Some variables in the data are suppressed for reasons of confidentiality. These include geographical variables with detail below government office region and urban/rural area indicators. Sample sizes for Wales would be too small for reliable information for individual minority ethnic groups to be derived.

Access to data: Data from the survey is published in ONS reports and microdata is available from the UK Data Archive (ESDS Government).

Longitudinal and cohort surveys

This section reviews the surveys available for Wales which incorporate a longitudinal design or which follow a birth cohort over time.

1970 British Cohort Survey (BCS70)

The 1970 British Cohort Study (BCS70) has traced the lives of children born in a particular week in 1970. There have been seven full data collection exercises in order to monitor the cohort members' health, education, social

and economic circumstances. These took place when respondents were aged 5, in 1975, aged 10, in 1980, aged 16, in 1986, aged 26, in 1996, aged 30, in 1999–2000, and aged 34, in 2004–2005. The initial aims of the study were to examine the social and biological characteristics of the mother in relation to neonatal morbidity, and to compare the results with those of the National Child Development Study (NCDS), which commenced in 1958. The scope of the BCS has broadened over time, to encompass physical and educational development at the age of 5, physical, educational and social development at the ages of 10 and 16, and physical, educational, social and economic development at 26 years and beyond.

The 'Thirty-eight year follow-up' in 2008–2009 was conducted via telephone interviews (CATI). The survey explored the following areas:

- housing;
- relationship history;
- births and other pregnancies;
- periods of lone parenthood;
- children and the wider family (social relationships and support);
- family income;
- employment status/employment history;
- academic education;
- general health;
- smoking.

BCS70 potentially yields extraordinarily rich lifetime data on relationships, family history and labour market history. It includes detailed family income data and a record of housing and housing mobility.

Access to data: Data from the survey is available via the UK Data Archive (in both End User and Special Licence versions) and the Centre for Longitudinal Studies at the Institute of Education, University of London, provides support for the BCS. See: <http://cls.ioe.ac.uk/MicrositeHome.aspx?siteid=46&siteid=46&siteid=46> [sh%20Cohort%20Study](http://cls.ioe.ac.uk/MicrositeHome.aspx?siteid=46&siteid=46&siteid=46). ESDS Longitudinal also provides a range of resources for using the data sets (<http://www.esds.ac.uk/longitudinal/access/bcs70/datasets.asp>). The survey data sets can be explored using Nesstar at: <http://nesstar.esds.ac.uk/webview/index.jsp>.

Infant Feeding Survey

The Infant Feeding Survey has been conducted every five years since 1975. It provides estimates of the incidence, prevalence and duration of breastfeeding and other feeding practices adopted by mothers in the UK, over the first 9 months or so of their baby's life. It asks about the sources of advice used by mothers and the practices followed by their friends. The survey also collects information on the smoking and drinking behaviour of mothers before, during and after pregnancy.

Methodology: The survey uses a longitudinal panel design, with three stages of data collection being carried out over a 9–12 month period in order to capture feeding practices at different ages. Stage 1 is carried out when the babies are approximately 6–10 weeks old; Stage 2, when they are approximately 4–6 months old; and Stage 3 when they are approximately 8–10 months old. The 2010 survey was based on an initial representative sample of 30,188 mothers of babies born in the UK between mid-August and late November 2010. In Wales and Northern Ireland, all births during the sample period (Wales: 9 August – 17 October 2010 and Northern Ireland: 9 August – 24 October 2010) were included in the sample. In England and Scotland, births in the most deprived neighbourhoods (the highest quintile of each country's Index of Multiple Deprivation) were over-sampled.

At Stage 1, a total of 15,724 mothers returned the questionnaire, representing a response rate of 52 per cent. Headline results from the 2010 survey were published in June 2011 (<http://www.ic.nhs.uk/statistics-and-data-collections/health-and-lifestyles-related-surveys/infant-feeding-survey/infant-feeding-survey-2010-early-results>). These focus on early results from Stage 1 of the survey only and cover two key topics, the initial incidence of breastfeeding and smoking during pregnancy. The report for the completed survey is expected to be published in the summer of 2012.

The survey does not collect data on income, but it does include questions on the demographic and socio-economic background of the mother. This includes age and family status, employment status and plans to return to work and ethnic group (using the 2001 Census ethnic group classification). The survey yields UK estimates, and the 2005 and 2010 surveys have also

provided estimates for the individual countries of the UK. In 2005, 12,290 mothers were surveyed across the UK, out of which 2,135 were in Wales, 2,194 in Scotland, and 1,886 in Northern Ireland.

