

Cambridge Centre
for Housing &
Planning Research

Measuring Housing Affordability: A Review of Data Sources

Measuring Housing Affordability: A Review of Data Sources

Christine Whitehead

Sarah Monk

Anna Clarke

Alan Holmans

Sanna Markkanen

December 2008

Cambridge Centre for Housing and Planning Research
Department of Land Economy
University of Cambridge
19 Silver Street
Cambridge
CB3 9EP

Tel: 01223 337118

Fax: 01223 330863

www.cchpr.landecon-cam.ac.uk

Contents

Section 1	Introduction	4
Section 2	Measuring affordability	5
Section 3	Understanding affordability	11
Section 4	Modelling affordability	14
Section 5	Key data sources	25
Section 6	Gaps in data sources	32
Section 7	Conclusions and recommendations	46
Annex 1	Key data sources	50
Annex 2	Overview of data sources by topic	71
Annex 3	Glossary	87
Annex 4	List of abbreviations	93
Annex 5	Private rented sector dwelling stock and buy-to-let loans	96
References		97

Section 1: Introduction

The purpose of this publication is to provide advice to the regions, and others working in this area, on the quality and availability of data relating to housing affordability. It is intended to be a resource, particularly for those with less experience in understanding the housing market as well as those making decisions about planning for housing.

Public Service Agreement 20 (PSA20) commits the Government to ‘increase long term housing supply and affordability’¹ and recognises the lead role of regional and local planning authorities in delivering 240,000 net additional homes per year, of which 70,000 should be affordable homes. In addition, Planning Policy Statement 3 (PPS3) on housing sets a new obligation on authorities to establish the impact of housing supply plans on affordability prospects for their communities over the medium and long term.

This publication brings together in one place information on the sources of standard data that planning authorities can use to understand their housing markets and specifically indicators of affordability. Where possible, it identifies which data sources are the most useful and / or robust at the different spatial scales.

This publication includes an evaluation of the strengths and weaknesses of the relevant data; sets out recommended standard national data sets; and relates the use of the data to the main modelling approaches currently being used.

The next three sections of this publication cover measuring, understanding and modelling market housing affordability. This is followed by a section on gaps and shortcomings in the available data and how to address them. Finally, conclusions are drawn about the use of data on affordability and housing markets more broadly.

Annex 1 gives details of the key data sources relevant for work on affordability. Annex 2 provides a more detailed list of data sources by topic. These two annexes can be used together when deciding which data are relevant for a given purpose. Annex 3 provides a glossary and Annex 4 a list of abbreviations. Annex 5 gives technical information relating to section 6.6, which discusses how the gaps in the existing data could be addressed.

¹ http://www.hm-treasury.gov.uk/media/9/D/pbr_csr07_psa20.pdf HM Treasury, 2007

Section 2: Measuring affordability

This section looks first at the two main approaches to measuring affordability: a general indication of whether house prices are affordable in relation to incomes; and a more specific measure of whether particular housing is affordable to certain groups e.g. first time buyers and key workers. In addition, there are a couple of more detailed ways of measuring affordability such as establishing the income required to purchase a home, looking at a household's 'residual income' or the income that remains once housing costs have been met and whether an individual or household has access to finance in order to purchase a house.

Which measure is used depends on the purpose. When looking at the impact of additional housing supply on affordability within a district, for example, it is common to use house price to income ratios, and this has the advantage that it enables comparisons between districts (see the CLG website's housing live tables 576 and 577 for district ratios going back to 1997). However, when making decisions about the proportion of new homes that need to be affordable, it is more usual to look at whether housing is affordable for an individual household rather than broad price to income ratios, and here the common approach is to establish the number of households unable to afford market housing, given that housing costs should not exceed a certain proportion of income.

The additional measures are useful for other purposes. For example, when deciding whether shared ownership is affordable to a key worker household, it is usual to look at the annual income necessary to purchase a home or to rent privately on the open market. Alternatively, when examining the extent of poverty in a district the residual income approach is more useful. Also, as market conditions change it may be useful to consider how access to finance affects housing affordability. All these measures are explained below, with examples and suitable sources of data.

2.1 House prices to income ratios

This approach to measuring affordability gives a general indication of whether house prices are affordable in relation to incomes. This indicator is often termed the house price to income ratio, and the government's preferred version is the ratio of the lower quartile house price to annual lower quartile earnings. This simple indicator can be used at national, regional, sub-regional and local spatial scales and is particularly useful in making comparisons over time or between areas.

The table below compares price to income ratios for Basildon, a district within the East of England with regional and national equivalents. The data show that Basildon is now (latest data 2005-6) less affordable than the rest of the region and the rest of the country according to the mean price /income ratio.

It is common to look at average incomes in an area (the mean) but as the table shows, in all cases in this example the median income is less than the mean. This is the usual situation and shows an unequal income distribution – more than half the households earn less than the mean, which is pulled upwards by some very high income households. In a more even income distribution, the mean and the median would be much closer together. Additionally, lower quartile earnings are well below median income. In order to understand an area's income distribution, it is clearly important to look at both measures, mean and median, as the mean alone will be misleading.

Using the lower quartile house price as an entry price for owner occupation, the price to income ratios, however measured, all show that housing is not affordable in England

generally, in the East of England and particularly in Basildon. People on or below the lower quartile income as measured by earnings are worst affected and would not be able to access owner occupation.

Table 2.1 Example of affordability analysis using price to income ratios

	Basildon	East of England	England
Mean income	£26,000	£25,800	£24,900
Median income	£18,900	£18,000	£17,400
LQ earnings	£11,898	£10,494	£11,159
LQ house price	£133,802	£130,000	£115,500
Mean income/ price ratio	5.15	5.04	4.64
Median income / price ratio	7.08	7.22	6.64
LQ earnings / price ratio	11.25	12.39	10.34

Sources: 2005-6 data

Mean and median income: Inland Revenue (more recent income data not yet available)

Lower quartile (LQ) earnings: Annual Survey of Hours and Earnings (ASHE)

LQ House price: Land Registry

CLG live table 576 provides ratios of lower quartile house prices to lower quartile incomes by district while live table 577 provides the ratio of median house price to median income by district. Incomes in both cases are measured by earnings from ASHE (available on the ONS website) and lower quartile house prices from the Land Registry (this is set out in PSA 20).

The advantages of using earnings data are that they include all earned income, although not income from other sources, and the sample is large enough to be robust at regional level, whereas the advantage of using Inland Revenue income data is that it includes incomes from all tax self-assessment as well as employees and is also a large sample (10%).

Neither source relates to household incomes however and the sources that do (Family Resources Survey, Survey of English Housing) are too small to be fully robust below national level² apart from the larger regions such as London and the South East.

Similarly, while there are several sources of house price data such as the Nationwide and Halifax price indexes, the Land Registry data cover all transactions (except Right to Buy sales and certain other categories) so are considered more complete and hence more accurate.

House price to income ratios provide an indicator of the location of the most and least pressured areas within a region or a sub-region. This information can then help the relevant authority to decide where new housing should be built, or which areas may be in need of regeneration, alongside other considerations such as environmental issues, infrastructure constraints, flood risk or planning matters that may also be important in determining the location of new homes.

² 'Fully robust' in this report means capable of complex disaggregated analysis. These national surveys are often fine for basic sub-national breakdowns but, for example, a breakdown of income by ethnic group and household size or type may produce too few numbers in each cell to be statistically significant.

2.2 Ability to afford home ownership/ private renting

The second approach to affordability is measuring whether housing is affordable to different household groups. This measure is used when determining the proportion of new homes that need to be affordable. Assuming that the local authority has details of the income distribution (e.g. from a household survey), then the proportion who cannot afford local market housing can be estimated. Government practice guidance on Strategic Housing Market Assessments (SHMAs) includes a section on how to assess affordability in this sense (CLG 2007, page 42).

Gross household incomes should be used to assess affordability. Benefits should be excluded from this as housing benefit is only paid to households who would otherwise be unable to afford their housing and other benefits are often intended to meet specific needs rather than housing costs (e.g. disability living allowances).

The estimate will be different for home ownership and private renting. The explanations below are taken from the SHMA Guidance.

Measuring the affordability of home ownership

A household can be considered able to afford to buy a home if it costs 3.5 times the annual gross household income for a single earner household or 2.9 times the gross household income for dual-income households. Housing market partnerships will need to consider that what is affordable may vary (e.g. a higher proportion may be affordable for someone on a higher income or the relevant ratio may change with interest rates). Local circumstances could justify using different figures. Where possible, allowance should be made for access to capital that could be used towards the cost of home ownership.

Measuring the affordability of private renting

A household can be considered able to afford market renting where the rent payable is up to 25 per cent of their gross household income. The 'Rent Payable' figure is defined as the entire rent due, even if it is partially or entirely met by housing benefit. Other housing-related costs, such as council tax and utility bills should not be included. As in the case of assessing home ownership, housing market partnerships will need to consider that what is affordable may vary (e.g. a higher percentage may be affordable for someone on a higher income). Local circumstances could justify a figure other than 25 per cent of gross household income being used.

Table 2.2 Data required for measuring affordability for an individual household

Data required	Source
House price	Land Registry mean or lower quartile house price
Deposit required	CML Regulated Mortgage Survey
Lender multiple	CML Regulated Mortgage Survey
Interest rate	CML Regulated Mortgage Survey
Period of loan	CML Regulated Mortgage Survey
Mortgage payment protection insurance	Nationwide website
Building insurance	Nationwide website
Household income	Newly commissioned household survey OR Family Resources Survey
Savings or capital available	Newly commissioned household survey

Three further common measures of affordability are:

- estimating the annual income necessary to purchase a house
- calculating the residual income remaining after the rent / mortgage and other essential costs have been paid and comparing this with the Income Support Applicable Amount.
- considering whether an individual or household has access to finance with which to purchase a home

In the first two cases, averages can be used. Worked examples of each measure are presented below.

2.3 Annual income necessary to purchase a house

This measure is a general indicator of affordability in a district or region and is useful for making comparisons. It is commonly used by HomeBuy agents and housing associations in setting the affordable price and rent levels for shared ownership products and assessing eligibility. Clearly housing associations do not wish to encourage home ownership for households on the very borders of affordability, because they will be vulnerable if interest rates rise. This measure helps to avoid such situations.

A worked example of purchasing a shared ownership flat is shown in table 2.3. There are two elements: the share that is purchased by the occupier, with a standard mortgage for (in this example) half the full market value; and the rental charge on the remainder. The assumption in this example is that total outgoings should not exceed more than 25% of gross household income.

In this example the rental rate (as a percentage of half the capital value) is varied to show the impact this has on affordability. It is clear that reducing the rental element makes the property more affordable to people on lower incomes.

Table 2.3 Annual income necessary to purchase 50% of a flat at different rental rates

Property Type

1 bed flat

Open Market Value	£152,000.00			
Share available %	50%			
Value of unsold equity £	£76,000			
Rental charge as % of value	3.75%	2.75%	2.00%	1.00%
Rent (£ monthly)	£237.50	£174.17	£126.67	£63.33
Service Charge £	£83.08			
Monthly Mortgage based on 5.5% over 25 years	£494.00			
Total outgoings per month £	£814.62	£751.25	£703.75	£640.41
Income required based on 25% gross	£39,101.64	£36,059.84	£33,779.84	£30,739.84

Source: HomeBuy data, East of England

The table shows that even for a one bed flat with a relatively minimal rental charge of 1% per annum of the unsold equity, an income of around £30,700 is needed to be able to afford to purchase a 50% share. This means the product is still unaffordable to those on lower incomes. In this case, intermediate rent may be a more affordable tenure for those in the lower income brackets.

The annual income required to purchase a house is also used to measure the proportion of households in an area that are at or below the threshold income and therefore are unable to

afford entry level housing. This requires data on local incomes, which are often collected in specially commissioned household surveys e.g. a housing needs survey. Where such survey data is not available, income distributions can be purchased from private data suppliers such as CACI (see Annex 1 for an evaluation of CACI as a data provider).

2.4 Residual incomes

The residual income is the income a household has left over after they have paid housing costs. It gives what is in some ways a more accurate picture of affordability than price-income ratios since it recognises that lower income households are only able to afford smaller proportions of their income on housing without facing difficulties

One way of establishing the “spare” money a household has is to reduce the residual income by an amount they are assumed to need for basic living expenses.

This can be worked out by using the formula:

Residual income = income minus rent minus Income Support Applicable Amount (ISAM) plus housing benefit

For housing to be affordable, income should not in principle fall below ISAM even if income support is the only income, and therefore residual income should always be greater than zero. The higher the residual income, the more affordable the housing, in the sense of having some income left over to buy other items. Where residual income is negative, there is a serious problem of affordability.

The example given below is taken from a study (Tang, 2008) that explored the extent of affordability problems of working families with children in terms of their dependency on Housing Benefit (HB). The study uses actual data from the COntinuous REcording (CORE) data set to illustrate how residual incomes in the social rented sector varied both over time and between different types of household. These households are therefore new entrants to the social sector, and the calculation is made on the basis that ideally working households should have minimum residual incomes at the level of 120 per cent of the Income Support scheme allowances (the Applicable Amount).

Table 2.4 Weekly residual incomes (GBP £) for working families in the social sector, 1999-2006

	1999-00	2002-03	2005-06	1999-00	2002-03	2005-06	1999-00	2002-03	2005-06
	non-HB dependent			HB dependent			All		
1 parent	55.84	69.80	73.07	- 25.45	- 28.13	- 29.71	3.28	56.05	61.28
2 parent	97.39	115.46	124.75	-13.00	- 32.32	- 41.01	54.03	109.51	120.34

Source: Tang, 2008, forthcoming.

This table shows that working families without Housing Benefit found their housing much more affordable than those who were dependent on HB, and couples with children had higher residual incomes than single parents. HB dependent households experienced negative residual incomes because they were, presumably, on the minimum level of Income Support. Affordability also worsened over time for those who were dependent on HB but improved for those not dependent on HB.

As this example shows, residual incomes are most relevant for analysing the affordability of housing for the poorest households. Research tends to show that these are mostly families with children where the parents are in low paid work, and/or where one partner may not be in

work or work part time. However, other groups are also vulnerable including single people, single income households and families with a joint income. Elderly retired and single unemployed people are also at risk.

2.5 Access to finance

When an individual or a household is making a decision to purchase a property one constraint may be their access to the mortgage they require. Mortgage lenders will factor in many individual characteristics when deciding whether or not to lend, but a key factor will be the size of the borrower's deposit and the associated loan to value ratio.

During periods of high house price inflation lenders have offered high loan to value ratios, with some lenders offering products of up to 100 per cent (and even above 100 per cent in some cases). This means that lenders are prepared to lend the whole value of the property, therefore reducing the deposit requirement to zero.

In this situation the size of the deposit required will not be the binding constraint in determining whether an individual can afford to purchase a property. In these cases other elements of affordability such as the income multiple or the proportion of the income a household can afford to be consumed by mortgage payments will be the binding constraints.

During periods of lower, or expected negative, house price growth lenders will not be willing to lend at higher loan to value ratios. In these cases an individual or a household will require a larger deposit to access mortgage finance. In such cases the level of deposit required may become a binding constraint for the household.

Analysis of an access to finance indicator of affordability has become increasingly important during the economic downturn as mortgage lenders have decreased lending to households and individuals with smaller deposits and borrowing at higher loan to value ratios. This change in lending practice means that the changes in deposit requirements are no longer related to changes in house prices and therefore the ratio of house prices to income.

Indicators of access to finance related affordability constraints include changes in the loan to value ratio, the size of deposit paid by first time buyers and the deposit as a proportion of annual income.

The Council of Mortgage Lenders (CML) collate data from mortgage lenders on a number of variables including purchase price and loan amount, and therefore the size of the deposit. The data also shows the loan to value ratio and the declared income on which the mortgage is made. The CML data indicates Government Office (GO) region and Local Authority (LA) code so can be used at a regional or local authority level. The data also gives post code so could be used to construct a data set for other geographical areas such as housing market areas.

The CML charges a fee for their data and an individual contract arrangement will be required to obtain the data.

In conclusion, it should be emphasised that there is no single 'correct' measure of affordability that will cover all aspects. Each measure has its strengths and weaknesses and the 'best' measure to use depends on the purpose. A 'basket' of measures is required.

Section 3: Understanding affordability

After looking at what is meant by the 'active' housing market, this section explores the drivers of affordability in the context of housing markets and the overall housing system, including the rented and affordable housing sectors. Further details are available in the Government's Practice Guidance on Strategic Housing Market Assessments (SHMAs).

3.1 The 'active' housing market

It is important to recognise that understanding the local housing market is not simply about measuring the total stock of housing on the supply side and estimating the number of households on the demand side. Only a proportion of dwellings are actually available for sale or for rent at any given time, and it is this plus newly built dwellings coming on stream that constitutes housing supply. Equally only a proportion of households plus newly forming households are actively seeking a home at any one point in time. These households constitute effective demand – those who want and are able to afford to rent or buy. In addition there may be 'pent-up' housing demand, from people who would like to rent or buy a home but are unable to do so. The main reason for this is usually affordability, although there can also be housing on the market that is affordable to a range of households but is not desirable or suitable. There can be a number of reasons for this, such as unfitness, poor location, or simply the wrong size e.g. one bedroom flats when demand is for larger units.

It is not possible to quantify precisely the extent of current demand and a range of indicators need to be examined and interpreted. Historical data may not be available at the local level. Available indicators can be supplemented with knowledge from relevant local stakeholders such as estate agents and developers. Where analysis suggests that there are issues of low demand within housing market areas (and therefore little market activity), additional research such as neighbourhood renewal assessments may be required.

The data used to understand housing demand for market housing differs from that used to assess need for affordable housing, since the needs of those who cannot afford market housing are not reflected in house prices. The accepted method for assessing housing need is covered in the section 4.3.

3.2 Drivers of the housing market

The key drivers of the housing market and therefore of affordability are:

- Demographic and economic changes
- Employment and incomes

Other drivers or indicators are:

- Second homes and vacant dwellings
- Accessibility and school catchment areas

These are discussed below. Data sources for these elements can be found in Annexes 1 and 2.

3.3 Demographic and economic changes

Demographic and economic changes are key drivers that underpin every housing market, influencing both demand and supply. Thus demographic trends such as ageing can rapidly

change the nature of housing markets. One example is rural areas where young people have been leaving, partly because they cannot afford to live there, and partly because of limited educational and employment opportunities. This leaves a population that is largely outside the labour market and increasingly dependent on public services such as health care. The relationships between the economy, household composition and housing are interactive and complex. Economic and demographic changes influence demand for housing (both quantity and quality). This is because income and household numbers are the key drivers of housing demand. As incomes rise, more housing is demanded – both in terms of larger dwellings and also more attractive locations such as more rural areas with good accessibility. In terms of demographic changes, as people live longer and women outlive their spouses, the number of one person households increases, changing the nature of demand for housing. If people remain in their family sized homes as they age, they effectively prevent new growing families from living there. However, if older people can be persuaded to downsize, the problem is eased. The housing market also influences household formation rates and the economy. For example, reduced non-household spending, borrowing against household equity in order to fund consumption, and new housing investment including spending on home improvements can all feed back into the economy both locally and nationally.

Changes in population and household profiles are key to understanding housing need and demand. Migration is an important factor driving demographic change. The age structure of the population is influenced by increasing longevity and by migration trends and in turn this affects new household formation and housing need. Gender and ethnicity also have a strong influence on the mix and location of housing need and demand. Different ethnic groups have different age structures, household formation rates, fertility rates, tenure and locational preferences. Female-headed households tend to have lower rates of home ownership and are over represented in older age.

Macroeconomic factors such as interest rates, inflation and national economic growth affect new house building and investment on the supply side and household finances on the demand side. For example, higher interest rates increase the cost of borrowing, thereby reducing demand for housing and affecting employment levels by increasing the cost of investment. Mortgage rationing by lenders (as during the current credit crunch) affects new house building and investment as potential buyers are unable to access funds. Government subsidies, grants and policies can also influence housing markets. Although new housing is only a small proportion of total stock (less than 2 per cent per year), some local areas may experience much greater increases in new housing supply. This is happening in parts of the Growth Areas and Growth Points, for instance, and should make access to housing easier for first time buyers.

3.4 Employment and incomes

The level and type of employment available in an area determine the levels of wages that will be paid and will influence the tenure, size and location of housing. The types of employment available may also influence household preferences for different housing locations; people with highly-skilled jobs tend to travel further to work than those with lower skilled jobs.