Millennium Cohort Study (MCS)

Also known as the 'Child of the New Century' study, this is a longitudinal survey of children born in the year 2000, managed by the Centre for Longitudinal Studies at the Institute of Education, University of London. It provides an extremely complete picture of the early lives of children, set within the context of their changing family circumstances and the characteristics of the area in which they live.

The sample comprises nearly 19,000 babies born between 1 September 2000 and 31 August 2001 in England and Wales, and between 22 November 2000 and 11 January 2002 in Scotland and Northern Ireland. Families living in areas of child poverty, in the smaller countries of the UK and in areas with high ethnic minority populations in England were over-sampled.

The first sweep (MCS1: 2001–3) interviewed both mothers and (where resident) fathers (or father-figures) of infants included in the sample when the babies were 9 months old, and the second sweep (MCS2: 2003–5) was carried out with the same respondents when the children were 3 years of age. The third sweep (MCS3) was conducted in 2006, when the children were aged 5 years, and the fourth sweep (MCS4) in 2008, when they were 7 years old. In Wales, surveys of teachers of the children in the survey were

undertaken when the children were aged 5 (in 2006) and 7 (in 2008), covering the child's behaviour and achievements.

The mother or main carer interviews covered ethnicity and language, the baby's father, lone parenthood, pregnancy, labour and delivery, baby's health and development, childcare, grandparents, friends and social support, parental health, education and training, employment and earnings, housing, community and local services, time with and without the baby and other interests. The father's or main carer's partner interviews covered similar issues, as well as father's involvement with the baby. The mother and father interviews also included a self-completion element covering baby's temperament and behaviour, relationship with partner, previous relationships, domestic tasks, previous pregnancies, mental health, attitudes to relationships, parenting, work, etc.

Either parent or carer could answer the household module questions, which covered household composition and relationships, children, employment status, caring responsibilities for the baby born in 2000 and language spoken at home.

The survey collects detailed information on sources of household/family income (including gross, net usual pay and benefits) and child support/maintenance. In the first wave of the survey, 2,760 interviews (14.9 per cent of the total) were undertaken in Wales. The ethnic group question was based on the 2001 Census question, the main difference being that there were four white categories, with 'White British' being divided into 'White-Welsh', 'White British' and 'Any Other White').

The structure of the data sets is explained in a guide to the MCS, available from:

http://www.esds.ac.uk/doc/4683%5Cmrdoc%5Cpdf%5Cmcs_guide_to_the_datasets_v5.pdf

Access to data: The data is made available by the UK Data Archive. Datasets can be browsed via the ESDS Nesstar software tool (nesstar.esds.ac.uk/webview/index.jsp). The MCS data currently available via ESDS (End User Licence) are surveys 1 to 4; a sub-study of mothers who received assisted fertility treatment; birth registration/maternity hospital episode data; a 3rd wave Teacher Survey and Foundation Stage Profile as well as a 4th wave Teacher Survey dataset. More detailed data available on a Special Licence basis includes details of the hospital of birth data and detailed geographical identifiers data. In November 2011, a 'First to Fourth Surveys, Linked Education Administrative Dataset' was made available via the Secure Data Service. These data sets are subject to strict conditions of access, concerning confidentiality and the request for access must be made by an Approved Researcher.

British Household Panel Study (BHPS) / Understanding Society (US)

Understanding Society (US) has now replaced the British Household Panel Study (BHPS), the two together comprising a longitudinal study conducted over more than 20 years. Households included in the BHPS have been incorporated into the US sample. Both have been funded by the ESRC and are managed by the Institute for Social and Economic Research at the University of Essex.

The BHPS was an annual survey which ran from 1991 to 2010. It was based on a nationally representative sample of around 5,500 households initially recruited in 1991, containing a total of approximately 10,000 interviewed individuals. A stratified clustered sample of addresses was drawn from the Postcode Address File, and all residents present at those addresses at the first wave of the survey were designated as panel members. These same individuals were re-interviewed in each successive year. If they left their original households to form new households, they were followed and all adult members of the new households were also interviewed. Similarly, new members joining sample households became eligible for interview. Children were interviewed when they reached the age of 16. From 1994, children aged 11–15 also completed a short interview.

The BHPS was initially conducted only in England, with Wales and Scotland following in 1994. The initial sample size for Wales was 509 in the main sample, with 35 in the Youth Panel. Boost samples were added for Scotland and Wales in 1999 of 1,500 households in each country. In 2001 a sample of 2,000 households was added in Northern Ireland. By the end of the BHPS in

2010, the total sample size was around 10,000 households across the UK. 18 waves of data from the BHPS are available via the UK Data Archive. The Welsh sample was 2,500 in 2008, with a Welsh Youth Panel consisting of 240 individuals.

The BHPS asked detailed questions about household structure, labour market participation, income and expenditure and geographical mobility (including motivation behind residential moves). It included questions about national identity, ethnicity and country of birth in Great Britain.

Data from the BHPS is available from the UK Data Archive. Less detailed information is available in the 'End User Licence' data sets. More confidential information has to be accessed via Special Licence arrangements or via the Secure Data Service (the latter applies to geographically detailed information in particular).

US is a much larger longitudinal study drawing upon the experience of the BHPS, but conducted on a larger scale and with much greater ambition. The sample size is 40 thousand households, incorporating the existing BHPS and Northern Ireland Household Panel Study (NIHPS) samples. The study has four sample components: the General Population component, the Innovation Panel, a boost sample of ethnic minority group members, and participants in the former British Household Panel Survey (BHPS). Continuous fieldwork for the survey commenced in Great Britain in 2009 (Wave 1).

The ethnic minority module aims to provide a sample of around 1000 adult respondents from the Indian, Pakistani, Bangladeshi, Caribbean and African

ethnic groups. Others are also included, and the large sample size across the study enables coverage of more dispersed and new migrant groups.

The first wave of US included data for 2,378 people in Wales, of which 135 were from minority ethnic groups (the BHPS sample was integrated into US at Wave 2). All residents in a household aged 16 or more were asked questions around: Initial conditions: e.g., place of birth, citizenship, qualifications, family background; ethnicity and national identity, language, religion and harassment; history of marriage and cohabitation, and fertility; health, disability and caring; employment, non-employment and discrimination; childcare, parents and children, and family networks; unearned income and state benefits; household finances; political affiliation and values; and environmental behaviour.

Access to data: Data from US is available from the UK Data Archive via: <http://www.esds.ac.uk/findingData/snDescription.asp?sn=6614>

Surveys that focus on Wales

Living in Wales (LIW) Survey / Welsh Household and Dwelling Survey

The Living in Wales (LIW) survey (also referred to as the Welsh Household and Dwelling Survey) collected information on housing condition and household structure. It was conducted between 2004 and 2008. It replaced the Welsh House Condition Survey (WHCS), which was last conducted in 1997 (household survey) and 1998 (property survey). The LIW survey

consisted of separate but linked household and property surveys, with the household survey being undertaken first.

The Household Survey was completed annually from 2004 and was conducted as a face-to-face interview with the household reference person (HRP), their spouse or another appropriate adult. It aimed to provide additional information to complement the property survey, but also to provide information about the community, the use of the Welsh language, the health of the members of the household, the values and opinions of the respondent as well as demographic characteristics. The property survey was conducted in 2004 and 2008 (it is not currently available from the UK Data Archive). The property survey comprised an internal and external assessment of the property which was completed by a qualified surveyor.

The topics covered by the survey were:

- household composition, ethnicity, religion, employment;
- disability and long-term limiting illness;
- travel and transport;
- quality of life;
- internet access and use;
- discrimination harassment and victimisation;
- housing history;
- tenure/area features;
- economic status;
- income, benefits and savings;
- provision of services/amenities;

- use of, and ability with, Welsh language;
- volunteering;
- environment;
- views on public services;
- overcrowding;
- fire and fire safety;
- heating and energy measures;
- unfitness and repair;
- housing adaptations;
- education and qualifications.

The 2008 household questionnaire included questions on the structure of the household, ethnic group (using the 2001 Census ethnic group classification), religion, national identity, labour market activity, volunteering, characteristics of the dwelling, housing tenure, access to services, housing costs, benefits and household income. In 2008, the survey was based on 7,721 successful interviews.

The results of the survey are published on a Welsh Government web page. Anonymised individual survey data has been deposited with the UK Data Archive.

National Survey for Wales (NSW)

The National Survey for Wales (NSW) is the replacement for the Living in Wales (LIW) Survey series (also known as the Welsh Household and Dwelling Survey) which ran from 2004 to 2008.

The topics covered by the survey are:

- overall satisfaction with the transport, health and education systems and the Welsh Government;
- use of the Welsh language;
- community cohesion (views about local area, perceptions of safety, discrimination);
- access to and satisfaction with GP and hospital services;
- experience of and attitudes towards schools and education;
- local authority services;
- wellbeing (including access to local amenities and questions around financial inclusion);
- use of the internet;
- smoke alarms.

The survey includes questions about the structure of the household, age and gender, country of birth, language, national identity, sexual orientation, religion, qualifications and labour market experience.

The results of the survey will be published on the Welsh Government website, from September 2012 onwards. The first publication will be of headline results, followed by detailed national findings and headline findings by local

authority district in summer 2013. From 2014, detailed annual results and local authority-level findings based on two years of survey data will be published each summer. The results of the survey are weighted to compensate for unequal selection probabilities and differential non-response, and hence the age and sex distribution of the final dataset matches that of the Welsh population.