Incomes and earnings are key drivers of demand. The evidence shows that households spend more on housing as incomes increase, although typically they do so at a rate less than their growth in income. Studies suggest that as incomes rise, demand for neighbourhood quality increases together with house size.

Ultimately demand for market housing is driven by how much housing people can afford, rather than how much they might be considered to 'need'.

3.5 Second homes and vacant dwellings

Second homes and vacant dwellings can both be drivers of demand. While second homes are quite a small proportion of the total stock at national level, in some areas the proportion is very high. The demand for second homes in desirable holiday areas fuels demand, pushing up house prices and creating affordability problems for local households.

Conversely, a large number of vacant properties is both an indicator of low demand and a driver of demand elsewhere, as any remaining potential demand will choose to go elsewhere, worsening the situation. Although high vacancy rates are usually an indication of low demand, vacancies can be high in high demand areas as a result of absentee landlords and 'buy to leave empty' on the part of investors. However, high turnover (as opposed to property remaining on the market and vacant) indicates a healthy housing market with active buyers and sellers.

3.6 Accessibility and school catchment areas

Accessibility, to employment, services, shops and schools, is an important driver of demand but is difficult to measure. Transport for London (TfL) has produced data measuring the Public Transport Accessibility Level, but such data is not available at national level. Within London, proximity to underground or railway stations adds to house prices while a lack of access to public transport reduces values. School catchment areas can also influence demand. Families are known to move home so that their children can attend popular schools. This puts additional pressure on house prices within the catchment area of the school. Again, poorly achieving schools can have the opposite effect, as people try to ensure that their children do not have to attend an unpopular school. Accessibility and school catchment areas are elements to consider in the local authority context but are difficult to quantify without local studies.

Section 4: Modelling affordability

This section outlines the data sources needed for the main models currently used for modelling affordability at national and regional level. It does not explain how to use these models. An important model is the CLG Reading Affordability Model and information about this is available from the CLG website and also from NHPAU who have used it to develop their supply ranges for regional housing requirements. Many regions undertake modelling of the housing market and of affordability. This section looks at the available data sources. The topics covered are:

1. *Projecting net future housing requirements* – new household formation, demographics, incomes and other determining factors.
2. *Estimating the relationship between supply and affordability* - the CLG Reading Affordability Model
3. *Projecting the future supply of new housing* – the main determinants of new supply, particularly in terms of land supply and planning evidence.
4. *Identifying potential housing need* - alternative approaches to identifying housing need, where market indicators do not help because those who cannot afford housing are outside the market
5. *Land supply and planning* –the data required to understand the scale and nature of the identified land supply and other potential sources of supply.
6. *The allocation of housing between tenures* – Includes data on the relative importance of the different tenures, how this has changed over time and what factors affect these trends. The impact of the affordability of private renting is also discussed.

4.1 Projecting net future housing requirements

This approach, commonly termed the ‘demographic projection model’, has been developed since the mid 1970s as a means of estimating housing and land requirements as well as the funding needed for affordable housing. It projects, rather than forecasts, demand and need on the basis of past trends. It therefore assumes that behaviour in the future will be the same as in the past. It includes an assessment of current unmet housing need (the ‘backlog’) because housing needs assessment guidance recommends that local authorities make an allowance to eliminate the backlog over five or perhaps ten years. Therefore this allowance is added to the estimate of the new housing that is required in the future.

The approach involves:

- demographic projections of total population and households produced by government and breaking these down into types and ages of household with their different probabilities of being in each tenure and rolling these forward into the future
- an estimate of current unmet housing need in terms of concealed and sharing households who on government standards require a separate home, plus other housing needs such as overcrowding which may or may not potentially be satisfied from the existing stock
- assumptions about vacancy rates, transfer of stock between tenures, second homes and demolitions based on past experience

One important benefit of this type of approach is that the results are in keeping with the way planners determine land allocations and the mix between market and affordable housing. It is thus fairly user friendly. The output measures and therefore the measure of success are also closely aligned with policy as they are numbers based.

In particular this approach addresses:

- the net new additions to stock required to meet projected demand and
- the net numbers of households currently in housing need that could additionally be housed.

It thus directly relates to the government's current target based approach to additional housing supply.

Data sources for projecting future housing requirements

One challenge when operationalising the demographic projection model concerns producing regional level projections because the use of the Survey of English Housing for household type and age (of household reference person) is not robust at regional level. There are 90 possible combinations of household type and age, and the sample size for the North East, for example, pooled over three years, is only 3,000 which is not sufficient for this amount of breakdown. Instead the practice is to use regional household projections from the Office for National Statistics to apportion national estimates between regions. This is possible so long as the changes in any region are not negative, for example net outward migration, as has been the case sometimes in the past.

The main data sources that are used for the projection of newly arising demand and need are given in table 4.1 below:

Table 4.1 Data sources for demographic projection model

Variable	Source	Notes
National population projections by age	Office for National Statistics (ONS)	
Sub-national (including regional) population projections	ONS	
Household projections by age and household type	CLG	
Tenure proportions by age and household type	Survey of English Housing (SEH) pooled over three years ³	
Change in number of second homes	As above	This plus vacants convert the net increase in households into a net increase in stock
Change in number of vacant dwellings	Housing Strategy Statistical Appendix (HSSA), CLG	As above
Losses from demolitions	Housing Statistics, CLG	This plus change of

³ The Integrated Household Survey (IHS) began in 2008. It integrates the following surveys:

- Labour Force Survey (LFS)
- General Household Survey (GHS)
- Expenditure and Food Survey (EFS)
- Omnibus Survey (OMN)
- English Housing Survey (EHS) – which launched in April 2008 (as part of the IHS) and represented a merging of the Survey of English Housing (SEH) and the English House Condition Survey (EHCS)

These surveys are referred to individually in this publication.

		use convert the net increase in stock into a gross dwelling requirement
Losses / gains from change of use	As above	As above
Number of Right to Buy (RTB) sales	As above	This is specific to the social sector

Note: The number of Right to Buy sales is required because when the household that purchased the property dissolves or moves away, the unit is bought by someone else in the market sector, but had it not been purchased through Right to Buy it would have been available for re-let. Therefore an estimate of the future need for social housing ought to include an allowance to replace such units.

Estimating current unmet housing need is more difficult because many of the original data sources have not been repeated and are now very out of date. These include a survey of local authority waiting lists undertaken by government in 1991. Given this data gap, much broader estimates have to be made, with caveats that actual unmet need is likely to be higher.

The main data sources for estimating current unmet housing need are given in table 4.2:

Table 4.2 Data sources for current unmet housing need

Variable	Source	Notes
Concealed households	Household projections, CLG on request	Detailed workings for the household projections
Households in temporary accommodation	HSSA on CLG website	Administrative data
Households in overcrowded accommodation	2001 Census	Increasingly out of date
Households in shared dwellings	Survey of English Housing, CLG	National figure only
Would-be couples living apart	Waiting list survey now out of date	Cannot be included in estimate
Single homeless and hostel residents	Survey of English Housing, CLG	National figure only

Because some of the survey data is now so out of date (1991 Waiting List Survey), the Barker Review of Housing Supply Interim Report included a figure which was the 1991 survey scaled up pro rata to the latest waiting list figures from the Survey of English Housing. This, however, is not robust at regional level.

Further information about the demographic projections approach is given in an annex to the NHPAU publication *Meeting the housing requirements of an aspiring and growing nation* (NHPAU, 2008a).

4.2 Estimating the relationship between supply and affordability

The CLG Reading Affordability Model has been developed in response to the Barker Review of Housing Supply in order to examine the market relationships and the impact of additional housing supply on prices and affordability at national and regional levels. This model is now being used by NHPAU as a key part of their work on affordability. This section aims to aid understanding of the model and its data requirements without going into technical detail. More information is available in the NHPAU publication *Meeting the housing requirements of an aspiring and growing nation* (NHPAU, 2008a), which includes an affordability analysis at regional level and extensive scenario modelling.

The CLG Reading Affordability Model includes a set of complex structural equations which aim to predict demand and house prices in the long run. The demand side directly models the relationship between demographic change and household formation, based on price and income elasticities. It also takes direct account of how changes in incomes and prices affect that demand with respect to tenure choice and willingness to afford market housing. It links the labour market to the housing market to help explain differences between regions.

One difference compared to the demographic projection model is that the demographic approach establishes a direct relationship between the numbers of new households and the net addition to the housing stock that will be required. The CLG Affordability Model, on the other hand, focuses on house prices, which are affected not only by demographic factors but also by economic influences such as income, mortgage interest rates, and the availability (or 'supply') of housing.

The model is essentially of the market sector with the social sector determined by supply. The relationship between the owner occupied and private rented sectors is assumed to be in equilibrium. There are also equations that estimate demolitions, second homes, vacancies and supply response at regional as well as national level.

The key relationships, however, are not in terms of numbers of dwellings. In principle demand is measured in terms of standard units of housing. It includes additional demand from existing households, e.g. as incomes increase, as well as demand from newly forming households. Therefore the model is more like a filtering model where demand enters at any point on the income scale, and those who move to a bigger property because their demand increases leave an empty home to be taken by someone else, and so the property filters down the market until it becomes vacant and eventually demolished.

Table 4.3 below sets out the data sources that are used in the CLG Reading Affordability Model.

Table 4.3 Data sources for the CLG Reading Affordability Model

Variable	Data source	Time, level etc.
Private housing starts 2001-2005	CLG live table 217 compiled from P2m returns from LAs plus returns from NHBC	Quarterly, by region and main tenures: private, social and local authority
Lower quartile earnings	Annual Survey of House and Earnings (ASHE) from ONS	Annual, by place of residence by local authority, mean, median and lower quartile earnings, gross weekly pay, individuals
Lower quartile house prices	CLG live table 583 compiled from Land Registry data	Quarterly by local authority. Excludes Right to Buy and sales below £1,000 and above £2m.
International migration gross inflows and outflows 2001/02	ONS 'Region in Figures' population chapter. 2003-04 are interpolated	Internal and International Inflows and Outflows by region – now very dated
International migration gross inflows and outflows 2005+	Direct from CLG	Interpolate for intervening years
National proportions of migrants by age	Calculated from above source	Mean proportion of males to females for each age range – still very dated, 2004
Inter-regional migration gross inflows and outflows	Same source as above	Now to 2006
Conception rates	ONS compiled from Health Statistics	Quarterly from 2003, women's conception rates by age, local

		authority level
Total live births	Same source as above	Quarterly from 2003, all live births, local authority level
Conception / birth ratio	Calculated from population projections direct from CLG	By region, entered as exogenous in model
Male/female death rates	Calculated from population projections direct from CLG	By region, age and gender
Population male/female 2001	Provided by ONS directly	Quinary age groups and gender, from mid-year population estimates, local authority level
Population male/female 2003	Calculated from above	
Population male/female 2005	Calculated from above	
Dependent children rate	Survey of English Housing	Annual, by region, gender (of HRP) age and marital / cohabiting status
Couple rate	Calculated from population projections direct from CLG	By age group, gender and marital status
Right to Buy average discount, as proportion of sale value	CLG live table 643 compiled from P1B returns from local authorities	By region, annual
Right to Buy sales	Same as above	As above
Total numbers of households in each tenure	Survey of English Housing	Annual, by region, owner occupiers, private renters, social renters, shared owners
Total annual supply by tenure	CLG live table 109 compiled from census and projected forward using completions data and annual net gain or loss due to conversions, demolitions and change of use	Annual by region and tenure as above – but vacant data from HSSA returns is not considered very accurate, therefore supply by tenure less accurate than total supply
Stock by age	CLG live table 110	Year Built, by region
Mean house price scalar	CLG live table 585 from Land Registry data	Annual, historical series, mean house prices by district
Average full time earnings	Annual Survey of House and Earnings (ASHE) from ONS	Annual, by place of residence by local authority, mean, median and lower quartile earnings, gross weekly pay, individuals
Median earnings	As above	As above
Unemployment rate	Table F.1 claimant count from ONS web site	Seasonally adjusted annual average claimant count figure by region
Regional house price index	CLG live table 591	Quarterly mix adjusted house price index by region
National house price index	As above	Calculated from the regional figures
Mortgage interest rate	Table 7.1L building societies interest rates from ONS web site (financial statistics)	Monthly, quarterly and annual averages are calculated from these. National data.
Consumer expenditure deflator	ONS Economic Trends table 2.4 income, product and spending per head	Divide household gross disposable income by real household disposable income times 100

The reason why this table is so large is that many elements which could be obtained from other official sources, such as new household formation, are estimated directly within the model. This is because worsening housing affordability will reduce or delay the formation of new households while improved affordability will increase new household formation rates.

Similar reasons apply for other variables. The aim of the model is to capture behavioural changes in response to movements in economic variables.

The data sources used by the CLG Reading Affordability Model are, quite simply, the best available for the purpose. If better data became available, for example information on international migration, then it would be appropriate to substitute such data

NHPAU expect to roll out copies of the CLG Reading Affordability Model to Regional Partners during 2009. Work will progress through 2009, by NHPAU and successful bidders, to develop a sub-regional housing market affordability model.

4.3 Projecting the future supply of new housing

This section identifies the main determinants of new supply, particularly in terms of starts and completions.

The main determinants of new housing starts are demand (as indicated by house prices, rents, rates of return on investment in the rental market, availability of mortgage finance including Buy-to-Let finance, number of transactions, and trends in all these indicators), availability of land, construction costs and the cost of borrowing. When house prices are high and rising, as in the period 2000-2007, developers are keen to build and are pushing for more land to be released for housing. Land availability includes not only land identified in local plans, but also 'windfall' sites, including sites currently in some other use such as manufacturing or retail where the owner decides to relocate and the site can be used for housing. Strategic Housing Land Availability Assessments are required to identify a five-year land supply for new housing, but by definition may not capture windfall sites. While high construction costs can push up house prices, so long as demand is buoyant this will not slow down the rate of new building, and the same is true for the cost of borrowing.

However some developments are highly sensitive to the cost of borrowing, so when interest rates rise the building rate slows down. House building is speculative and highly risky, so lack of confidence from potential purchasers will rapidly affect developers' activity. Some sites will be mothballed, and some schemes that are in the pipeline will not go ahead. Build-out rates will begin to slow down, so that the gap between starts and completions grows larger. This is what has happened since late 2007, as the loss of confidence in the banking sector led to a withdrawal of mortgage credit which had an even worse impact on new housebuilding than rising interest rates.

Data on housing starts and completions comes from inspections undertaken for Building Regulations purposes (CLG live table 253) and it is less robust than it used to be. This is because of a large increase in mixed tenure developments that has arisen in response to the Government's mixed communities agenda. In most cases of sites above 10 dwellings, social rented housing is pepper potted among the private housing. This makes it very difficult to assess the tenure of the dwellings. Even for private starts, it is not known from inspection whether the occupant is an owner or a tenant, and it is increasingly difficult to separate social housing (largely housing association) from market housing. CLG is considering using data from the new Homes and Communities Agency on social housing completions to provide an indication of tenure. However, this would not capture affordable dwellings provided without grant – such as low cost home ownership built with free land from the developer and financed by the housing association from borrowing against the future proceeds of the sale along with rental stream from the unsold share.

If local authorities were to establish some really effective monitoring systems this situation could change. Most authorities employ enforcement officers who ensure that the

development meets the requirements of the planning consent, but they are often extremely stretched and there is no suggestion that they are currently producing adequate data. Local authorities do contribute to regional Annual Monitoring Reports although this source is not perfect for completions as it depends on individual responses. It includes conversions and changes of use but not demolitions, so it is neither a record of new housebuilding nor of net additions to stock. One alternative might be to use the planning consent, which states the proportion of private, intermediate and social rented units, on the assumption that these will be delivered to plan. Certainly recent research on S106 found that most S106 agreements were delivered as planned (Monk *et al*, 2006).

4.4 Identifying potential housing need

Identifying potential housing need from households that cannot access market housing is covered in detail in the Government's Practice Guidance on Strategic Housing Market Assessments. It covers current housing need (often termed the gross backlog), future housing need, affordable housing supply and the housing requirements of households in need.

Current housing need refers to existing households who are either homeless or living in unsuitable housing. The latter can involve a mismatch of housing need and dwelling or dwellings in poor condition. The main data sources for estimating the backlog of current unmet housing need are given in table 4.4 below (there is some overlap with the backlog as estimated in the demographic projections model but the two are not the same).

Table 4.4 Current housing need

Category of need	Data source
Homeless households	Local authorities hold data on homeless households and those in temporary accommodation
Overcrowded and concealed households	Census plus SEH to cover trends between censuses plus CLG household projections (for concealed)
Households in insecure tenure	Housing registers (reported by DCLG in live table 600) or local survey
Overcrowded households	Census plus SEH as above or local survey
Too difficult to maintain even with equity release*	Census plus SEH or local survey
Sharing	Census plus SEH or local survey
Mobility impairment or other specific needs	Census plus SEH or local survey
Unfitness	Census plus English House Condition Survey (EHCS) to cover trends between censuses or local survey

Source: SHMA Practice Guidance, 2007

* In other words, even if a householder is able to take out a second mortgage (equity release) there is not enough value in the property to cover essential repairs.

While many practitioners use housing register (waiting list) data as a proxy for housing need, it is important to note that there are risks attached. Some authorities clean up their registers regularly, others less often. There is a lack of consistency between local authorities over definitions of priority need, and there may be double counting where households are registered on more than one waiting list.

Future housing need is an attempt to project the housing needs that will arise in the future. The demographic model described above in section 4.1 produces a projection of newly arising housing need based on recent past trends. Table 4.5 shows the main data sources

used to estimate future housing need. Further detail is available in the Strategic Housing Market Assessment Practice Guidance (CLG, 2007).

Table 4.5 Future housing need

Category	Data source
New household formation (gross p.a.)	Census, SEH
Proportion unable to buy or rent	Entry level rents / house prices, SEH, CML
Existing households falling into need	Housing register, household surveys

Source: SHMA Practice Guidance, CLG, 2007

New household formation includes potential first time buyers as well as potential private and social renters, and is itself sensitive to housing affordability. Young people, for example, may delay leaving the parental home to form a new household if they cannot afford to live independently.

It is observed that CLG has commissioned a model to estimate housing need. This may provide a more thorough understanding of the drivers and types of housing need.

4.5 Land supply and planning

This section covers the data required to understand the scale and nature of the identified land supply and other potential sources of supply. As noted in section 6 on gaps, this is an area where data are often missing or unreliable. The key elements that are required are:

- 'strategic land' often held on option by developers who then pursue its allocation for housing use and eventual planning permission
- land allocated for housing by the planning system
- land coming through the planning system (from outline application to detailed consent)

In each case, the size of the site as well as the number of units it is expected to deliver at a given density should be available. In an ideal world, an indication of ownership and whether the land is 'greenfield' or 'brownfield' should also be provided.

In practice there is no information about strategic land (i.e. land with potential for housing development but not identified as such in any development plan) other than broad indications in the annual reports of the large house builders. These may be overstated to reassure shareholders that the company has an adequate land supply to continue operating at their current level over the next few years. To collate this material would be a major task, however, and it would not provide sizes or numbers of units.

Information about land allocated for housing is available from Regional Spatial Strategies (RSS) and their associated Annual Monitoring Reports (AMR). At district level it is available in the Local Development Framework (LDF). The information is presented in terms of the number of dwellings the land is expected to deliver on the basis of a broad assumption on density. Strategic Housing Land Availability Assessments (SHLAA) are required in order to identify a five year supply of land for new housing development. As these are completed across the country, they will provide a much better indication of the available land supply at local authority level than previously. Again, this information is likely to be collated at regional level, at least to inform Annual Monitoring Reports.

Land coming through the planning system is available from Emap Glenigan. Again it is in terms of housing developments and their size in terms of numbers of dwellings rather than

size of site. It is also expensive as it is gleaned from local planning authorities and other 'sources' of information on proposed schemes. There is no way of knowing whether it is fully comprehensive (see section 5) although it probably covers all major schemes.

All this means that there is no information on the length of time taken from land purchase to completed development. The widespread use of options by developers makes the date of purchase somewhat ambiguous. The eventual price paid for the land may have been determined years before, according to a formula set out in the option contract. Once land has planning permission, it may be profitable to sell it on quite quickly at a highly enhanced price.

Information is also lacking on the degree of constraint that is placed on development by the local planning regime. This gap is discussed in more detail in section 6.