Access to data: The survey datasets will be made available via the UK Data Archive from December 2012. Data from a pilot survey undertaken in 2009 is currently available from the UK Data Archive. This experimented with the survey methodology, notably a switch from a household to an individual focus for the survey.

Sampling: The survey is based on a sample of approximately 25,000 addresses per year, drawn at random from the Royal Mail Postcode Address File, with the aim of achieving a successful sample size of 14,500 households across Wales (660 per Local Authority). If more than one person lives at the address, one person aged 16 or over is randomly selected to take part in the study. If a household refuses to take part, the address is not replaced and is dropped from the sample. In the January–March 2012 quarter, a total of over 3,550 interviews were completed, with a response rate of 66 per cent.

Welsh Health Survey (WHS)

The WHS collects information about the health of people living in Wales, the way they use health services, and the factors that can affect their health.

Introduced in 2003/4, it is the successor to a previous WHS (conducted in 1995 and 1998), which itself replaced the Health in Wales Survey (conducted in 1985, 1988, 1990, 1993 and 1996).

The WHS is designed to:

- provide national estimates of health and health-related lifestyle;
- provide data for setting and monitoring targets and indicators;
- examine differences between population sub-groups (such as sex, age, social class) and local areas;
- provide evidence to inform and monitor policies and strategies for promoting better health;
- provide local health board- and local authority-level information for the development of joint local health, social care and wellbeing strategies.

The main topics covered for adults are: health service use, health status, medicines, illnesses and other health problems, and health-related lifestyle (including smoking, alcohol, diet and exercise). For children, the survey covers health status, health service use, accidents, illnesses and other health problems, eating habits and physical activity.

In addition to questions related to health, the survey includes questions on demography (age and gender), ethnic group, national identity, qualifications, housing, car ownership and socio-demographic group. However, the sample size is too small to yield robust information on ethnic group or national identity (or housing and car ownership). The survey is weighted to match the age and sex distribution of the population of Wales.

Sample design: The WHS is based on a representative sample of people living in private households in Wales. Information is collected at two levels: household (through a short interview) and individual (through self-completion questionnaires). First, a short interview with a responsible adult in the household is conducted to enumerate the household, and to record brief demographic information for each household member and the socio-economic classification of the household reference person. Self-completion questionnaires are then left for completion by each adult aged 16 years and over in the household, for later collection by the interviewer. In households with children, a maximum of two children are randomly selected for inclusion in the survey. Since 2007, three age-specific questionnaires for children have been used; 0–3, 4–12 (both completed by parents) and 13–15 (self-completion). In addition, height and weight measurements are taken by the interviewer for selected children aged between 2 and 15 years. The survey relies on a self-completion questionnaire. The results, therefore, reflect people's own understanding of their health rather than a clinical assessment of their medical condition, and their own interpretation of the health services they have used. In 2011, the survey was based on a sample of 14,775 addresses, with an eventual sample size of 10,319 households. A household interview was obtained with 78 per cent of eligible households in the sample. Self-completion questionnaires were obtained for 83 per cent of adults (16 thousand) and 79 per cent of selected children (three thousand) in participating households.

Access to data: The results of the survey are presented in a published report. More detailed age and sex tabulations are made available on the Welsh

Health Survey theme page of the Statistics for Wales website and on StatsWales. Key findings for local authorities and local health boards (averaged over two years) are published in a bulletin. An anonymised version of the main dataset is deposited with the UK Data Archive annually.

The Wales Life and Times Studies (WLTS)

The WLT surveys are conducted by the Centre for Research into Elections and Social Trends (CREST) in collaboration with the Institute of Welsh Politics at the University of Wales, Aberystwyth. It is the Welsh component of a survey programme funded by the Economic and Social Research Council's (ESRC) Devolution and Constitutional Change Programme. Surveys were conducted in 1999, 2001, 2003 and 2007 following elections to the Welsh National Assembly.

The surveys are concerned with newspaper readership and internet usage, party identification and political views, attitudes towards public services, voting in the Welsh National Assembly election and national and class identity. Information was also collected on health, housing tenure, Welsh speaking, religion, ethnicity, national identity, education, economic activity, job details, and income. The income questions have become more detailed over time, with the 2007 survey asking about all sources of household income. There was also a self-completion questionnaire which included questions on public services, voting and proportional representation.