At present, probably the best way to understand the nature and scale of the land supply is to collate information from Regional Spatial Strategies and Annual Monitoring Reports which will presumably incorporate information from local authority SHLAAs. The result will be an estimate, but until the dwellings have been completed, any information will be subject to revision. Data on planning permissions can only be used as a guide, because in practice not all planning permissions are implemented.

4.6 The allocation of housing between tenures

As part of their statutory obligations, regions and sub-regions need to understand how their housing markets are changing, and this includes knowing about the relative importance of the different tenures. As owner occupation has risen, the other tenures have decreased in importance, although this will vary from place to place.

Some areas have a larger private rented sector than others, such as university towns, seaside towns and the major cities. Private renting provides short term accommodation for more mobile households, although a significant minority of households end up renting privately all their lives. The recent introduction of Buy-to-Let (BTL) mortgages has encouraged a renaissance in private renting in certain parts of the country, again mainly the large cities. Although the demand for BTL mortgages has added to total housing demand, the provision of rental units via BTL mortgages has contributed to an increase in the private rental sector, helping to keep rents low. It has therefore contributed to the provision of more affordable housing for those who cannot afford home ownership (NHPAU, 2008b).

However, data on the private rented sector is incomplete. A limited set of rents data is available from the Rent Service and provided by Dataspring, along with the comparative costs of owner occupation and social renting. Information on the size of the private rented sector is available from the census and, at national level, the Survey of English Housing (now the English Housing Survey, a module in the new Integrated Household Survey). The latter has the advantage of being linked to a range of other data about households and their housing and is robust at regional level by pooling two or three years together.

Equally, some areas have much larger social rented sectors than others, particularly large cities. Information about the social rented sector, comprising both council housing and housing provided by housing associations, is available at local authority level from a number of sources. One is local authorities themselves, in their annual returns to the CLG in the form of the Housing Strategy Statistical Appendix, which gives information about the total number of units and the supply of re-lets. CORE and the Regulatory and Statistical Return (RSR) both provide data on registered social landlords, including total stock and new lettings, plus

rents by size of unit. Equivalent information for council housing is more limited, as LA CORE is only completed by a sample of districts.

At national level, the Survey of English Housing/ English Housing Survey can provide more detailed information. This data set, however, is not wholly reliable because some tenants of stock that has been transferred from the local authority to a registered social landlord do not realise this and still say they are council tenants. Similarly, some tenants in receipt of full housing benefit say that they live 'rent free'.

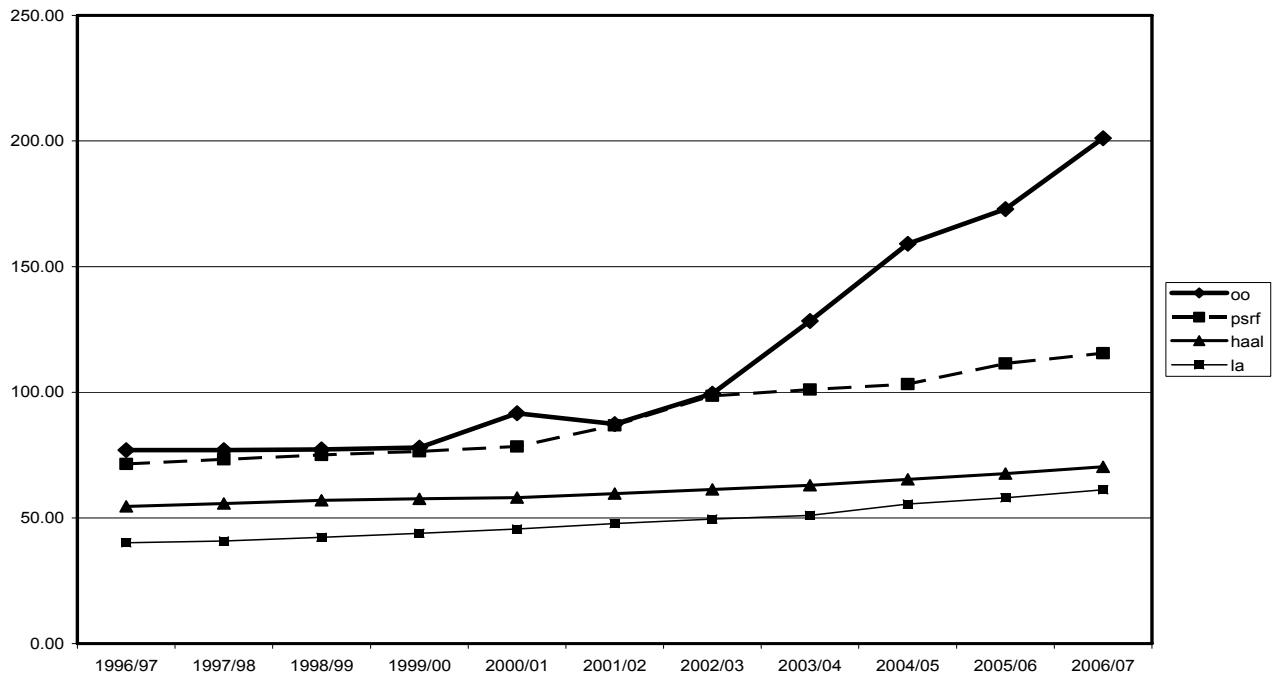
As home ownership has become less affordable in recent years, there has been an increase in the number of households who are able to afford to pay more than the social sector rent for a suitable property but cannot afford home ownership. An intermediate housing market has been developed to meet the needs of these people, including forms of shared ownership and shared equity designed to assist them onto the 'housing ladder'. However, while some information is available via CORE, the stock of low cost home ownership (LCHO) remains a very small proportion of the total largely because the dwellings tend to be sold on to the open market. This makes it difficult to find adequate data to 'project' demand for this housing tenure into the future (see section 4.1 above).

Evidence from the government's HomeBuy scheme, run by HomeBuy agents, suggests that those who successfully purchase a home through one of the available products are mainly couples or single people with no children, buying a one or two bedroom property. In London, the South East and East, they are mainly key workers (as defined by government) who appear to prefer the shared equity product for the purchase of existing properties, whereas in the East Midlands, for example, the preference among purchasers is for shared ownership of a newly built property.

Private renting is generally more affordable than buying with a mortgage. This has been true during recession as well as boom. The growth in the private rented sector in recent years, fuelled in part by the Buy-to-Let market, has made private renting even more affordable than in the past. Further information about Buy-to-Let is available in a research publication produced by NHPAU (NHPAU, 2008b).

Graph 4.1 shows the comparison over time of the costs of owning, renting privately, renting from a social landlord and renting from a local authority for the period 1996/7 to 2006/7 (the most recent available data). It demonstrates how private renting has become significantly more affordable than house purchase for first time buyers, and how steeply first time buyer housing costs have risen since 2002/03.

Graph 4.1 Comparison of the user costs of owning, renting privately, renting through a social landlord and renting from a local authority, £ per week, England 1996-2006



Key

- oo: user cost of owner occupation
- psrf: private sector referred rents
- haal: housing association all lets (from CORE)
- la: local authority rents

Sources:

Dataspring user cost of owner occupation www.dataspring.org.uk
 Private rents are Housing Benefit reference rents from The Rent Service website
<http://www.therentservice.gov.uk/corporate-publications/index.asp>
 CORE data from the University of St. Andrews www.core.ac.uk
 Local authority rents from CLG, live table 702

Section 5: Key data sources

This section sets out the most appropriate data source or sources for the elements identified earlier.

It focuses on the data sources themselves, rather than the ways to use them. However, it is important to consider the context in which the data sources are to be used, as this can directly affect the choice of data source. For instance, the British Household Panel Survey (BHPS) may not be the best choice for income data for most purposes, but would be used in the context of a model that required panel data to model tenure choices over time. The spatial scale at which the data is required is also crucial to the choice of data source. The main focus here is on data that can be used at regional and sub-regional (e.g. housing market area) level.

5.1 Key data sources and key variables

The table below gives details of the key variables that are available from the most useful data sources for any affordability analysis, whether descriptive or statistical. Annex 1 gives more detail about these data sources, including their strengths and weaknesses.

Table 5.1 Key Data Sources

Source	Key Data
Annual Population Survey (APS) Largest sample size, collated from other sample surveys plus own interview survey giving 215,000 households p.a.. Robust at LA level Source: ONS	Household composition and relationships Housing tenure Nationality Ethnicity and residential history Employment and training Workplace and location Job hunting Educational background Qualifications
Annual Survey of Hours and Earnings (ASHE) Large sample size, best source of earnings data (by place of work and by place of residence) Robust at regional and sub-regional level. Source: ONS website	Gross weekly pay Weekly pay excluding overtime Basic pay including other pay Overtime pay Gross hourly pay Hourly pay excluding overtime Gross annual pay Annual incentive pay Total paid hours Basic paid hours Paid overtime hours
CLG live tables Large range of relevant data accessible on line Robust at the lowest level provided in each case Source: CLG website	Housing stock data Housing market and house price data Rents, lettings and tenancies Housebuilding Homelessness Household estimates and projections Planning data
Continuous REcording (CORE) Data on new lettings by registered social landlords and, increasingly, local authorities.	Household characteristics of new tenants House type, tenure and size Incomes Rent paid

Complete count so robust at all levels	
<p>Council for Mortgage Lenders 90% coverage of all regulated mortgages but only available at a cost Robust at level provided. Source: CML direct</p>	Mortgage provider Month of mortgage FTB or not Loan amount Initial mortgage rate Type of mortgage Repayment method Joint or sole income Age of main borrower Total gross income Purchase price (for sample only, c 60% coverage) New or old dwelling Type of dwelling Number of rooms Number of bedrooms Mortgage term LA area, Government Office region
<p>English House Condition Survey (EHCS) Small sample only robust at national level but data can be pooled across years to support sub-national analysis Only source for some of the data collected. Is now combined with SEH to form the English Housing Survey, a module within the Integrated Household Survey. Robust at levels provided. Source: CLG</p>	Property size (number/type of rooms and floor space) Property type and construction Household characteristics including detailed income/equivalised income, breakdown of benefits, ethnicity, economic status Detailed housing costs – rent, housing benefit and mortgages Energy efficiency indicators including CO2 emissions and SAP Tenure and occupancy State of repair and repair work required Value of housing stock for all tenures/council tax band/equity Attitude and satisfaction questions
<p>Family Resources Survey (FRS) Designed with the Department for Work and Pensions (DWP) in mind, large sample size so robust at national and regional levels and includes details of income and benefits Source: Essex Data Archive</p>	Tenure Household characteristics Rent or mortgage paid Council tax and other property charges Insurance policies Ownership of household durables Ownership of vehicles Welfare services used Childcare arrangements Help given or received by household members Health and ability to work Education, grants and loans Employment status Job description Travel to work Pensions State and other benefits Income from other sources Savings and Investments Assets
<p>Hometrack Private supplier, robust at local authority level. Not free. Source: Hometrack direct</p>	Household income House prices Sales to asking price as a % Viewings per sale Time taken to sell Price by bed count and type

	<p>Number of newly listed buyers Number of newly listed sellers Trends in the housing market</p>
<p>Housing Strategy Statistical Appendix (HSSA) Brings together statistical information from a range of sources relating to local authorities and to inform Regional Housing Strategies. For some variables this is the only source of data. Robust at LA level. Source: CLG website</p>	<p>Housing Dwelling stock in local authority at 1 April Condition of private sector housing Housing waiting list and choice-based lettings Lettings, nominations and mobility schemes Homelessness Household numbers at 30 June Houses in Multiple Occupation Local authority finance Housing capital expenditure Housing capital receipts Cash incentive scheme grants Private sector renewal assistance Private sector demolitions Mandatory disabled facilities grants Provision of affordable housing through S106</p>
<p>Land Registry (LR) The House Price Index is the most comprehensive available. It is based on completed sales. It is the only index based on repeat sales. It is robust to LA level. Source: Land Registry website</p>	<p>Mix adjusted house price index (HPI) Average house price (mean, median, LQ) Monthly change in house prices Annual change Sales volume Property type Excludes Right to Buy sales, all commercial sales and sales under a compulsory purchase order</p>
<p>Labour Force Survey (LFS) The largest regular sample survey available. Robust at regional level, and has a longitudinal element for studying changes over time. Will become part of the Integrated Household Survey and will be robust at the level provided Source: ONS</p>	<p>Age Sex Marital status and cohabitation Ethnicity Nationality and place of birth Housing tenure Employment details Individual earnings</p>
<p>Office for National Statistics (ONS) Responsible for a range of national statistics and provides a wide range of data via its web site. Data robust at level provided. See above for individual data sets</p>	<p>Census Longitudinal study (linked census data) Neighbourhood statistics Economic data e.g. consumer and producer prices and inflation Labour market data e.g. employment, unemployment and vacancies Pensions data Social capital</p>
<p>Survey of English Housing (SEH) This survey provides the most comprehensive information about households and their housing, including basic income and rent / mortgage information. But the sample is small so it is not robust for detailed analyses at regional level. It also includes attitudinal and satisfaction questions. It combined with the EHCS in April</p>	<p>Information about individuals Age, sex, legal marital status and whether cohabiting Relationship to household reference person Ethnicity Employment status Occupation Length of residence at interview address Previous residence if moved in last three years Tenure at previous address Income of reference person and partner Information about dwellings</p>

<p>2008 to form the new English Housing Survey, a module within the Integrated Household Survey</p> <p>Source: CLG</p>	<p>Dwelling type Year built Number of rooms Number of bedrooms Central heating Garden or patio Council tax band Tenure of dwelling (and whether Right to Buy purchase) Whether first dwelling purchased Information about households Source of finance for purchase Category of lender and type of mortgage Whether difficulties paying mortgage Rent Whether Housing Benefit received Type of tenancy Type of landlord Standard derived variables Size of household Type of household Number of bedrooms required to meet the bedroom standard</p>
<p>Integrated Household Survey Will be robust at regional level but not for very detailed analyses. Source: ONS</p>	<p>This new project embraces the following survey modules:</p> <ul style="list-style-type: none"> • Labour Force Survey and associated boosts • Annual Population Survey • General Household Survey • Expenditure and Food Survey • National Statistics Omnibus Survey • English Housing Survey <p>The English Housing Survey represents a merging of the SEH and EHCS. The sample size is smaller than SEH, however, it will achieve the same degree of robustness as the SEH because the sampling method is unclustered – the SEH sample was clustered.</p>

These data sources are recommended on the basis that the main use will be to measure, model, and understand affordability and related issues. Measures of affordability as discussed above include the cost of housing and the income required to afford it (or the income below which it is not affordable). The cost of housing varies by tenure, house type and size (house, flat, number of bedrooms), location, method of payment (mortgage, rent). Income varies by employment, hours worked, unearned sources of income, benefit income, number of earners in household. Affordability adds the nature of the household, number of dependents and whether in receipt of care. Data are available relating to individuals, households, dwellings and local authorities.

In all cases the best source of data depends on what it will be used for. While the census may be the ‘best’ source if data are needed at local authority level or below, it is quickly out of date so if data are only required at regional or national level, survey data will be ‘best’ because they are up to date. Equally, the Labour Force Survey has a larger sample than the Survey of English Housing, but the latter is preferable for cross-tabulation of a range of variables including household and housing characteristics.

When results from the Integrated Household Survey are available, they will replace the other national surveys as set out in table 5.1 above, making it the best data source for a wide range of uses, particularly those relating to affordability.

5.2 Recommended data sources for individual variables

Another way of thinking about the way people might use data is by identifying some key variables that they wish to use rather than necessarily knowing which data source they wish to use. The table below lists some of the key variables and gives the recommended data sources for each one, depending on how it will be used. Annex 2 provides a comprehensive list of data sources.

Table 5.2 Recommended data sources for key variables

Variable	Data source
Information about Individuals	
Age, sex, legal marital status and whether cohabiting	Census for small spatial scale APS for up-to-date information LFS
Ethnicity and nationality	Census for small spatial scale APS for up-to-date information LFS
Employment status and occupation	Census for small spatial scale LFS for up-to-date information
Length of residence and previous address if moved	Census for small spatial scale SEH for up-to-date information
Income	LFS for earning at smallest scale ASHE for earnings by place or work and place of residence FRS for up-to-date information SEH for cross tabulating with other individual and housing data
Information about dwellings	
Dwelling type	Census for small spatial scale SEH for cross tabulating with household and individual data EHCS for cross tabulating with other dwelling stock data (but small sample, national only)
Year built	SEH and EHCS (EHS)
Number of rooms	SEH and EHCS (EHS)
Number of bedrooms	Census SEH and EHCS (EHS)
Central heating	Census SEH and EHCS (EHS)
Garden or patio	SEH and EHCS (EHS)
Construction type	EHCS (EHS)
Energy efficiency indicators	EHCS (EHS)
State of repair and repair work required	EHCS (EHS)
Value of housing stock	EHCS (EHS)
Council tax band	SEH and EHCS (EHS)
Tenure (and whether Right to Buy purchase)	Census SEH and EHCS (EHS)
Whether first time buyer	SEH (EHS) CML regulated mortgage survey
Information about households	
Source of finance for purchase	SEH (EHS)

	CML
Category of lender and type of mortgage	SEH (EHS) CML
Whether difficulties paying mortgage	SEH (EHS)
Rent paid	SEH (EHS)
Whether Housing Benefit received	SEH (EHS) FRS
Type of tenancy	SEH (EHS) EHCS (EHS)
Type of landlord (social, council, private)	SEH (EHS) EHCS (EHS)
Whether household receives or gives care	FRS
Pensions	FRS
State and other benefits	FRS
Income from other sources	FRS
Savings and investments	FRS
Assets	FRS
Household income	FRS SEH, EHCS (IHS) CML for purchasers with mortgage
Information about housing costs	
House prices	Land Registry Index Mean, median and LQ prices Hometrack
Sales volume	Land Registry Hometrack
Time taken to sell	Hometrack
Private rents	Hometrack
Housing association rents	CORE
Local authority rents	HSSA LA CORE
Information at local authority level	
Dwelling stock in local authority	HSSA
Condition of private sector stock	HSSA
Housing waiting list and choice-based lettings	HSSA
Lettings, nominations and mobility schemes	HSSA
Houses in Multiple Occupation	HSSA
Household numbers	HSSA
Private sector renewal assistance	HSSA
Private sector demolitions	HSSA
Provision of affordable housing through S106	HSSA
Live tables	
Housing stock data	CLG live tables
Housing market and house price data	CLG live tables
Rents, lettings and tenancies	CLG live tables
Housebuilding	CLG live tables
Homelessness	CLG live tables
Household estimates and projections	CLG live tables
Planning data	CLG live tables
Information from ONS	
Longitudinal data from linked censuses	ONS
Neighbourhood statistics	ONS
Economic and labour market data	ONS

More information is available from the websites. To take the ONS as an example, the ONS longitudinal study contains linked census data and data on births, deaths and cancer registrations for one percent of the population of England and Wales. This longitudinal data set covers the 1971, 1981, 1991 and 2001 censuses.

The Neighbourhood Statistics part of the ONS site contains information from a variety of sources at super output level, ward, parish and local authority levels as well as other geographies. For housing, there are 33 datasets mostly from the census, the HSSA and Council Tax records. A useful PDF file for each dataset gives full details of the data including definitions and methodology where relevant.

Data on the economy includes prices and inflation, labour market indicators (employment, unemployment and vacancies), national accounts and economic activity, the balance of payments and balance of trade, plus short term indicators of retail sales, the index of manufacturing, production, productivity and services. Most of these are only available at national level, although labour market indicators are included in Neighbourhood Statistics along with other indicators such as quality of life and indices of deprivation.

Section 6: Gaps in data sources and suggestions on how to address them

This research has demonstrated that there is a huge quantity of data available on housing and affordability. These data are of varying quality. Coverage and accuracy are typically much better at national and, to a lesser extent, regional level. More information is available about demand than supply.

This section identifies gaps in the data in relation to the basic analyses of market affordability that might be undertaken at different levels of detail and spatial scale. In some cases no data are available that would be fit for purpose; in other cases data are available but not at the right spatial scale. This is particularly true of some of the major national surveys, where sample sizes are too small to be robust for all purposes even at regional level, especially in small regions such as the North East where numbers make detailed analysis less reliable.