Sample size: The survey yielded response from 522 individuals in 1999, 1,085 in 2001, 988 in 2003 and 884 in 2007. The question on ethnic group was asked in 1999, 2001 and 2007 (using the most recent Census of Population ethnic group classification) but the number of respondents from minority ethnic groups was small – 10 in 1999, 12 in 2001 and 18 in 2007.

Access to data: Anonymised individual responses to the survey have been deposited with the UK Data Archive. Data is available for the 1999, 2001, 2003 and 2007 surveys.

Omnibus surveys

There are three regular omnibus surveys undertaken by market research companies in Wales. Their topic focus is on public opinion and behaviour. Beaufort Research Ltd. undertake the Welsh Omnibus Survey quarterly (in March, June, September and November of each year) using face-to-face interviews and CAPI (Computer Assisted Personal Interviewing) data collection software in 68 locations across Wales. It is based upon a representative quota sample (stratified by age, gender, region and social class) of at least 1,000 Welsh residents aged 16 or above. This company also undertakes the Omnibus Survey of Welsh Speakers. This is based on face-to-face interviews (in March and October) with a representative quota sample of at least 500 Welsh residents aged 16 or above who can speak Welsh. Interviews can be conducted in Welsh. Opinion Research Services (ORS) undertake the Welsh Monthly Omnibus Survey. This focuses on monitoring

change in public opinion, using 1000 telephone interviews (conducted in either Welsh or English) of a representative sample of adults living in Wales.

Data from all three surveys is available to subscribers to the survey only. It is possible to commission questions to be included in the survey. It is not clear whether these surveys include questions on ethnicity, country of birth and migration.

Administrative and related data sets for Wales containing information relevant to ethnicity and poverty

This chapter changes the focus of the report to consider a number of regularly collected sources of official statistics which include information relevant to the study of minority ethnic and migrant groups in Wales.

Table 2 summarises these sources and outlines their broad characteristics. The ONS Longitudinal Study is included in this section because it includes linkages to administrative sources – notably registers of births and deaths, the cancer registration system and the NHS Central Register (which is the source of information on international migrants). The other sources differ in being regularly produced as an outcome of administrative systems.

Table 2: Summary of administrative data sources relevant to ethnicity and poverty in Wales

Data set	Type	Contains ethnicity information	Contains migrant information	Contains information on religion	Poverty-relevant	Current
ONS Longitudinal Study	1% sample of census linked to administrative data	Yes	Yes	Yes	Yes	Yes
Wales Schools Census	Administrative	Yes	No	Yes	Yes	Yes
Unemployment claimant count	Administrative	Yes	No	No	Yes	Yes
Annual Civil Service employment survey	Administrative	Yes	Yes	No	Yes	Yes

The ONS Longitudinal Study

Census longitudinal studies provide powerful resources for monitoring socio-economic change through time by linking the details of sample members taken from successive Censuses of Population to demographic events affecting these individuals in the years between censuses. The first Longitudinal Study in the UK was established for England and Wales following the 1971 Census of Population, in response to the need for better information on fertility patterns and occupational mortality. A 1 per cent sample of individuals born on one of four dates and living in England and Wales was drawn from the 1971 Census. These individuals were then traced in the National Health Service Central Register, in order to identify their birth, death and cancer records. Data on individuals living in the same household as the study member was added from the census. These individuals were traced in the 1981, 1991 and 2001 Censuses and will be traced in the 2011 Census. Study members who die are replaced by new individuals with one of the sample birth dates. Births, deaths, marriages, cancer registrations and international migration affecting study members between the census years are recorded.

With c. 500 thousand people in the sample, the ONS Longitudinal Study has proved to be an exceptionally powerful tool for investigating socio-economic trajectories for individuals and households over a 30 year time period. It has had particular value in the study of health conditions. Because it was based on a probability sample taken when the minority

population of England and Wales was small, the number of observations for minority ethnic groups and international migrants is relatively small, limiting the analysis which can be undertaken. This problem is magnified for Wales, in which the total sample size is around 25 thousand.

Access to data: The Longitudinal Study is managed by the ESRC-funded Celsius (Centre for Longitudinal Study Information and User Support) research centre, now based at University College London (<http://celsius.lshtm.ac.uk/>). Users are required to produce a research proposal which is approved by the support team. Access is via a safe setting and users are advised by specialist researchers. Outputs have to be approved by the ONS, in order to preserve confidentiality.

Annual School Census / Pupil Level Annual Schools Census (PLASC)

All maintained sector primary, secondary, nursery and special schools in Wales provide electronic pupil and school level data to the Welsh Assembly Government in January of each year. The PLASC data contains the details of each individual pupil, including age, gender, ethnicity, national identity, eligibility for free school meals, special education needs status, Foundation and Key Stage assessment, GCSE and A level or equivalent results and home postcode. The analysis of pupil characteristics provides an indication of the changing size of an ethnic group, its relative disadvantage and concentration into geographically

disadvantaged locations. A two-tier ethnicity classification is used. The top tier ('main codes') is equivalent to the ONS ethnic group classification, with the addition of 'Traveller' and 'Gypsy/Gypsy Roma' to the 'White' categories. The lower tier ('extended codes') includes a wider number of groups than the census ethnic group classification, enabling the emergence of new ethnic groups to be identified. The Schools Census also collects information on individual examination performance and destinations of school leavers. Thus, the data set enables trends in educational achievement/disadvantage and access to the labour market to be monitored by ethnic group. Further details of the PLASC can be found at the education section of the Welsh Government web site: <http://wales.gov.uk/topics/educationandskills/schoolshome/schooldata/ims/datacollections/pupillevelannualschoolcensus/?lang=en> and from the Administrative Data Liaison Service web site: <http://www.adls.ac.uk/welsh-government/welsh-pupil-level-annual-schools-census-and-pupil-attainment-dataset/?detail>.

Access to data: Statistics derived from PLASC are made available from the <http://wales.gov.uk/topics/statistics/theme/schools/> for Wales (<http://wales.gov.uk/topics/statistics/theme/schools/;jsessionid=056137DF1D87E91BCF83A29E857405A8?lang=en>) and StatsWales websites. Requests to access raw PLASC or Pupil Attainment data should be made to the Welsh Government School Statistics team (school.stats@wales.gsi.gov.uk or 029 2082 5060), providing details of the

research and the information required. The ultimate decision on whether or not to release the data rests with the Chief Statistician of the School Statistics team. All data releases are subject to a data access agreement, and data is generally provided within 10 to 15 working days. However, the PLASC and Pupil Attainment datasets are highly sensitive. Hence a strong business case must be made in order to be granted access (particularly where potentially disclosive data such as home postcode is requested).

Claimant Count

The claimant count is measured by the JobCentre Plus administrative system. The count is of the number of people claiming unemployment-related benefits (since October 1996 this has been Jobseeker's Allowance). Those who are not entitled to/do not claim such benefits are not included in the count, and hence the ILO measure of unemployment (self-reported unemployment, as used by the Labour Force Survey) is much higher. Since mid-2005, JobCentre Plus has classified Job Seeker's Allowance claimants by (self-reported) ethnic group, using the 2001 Census ethnic group classification.

While trends in the numbers on the claimant count can be analysed by ethnic group, there is no information on the proportion of people from minority ethnic groups who are claiming unemployment benefit. Nevertheless, comparing trends in unemployment is a useful way of

identifying relative disadvantage. It is possible to calculate the number unemployed as a percentage of the number of people aged 16–64 for each ethnic group, using the ONS experimental estimates of the population by ethnic group. However, these estimates are 2–3 years out of date.

Access to data: The numbers currently unemployed, joining and leaving the claimant count are published for local authority districts and electoral wards/communities on the NOMIS website (www.nomisweb.co.uk).

Annual Civil Service Employment Survey

The Annual Civil Service Employment Survey (ACSES) started in 2007. Statistics are published on the age, gender, ethnicity, nationality and disability status of the Civil Service workforce. The survey covers employees with an employment contract who are being paid by the organisation. The survey yields information on the job grade and pay level of civil servants and hence enables the earnings distribution and occupational distribution of individual ethnic groups (and UK national/non-nationals) to be compared. The survey also collects information on joiners and leavers and hence retention rates can be compared by ethnic group. However, small numbers limit the detail of analysis which can be undertaken for individual ethnic groups. Numbers are rounded to the nearest 5, and cross-tabulations based on small numbers are suppressed.

Access to data: The data is available via the NOMIS website, where each variable can be cross-tabulated against each other. The results are also published in the ONS Civil Service Statistics Statistical Bulletin and associated data tables (<http://www.ons.gov.uk/ons/rel/pse/civil-service-statistics/2012/stb-civil-service-statistics-2012.html>).

Assessment of the availability of data on minority ethnic and migrant groups and poverty in Wales

A number of surveys have been identified in this report which can yield information on ethnic minorities in Wales and their experience of poverty. Though the number of Wales-specific statistical surveys is not large, Wales is a component part of the main UK- and Great Britain-wide surveys run by the ONS. Published reports from these surveys often contain little detail for Wales, but the availability of microdata makes it possible to produce bespoke tables for sub-groups of the population. Unfortunately, if the sample size is proportionate to the UK population, the resulting sample size can be too small for analysis by ethnic group to be undertaken. The Labour Force Survey is the largest regular general population survey, but the sample size is too small to permit data for local authorities within Wales to be generated. With data collected every quarter, it is possible to create a larger sample by aggregating successive surveys. Additionally, the Annual Population Survey is a boosted version of the LFS, enabling local authority data to be generated for key indicators. However, even with this boost, it is necessary to aggregate APS data for a series of years in order to generate robust information for minority ethnic groups at the local scale, and it may be necessary to group smaller local authority districts together.