The gaps fall into four areas, which are dealt with in turn in the remainder of the section:

1. Housing demand, including household characteristics, the backlog of unmet housing need, and migration statistics
2. Housing supply, covering tenure (Buy-to-Let, Intermediate and Social), housing stock (value, size, conversions and change of use, second homes and vacant homes, and newly built but unoccupied), split between owner occupied and private rented housing, and planning and land data
3. Housing market processes, including house prices by property size, private sector rents, incomes, funds for house purchase and user cost of house purchase
4. Context, such as school catchment areas and density of development

The gaps arise for a variety of reasons. In many cases, the national government sample surveys are an excellent data source, but are not suitable for sub-regional use or complex disaggregations at regional level because of sample size and technique. Household incomes are a good example of this. In other cases, suitable records are being collected, but the way in which specific data are recorded places significant limitations on analysis; the absence of property size information in the Land Registry dataset is a notable example here. Lastly, there are some data which are inherently difficult to gather; the size of developer land banks, information about options in land, and data on parts of the private rented sector would be good examples.

The research has drawn upon the knowledge of data providers, data users and experts in the field to suggest ways in which existing gaps can be addressed. There are four main ways in which a gap in the data might be addressed:

1. Using a proxy variable; for example, using Rent Service reference rents as a proxy for private sector rents as a whole.
2. Using a modelled variable; for example, estimating household income on the basis of individual income, age, household type and occupation.
3. Collecting data by adding questions to existing surveys or commissioning a new survey.

4. Collecting data by adding questions to collections of administrative data, for example Housing Strategy Statistical Appendix (HSSA) returns or Annual Monitoring Reports.

6.1 Housing demand

This group of data gaps is concerned with the drivers of demand for housing – that is, the longer-term trends in the population that create the need for new housing. Central problems here centre around verifying population changes between census years and judging the number of households whose housing needs are not satisfied in their current situation.

Household characteristics – age and type

Robust information with which to estimate the number of households by household type and age of the household reference person (HRP) is not available outside of the census years. This information is required to check whether the actual increase in households is in line with the household projection, and if it is not, where the differences are. There is strong evidence that between 2001 and 2006 the actual increase in households was about 30,000 a year less than the projected increase.

To understand the reasons for this, information is needed about which categories of the actual increase have fallen short. For example, middle aged people living alone might be a candidate, a group for which a particularly steep increase was projected. Evidence about this can be produced only from a large survey. The Annual Population Survey (special licence) dataset includes over 500 respondents each year in each district. This should be sufficient to provide counts by household type or age of household reference person. However, it does not provide the same level of confidence as the Census.

An important issue is the question of whether official household projections might need to be modified and if so by how much. The new Integrated Household Survey (IHS) could be used – possibly two years combined – to produce estimates of household numbers according to household type making it possible to address this issue.

Migration statistics

The lack of accurate data on migration affects the validity of overall household projections and also limits knowledge as to the profile of households moving into different areas. Data are not recorded on

- New non-British entrants to social housing by date of entry into country
- Overseas investors in the property market.
- Illegal aliens / undocumented migrants (by definition impossible to record)
- Characteristics of the immigrant population (the International Passenger Survey is very small)

Almost by definition, this information would be very difficult to obtain. In recent CORE datasets, the nationality – though not entry date – of new tenants is collected. However, an inter-departmental task force led by the Office for National Statistics is looking at ways of improving statistics on immigration, so improved data may be available at some time in the future.

Backlog of unmet housing need

There is no universally agreed method of measuring the backlog of unmet housing need, but one definition that is widely used – notably by the Barker Review of Housing Supply – is that developed in work by the Cambridge Centre for Housing and Planning Research (CCHPR). This follows on from work for the Joseph Rowntree Foundation (1995) and the Town and Country Planning Association (2000) by Alan Holmans. Here the backlog has three parts:

1. Households, and would-be households, without self-contained accommodation
2. Owner-occupiers and private sector tenants with self-contained accommodation that is unsatisfactory for their circumstances
3. Social sector tenants whose accommodation is unsatisfactory (e.g. too small).

Some of the groups in (1) and (2) were identified and estimated from a 1991 survey published in 1994 by the Office of Population, Censuses and Surveys (Routes into Local Authority Housing). This included interviews with a sample of people on local authority housing waiting lists. Categories of housing need identified from this survey included:

1. Would-be couples living apart
2. Households unable to afford mortgage payments
3. Private sector tenants applying for social housing for age or medical reasons
4. Tenants facing end of tenancy, expiry of lease, or unable to afford the rent.

The 1991 survey has not been repeated at national level, although individual authorities may have done similar work as part of their housing needs assessments. The 1991 survey has now become very dated yet no new information has since been collected. In some years the Survey of English Housing (SEH) has asked about household members on housing waiting lists. This can provide information about characteristics such as age and household circumstances, but not about the reasons for being on the list. Overall, information about the size of these groups and their reasons for being on the list can only come from a survey of housing waiting lists.

For households and would-be households living involuntarily in someone else's household, the SEH collected information which could be used to estimate numbers of couples, lone parents with children, and other adults living in someone else's house. But whether they live that way involuntarily (e.g. cannot afford separate accommodation) would require questions too detailed to include in SEH. This subject was last studied in 1990 by a follow-up survey from the Labour Force Survey housing trailers (i.e. supplement) (OPCS, 1992).

This gap could be filled by a new survey. Alternatively the new English Housing Survey could be lobbied to collect information on concealed households who would rather live independently.

6.2 Housing supply

Total overall supply includes existing stock and new build. There are data concerns in relation to both of these. As elsewhere, there is a lack of data about the size and composition of the private rented sector, and a want of detailed information about the increasingly important intermediate sector. Available sources on the characteristics of existing stock as a whole are not suitable for use at smaller spatial scales. Information about the new supply pipeline is also rather thin.

Housing stock

Up-to-date information about the existing stock of housing could be used for many purposes. One example would be to see whether better use could be made of the existing stock, which might reduce the pressure on new construction and / or on house prices. However, data on the existing stock is very limited. Below we list some of the most obvious gaps.

The value of existing stock

- Lower and upper quartile of value of the existing stock
- Median value of the existing stock
- Other measures of dispersion of value of the existing stock

The size of existing stock

- Size of existing stock by number of bedrooms and/or floorspace
- Size of the existing stock by the age of the property

Conversions and change of use

- Total net addition or losses to the housing stock by conversion
- Characteristics of property before and after conversion
- Number of housing units before and after conversion
- Tenure category before and after conversion

The priority is to get correct information about the number of flats provided by conversions and the number of houses and flats converted; and the number of flats provided by change of use. A non-residential building such as an office block 'converted' into flats would be classified as a change of use. Accurate figures for demolitions, for example, houses pulled down to make way for apartments or town houses, should also be included.

Planning permissions are the most promising source, albeit incomplete. Local authorities would have to assemble the information. This has been achieved in the South East by the Regional Assembly working with local authorities – county councils, unitary authorities and districts. The information collected included the number of units before and after conversion, which is vital. The only other information that might be collected is the number of units by tenure before and after. Design quality and energy efficiency would require a special study but ought to be assessed and monitored at local authority level by the authorities themselves.

Second homes and vacant homes

Data on second homes and vacant homes are another aspect of the existing stock. Data on both second homes and vacant homes are available from Council Tax records and are collated in the HSSA from CLG but they are not considered very reliable. This is because subsequent surveys conducted by researchers have produced different responses to the same questions, and also because some authorities simply do not complete these sections and CLG have to interpolate. Nevertheless this is the best available source at present.

The English House Condition Survey can provide estimates at national level and, with reservations about sampling error, regional level too. It cannot be used at local level.

CLG has begun to compile Council Tax second homes information, which it uses to answer questions about second homes locally. It has begun (from 2002 onwards) to produce regional and national totals. Since 2003, however, local authorities have had discretion to remove the discount on second homes. This has provided an incentive to second home owners not to register their property as such. For example, after this change took place the number of second homes in Cornwall ceased to rise while the number of holiday lets (which pay less business rates than second homes would pay council tax) rose. Yet in many cases second homes and holiday lets are the same thing (Taylor Review of Rural Economy and Affordable Housing, 2008). This casts further doubt on the reliability of council tax data on second homes.

The census figures for second homes are improbably low and the reason for this is understood to lie in the fact that the primary aim of the census is to enumerate people and therefore occupied dwellings. As a result the enumeration of vacant dwellings was variable as it relied on the enumerator deciding whether a house was vacant, a second home or a holiday home for rent. The previous 1991 census placed greater emphasis on dwellings.

Effective additional supply

The effective additional housing supply is over-stated to the extent that newly completed dwellings that remain empty are counted in local authorities' returns of completions. This is reported to be important in respect of over-supply of new city centre flats in cities and towns of the North and Midlands. It is reported that some authorities attempt to get an idea of the number of unoccupied dwellings by observing how many are without lights after nightfall. A more comprehensive and less hazardous procedure would seem needed. A post-completion survey of new dwellings would be one way forward.

Housing supply by tenure

New supply to the private rented sector

The private rented sector is a difficult sector to monitor. There is no overall data source on who buys new housing, in particular on the proportions that are bought for owner-occupation and by Buy-to-Let (BTL) investors. Any privately owned property can be rented out, either on a long or short term basis. New supply of course also includes second hand property changing tenure from owner-occupation to private rental. Nonetheless, this sector is seen as an important part of the overall housing market, as it enables greater mobility between and within regions and labour markets.

One possible proxy might be the number of BTL mortgage loans advanced or outstanding, as recorded by the Council for Mortgage Lenders (CML). The gap in the data, however, goes wider than this. No information is available about the properties purchased with BTL loans, in particular no distinction is made between new and second hand properties. The number of BTL loans issued far exceeds the increase in the number of dwellings in the private rented sector (see Annex 5), and the number of BTL mortgages terminated indicates that many BTL loans replace existing loans and thus do not finance additions to the private rented stock. To find out about additions to the private rented sector financed by BTL loans would require a survey of new BTL loans to collect information from lenders about the properties and the loans. That would be feasible, given a willingness by lenders to participate. In order to find out about additions to the private rented supply not financed by BTL loans, a research project would be required to ask a sample of landlords about properties that they had acquired for letting and how they were financed.

To find out about newly built dwellings acquired for letting, i.e. how much of the private sector's new build supply goes into the private rented sector, one possibility would be a post-completion survey of new dwellings, possibly on a sample basis. This would be a large undertaking best attempted by a national agency. It could, however, be able to address a number of currently unmet data needs; see below for further examples. A more limited and less timely option would be to use SEH (or the EHS module of the new Integrated Household Survey) to ask households moving to private rented accommodation whether they were the first to occupy the accommodation. Most households know this.

Intermediate housing

There is no single data source on the size of this sector, yet it is seen as increasingly important in terms of government policy, particularly policy on affordable housing provision. For monitoring and analysis purposes, what is needed are figures for flows of units into the intermediate housing stock by completions of shared ownership units and Homebuy purchases, and for flows out of the stock through purchasing the final share of shared ownership, and sale and repayment of equity loans on dwellings bought through Homebuy. Lastly, the market for these products would be better understood with more information on the size of the initial share being sold, and whether the property being sold was an addition to the intermediate stock, or the resale of a property previously owned as shared ownership.

For shared ownership products which involve government grant, the HomeBuy Agents collect this information from developing Registered Social Landlords (RSLs). The Regional Assemblies are trying to ensure that it is collated in order to include it in their Annual Monitoring Reports, and so it should not be too difficult to assemble. What is not currently collected is information on staircasing including final share bought. This information exists, and could be collected by means of the Housing Corporation's RSR⁴. The London Home Ownership Group currently collects a lot of this data, but only for the London area. Their approach could be rolled out nationwide.

Data on the numbers of new shared ownership units built is somewhat better. However, some shared ownership is provided by RSLs without grant, and currently this is not accurately recorded, although local authorities may collect some information as part of their monitoring of Section 106 (S106) agreements. Local authority monitoring is known to be patchy. What is not recorded is the extent of developer provided products – for example, Bellway offers shared ownership or shared equity (depending on customer preference) through its affordable housing arm, Bellway Housing Trust. Details of these schemes are not currently recorded on a systematic basis and definitions are varied. Such information could be collected through a post-completion survey as suggested above, but this would require a more significant government initiative as well as consensus on definitions of low cost home ownership built without grant.

Social sector supply

The supply of social rented tenancies comprises, firstly, additions to the stock and, secondly, units available for re-letting because the previous tenant has died or moved out of the sector. New supply in terms of additions to the stock is available from the CLG web site (live tables on housebuilding), by region and tenure. They are also available from the HSSA returns on the CLG website.

The annual supply of relets, net of transfers, in the local authority sector is available from the HSSA returns on CLG website. This is not necessarily entirely accurate, but it is the best available. The Housing Corporation produces this information for the housing association sector from the RSR. The new regulatory body will presumably have to establish a new return in order to regulate RSLs. They could also ask the right questions to produce net supply of relets. This could potentially provide more detail on the characteristics of social housing coming up for relet, for instance, distinguishing between flats and houses.

Tenure split between owner occupation and private rental housing

The split between owner occupation and private renting at LA level is a gap because estimates from the SEH are not robust at local level.

Current data on tenure comes from HSSA, which is a return completed by LAs. They do not know which dwellings are owner occupied and which are rented out (see section 4.5 above) apart from registered dwellings in multiple occupation (HMO), which are recorded separately. The census provides a tenure split – as do construction statistics on new build – but the census is quickly outdated while new build is a tiny proportion of the total stock. Current statistical practice is to estimate the size of the private rented sector from the number of private sector tenant households. The 'statistical' private rented sector includes accommodation rented from employers, including public sector employers (notably Service family quarters) and rented with business premises as well as private rented accommodation in the ordinary sense of the term.

Finding out about sale of new properties for renting would have to be part of a survey of newly built properties (see section 6.2 above), to find out when they are occupied, who

⁴ The RSR will be taken over by the Tenant Services Authority from 2009

occupies them and who owns them. This would also find out about completed but unoccupied dwellings, an important gap in information about housing supply. This, nevertheless, would be a major undertaking. It has been completed for London but only as a one-off survey. London is very different from other regions in many ways, so it would be dangerous to use London proportions as a proxy for actual data elsewhere.

Land and planning data

The land supply is extremely important for planning purposes. It is particularly significant given current government policy to increase the rate of new house building in order to reduce demand pressures and improve the affordability of housing for owner occupiers.

Data on land supply and the development pipeline, however, is poor. There is also a lack of data on developer's land banks – both strategic land and land ready for development (i.e. with outline planning permission). There is a lack of comprehensive data on developable land available by size of site. Emap Glenigan collect some data on this but it is incomplete because it is based on researchers 'picking up' information about new planning applications from local planning authorities. Emap Glenigan admit that they cannot know whether they have actually picked up all of them (although they are confident about large schemes). Their data collection is not aimed specifically at housing as its main purpose is to provide marketing opportunities for subscribers who come mainly from the construction and civil engineering sectors. This is very different from producing information directly on the housing supply pipeline. Their data starts from the initial planning application, so the time between land purchase and initial planning application is not known.

There is information about the planned land supply in both local authority development plans and in Regional Spatial Strategies and Annual Monitoring Reports. However this is currently not available nationally in a single, readily accessible place. The Strategic Housing Land Availability Assessments that local authorities are currently undertaking will help to fill this gap, but again they will need to be collated.

The confidential nature of private developer land banks makes it almost impossible to address these gaps. Local authorities, nevertheless, know about land with planning permission, and this is reflected in their HSSA returns, although the accuracy of this information is in doubt because of the high potential for double counting (an outline permission can persist for five years before it lapses). It is also not given by size of site but by number of dwelling units.

There is good data on previously used land from the National Land Use Database (NLUD), but it does not include whether that land has been allocated for development. Nor does it include ownership (apart from publicly owned land), which would be needed to explore whether the land is part of a private developer land bank, for example.

In practice, probably the best that could be done on the new housing supply pipeline is to put together outstanding permissions, new permissions issued and permissions used, cancelled or expired. This could be done using the SHLAAs that are currently being undertaken by local authority sub-regional partnerships and which, as a minimum, will identify a five year supply of available and deliverable land plus a further five or ideally ten year supply of developable land to enable the five year supply to be topped up. Collating this information could be done by CLG but would take additional resources.

There is no information about the degree of constraint placed on development by the planning system. Given that this is a key form of market regulation, this is a significant gap. Previous research has developed measures of constraint such as the ratio of agricultural land prices to housing land prices (Cheshire, 2005), or numbers of planning permissions, amount of land allocated and other local measures (Bramley, 1993a, 1993b). A more case

study approach selected 'tight' and 'less tight' local planning contexts in order to test whether a restricted land supply pushed up house prices (Gerald Eve/ Property Research Unit, 1992; Monk *et al*, 1996). To produce a comparative national indicator, more research would be required. This could be commissioned by CLG or the Economic and Social Research Council, or perhaps both.

6.3 Housing market processes

This group of data items all concern the transaction that underpin housing markets. Aside from sample surveys, relatively little is known about how housing costs relate to housing characteristics, and what resources households draw upon to meet these costs. The key gaps concern house prices by property size, private sector rents, household incomes, the funding of house purchase and the user cost of owner occupation. They are discussed below.

House prices by size of property

In any analysis of house prices and affordability, it is important to be able to compare like with like. Land Registry house prices are commonly considered to be a robust source of data because they reflect actual transactions, as opposed to asking prices or valuations. However the data are not available by size of property – either number of bedrooms, number of habitable rooms, or – as in Europe and USA – floorspace. House prices per square metre would enable international comparisons to be made. Land Registry produce a mix-adjusted index but this is still just by property type – detached, semi-detached, terrace, flat / maisonette – not the actual size of the dwelling. The Land Registry refused to ask conveyancers for the number of rooms which prevents its data being used to produce proper mix-adjusted house price indexes at local level. Sources that do collect information about property sizes, such as the English House Condition Survey (EHCS) or SEH are not robust at local level or even regional level without pooling.

The voluntary part of the Regulated Mortgage Survey (as successor to the Survey of Mortgage Lending) has numbers of habitable rooms as well as information about dwelling type and when built. Mix-adjusted house price indexes comparable to the national and regional mix adjusted index which CLG produce would technically be possible at local authority level. CLG is understood to have bought the Regulated Mortgage Survey from CML which collects it. Negotiations with CLG and CML would be required. This data, however, would only cover purchases made with a mortgage loan.

In the absence of such data, one approach is that trialled for the Joseph Rowntree Foundation by Steve Wilcox, who purchased Hometrack data which covers number of bedrooms. The drawback of this used to be that Hometrack data is based on surveyor valuations rather than actual transactions, but Hometrack now calibrate their valuations data against Land Registry data to ensure greater accuracy.

Private sector rents

Analysis of private sector rents is an important aspect of affordability. Indeed recent monitoring has shown that private renting is now generally cheaper than entry-level owner occupation. Rents data has been generally lacking until recently, aside from private sector providers such as Hometrack. However, there are gaps in the Hometrack data where districts have a very small private rented sector. The data come mainly from a survey of 8,000 estate agents. It provides a limited picture of the distribution of rents, different sizes of property, availability, etc. Hometrack data typically cost £8,000 for up to nine adjacent LAs, which may be quite expensive for a small authority. Hometrack provide a service that is tailored to the needs of local authorities. The main strength of Hometrack is that it includes a range of relevant housing market information.

In the short term, a proxy could be provided by the Rent Service's 'referred rent' (including service charges eligible for Housing Benefit) proposed by the landlord and referred by the local authority to the Rent Service for Housing Benefit determination. The Rent Service collects average rents by local authority level by the number of habitable rooms, including bedsits. The data include unfurnished and furnished assured shorthold tenancies and secure tenancies. Assumptions have to be made about the relationship between the total number of habitable rooms and the number of bedrooms in a property. As with housing for sale, it is not possible to disaggregate by size of dwellings in terms of square footage. Another drawback in using this data is that rents where Housing Benefit is required are likely to relate only to the lower end of the private rented market.

While the Rent Service data are not readily available to the public, the Dataspring website (dataspring.org.uk) provides a freely available derived dataset of private rents – and comparable costs in other tenures – by number of bedrooms at local authority level.

In the longer term, the simplest way of addressing this gap would be for the Rents Service to produce publicly accessible data on all private sector rents, not just those required for Housing Benefit cases. This would be a major undertaking, which the Rent Service might be reluctant to do, even if funded.

Incomes

Household income data or robust estimates at local authority (LA) and smaller area level, giving distributions by household types or ages, are a vital part of any kind of affordability analysis.

However, household incomes are only collected in household surveys, which are not robust at regional scale except for broad cross-tabulations. The best source of data on household incomes, the Family Resources Survey, cannot give LA level incomes directly.

Individual earned incomes are available at a lower spatial scale from HM Revenue and Customs. Therefore one way to overcome the absence of household incomes is to use modelling techniques based on individual incomes. A model developed by Professor Glen Bramley can be run for sub-LA zones, e.g. built up from wards, as in studies by Bramley and others in Scotland and nationally (see for example Bramley and Smart, 1998; Bramley and Lancaster, 1996). However, Bramley notes that Annual Population Survey (APS) and Annual Survey of Hours and Earnings (ASHE) data are subject to sampling error even at LA level.

It is possible to model household incomes from individual earned incomes (collected by Her Majesty's Revenue and Customs (HMRC) and available via their web site or Dataspring) or from individual income from earnings (ASHE). Hometrack provides modelled household income data which have been used in affordability analysis with similar results to incomes modelled from ASHE (Wilcox, 2007). Recent work supported by CLG explored the use of simulation methods to create income estimates for small areas, with quite encouraging results (Anderson, 2007). Some regions already model incomes at sub-regional level; while others have found Hometrack's local authority package useful.

One might therefore argue that the Hometrack data could fill this gap. However, the information that Hometrack provides about its sources (including sample sizes and confidence levels) is limited and the data sets are not freely available to all. CLG simulation estimates would be preferable, not least because they would be consistent with the regional income distributions currently available from CLG. Hometrack use CACI data for their income components and of course CACI could be used as an independent source. Again, the data are not free.

Funding of house purchase

Current information cannot identify instances where house buyers are able to afford prices that are larger than normal in relation to income because of gifts or loans from family or friends, or capital receipts such as inheritance. The Regulated Mortgage Survey (RMS) managed by CML provides information about the size of deposits (defined as that part of the purchase price not met from the mortgage loan), but does not identify where this deposit comes from. Information about buyers with large deposits consists of their income, age and whether they are first time buyers. Information about the sources and amounts of finance from sources other than the first mortgage is a very important gap. The SEH collects information about the use of sources of funds other than the main mortgage, but not the amounts. This could be asked in the future. A more ambitious approach would be to issue self-completion forms with mortgage documents, on a sample basis. Such a procedure was used successfully by the Ministry of Housing and Local Government in co-operation with the Abbey National Building Society in 1969.

There is also no way of knowing which members within a household are responsible for paying the rent/mortgage when payment is shared by two or more people. This is currently only collected in surveys which are not robust at LA level.

User cost of house purchase

It is one thing to analyse house prices and incomes to produce an affordability ratio, but quite another to be able to compare the relative cost of owning and renting. At the present time there is no standard approach to estimating the user cost, or even agreement as to which elements should be included.

The financial user cost of owner-occupation – that is, the total cost that must be met out of a household's budget – can be estimated given a purchase price. This is done by making assumptions about the interest rate, size of deposit and repayment term – which together determine the size of mortgage payments – and about the cost of insurances and the earned interest forgone on the deposit sum.

An alternative approach is to include the investment value of the property in the calculation of the cost. That involves deducting the capital appreciation – or adding any capital depreciation, in times of falling prices – to the financial cost of the property.

However, some of the major determining elements of these user costs – such as the cost of available credit and size of deposit put down – are likely to vary systematically between households of different types and in different starting positions in relation to the housing market. For example, older higher-income households with substantial existing equity in property are more likely to be able to secure cheaper mortgage credit and place a larger relative deposit than first-time buyers.

Dataspring calculates the user costs of owner occupation for the lower end of the market using financial comparisons. This is available at dataspring.org.uk. However, this only partially fills the gap because it uses a narrow version of user cost and only looks at the bottom end of the market.

It would be difficult to produce a national user cost variable for owner occupation, because different versions will be needed depending on the scale and purpose. There are also significant differences between different households and different starting positions in the market, especially in the current credit crunch context. Therefore research would be needed. Such research could be funded by government or the Economic and Social Research Council.

6.4. Context

School catchment areas

It is known that school catchment areas can affect the price of market housing. A location in the catchment area of a good school influences house prices at local level and is one of the variables to be included in an analysis of spatial variation in house prices. However there are difficulties in quantifying the extent of this at national level. At present the definitions of school catchment areas are held by individual LAs only, where they typically exist in the form of maps and lists of streets, or villages in rural areas. Previous studies of the relationship between house prices and catchment areas have tended to work within a single education authority area.

One way to address this gap would be for central government to collate all catchment areas. This would be a major undertaking given the fact that the definitions may well not be in terms of other boundaries (for example, wards) and so would entail the creation of an entirely new spatial data set. Since this would also rapidly become dated, the policy justification would have to be very strong indeed.

An alternative would be to use the actual locations where current pupils live to approximate the catchment areas of their schools. This could be done by Geographic Information System (GIS) analysis of existing administrative data, the Pupil-Level Annual School Census (PLASC), held by the Department for Education and Skills (DfES). This dataset, which is restricted for obvious reasons, identifies the location of each pupil and school by postcode, and hence cases can be matched to any relevant geography.

Density of new development

The main source for density figures is the Land Use Change Statistics (LUCS). Tables from this source are produced and available on the CLG web site giving annually updated information on residential development in England. The data includes an estimate of the proportion of new dwellings provided on previously developed land, and also includes density of new dwellings. However, the information is only provided regionally. Since the data are collected at a local level, LUCS could be asked to produce tables at local level. The alternative source – Emap Glenigan – is incomplete as well as expensive (see section 4.5 above).

Index of Multiple Deprivation

Other aspects of the context for housing markets can be explored using the Index of Multiple Deprivation (IMD) which is available from the CLG website for lower super output areas as well as summary scores for districts, counties and regions. The IMD identifies the 'poorest' and 'wealthiest' areas according to seven domains – income, employment, health, education, housing and services, living environment and crime. However, the Index is not itself a data source: it merely uses data from a wide range of sources in order to create a measure that can be used to identify areas of deprivation or to make comparisons between areas. It is also an example of how gaps in data can be overcome as it models data to smaller spatial scales than are actually available.

The ACORN classification produced by CACI is a system for classifying the population at postcode level into 56 types which can be aggregated up to 17 Groups and (ultimately) five basic categories: wealthy achievers, urban prosperity, comfortably off, moderate means and hard pressed. Unlike the IMD it is not designed to rank areas. It includes a measure of prosperity. MOSAIC is a similar classification produced by Experian although it classifies households rather than types of area. Both systems are widely used for marketing purposes.

Access story

Earlier the drivers of housing demand were discussed, and accessibility to employment and services were identified as influencing demand, house prices and hence affordability. Information is poor on all aspects of accessibility. Transport for London produce a Public Transport Accessibility Level dataset which is not available nationally. The Department for Transport has commissioned research to estimate the impact of transport improvements on housing values (Gibbons and Machin, 2003) but there are no obvious data sources to measure access and studies tend to define it differently. Travel to work areas reflect actual accessibility of housing to employment. It is well documented, however, that poorer households have smaller travel to work areas than the better-off. This is a gap although it is not clear how it could easily be addressed other than through one-off studies.

6.5 Summary table of gaps and how to address them

Area	Data area	GAP	Suggested work	What to use that is currently available
Demand	Household characteristics	Age and type outside census years	English Household Survey	Census only
	Migration	Information about migrants	Government working party is looking at this	Integrated Passenger Survey
	Housing need	Current unmet need	Integrated Household Survey or a new survey	Waiting lists
Supply	Housing stock	Value of existing stock, size of existing stock, conversions and change of use	Local authorities would have to assemble this information	This has been done in the South East. CLG publish net additions to stock only
	Second homes and vacant homes	Data unreliable (census, HSSA)	LAs to improve council tax data	Use HSSA data
	New supply by tenure	New supply to private rented sector	Post-completion survey of new dwellings IHS	CML data on Buy-to-Let mortgages
		Intermediate housing	Post-completion survey	RSR
		Social sector supply	Post-completion survey	RSR but does not include LA housing
Tenure	Split between owned and rented	Estimates from SEH not robust at local level	Post-completion survey	HSSA data Survey has been done for London
Land and planning data	Land supply	Emap Glenigan not complete, costly	Improved local authority monitoring	Use data on land with planning permission (watching for double counting)
Housing market processes	House prices	By size of property e.g. sq. metres	Land Registry could record this information	Hometrack data on number of bedrooms
	Private sector rents	Only available for Housing Benefit	Rents Service could produce	Use Rent Service reference rents

		cases	data on private rents used to calculate local housing allowances	(but only cover benefit cases)
	Incomes	Household survey data not robust at local level	Model household incomes or earnings or use CACI or Hometrack	Use ASHE or HM Revenue and Customs data on individuals not households
	Funding of house purchase	Cannot identify gifts or inheritance etc.	Integrated Household Survey could ask?	Regulated Mortgage Survey provides data on deposits
	User cost of house purchase	No standard approach or agreement on what elements should be included	Integrated Household Survey	Use method described on Dataspring website?
Context	School catchment areas	Held only at LA level, constantly changing	Collate information from LAs	GIS analysis of existing administrative data from DfES
	Density of new development	LUCS on CLG web site but only produced regionally	LUCS to produce tables at local level	Emap Glenigan but incomplete and costly
	Accessibility	Poor information on accessibility	Department for Transport could produce a public transport accessibility data set	Transport for London already produce this for London only

6.6 Addressing the gaps

To summarise, gaps that might be readily addressed, and suggestions on who might address them, include:

- better utilisation and collation at national level of administrative data such as Council Tax records on vacant and second homes, and Strategic Housing Land Availability Assessments on the land supply pipeline (once they are complete) – CLG, individual regions
- collation of planning data for conversions and change of use by tenure –CLG, individual regions

Important gaps that are more difficult to address are:

- income distributions at local level based on government simulations – further research would be required if HM Customs and Revenue cannot provide the data
- private sector rents, from the Rent Service –lobby for release of the data

- a survey of newly completed dwellings to collect Information on the private rented sector, ideally repeated annually to monitor change over time – lobby to include in IHS
- a repeat of the 1991 survey of waiting lists, to provide details of households wanting social housing – either through a new survey (CLG/ ONS) or through extra questions in the EHS (CLG)
- house prices by floorspace area, from the Land Registry – lobby for this data to be collected and released
- a repeat of the 1990 survey of sharing and concealed households, to find out how many of them would prefer to live independently – either through a new survey (CLG/ ONS) or through extra questions in the EHS (CLG)

Gaps that require more research are:

- user cost of owner occupation – research would be required
- degree of planning constraint – research would be required

Section 7: Conclusions and recommendations

7.1 Summary

This report has aimed to provide advice to the regions and others on the quality and availability of data relevant to the use, measurement and understanding of housing affordability. By bringing together information in one place, it will assist the regions to fulfil their statutory obligations under PPS3 and PSA20.

The research consisted of four elements:

- (i) a survey of the English regions and follow-up telephone interviews about how they use data and what problems they find when doing so
- (ii) interviews and email exchanges with data suppliers and providers
- (iii) interviews with academics working in the area of housing affordability
- (iv) evaluation of all relevant data sources and identification of gaps.

Regional representatives saw the need for more information and guidance on what data to use, and where to obtain it, in relation to affordability. While some regions already use data on affordability in a range of contexts, the aim of this report is very much to assist those coming to the topic for the first time, or wanting to make sure that they were not re-inventing the wheel.

Section 2 of the report sets out the different ways of measuring affordability, each of which is suitable for particular purposes. Thus price to income ratios are important when determining where additional housing should be located (in or near the areas where affordability is a problem) while the income required to purchase an entry-level home is relevant when determining whether new housing is affordable to first time buyers. It emphasises that there is no single 'correct' measure of housing affordability, and a 'basket' of measures may be required.

Section 3 covers the main drivers of the housing market as a background to the specification and assessment of data sources in Sections 4 and 5. These include demographic, economic and financial changes, employment and incomes. In addition, it discusses other drivers or indicators of housing demand, such as vacant dwellings and second homes, accessibility and school catchment areas. These indicators help regions and local authorities – and local housing market partnerships – to have a better understanding of the pattern of affordability in their area.

Section 4 does two things. First it outlines the data sources currently used in the different approaches to estimating future demand for housing. Second it covers a range of additional material relating to the wider aspects of demand, supply and the market. The approaches to estimating future demand include the trend-based 'demographic' method as well as more formal economic modelling approaches. The recent NHPAU publication on housing supply ranges for the regions is based on these two approaches. Section 4 explores the data sources that have been used and concludes that these are 'fit for purpose' in their different contexts.

The additional material on supply and the housing market cover:

- (i) the data required in projecting the future supply of new housing, particularly in terms of land supply and planning evidence
- (ii) alternative approaches to identifying housing need, where market indicators do not help because those who cannot afford housing are outside the market
- (iii) the data required to understand the scale and nature of the identified housing land supply and other potential sources of supply

- (iv) data on the relative importance of the different tenures, how this has changed over time; and what factors affect these trends
- (v) the impact of the affordability of private renting

Section 5 covers all the data sources relevant to housing affordability and presents the main data evaluation. This is presented in two ways. Table 5.1 starts with the data sources and lists the key variables they contain, while Table 5.2 starts with the variables and lists the best data sources that provide this information. Together this information should enable people without expert knowledge to identify the best data source for their purpose, whether they start with an idea of the data source or the variable that they need. More detail on the data evaluation is given in Annexes 1 and 2.

Section 6 looks at the gaps in the available data, including instances where there is no relevant statistical source and where existing data sources are less than adequate for affordability and related analyses. It discusses what to do when data are imperfect or unavailable and the implications for analysis. It concludes by identifying the most important gaps and ways forward for filling these gaps, indicating who might take responsibility for action.

7.2 Conclusions

In conclusion, the regions are tasked with a set of responsibilities under PPS3 and PSA20. These involve:

- identifying what is needed in order to maintain or improve affordability in general terms
- examining and monitoring relativities across their region
- collecting an evidence base to underpin their policies on planning and housing
- understanding that evidence, particularly how it is changing over time
- assisting sub-regional partnerships and local authorities to develop policies and practices that will address housing affordability over the longer term.

This report provides a useful resource for regional practitioners that will help them to meet these responsibilities. It will also be helpful to all those who are making decisions about planning and housing, whether new supply or housing affordability problems, or indeed regeneration needs. Fundamentally, this means trying to measure demand, supply and price – but affordability measures can provide a simple short cut.

Ultimately, the choice of data source and variable depends on the question being asked. So too does the degree of accuracy or robustness of the data. The data sources examined here are all of potential use, but some may be preferred to others depending on the purpose for which they are needed. Thus for some purposes national data are sufficient, and most data sources are available, accurate and robust at national level. For other purposes regional data are required, and here caution is needed when using national survey data because the sample sizes may not be large enough for detailed analysis of individual categories. Data at lower spatial scales is the most difficult, but this report will help to overcome these problems.

As we have seen, there are ways of addressing shortcomings and gaps in the available data. In some cases, local authorities themselves could improve their own data collection and monitoring. Regions are already making strides in terms of better data sources, for example setting up Regional Observatories which are constantly being updated. And central government is looking at the data it collects, both in order to rationalise and to achieve greater accuracy and accessibility.

7.3 Recommendations to address important gaps in the data

The main gaps that have been identified in relation to affordability and the requirements for additional affordable housing are set out below, together with our recommendations as to how they should be addressed:

Better collation and use of existing administrative data would meet many gaps that are considered important:

- Second homes and vacant homes

Council Tax records could be better utilised, particularly for second homes and vacant homes. While this information is published via the HSSA dataset by CLG, coverage varies between local authorities and therefore a study is required to identify how best to ensure robust reporting.

- The housing land pipeline

The Strategic Housing Land Availability Assessments could be collated nationally and regionally to provide information about land for housing

- Conversions and change of use

Planning data could equally be collated with respect to conversions and change of use by tenure. This has been done in the South East and should be repeated across all the regions.

Other gaps are more difficult to address:

- Local incomes

For local household incomes, the most promising way of filling this gap is either improving existing surveys or continuing the modelling of local incomes that CLG has already begun. This is not a gap in the sense that CACI provides local income data but at a cost. However, it is surprising that other countries freely publish these data which are effectively already collected through the tax system but the UK does not.

- Private sector rents

The private rented sector has become increasingly important in recent years and data here are poor. One improvement would be for the Rent Service to release the private rental information that they use as a basis for determining the reference rent. Private sector data such as Hometrack could be used, although this would be costly. Alternatively resources could be allocated to collect data separately. Regional Observatories could play a role here, providing consistent information that can be used across the region with easy access. Some regions are already doing this, others could usefully follow suit.

- How much new private supply is being let
- Completed but vacant dwellings

For the size of the rented sector a survey of newly completed dwellings could find out whether they were acquired for letting, who owns them, when they were first occupied and whether they are completed but vacant dwellings. This has been done

for London but only as a one-off survey. It could be repeated either nationally or in each of the other regions.

- Characteristics of households on waiting lists

A survey of waiting lists would provide details of the characteristics of households wanting social sector housing. The last one was undertaken in 1991 and provided useful insights. It should be repeated.

- House prices by size of property (floorspace area)

House prices by size of property in terms of floorspace is a major gap that makes comparisons between the UK and European countries or the USA difficult. Existing sources are partial. One solution would be to ask the Land Registry to collect, collate and publish data in this form although this would be a large task. The Regulated Mortgage Survey gives prices by property size but this only covers homes bought with a mortgage and as the numbers bought without a mortgage are significant this distorts the overall picture.

- Whether sharing and concealed households wish to live separately

A one-off survey last conducted in 1990 filled an important gap but is now out of date. A new survey would clarify the proportions of sharing and concealed households which would prefer to live independently.

Annex 1: Key Data Sources

This annex provides details of the main national data sets that are relevant in the context of housing affordability. In the interests of accuracy, much of the text has been taken directly from the data source web sites.

Annual Population Survey (APS)

The Annual Population Survey (APS) is a combined survey of households in Great Britain. The APS combines results from four different sources: (1) the Quarterly Labour Force Survey (LFS) (waves one and five for the UK funded by the ONS); (2) the English Local Labour Force Survey; (3) the Welsh Labour Force Survey (WLFS); (4) the Scottish Labour Force Survey (SLFS). It has been available since 2005, published on a quarterly basis, but each publication covers a year's data.

Topics covered include: household composition and relationships, housing tenure, nationality, ethnicity and residential history, employment and training (including government schemes), workplace and location, job hunting, educational background and qualifications. Many of the variables included in the survey are the same as those in the LFS.

APS aims to achieve interviews with approximately 65,000 households each year. This means that, in combination with data from the main LFS and its local boosts, the APS will be able to provide information on key social and socio-economic variables from approximately 215,000 households per annum.

The APS aims to provide enhanced annual data for England, covering a target sample of at least 510 economically active persons for each Unitary Authority (UA)/Local Authority District (LAD) and at least 450 economically active persons in each Greater London Borough. In combination with local LFS boost samples from Wales (WLFS) and Scotland (SLFS) the survey provides estimates for a range of indicators down to Local Education Authority (LEA) level across the UK.

At a local level, the APS survey will provide information from at least 450 economically active adults per year for London Boroughs and at least 500 economically active adults in each Local Authority District in the rest of England, and Wales. Since some districts already meet sample size requirements, addresses for the APS will only be required in certain areas of England, mainly the Shire counties.

APS datasets are produced quarterly with each dataset containing 12 months of data. More robust local area labour market estimates are available from the APS than from the main LFS.

APS is a repeated cross-sectional study, which uses multi-stage stratified random sample. The survey has a panel element, in that the households at selected addresses are interviewed annually over four waves, in the same way as for the LFS boosts. They then leave the survey and are replaced by other households.

Topics covered include: household composition and relationships, housing tenure, nationality, ethnicity and residential history, employment and training (including government schemes), workplace and location, job hunting, educational background and qualifications. Many of the variables included in the survey are the same as those in the Labour Force Survey.

Strengths: The APS is the best source presently available for robust local area labour market estimates. This is because of its large sample size and the variables covered. This

means that it is more robust at small spatial scales than the LFS yet it covers the same variables plus some additional ones e.g. relating to health.

Weaknesses: Changes to variables are made annually, making it difficult to pool data together from adjacent years. It only samples the economically active, so cannot provide details of those who are not economically active in the labour market, e.g. long term sick, or otherwise unwilling to actively seek work. The full APS is only available back to 2004, although the first annual dataset for England goes back to 2000/01. A further boost was introduced in 2004 but funding ceased in 2006.