Wales-specific surveys tend to have sample sizes which are larger than those of national surveys, but these are probability samples of the entire population and do not include sample boosts for minority ethnic groups. Hence, their ability to yield reliable data for minority ethnic groups is limited by their small sample size. The geographical concentration of minority ethnic groups into the cities and larger towns of Wales also means that there will be too few observations for robust estimates of minority characteristics outside these areas to be made. Another problem is that Wales-specific surveys usually do not include questions on country of birth and international migration, meaning that they are not able to identify the characteristics of new migrant groups (such as eastern Europeans).

The most reliable indicators of changing population come from administrative sources such as the Schools Census and unemployment claimant count. However, these provide a partial picture of the population. The Census of Population provides the most comprehensive data on the population by ethnic group and local area. Publication of the results of the 2011 Census offers the potential for much more detailed analysis of the living conditions of minority groups and new minorities in Wales.

References

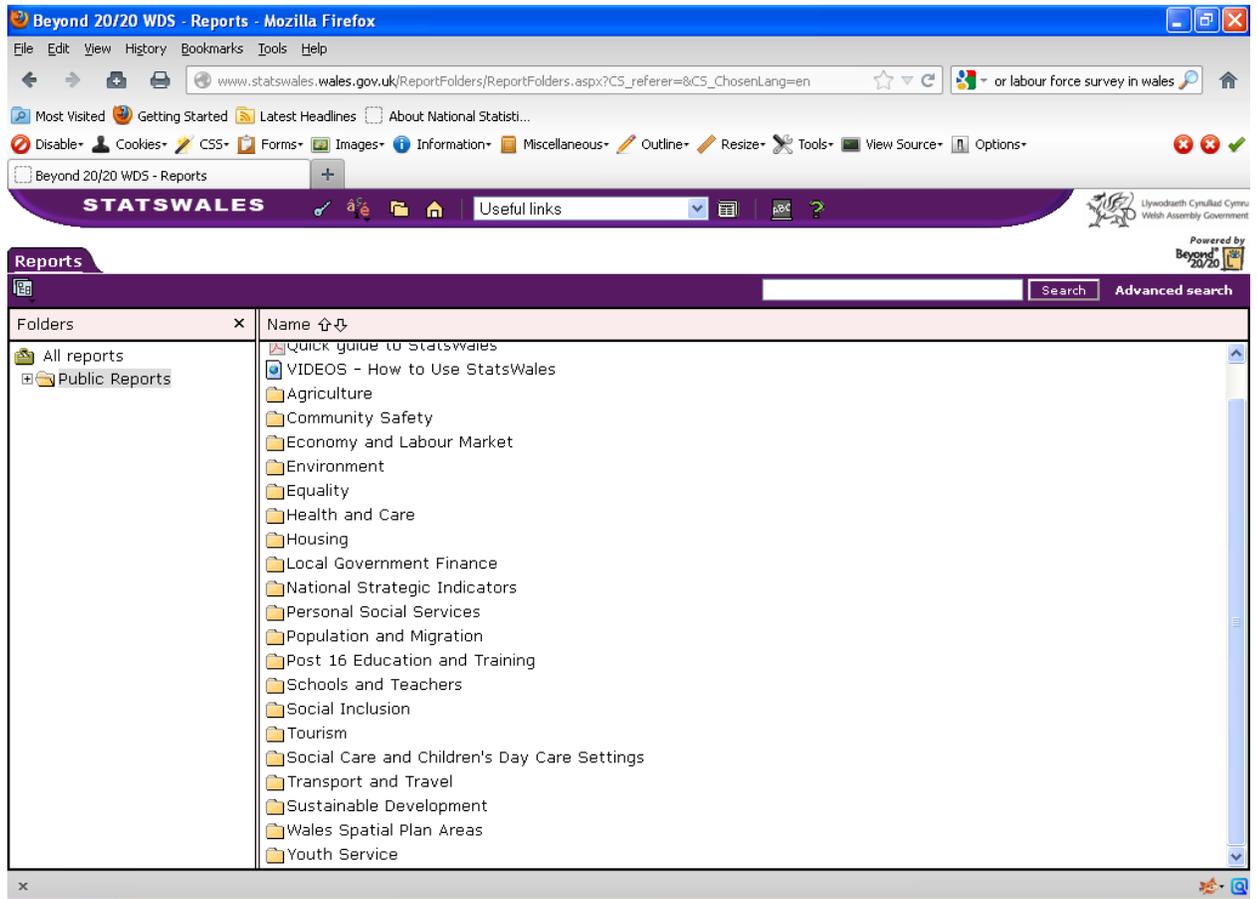
Brown, J., Abbott, O. and Smith, P. (2011) 'Design of the 2001 and 2011 Census Coverage Surveys for England and Wales' *J. R. Statist. Soc. A*, 174 (4), pp. 881–906