The main dissemination route for labour market data from the APS is Nomis®. Other data from the APS, for example, on housing, education, employment, ethnicity and health is available from the Neighbourhood Statistics and National Statistics websites. Datasets are also available at the UK Data Archive and from the LFS Customer Service (SOS) team in the ONS.

Local area labour market data previously published from the Annual Local Area LFS (ALALFS) is now published from the APS. Other data from the APS is also published on the Neighbourhood Statistics website.

Tables using APS data can also be requested from the Labour Force Survey data service.

Free of charge.

Annual Survey of Hours and Earnings (ASHE)

The Annual Survey of Hours and Earnings (ASHE) provides information about the levels, distribution and make-up of earnings and hours paid for employees within industries, occupations and regions. It started in 2004 and is available annually, but the dataset has been backdated to 1998/99 using New Earnings Survey (now discontinued) data.

It is a 1% sample (approximately) of employees on the Inland Revenue PAYE register for February, with additional samples from PAYE register in April (roughly 150,000 in total). When weighted up, the total survey population is around 26 million.

The sample is based on NI letters, using the last two as the basis for identifying employees whose employers should complete the questionnaire. In the most recent year the sample size has been reduced by approximately 20%; this has been done on an industry basis to retain good coverage.

It is robust down to Local Authority or Parliamentary Constituency level but there are insufficient cases for smaller geographic areas.

Imputation and weighting is carried out on the published results. Weighting is carried out on the basis of occupation, gender, age and region to a profile from the Labour Force Survey.

Individuals with periods of no-pay are excluded from some analyses, and annual figures exclude those who have been in their job less than one year. This reduces the total weighted-up population by around 2 million.

The ASHE tables contain UK data on earnings for employees by sex and full-time/part-time workers (to LA level). Further breakdowns include by region, occupation, industry, region by occupation and age-groups.

These breakdowns are available for the following variables:

- gross weekly pay
- weekly pay excluding overtime
- basic pay including other pay
- overtime pay
- gross hourly pay
- hourly pay excluding overtime
- gross annual pay
- annual incentive pay
- total paid hours
- basic paid hours
- paid overtime hours.

Also includes percentiles (10 to 90)

Data are available by area of residence and work.

Strengths: Large sample size so robust at sub-regional level.

Weaknesses: ASHE is an individual rather than household survey, so no data are available on housing tenure, household composition and so forth. Instead these may be obtained from the Labour Force Survey (the next largest survey). Microdata are not available outside government.

The survey is available from National Statistics Online for basic LA level data and more detailed national data (e.g. by Occupation):

Microdata are not generally available, but custom analyses can be conducted by the ASHE team. New analyses involve a cost for preparation.

Free for published tables; cost for custom cross-tabulation.

CACI

CACI PayCheck profiles all 1.6 million postcodes in the UK using information on over 4 million households from lifestyle surveys and Census and Market Research data. It is available as a mean, median and mode figure for each postcode or as a PayCheck type.

A 4 page area report is provided that includes the numbers of households for one area. There are 21 income bands and this area report gives the numbers in each band. The report includes upper and lower confidence limits.

They also produce StreetValue, which they claim is the only property pricing model to be based on real information from lenders. StreetValue provides the users with the average house value; the number of properties; the predominant type of property, and the number of rented and owner occupied properties for each postcode.

StreetValue is based on previous sales in the area – if there were none within that postcode they look wider. It is about interpreting the data as much as anything.

Mostly the information is provided as a one-off payment but people also subscribe i.e. pay an annual fee.

Strengths: CACI claim that PayCheck this is the only source of income data that is considered robust at local level. The income distributions go down to £5,000 bands thus

providing much greater detail than other sources. The mean, median and mode is also an important indicator of the overall income distribution in an area. PayCheck is the only tool that takes account of state benefits and incomes from savings and investments as well as income from employment and so provides a picture of overall income flows. It is used as the basis for Hometrack's affordability analysis. Hometrack data, including CACI incomes, has also been used successfully by Professor Steve Wilcox in his analysis of the intermediate housing market undertaken for the Joseph Rowntree Foundation.

Weaknesses: Cost. There is a one-off fee for a single area analysis, or an annual subscription. A further weakness is that much of the information about the service is not transparent and only becomes available once you have subscribed.

Continuous Recording system (CORE)

The CORE dataset is created by the continuous recording of lettings and sales by housing associations and local authorities in England. It covers all housing associations with more than 250 units and over 80% of local authorities. It includes details of the dwelling and of the tenant household. A CORE log is completed by a housing association or participating local authority each time a letting or sale is made. The data on general needs housing goes back to 1989/90 while Supported Housing was added in 1996.

Variables and their definition – Lettings Log for General Needs and Supported housing:

Landlord/lettings details

- General needs
- Tenancy details
- Household details (characteristics)
- Ethnic group
- Nationality
- Disability
- Tenant's or tenant and partner's net weekly income
- Sources of income for tenant or tenant and partner
- Housing benefit eligibility
- Main reason the household left their last settled home
- The housing situation for the household immediately before this letting
- Location of previous accommodation
- Post code if moved previously from settled accommodation
- Homelessness status immediately prior to letting
- Allocation systems
- Source of referral
- Rent and other charges
- Charges for care homes
- Period of rent and other charges
- Key dates (void, repairs)
- Tenancy start date
- How many times unit has been previously been offered since becoming available for letting
- NPLG UPRN

Lettings in supported housing schemes:

- Property details
- Number of bedrooms
- Type of unit

- Type of building
- Adapted to wheelchair user standards
- Reason for vacancy
- Location of property – ONS LA code and full post code

CORE Sales Log –

- Household details (characteristics)
- Gross annual income
- Net income of purchaser(s)
- Weekly amounts of benefits
- Total savings of purchaser(s) before any deposit is paid
- Ethnic group
- Nationality
- Housing benefit eligibility
- Purchasers previously owned property
- Homelessness status immediately prior to sale
- Previous tenure
- Previous location ONS LA code and post code
- Organisation the purchaser(s) registered with for rehousing
- Disability
- Wheelchair user

Property details

- Number of bedrooms
- Property type
- Building type
- Property location ONS LA code and postcode
- NPLG UPRN
- Property adapted to wheelchair user standards

Section 1: Shared ownership/shared equity HomeBuy schemes

Section 2: Homebuy/voluntary purchase grant/Right to Buy

Section 3: Outright sales/other sales

Standard tables and analysis are available on line. These include amongst others:

- The new LA CORE Annual Digest 06/07 tables and commentary are available online. This publication contains a full summary of CORE lettings information submitted by participating local authorities during the 2006/07 year
- The 2006/7 Key Findings Digest for housing association lettings and sales is available online. The publication highlights key trends in CORE data over the past 5 years and uses CORE data to explore two topics in more detail: 1) worklessness among social housing tenants and 2) ethnicity and low cost home ownership sales.
- The CORE Annual Digest 06/07 tables and commentary are available online. The Digest includes approximately 90 summary tables including regional and trend information.
- The new CORE Strategic Summary Landlord Reports are available online. These reports draw together the key statistics from the full annual reports and include national benchmarking information
- Bespoke CORE analysis can be obtained from 'HousingFigures' - the National Housing Federation's ad hoc data supply and analysis service.

Strengths: This data set is a complete count, not a survey and is therefore robust at local authority level and below. It is the best source of information about housing association and local authority lettings.

Weaknesses: Although it covers over 80% of local authority lettings it is incomplete. It only relates to new lettings and so does not provide information about existing tenants. Some variables are under-recorded e.g. ethnicity.

The whole CORE dataset is sent by disc to researchers providing they can confirm why they want the dataset (must be for research purposes rather than commercial) and then obtain a CORE data licence and sign up to the principles of the CORE Data Sharing Protocol. A charge of £70 is made for the cd.

Council of Mortgage Lenders (CML)

CML compiles and publishes a range of statistics on the UK housing and mortgage markets including key data on mortgage lending, arrears and possessions, payment protection insurance and buy-to-let.

The Survey of Mortgage Lenders (SML) was for years one of the primary sources of the CMLs mortgage market information, including mortgage lending data. However, with effect from October 2005, the SML has now been replaced by the Regulated Mortgage Survey (RMS). However, time series data are available back to 1992 based on the SML.

The Regulated Mortgage Survey (RMS) is based on lenders providing CML with a copy of their statutory product sales data (PSD) which they also submit to the Financial Services Authority (FSA). The RMS dataset includes data from a subset of UK mortgage transactions collected by CML through RMS for Financial Services Authority (FSA). The RMS results are published quarterly.

The RMS has around 90% of market coverage (regulated lending only) therefore the data are robust at local authority level.

The questionnaire on which the RMS is based has been revised from time to time as the survey has been running in various forms since the 1960s. From 1981 sales to local authority sitting tenants were excluded, and from 1982 all sales to sitting tenants were excluded. From 1983 the standard mix of dwellings was updated annually to reflect the types of dwellings purchased during the previous three years.

The RMS, successor of the SML, has greater market coverage than the SML – around 90% compared with less than 50% for the SML, has more contributing lenders (including many of the niche market participants e.g. sub-prime and lifetime mortgages) and is likely to be more accurate (statutory backing & wider coverage). House price information, however, is provided voluntarily for RMS and is roughly 60% coverage.

The RMS, as SML before it, is based on a sample of lenders. In the figures CML makes publically available adjustments have been made to the data in order to present an estimate of the total of market lending and not just that of the sample of reporting lenders. The RMS is also subject to random sampling errors. Approximate coefficients of variation of the average dwelling prices, advances and incomes in each series vary.

The RMS excludes some types of loans that were included in SML – most notably buy-to-let and further advances. So whilst house purchase and remortgage are explicitly identified, all other loans are grouped together as 'other' in the data that is made available by CML.

The variables covered include:

- Mortgage provider
- Month of mortgage

- FTB or not (first time buyer / former owner occupier)
- Loan amount
- Initial mortgage rate
- Type of mortgage
- Repayment method
- Joint or sole income
- Age of main borrower
- Total gross income
- Purchase price (this information submitted voluntarily – lower market coverage)
- Dwelling new or old
- Type of dwelling
- Number of rooms
- Number of bedrooms
- Garage or not
- Post code of property
- Mortgage term
- Local authority area
- Government office region area

Strengths: Robust, covering all UK. A large and diverse sample of lenders submit data to RMS. Niche market mortgage providers represented. The PSD dataset on which RMS is based, is more clearly defined and structured than SML. This means that there is less ambiguity in what the data mean, and lenders submit data that are more complete and less subject to error.

Weaknesses: Only regulated mortgages are included. Some records have missing data. The number of rooms variable is defined differently by lenders. Households returning to owner-occupation after a gap are classed as first time buyers even though they have owned a property before. There may be some duplicates due to some lenders putting in multiple entries for group mortgages. Some data is input wrongly. At the moment robust data available in quarterly basis only (although this is subject to change in the future). Some lenders only provide data quarterly (although with monthly identifier). Raw data are available through the CML for a fee. Some tables are available publically from <http://www.cml.org.uk/cml/statistics>.

English House Condition Survey (EHCS)

The information collected by the English House Condition Survey (EHCS) is the main source of information on the condition and energy efficiency of housing in England. The survey builds a picture of all types of housing, whether owner-occupied or owned by local authorities, housing associations, or private landlords. From 2008 the survey has merged with the Survey of English Housing to form the English Housing Survey. The content and structure of the house condition part of the new survey will continue largely unchanged.

Annual reporting on EHCS goes back to 2003 (based on data gathered April 2002 to March 2004), but datasets are available back to 1996 (although not annually). Data are available for 1996, 2001, 2003, 2004, and 2005 at present (February 2008).

The aim is to achieve a sample of 8,000 properties each year but data are analysed on a 2 year rolling sample of about 16,000 cases to provide robust estimates at national level. Some regional analysis is also undertaken. Two different grossing factors are used, one for dwellings and another for households. The two grossed totals differ because of vacant dwellings (no household present) and a relatively small number of households in multiple occupation (HMOs) where there is more than one household per dwelling.

Vacant properties are eligible for the physical component of the survey. Second homes are considered ineligible and no interview or physical inspection is undertaken.

EHCS data is fully robust at a national level only, and use of EHCS data on a regional level can only be carried out with caution because of the limited sample size at regional level.

Extensive measures are in place to ensure data integrity and robustness. This includes comprehensive training of both interviewers and surveyors; a robust weighting strategy that takes into account the impact of non-response to all stages of the survey and extensive post fieldwork data validation and editing.

A market valuation of each of the sampled properties is also undertaken through a desk-based exercise completed by the Valuation Office Agency (VOA). This is undertaken through a dedicated web site where valuers can view a description of each property including the costs of necessary repairs together with photographs and use this to information to formulate their valuation. The VOA undertake their own Quality Assurance checks on the data. 100% response is achieved.

Topics include dwelling characteristics such as property type and size, construction type, household characteristics including detailed housing costs and income, energy efficiency indicators, tenure and occupancy, state of repair and repair work required, value of the housing stock, area description, and heating and energy.

This is collected through the three components of the EHCS:

- an initial interview survey with the occupants of these properties – to collect information on household composition, housing costs, income, attitudes to landlord and views on the local area, work done to the home, adaptation of the home for disabilities, presence of damp and mould
- a follow up physical inspection of the sampled properties (including both occupied and vacant properties) by a professional surveyor who inspects the inside and outside of the property
- a market valuation of these properties undertaken by the Valuation Office.

There is also a follow up survey with private landlords identified through the initial household interview when private renters are asked to supply the contact details of their landlord. The PLS collects information on the landlords property portfolio, their approach to property maintenance and lettings. Their views on and awareness of recent government initiatives are also sought.

The core data collected in the survey are then used to generate a number of derived variables that form the basis of most of the internal analysis and external publications. These include energy efficiency ratings (SAP); household and equivalised income measures; mortgage and rent payments; housing equity; measures of the four components of the Decent Home standard; assessment of risks under the Housing Health and Safety Rating system and costs to undertake required repairs, costs to make a property decent and costs to make a property 'safe'.

The EHCS dataset therefore consists of a set of originally collected data, and a set of derived variables that have been created by a complicated programming process. Although original data can be accessed in SPSS format by data users, the use of derived variables (such as 'household income') is recommended. The data that is presented in the derived variables has been fully cleaned.

Strengths: The main strength of the EHCS is its uniqueness. It is the only dwelling-based survey assessing dwelling conditions. It has depth of information, and it makes it possible to link the dwelling data with that of the household occupying the property (this data can be used to measure things such as fuel poverty).

The EHCS is also continuous, enabling the monitoring of trends over time. Different datasets within the survey (household and dwelling data) data can be linked together and used for cross-tabulation (for example, to calculate the price of a square metre of floorspace (£/m²) or to cross-tabulate household income with different aspects of the dwelling). Data can be used at regional level for snap-shot picture of stock condition - although analysis of trends at regional level has to be undertaken with some caution as it may o be influenced by the surveyors' individual judgments.

At the moment, the EHCS collects very detailed income data (superseded in quality only by the Family Resources Survey, FRS), mainly because of the survey's interest in assessing grant eligibility for repairs and for energy efficiency / fuel poverty. In addition to the derived 'household income' variable, EHCS also has an 'equivalised income' variable (based on FRS definition) that is available for both before and after housing costs. The income that is used for these derived variables is the net income, although the income measure used for measuring fuel poverty is different and defined separately.

Weakness: The small sample size of the EHCS restricts the potential for analysis below national level (to a great extent due to the high costs of conducting a visual dwelling condition inspection). Because of its small sample size, data is routinely analysed on a rolling 2 year sample. This sample size has not improved with the English Housing Survey (EHS), a module of the new Integrated Household Survey(IHS). Sub-national analysis is possible by pooling of data over three years although results will need to be interpreted with care, particularly for smaller regions where sampling variation is likely to be highest

The EHCS is a public dataset, and access is provided to anybody who wants it. User guide is also made available free of charge. Standard tables can be downloaded from the CLG web site free of charge.

English Housing Survey (EHS)

This survey combines the Survey of English Housing with the English House Condition Survey. The survey is based on an unclustered sample which means a smaller sample size can deliver the same or higher levels of precision as the previous larger samples. Therefore similar analyses to those using SEH or EHCS will be possible.

The EHS has three component surveys: the interview survey; the physical survey; and the market value survey.

The **interview survey** is conducted with all householders in the sample (around 17,000 households per year). It covers all the main topics from the two previous surveys:

- household composition, ethnicity, faith, nationality, economic status, education and health
- household accommodation and length of residence
- housing history and aspirations
- tenancy deposits
- disability and adaptations to the home
- rent and mortgage payments
- satisfaction with landlord/ attitudes to neighbourhood
- second homes

- work done to the home/damp
- detailed income module

The survey also includes one or two rotating modules each year on topics such as rent and mortgage arrears and waiting lists.

The **physical survey** involves a physical inspection (as conducted for the EHCS) by qualified surveyors of a sub sample of around 8,000 properties per year. The surveyor conducts an internal and external inspection of the property and plot. This data is used to derive a number of key variables for analytical purposes covering energy efficiency, carbon emissions, energy costs, local environmental quality and condition of the dwelling stock. New questions were added in 2008/09 around accessibility and cavity wall insulation to improve data modeling in these areas.

The **market value** survey comprises an independent desk based valuation for the 8,000 sub sample where a physical survey is undertaken. The sample covers all tenures and will include for example properties that rarely come onto the housing market. The data is used with the household information to derive an equity variable.

A periodic follow-up **Private Landlord Survey** is also undertaken to collect information on landlord experiences and attitudes, what type of landlord they are and why and how they became landlords.

The EHS will provide a high level of continuity with the former SEH and EHCS and so permit a number of key on-going time series. The unclustered sample will also deliver greater precision for a number of measures from the household interview. The rotating module will provide flexibility in introducing emerging policy topics to the survey. The large household interview sample will continue to enable detailed analysis of sub-groups as for the SEH.

Family Resources Survey (FRS)

The Family Resources Survey collects information on the incomes and circumstances of private households in the United Kingdom (or Great Britain before 2002-03). It is sponsored by the Department for Work and Pensions. Whilst it was designed with the needs of DWP specifically in mind, it contains information which is of interest to researchers and analysts from a wide range of disciplines in both the public and private sectors. It is a continuous survey with data published annually. It was launched in 1992 and the data go back to 1993/94.

The target sample size is around 28,000 households per year. It is robust at national or regional level, depending on the inquiry. For regional level analysis of population sub-groups (such as lone parents or minority ethnic populations) etc., data from adjacent years must be pooled together to achieve sufficiently large sample size.

The questionnaire is divided into three parts:

- Household schedule: asked of one household member, preferably the Head of Household;
- Individual schedule: asked of each unit (comprising parent/s and dependent child/children) in the household; and the
- Assets Questionnaire: asks about the total value of respondents' investments if relevant.

The FRS is a large survey both in terms of the number of respondents and the length of the questionnaire.

Main Topics:

- Tenure and Household Characteristics
- Details of Rent or Mortgages paid
- Council Tax and other Property Charges
- Insurance Policies
- Ownership of Household Durables
- Ownership of Vehicles
- Welfare Services used
- Childcare arrangements
- Any help given or received by members of the household
- Health and Ability to work
- Education, Grants and Loans
- Employment Status and Current Situation
- Job Description
- Travelling to Work
- Pensions
- State and Other Benefits
- Income from Other Sources
- Savings and Investments
- Assets

FRS is a specialised survey of incomes and some categories of expenditure set up initially to study eligibility and take-up of State benefits. Income information is collected in considerable detail, as is information about assets relevant to eligibility for Income Support. Information is collected about mortgage payments in sufficient detail to estimate eligibility for Income Support for Mortgage Interest (ISMI). FRS is a better source of information about income according to housing tenure than SEH owing to more detail and thoroughness.

Strengths: The main advantages of the FRS are the large sample size (This sample size is considerably larger than the Expenditure and Food Survey and so allows for more reliable analysis of particular subgroups of the population, such as lone parents, as well as regional level data), the ability to identify benefit units and the inclusion of detailed questions relating to social security benefits.

Weaknesses: The main weakness of FRS is that data must be pooled from adjacent years for robust regional level analysis for population sub-groups. The data source is not recommended for sub-regional analysis. The FRS is also known to under-report receipts for most benefits when compared with administrative data on benefits.

Each year the FRS database is deposited at the Data Archive at Essex University. In order to access the data, an Athens password is required. This can be done by filling in an application form and is free of charge.