Appendix: Websites providing access to data

Wales websites

StatsWales

This is a website run by the Statistical Directorate of the Welsh Assembly Government. It presents statistical data sets for a range of key social, economic and environmental topics, illustrated in the screen shot below:



Clicking on a 'folder' lists the data sets available, presented as reports which can be customised by the user to present the information in different ways. The website includes both statistics generated by departments of

the Welsh Government and data for Wales taken from UK- or GB-wide surveys (e.g. the Annual Population Survey).

UK websites

Office for National Statistics (ONS) website

This provides access to information about official statistics products in the UK. Methodological reports, reports on data and datasets derived from analysis of surveys and administrative data are available. Northern Ireland-specific data sources are not available via this website. Data for Northern Ireland is available in aggregate where the province is covered by the survey if sample sizes are large enough for statistically robust data to be released at the geographical level of the nations of the UK and regions of England (and sub-national data may also be available).

ESDS (Economic and Social Data Service) Government website

This is the prime repository of official survey data deposited with the UK Data Archive. It holds time series for all the major government surveys, and contains metadata, questionnaires and codebooks for each individual survey. There is a Nesstar viewer which allows users to identify which datasets are available and the variables which they contain (<http://nesstar.esds.ac.uk/webview/index.jsp>).

In order to order data from ESDS Government, users have to register with the UK Data Archive (<http://www.data-archive.ac.uk/>). People from the academic and public/voluntary sectors can use the service for free, but commercial users have to pay. Users have to register a research project to order data and must conform with the conditions of access. There are three routes to accessing data – via the Nesstar website (for basic cross-

tabulations and frequency counts), downloading 'End User Licence' data sets (these have minimal constraints on usage but many variables are suppressed) and 'Special Licence' data sets. The latter are the most detailed, but users have to become an ONS 'Approved Researcher' in order to access them. The more detailed and disclosive the data involved, the more onerous the conditions which apply to accessing the data and publishing results.

NOMIS (National On-line Manpower Information System)

The NOMIS website (<http://www.nomisweb.co.uk/>) is run by Durham University on behalf of the ONS. It provides access to labour market data sets (e.g. claimant unemployment and employment data), and related information on the population. This includes the Annual Population Survey and Census of Population data. Wales is represented in most of the datasets held by the website. Access is freely available to most of the data sets (the main exception being employment data, which requires a payment). The website provides extremely flexible access to data and customised queries can be created. Data is downloadable in Excel, .CSV or database formats. There is also an API interface to the NOMIS databases.

Data.Gov.uk (www.data.gov.uk)

This website promotes the UK government's Open Data initiative, where data sets created by government departments are made freely available for download and re-use (under the terms of the 'UK Open Government Licence'). The service is predominantly focused on the data provided by government departments serving England. However, many Welsh Government statistical reports and data sets appearing on the StatWales

website are included. The list of data sets can be out of date and the links are not always updated when a website is revised.

This paper was commissioned as part of the JRF programme on [poverty and ethnicity](#), which aims to understand the underlying reasons for variations in low income and deprivation among different ethnic groups in the UK and the problems caused.

The Joseph Rowntree Foundation has funded this research paper as part of its programme of research and innovative development projects, which it hopes will be of value to policy-makers, practitioners and service users. The facts presented and views expressed in this paper are, however, those of the author and not necessarily those of JRF.

Joseph Rowntree Foundation
The Homestead
40 Water End
York YO30 6WP
www.jrf.org.uk

This paper, or any other JRF publication, can be downloaded free from the JRF website (www.jrf.org.uk/publications/).

© University of Warwick

First published 2013 by the Joseph Rowntree Foundation
All rights reserved. Reproduction of this report by photocopying or electronic means for non-commercial purposes is permitted.
Otherwise, no part of this report may be reproduced, adapted, stored in a retrieval system or transmitted by any means, electronic, mechanical, photocopying, or otherwise without the prior written permission of the Joseph Rowntree Foundation.

ISBN: 9781859359723 (pdf)

Ref: 2884

Contact:
Helen Barnard
helen.barnard@jrf.org.uk