Hometrack

Hometrack is a private sector data provider. It provides affordability data as part of a package of housing data. Incomes are updated annually and house prices are updated monthly. It also undertakes analysis on behalf of clients including local authorities or partnerships.

The data are compiled from around 8,000 estate agents on a monthly basis. Data are available at LA, ward, statistical boundary level, and lower output level. While the dataset is based on valuations data, it is calibrated to be consistent with Land Registry data.

Variables:

- Household income (from CACI Paycheck).
- House prices per sq. ft. (from estate agents and the Land Registry).
- Private sector rents (from estate agents)
- Sales to asking price as a %
- Viewings per sale
- Time taken to sell
- Price by bed count and type
- Number of newly listed buyers
- Number of newly listed properties
- Trends in the housing market

Strengths: The main benefit is that it provides not only affordability data but other housing data. It is very well targeted at the data requirements of local authorities. The analysis provided is very comprehensive and high quality e.g. mapping data. The housing intelligence service is available by subscription.

Weaknesses: The cost of the data (it can be very expensive for small authorities to purchase) but perhaps more importantly the lack of transparency about how robust the data are (e.g. is 8,000 estate agents a representative sample or a population?). There is no information about coverage or details of the methodology (unless you are a subscriber to the Housing Intelligence System).

Housing Strategy Statistical Appendix (HSSA)

The purpose of the HSSA is to bring together statistical information relevant to the formulation of Housing Strategies. Presentation of the data is intended to enable local authorities, partners and other interested parties to gain a quick overview of each local authority's position and to inform development of the Regional Housing Strategy. It brings together data items from many different housing areas in a concise reference document.

The data are available annually from 1st April 2005 to 31st March 2006. They are also available on line back to 2001/02, and earlier years (as Housing Investment Programme or HIP) are available from CLG on request.

All of the 354 local authorities in England complete Sections A through N (see below) for all properties contained within them. It is intended to represent the total population of homes by local authority.

However, some sections are more robust than others. Some data are included from other sources where these are seen as more robust. For example, the data on housing association stock is variable and CLG has published an Advice Note recommending using the RSR (Regulatory and Statistical Return) produced for the Housing Corporation as a more accurate source.

CLG have system validations – totals add up, correct units of measurement etc. – as 'hard' checks within the on-line form. There are more checks once the data have been collected but these are only manual checks by individuals using their judgement, they are not systematic, so it depends on who is doing them.

CLG also do imputations for missing data in order to produce national and regional totals, but this data is not published, only the totals. There are mechanisms for imputing missing data: use last year's figure, or use regional proportions and major urban categories

Variables:

- Section A: Dwelling stock position in authority area at 1 April 2006
- Section B: Condition of private sector housing
- Section C: Housing waiting list and choice-based lettings
- Section D: Lettings, nominations and mobility schemes
- Section E: Homelessness
- Section F: Household numbers at 30 June
- Section G: Houses in Multiple Occupation (HMOs)

Financial Sections:

- Section H: Housing capital expenditure
- Section I: Housing capital receipts
- Section J: Cash incentive scheme grants
- Section K: Private sector renewal assistance (expanded to remove the need for P4 return)
- Section L: Private sector demolition
- Section M: Mandatory Disabled Facilities Grants

- Section N: Provision of Affordable Housing

Strengths: HSSA has complete coverage of a wide range of housing related data at local authority level. It is the only source for much of this data (although not all). It can be compared over time (although it is not a time series and the variables and definitions can and do change).

Weaknesses: It is dependant on who in the LA is filling in the data. It is not a live system. Some sections are better than others. The mandatory cells are the most accurate. Sections A, B and N are of reasonable quality. Some of the financial data should be used with caution because it is harder for CLG to check its accuracy. The rest is relatively consistent over time

Land Registry (House Price Index)

The Land Registry registers title to land in England and Wales and records dealings (for example, sales and mortgages) with registered land.

The Land Registry House Price Index (HPI) is the most comprehensive available. Using Land Registry's data set of completed sales, it is the only index to be based on repeat sales. The index includes figures at national, regional, county and London borough level. It is a repeat sales regression (RSR) index, measuring average price changes in repeat sales on the same properties, ensuring a like for like comparison.

The House Price Index (HPI) is published on the 20th working day of every month. The report will include data up to the final day of the preceding month. Previously this was published quarterly.

The HPI is available from April 2000. This is the earliest that Land Registry's price paid dataset is currently available. The previous residential property price reports did go back to 1995 but this data is not in a format that can be analysed.

The HPI is calculated by using Land Registry's own 'Price Paid Dataset'. This is a record of all residential property transactions made in England and Wales since April 2000. At present

it contains details on over seven million sales. Of these, just under 1.5 million are identifiable matched pairs, providing the basis for the repeat-sales regression analysis used to compile the index.

The HPI is robust at Local Authority level (there is insufficient data for post code level).

The HPI is calculated by using Land Registry's own 'Price Paid Dataset'. This is a record of all residential property transactions made in England and Wales since April 2000. At present it contains details on over seven million sales. Of these, 1.5 million are identifiable matched pairs, providing the basis for the repeat-sales regression analysis used to compile the index. The Land Registry HPI is designed to remove the influence of seasonal variations on house price changes.

In addition the following variables are available:

- Average house price (mean, median, LQ)
- Monthly change
- Annual change
- Sales volume
- Property type (Detached, Semi-detached, Terraced, Flat/maisonette).

The Land Registry HPI is derived from all residential property transactions registered with Land Registry since April 2000.

The following are excluded:

- all commercial transactions
- transfers, conveyances, assignments or leases at a premium with nominal rent which are:
 - 'right to buy' sales at a discount
 - subject to a lease
 - subject to an existing mortgage
 - to effect the sale of a share in a property
 - by way of a gift
 - by way of exchange
 - under a compulsory purchase order
 - under a court order
 - to Trustees
 - vesting deeds
 - transmissions or assents of more than one property
 - leases for seven years or less

Strengths: The Land Registry House Price Index (derived from the Land Registry's Price Paid dataset) is the most comprehensive available. Using Land Registry's data set of completed sales, it is the only index to be based on repeat sales. The index includes figures at national, regional, county and London borough level. HPI uses a sample size that is larger than all other statistical measures available. The Land Registry data set is the only complete record of residential property transactions in England and Wales.

Weaknesses: Some of the data is available only for sale. No part of the data can be reproduced without permission from the Land Registry. House prices are not available by size of property, only by house type (see above).

Labour Force Survey (LFS)

The LFS is a quarterly survey of private households, with questions concentrated on labour market activity. It is quarterly, though much of the analysis is annual. It has been collected since 1979. It became an annual survey in 1985, and quarterly in 1992. 'Housing Trailers' were attached to the LFS in 1981, 1984, 1988 and 1991 (see notes on Survey of English Housing). Its use as a housing data source is more recent (*Housing Statistics 2002*).

The main LFS has a sample of about 60,000. The sample is of addresses and households; it does not cover communal establishments. Recruits to the survey are re-interviewed for a further four successive quarters, which enables the use of the LFS as a short-term longitudinal source.

The main LFS is robust at regional level, though for very small population categories grouping of adjacent years would probably be desirable.

In the medium term the LFS will become part of the Integrated Household Survey.

LFS collects the basic demographic information about age, sex, marital status and co-habitation and ethnicity. It also asks about nationality and place of birth. Where the place of birth is outside the UK, the year in which the sample member came to the UK is asked for, which allows household formation by recent inward migrants to be compared with less recent in-migrants and the UK-born population. Housing tenure is asked for; but no housing detail.

Most of the specific detail in the survey relates to employment: occupation, industry, hours worked, previous work, and so forth. Individual income earned in the course of employment is collected, but not total household income, including that from benefits, pensions and so forth.

Strengths: The LFS is the largest regular sample survey available. It provides a relatively long time-series, and headline results are published in a timely way. It is household-based and includes information about age, composition and so on. It has a longitudinal version (over three or five quarters) although the data for this are somewhat more limited. It is a robust source of data on the labour market and particularly, ethnicity and in-migration.

Weaknesses: The coverage of some non-labour-market topics (such as housing) is not as detailed as some other survey sources, such as the Survey of English Housing. As a result, the SEH is often the preferred data source for housing-related analysis.

Other: Office for National Statistics (ONS)

The Office for National Statistics (ONS) is the government department responsible for collecting and publishing official statistics about the UK's society and economy. It provides access to a number of other data sources, as well as statistics that are available only via ONS.

ONS carries out the ten-yearly census and reports on vital events. Its website provides access to data that is made available in many different formats.

The ONS was created by merging the Central Statistical Office, the Office of Population Censuses and Surveys and the statistics division of the Department of Employment. In 2000 the Statistics Commission and 'National Statistics' were established, but in 2008 the

Statistics Commission was replaced by the UK Statistics Authority, an independent body. ONS is the executive office for the UK Statistics Authority.

ONS produces several publications including Social Trends, Regional Trends, Family Spending, Economic and Labour Market Review and Pension Trends, all of which are available on line as well as in hard copy. These provide analysis as well as data.

There are links to other web sites such as the General Registrar (registration of births and deaths) on the Home Office site and Nomis which provides customised on-line data on the labour market at various spatial scales.

Access is easy as most data can be downloaded on line.

UK Snapshot provides snapshot statistics about:

- Economy
- Environment
- Health
- Labour market
- Population
- Society
- Travel and Tourism

Neighbourhood statistics

Statistics for an area by Local Authority, ward, New Deal for Communities and Super Output Area / Neighbourhood Summary by postcode

Economy statistics

- Prices and inflation
- Labour market
- National accounts / Economic activity
- Balance of payments and trade
- Short term indicators

Census data.

2.1. Guide to census data

An introduction to accessing 2001 Census data.

2.2. Data sources

Census data sources for those familiar with 2001 Census products, definitions and topics.

2.3. Catalogue

A full list of products and services from the 2001 Census.

2.4. Census data 1801 - 1991

A list of the data available, online and in printed volumes, from censuses up to 1991.

2.5. Data comparability over time

Explanation of the difficulty in making comparisons between data from different censuses.

2.6. Help

Contact information for Census Customer Services.

2.7. 2001 Census

More information about the conduct and results from the last census.

Time Series Data is a free-of-charge service that gives access to a comprehensive database of more than 40,000 time series from major ONS economic and socio-economic releases. The complete histories of the time series are available in the database.

Virtual Bookshelf provides access to abundant data from numerous different sources. This data is classified by theme.

Index of Services

The experimental monthly Index of Services is currently being developed - this guide includes definitions, progress reports and methodology.

Labour Market

Labour market statistics are structured around a supply and demand model known as the labour market framework. Descriptions of the concepts used in the framework, the sources behind the data, the methods used in data production and analysis, and the publication of labour market statistics are available on-line at www.nomisweb.co.uk

Longitudinal Study

The ONS Longitudinal Study (LS) contains linked census and vital event data for one per cent of the population of England and Wales. Information from the 1971, 1981, 1991 and 2001 Censuses has been linked across censuses as well as information on events such as births, deaths and cancer registrations.

Pensions data sources

Describes the main data sources that deal with pension statistics. The section gives information about the main sources of UK pension statistics, with links to data and references to relevant tables.

Social Capital

Information about social capital, including any developments in progress. This includes details of ONS's Social Capital Project and a bank of questions from the 15 social capital data sources.

Social capital describes the pattern and intensity of networks among people and the shared values which arise from those networks. While definitions of social capital vary, the main aspects are citizenship, neighbourliness, trust and shared values, community involvement, volunteering, social networks and civic participation.

The interactive question bank builds on the social capital survey matrix, allowing on-line users to compare and contrast the wording of the questions from fifteen surveys that measure aspects of social capital.

There is an introductory guide to large scale surveys providing social capital data. This provides details of surveys which include any questions on aspects of social capital and links to related documents. This guide extends the information in the interactive question bank.

Strengths: The National Statistics website is easy to use; an extremely large amount of data is available on line; there are handy pdf files explaining the data sources and methodology for most of the tables; and there is a strict code of practice ensuring reliability and transparency of the data provided.

The Neighbourhood Statistics part of the website is very useful in terms of the scale at which analysis is possible.

Weaknesses: Although many variables and tables are right up-to-date e.g. retail price index by month, others are out of date, e.g. benefits payments and the retail price index relates to 2006. Some parts of the web site are easier to use than others.

Other: Communities and Local Government (CLG)

In addition to being responsible for the Survey of English Housing, the English House Condition Survey and the Housing Strategy Statistical Appendix, CLG is the Government Department responsible for planning policy and building regulations in England. It also provides statistics on planning and housing.

Housing and planning research and statistics are available on-line from (<http://www.communities.gov.uk/housing/housingresearch>). This website provides access to a range of information on the following topics:

Housing statistics
Housing data projects
Live tables

The best source for CLG data on housing is the live tables, which contain data specially chosen and collected from a variety of sources such as administrative returns from local authorities, Communities and Local Government and external surveys and other government departments. The tables present the latest available summary or detailed data relating to an important or interesting aspect of a housing statistics topic.

These tables contain either basic essential data or specific data for trend analysis. They are updated as soon as new or revised data become available (hence the word "live" is used).

Tables and detailed notes and definitions are grouped and can be found by housing topic area. The contents of *Housing Statistics Annual* are directly produced from these tables and notes.

Live tables provide data on the following:

- Housing Stock
- Housing Market and House Prices
- Rents, Lettings and Tenancies
- Housebuilding
- Homelessness
- Household Estimates and Projections
- Planning statistics

All live tables, as well as an index and a list of definitions are available from the CLG Housing website.

Strengths: It is obviously extremely useful to have tables ready to download in excel spreadsheet or other formats on such a wide range of topics. In most cases time series are available, and in many cases this is the only source of the information on a particular topic. Where the tables provide regional or sub-regional data, they are robust – if not robust they are not published.

Weaknesses: The CLG website is difficult to navigate and the live tables seem to be accessible by several different routes, producing tables in a slightly different format. A problem with live tables is that they are updated as and when better information becomes available. This means that the figures can vary week to week.

The timeliness of the data varies between data sets. Some variables are not available at sub-national level.

It would be useful to have an overview or user guide to some of the sets of tables, for example, house prices or starts and completions, where there is a whole set of tables and it is not clear which are the best to use. It is not always clear whether the data are for England, UK, etc.

The live tables do not include all of the data items CLG collect. This information can be found in the *Overview* for each topic or in the *Data Collection and Forms* section, accessible from the *Housing Statistics* Homepage

All other topic based information such as statistical releases and reports can be found from the *Housing Statistics by topic* homepage. The CLG website also provides links to other Communities and Local Government (e.g. *Survey of English Housing*) and external sources of statistical information. The *Index of Data* is also a useful guide to help to locate the information.

Survey of English Housing

The Survey of English Housing is a continuous survey with annual reports giving key tables and analysis. It is available annually from 1993/94 onwards. Some key variables may also be available from earlier surveys, such as: the 1977/8 National Dwelling and Housing Survey (NDHS) and the Housing Trailers to the 1981, 1984, 1988 and 1991 Labour Force Survey (LFS).

The achieved sample size around 19,000 to 20,000 annually. It is a sample of the private household population. Residents of institutions (primarily residential care homes; nursing homes; and long-stay hospitals) are therefore not included.

This small samples size makes the survey robust at national level. At regional level grouping of adjacent years may be necessary if information is required about groups that are small in proportion to the total of households, e.g. households that moved during the year, or for three-way tables, e.g. housing tenure x age of household reference person x type of household.

Regional level robustness is dependent on the size of the region, and questions such as how to deal with weighting and grossing when pooling data from several years can be problematic. Caution should thus always be exercised when pooling data, and one should also ensure that the questions that are used from adjacent years are phrased in the same way.

SEH collects a very large amount of information, too much to describe fully here. The information is about the private household population and their housing, provided they are occupying their usual place of residence. Interviews are not conducted of people living (temporarily) in a second home. Vacant dwellings cannot be covered.

Information is collected individually about all household members, defined by the usual residence and including persons usually resident but not present at the time of the interview. The information collected allows the 'household reference person' (formerly household head) to be identified. Information about individuals, and hence about the household reference person, comprises:

- Age, sex, legal marital status and whether or not cohabiting
- Relationship to the household reference person
- Ethnicity
- Employment status – full time work, unemployed, retired etc.

- Occupation – used to code into National Statistics Socio-Economic Classification
- Length of time resident at the interview address
- For household reference person, previous residence (if they last moved within the past three years), their housing tenure at the previous address, including not being the reference person at previous address (i.e. new household).
- Income of the reference person and spouse or partner (if there is one, otherwise income of reference person only)
 - Information collected about the dwelling comprises:
 - Dwelling type – detached, semi-detached, etc. or not self-contained
 - Year dwelling was built (as a range)
 - Number of rooms
 - Number of bedrooms
 - Whether there is central heating (questions are no longer asked about other ‘amenities’ such as a bathroom or WC inside the dwelling because the number without them is very low).
 - Whether there is a garden or patio (not asked in all years)
 - Council Tax band
 - Tenure of the dwelling – owned with a mortgage, rent from RSL, etc. If owner occupied, did the reference person rent it before buying it, and if so was it rented from a local authority or RSL (to identify Right to Buy purchasers)
 - For owners, is the present dwelling the first that the household reference person has owned – so first-time buyers can be identified.
 - Information collected about households (by specific question as distinct from inference from the information about individuals) comprises:
 - Source of finance for the purchase, main and other
 - Where the main source is a mortgage, category of lender and (in recent years) type of mortgage
 - Whether difficulties about mortgage payments
 - For private sector tenants
 - rent
 - whether Housing Benefit received
 - type of tenancy – assured, assured shorthold, regulated etc.
 - type of landlord
 - There are several standard derived variables, including:
 - Size of household
 - Type of household
 - Number of bedrooms required to meet the bedroom standard (the benchmark for determining overcrowding)

Strengths: One of the main strengths of the survey is its flexibility, i.e. the ability to add/take out questions, and change questions year by year. This, however, may be compromised when the SEH and the EHCS are merged to become a single module (EHS) within the Integrated Household Survey (IHS). In future there may be less flexibility than at present.

At the moment, the SEH consist of three different data ‘sets’: Case-by-case data on household level, as well as individual data on members of the household (such as age, relation to HRP); tenancy data set (one ‘case’ per tenancy); family unit data (separating family units who live in one household from one another). Both the tenancy and the family unit data can be seen a strength of the SEH. The tenancy data set may not be available from the EHS.

Weaknesses: The main weakness of the survey is the sample size (which is too small to produce robust data at local level). Sampling is done by using cluster sampling, which can be seen as another weakness.

Another weakness of the SEH is the housing association tenant / local authority tenant issue, though this problem affects all household surveys that seek to distinguish between tenants of local authority housing and housing association housing. The problem is that a number of social tenants living in HA accommodation think they are still living in council housing.

A set of live tables can be downloaded from the web for recent years. These vary in how far back in time they go depending on the topic. Most go back to 1999/2000.

Following publication of the main report (produced annually) the full datasets are deposited with the Data Archive at Essex University.

Annex 2: Overview of data sources by topic

This section sets out in tabular form variables identified by NHPAU and the research team that are of use in measuring, understanding and modelling affordability. The reader can look up the variable they are interested in to help identify likely data sources. More detail on the pros and cons of the main data sources are discussed in Annex 1.

How to use the table

There are nine sections to the table:

- 1 Data on households/properties/HRP
- 2 Data on individuals
- 3 House-building
- 4 House Prices (Flow data – coverage: properties sold/let)
- 5 Land data
- 6 Demolitions
- 7 Conversions
- 8 Change of Use
- 9 Demographics

The table shows each individual variable. The data sources are ranked in order of recommendation, taking into consideration

- ease of use
- quality of data
- assuming that one wants to look only at that one variable at regional level.

There are several important factors to consider in choosing data sources:

In many cases it is necessary to “cross-tab” the data. This can be done for data sources where unit level data is recorded, *within the data source itself*. These data sources are shown with an asterisks (*) within the table. This will, however depend upon the data source and the way in which data is released. Published reports and on-line data suppliers will generally give most of the popular cross-tabs (usually just two-way ones). For data not published it will be necessary to obtain the full data set. Some further information on the availability of full data sets it contained within Annex 1/ NHPAU or the data owners/providers should be able to provide fuller information.

It is never possible to cross-tab between data sources, because (with the exception of Land Registry house price data) the data is always anonymised when released at unit level.

For some data sources, the only published data may be in aggregate form. It is not possible to “cross-tab” this data.

Example

If household income by tenure is required, you would look up the two variables separately. This shows that SEH, GHS, EFS, EHCS and FRS are all listed under both variables so could all provide this data.

When data is used in this way the spatial scale at which this is robust will vary depending on sample size in the chosen variable.

Secondary suppliers of data

In many cases survey data can be most easily accessed through data suppliers – Neighbourhood Statistics, the Office of National Statistics, the Local Housing Markets database, the National Housing Federation, Hometrack and Nomisweb are the main ones used. These have not been listed separately within this table except in cases when they also supply information not found within separate data sources – such as population projections, or aggregate data not published elsewhere.

Local authority data, when collated centrally, is listed under that central source (e.g. HSSA or CLG live tables) as this is the easiest way to access such information. In cases where the tables refers to LAs or individual RSLs this is because no central source is known to exist and the data may only be held locally.

The Integrated Household Survey is being rolled out over the course of 2008 and 2009. It integrates the following surveys:

- Labour Force Survey (LFS)
- General Household Survey (GHS)
- Expenditure and Food Survey (EFS)
- Omnibus Survey (OMN)
- The English Housing Survey (EHS) – a new survey which brings together the Survey of English Housing (SEH) and the English House Condition Survey (EHCS)

For the purposes of this table, however, these surveys have been identified separately as that is currently the form in which the data exists. Similarly, NROSH and the Blue Book have not been included in this list as they are not yet in operation.

The National Statistics omnibus survey is not included in the table as the data it collects varies with each wave.

Table A1

1. Data on households/properties/HRP	Source
Age of household head	Census* APS* SEH* GHS* EHCS* FRS* BHPS*
Ages of all household members	Census* FRS* EHCS* APS* ⁵ SEH* ⁶ GHS* ⁷ BHPS* ⁸
Cost of property (when bought)	EHCS* BHPS* ⁹
Council tax band	Valuation Agency ONS SEH* EHCS* FRS* BHPS*
Current interest rate on mortgage	SEH* ONS Table 7.1L building societies interest rates ¹⁰
Current share owned of shared ownership properties	Individual RSLs SEH
Current value of property	EHCS* ¹¹ SEH* ¹² BHPS* ¹³
Date built/ age of property	CLG live tables EHCS* ¹⁴ SEH* ¹⁵
Date of entry into country/ international in-migrant within last 12 months	LFS*
Date purchased	SEH* ¹⁶ EHCS* BHPS*
Date current home valued	EHCS* ¹⁷ SEH* ¹⁸ BHPS* ¹⁹

⁵ Age of adults and the youngest and oldest children only.

⁶ Age of youngest child, interviewee, HRP and oldest married/single person only

⁷ Age of the adults and youngest child only

⁸ Age bands of children and adults into pensionable and over 75s only.

⁹ If bought during period when panel member partaking in BHPS

¹⁰ Monthly, quarterly and annual averages are calculated from these. National data.

¹¹ Independent valuation of all sample properties by the VOA

¹² Occupiers estimate. Not full data. Lots of missing data

¹³ Banded data

¹⁴ Banded data

¹⁵ Banded data. Exact ages largely unknown

¹⁶ Collected some years only. Banded data. Exact date unknown

¹⁷ Independent valuation by VOA of all sample properties as at 1st October each year

¹⁸ Occupiers estimate. Not full data. Lots of missing data

Deposit paid on previous rented home	SEH*
Difficult to let dwellings	HSSA
Dwelling type	EHCS* Census* CLG Housing (Stock live tables) SEH* FRS* GHS*
Economic status e.g. full-time employed/ part-time employed/ unemployed/ retired/ full-time student and employed/ full-time student and not employed/ permanently sick/disabled etc. ²⁰	Census* LFS* FRS* SEH* GHS* EHCS* GLA* ²¹
Ethnicity	Census* SEH* APS* GHS* LFS* EHCS* FRS* BHPS*
Family type	Census* SEH* APS* GHS* FRS* ²² EHCS*
Family size	Census* APS* SEH* GHS* EHCS* FRS* ²³
Homeless households in temporary accommodation (statutory)	CLG Housing (Homelessness) HSSA
Household composition - number of adults, number of dependent children	Census* LFS* APS* SEH* FRS* GHS* EHCS*
Household composition - number of families ²⁴	Census* SEH* APS* GHS* FRS* ²⁵ EHCS*
Household composition - number of full-time employed/ part-time	Census*

¹⁹ Banded data

²⁰ Different data sources have varying definitions

²¹ London only

²² Defines as “benefit units”

²³ Defines as “benefit units”

²⁴ The definition of a family varies between data sources.

²⁵ Defines as “benefit units”

employed/ unemployed/ retired/ full-time students/ permanently sick/disabled etc.	APS* EHCS* SEH* ²⁶
Household income/earnings	FRS* EFS* SEH* GHS* EHCS* CACI ²⁷ Experian ²⁸ ONS ²⁹
Household type, e.g. one-person household, couple household with no children, couple household with dependent children, couple household with non-dependent children etc. ³⁰	Census* SEH* APS* GHS* LFS* FRS* EHCS* BHPS*
Households on housing register	HSSA ³¹ SEH*
Income of those paying the mortgage/rent	FRS* ³² SEH* ³³
Initial share purchased of shared ownership properties	Individual RSLs
Landlord's age(s)	SEH* ³⁴ BHPS* ³⁵
Landlord's income	SEH* ³⁶ BHPS* ³⁷
Landlord's other sources of income e.g. benefits, employment, pension etc.	SEH* ³⁸ BHPS* ³⁹
Landlord's own tenure	SEH* ⁴⁰ BHPS* ⁴¹
Landlord's country of residence (if don't usually live in England)	No known data source
Location - Region	Census* SEH* EHCS* FRS* LFS* APS* BHPS*

²⁶ Number of employed people, students and number of economically inactive only

²⁷ CACI sell data tailored to client requirements – some local authorities purchase their income data

²⁸ Experian sell a range of data, including income data

²⁹ Experimental ward-level estimates

³⁰ Different data sources have varying definitions

³¹ Also indicates the number of bedrooms needed

³² Asks whether non household-head contributes, and asks their income

³³ Has income of HRP and income of HRP & partner as well as household income

³⁴ This would be based on using the data on whether the households interviewed own property elsewhere. It is therefore not a sample of all landlords as it excludes landlords not living in England and commercial landlords.

³⁵ See footnote 34

³⁶ See footnote 34

³⁷ See footnote 34

³⁸ See footnote 34

³⁹ See footnote 34

⁴⁰ See footnote 34

⁴¹ See footnote 34

	GHS*
Location - Housing Market Areas	Census* ⁴²
Location – local authority	Census* SEH* EHCS* FRS* LFS* APS* BHPS* GHS*
Location - ward	Census*
Location - Lower Layer Super Output Areas	Census* BHPS* ⁴³
Location - Middle Layer Super Output Areas	Census* BHPS* ⁴⁴
Location - Output areas	Census* BHPS* ⁴⁵
Low demand dwellings	HSSA
Mortgage - monthly repayments on mortgage (total and interest only)	SEH* FRS* ⁴⁶ EFS* ⁴⁷ EHCS*
Mortgage - term of initial mortgage	FRS*
Mortgage - term remaining on current mortgage	FRS* SEH* ⁴⁸ BHPS*
Mortgage – amount outstanding	SEH* FRS* EHCS*
Mortgage - whether household in arrears	SEH*
Mortgage - whether difficulties making mortgage repayments	SEH* BHPS*
Mortgage information - other	SEH* ⁴⁹ FRS* ⁵⁰ EHCS* ⁵¹ BHPS* ⁵²
Number of properties landlord owns (and tenure of other properties)	SEH* ⁵³ Paragon ⁵⁴
Quality measures e.g. central heating	EHCS* ⁵⁵

⁴² Assuming that HMAs are defined as entire SOAs

⁴³ Not publicly available but analysis can be done at this level on request.

⁴⁴ Not publicly available but analysis can be done at this level on request.

⁴⁵ Not publicly available but analysis can be done at this level on request.

⁴⁶ Weekly housing costs by tenure

⁴⁷ Can be worked out from weekly data

⁴⁸ Could calculate this because the date it was taken out and the date it runs to are both recorded

⁴⁹ Whether household has used current home as security on loan or remortgaged, eg equity release.

Whether h/h has mortgage. Who mortgage is from. Type of mortgage. How expect to repay mortgage. Whether discounted.

⁵⁰ Type, details of mortgage protection policies, year took out mortgage, second mortgages

⁵¹ Type, source, length, present monthly repayments, original deposit for mortgage, whether still original mortgage, whether mortgage reduced by a lump-sum payment

⁵² Asks how home came to be owned outright. Also asks amount borrowed at start of mortgage and additional borrowing since

⁵³ This would be based on using the data on whether the households interviewed own property elsewhere. It is therefore not a sample of all landlords as it excludes landlords not living in England and commercial landlords.

⁵⁴ Calculates the average number of properties landlord has using Paragon mortgages.

	Census* ⁵⁶ SEH* ⁵⁷ GHS* ⁵⁸ BHPS* ⁵⁹
Overcrowding	SEH* ⁶⁰ GHS* ⁶¹ Census* ⁶² BHPS* ⁶³
Payment shared by one, two or more people	FRS* ⁶⁴
Percentage deposit for initial mortgage	EHCS* ⁶⁵
Presence of garden, garage, conservatory etc.	EHCS* BHPS* ⁶⁶ SEH* ⁶⁷
Previous tenure (if any)	SEH* ⁶⁸ BHPS*
Reason for tenure change e.g. repossession, first-time buyer, relationship breakdown	SEH* ⁶⁹ BHPS* ⁷⁰
Rent	CLG Housing (Rents, lettings and tenancies) FRS* EFS* EHCS* SEH* ⁷¹ BHPS* RSR ⁷²
Savings	FRS* EHCS*
School catchment area	Individual LAs
Second home – whether property is a second home	HSSA Census*
Second-home owners usual place of residence	SEH* ⁷³
Second home - reason for second residence e.g. holiday home, work away from home, etc.	SEH* ⁷⁴
Size - floor space (sq. m)	EHCS*
Size - number of bedrooms	EHCS* SEH*
Size - number of rooms	Census*

⁵⁵ Whether central heating or not, heating fuel, heating system

⁵⁶ Whether household has central heating and also whether has sole use of bath/shower and toilet

⁵⁷ Whether household has central heating

⁵⁸ Whether household has central heating

⁵⁹ Whether household has central heating

⁶⁰ Can use the bedroom standard or number of persons per room

⁶¹ Persons per room and bedroom standard

⁶² Can use number of persons per room. Using the bedroom standard involves estimates as to the number of bedrooms.

⁶³ Asks whether there is a shortage of space

⁶⁴ Whether non household-head and outsiders contribute to mortgage

⁶⁵ Can be derived

⁶⁶ Whether has own or shared terrace or garden

⁶⁷ Collected some years only

⁶⁸ If moved within last 5 years

⁶⁹ If moved within last 5 years

⁷⁰ Asks reason for move of panel members, rather than HRP

⁷¹ Social tenants only

⁷² HA housing only

⁷³ This question is asked in relation to second homes that the household own/rent elsewhere.

⁷⁴ This question is asked in relation to second homes that the household own/rent elsewhere.

	EHCS* SEH* BHPS*
Source of deposit e.g. gift from family, equity from another property, savings etc.	SEH*
Source of income e.g. benefits, employment, pension, property investment etc.	FRS* APS* EHCS* FRS* LFS* BHPS*
Source of valuation of property e.g. sale, re-mortgage etc.	EHCS* ⁷⁵ SEH* ⁷⁶ BHPS* ⁷⁷
State of repair	EHCS* HSSA ⁷⁸ ONS ⁷⁹ BPSA* ⁸⁰
Shared ownership - stepping up over time and time between steps of shared ownership properties	Individual RSLs
Shared ownership - whether property was previously owned as shared ownership	No known data source
Tenure (inc detailed split – shared ownership, owned with and w/o mortgage, HA or LA)	SEH* EHCS* FRS* Census* HSSA ⁸¹ BHPS* ⁸²
Time in current home	SEH* ⁸³ FRS* BHPS*
Time in current tenure	BPHS* ⁸⁴
Transport links (availability of public transport/ travel times)	Census* ⁸⁵
Travel to work area	ONS (travel to work areas)
Type of Intermediate housing e.g. Home Buy, Key Worker	RSR
Type of landlord – private, friend, employer or commercial	Census* SEH* EHCS* BHPS*
Type of mortgage e.g. fixed rate, variable rate, tracker etc.	SEH* EHCS*
Type of mortgage that landlord has e.g. commercial, buy-to-let, homeowner	No known data source
Vacant dwellings - whether property vacant	HSSA ⁸⁶ ONS ⁸⁷ CLG Housing (Stock) ⁸⁸ Census*

⁷⁵ Independent valuation by VOA of all sample cases as at 1st October each year

⁷⁶ Occupiers estimate. Not full data. Lots of missing data

⁷⁷ Banded data

⁷⁸ Stock condition

⁷⁹ Modelled estimate of disrepair

⁸⁰ Condition of stock (unfit, in need of major/minor repairs) by tenure and location

⁸¹ Private ownership, LA and HA only

⁸² Data on shared ownership problematic

⁸³ Banded data

⁸⁴ For Panel members only, not necessarily the HRP

⁸⁵ Records way in which each household member travels to work

	EHCS
Vacant dwelling - reason for vacancy e.g. probate, for sale, investment, between tenants etc.	Individual LAs ⁸⁹
Vacant dwellings - length of time vacant	CLG Housing (Stock)* ⁹⁰ HSSA* ⁹¹
Whether let through letting agency	SEH*
Whether property purchased under RTB	SEH* CLG ⁹²
Whether problems paying for housing in last year	BHPS*
Aggregate data:	
Lower and upper quartile of value of <i>existing stock</i>	Some data available from EHCS
Mean of <i>existing stock</i>	VOA ⁹³ CLG Housing (Stock) ⁹⁴ Some data from EHCS
Median value of <i>existing stock</i>	No known data source
Other measures of dispersion of value of <i>the existing stock</i>	No known data source
2. Data on individuals	
Age	Census* APS* ASHE* LFS* BHPS* HMRC* ⁹⁵
<i>De facto</i> marital/ civil partnership status e.g. single, cohabiting, married etc.	Census* LFS* APS* BHPS*
Economic status e.g. full-time employed/ part-time employed/ unemployed/ retired/ full-time student and employed/ full-time student and not employed/ permanently sick/disabled etc.	Census* LFS* APS* BHPS* ASHE* ⁹⁶
Ethnicity	Census* LFS* BHPS* HMRC ⁹⁷
Income	FRS* APS* LFS* BHPS* ASHE* ⁹⁸ HMRC* ⁹⁹

⁸⁶ Gives numbers of HA, LA and private vacant dwellings, derived from council tax records

⁸⁷ From council tax records, to ward level

⁸⁸ Regional

⁸⁹ Some individual LAs, or HMAs have carried out research in this area, but no nationwide data is believed to exist

⁹⁰ Records whether private sector homes have been vacant over 6 months or not

⁹¹ Records whether private sector homes have been vacant over 6 months or not

⁹² Total number of sales by district per year, so can derive total number of properties that have been bought in this way

⁹³ Average prices, by property type for 124 locations in England and Wales

⁹⁴ 'Simple average'

⁹⁵ People for whom HMRC hold data only

⁹⁶ Identifies full-time and part-time employed people

⁹⁷ People for whom HMRC hold data only

Income source e.g. benefits, employment, pension, property investment etc.	APS* DWP LFS* BHPS* HMRC* ¹⁰⁰
Location	Census* BHPS* ¹⁰¹ HMRC* ¹⁰² ASHE* ¹⁰³
Number of dependent children	Census* APS* LFS* BHPS*
Previous tenure (if any)	BHPS*
Reason for tenure change e.g. repossession, first-time buyer, relationship breakdown	BHPS*
Savings/ assets	BHPS* ¹⁰⁴
Sex/gender	Census* APS* LFS* BHPS* ASHE*
Tenure e.g. social rented, private rented, private rented (with housing benefit), owned outright, owned with a mortgage, living with parents/friends but not responsible (or jointly responsible) for payment of mortgage/rent	Census* ¹⁰⁵ BHPS*
Time in tenure	BHPS*
Aggregate data:	
Lower and upper quartile earnings	ASHE HMRE LFS
Median earnings	ASHE HMRE LFS
Mean earnings	ASHE HMRE LFS
Other measures of dispersion of earnings	ASHE
3. House-building	
Assessment of the housing market and building trends	HBF
Brownfield/ greenfield site	Regional AMR LUCS
Commitments	Regional AMR
Completions	CLG live tables on housebuilding

⁹⁸ Income earned from employment (not self-employment) only – mean, median, quartiles by local authority

⁹⁹ People for whom HMRC hold data only

¹⁰⁰ People for whom HMRC hold data only

¹⁰¹ Region

¹⁰² People for whom HMRC hold data only, by region and district

¹⁰³ District and region

¹⁰⁴ Asks how much saves each month from earnings and also value of lump sums, e.g. inheritance

¹⁰⁵ Does not record whether receives housing benefit

Density	Regional AMR LUCS
Details of sites/schemes	Emap Glenigan ¹⁰⁶
Location	CLG Housing (Housebuilding) ¹⁰⁷
New supply	Regional AMR
Property characteristics	CLG Housing (Housebuilding) CLG Live tables ¹⁰⁸
Shared ownership - initial share purchased	Individual RSLs only
Shared ownership - final share bought	Individual RSLs only
Source of funding, e.g. S106, housing corporation	HSSA ¹⁰⁹
Starts	CLG live tables on housebuilding
Tenure category e.g. social rented, intermediate housing, private rented/owner-occupied etc.	CLG Housing (Housebuilding) ¹¹⁰ HSSA ¹¹¹ Regional AMR ¹¹² Housing Corporation ¹¹³ RSR ¹¹⁴
Total provision	Regional AMR
Type of Intermediate housing e.g. HomeBuy, Key Worker	Housing Corporation ¹¹⁵ English Partnerships ¹¹⁶
• source of land e.g. public, private	No known data source
• design quality e.g. good, excellent etc.	Housing Corporation? ¹¹⁷
• whether in flood plain/greenbelt	LUCS
4. House Prices (Flow data – coverage: properties sold/let)	
Date purchased	Land Registry* ¹¹⁸ CML ¹¹⁹ Halifax House price index ¹²⁰ Nationwide House price index ¹²¹ CORE* ¹²²
Discount received under RTB	CLG ¹²³
Location - region	Land Registry* CLG Housing (Housing market & House prices) CML ¹²⁴ Halifax House price index Nationwide House price index CORE* ¹²⁵

¹⁰⁶ Emap Glenigan for a fee will provide information on schemes, sites, builders etc

¹⁰⁷ District and region

¹⁰⁸ Number of bedrooms, by tenure and type, and by region

¹⁰⁹ All funded housebuilding, and grants for private sector

¹¹⁰ Social rented and private ownership only at district level

¹¹¹ Identifies affordable (rented and shared ownership) and private sector completions

¹¹² Identifies affordable (not specified) and market

¹¹³ Collects data on social rented and shared ownership built with grant funding

¹¹⁴ HA housing only

¹¹⁵ Key worker living and other LCHO

¹¹⁶ First-time-buyers initiative

¹¹⁷ If grant funded

¹¹⁸ Records date that Land registry records were altered, usually very close to the date of purchase

¹¹⁹ Gives numbers of mortgages approved monthly

¹²⁰ Monthly data

¹²¹ Quarterly data

¹²² Shared ownership only

¹²³ Average discounts by number of sales in each year

¹²⁴ Average prices by region.

¹²⁵ Shared ownership only

Location – district	CLG Housing (Housing market & House prices) Land Registry* CML CORE* ¹²⁶
Location – address	Land Registry*
Mortgage information	CML ¹²⁷
New build/ existing property	Land Registry* Nationwide House price index
Property type (detached, semi, terrace, flat)	Land Registry* Nationwide House price index CORE* ¹²⁸
Previous tenure e.g. social rented, private rented, private rented (with housing benefit), owned outright, owned with a mortgage, living with parents/friends but not responsible (or jointly responsible) for payment of mortgage/rent	SEH* ¹²⁹ BHPS* ¹³⁰ CORE* ¹³¹
Purchase price	Land Registry* CLG Housing (Housing market & House prices) ¹³² CML ¹³³
Purchase finance e.g. cash purchase, with mortgage, shared ownership etc.	CML ¹³⁴ Land Registry ¹³⁵ Individual RLS LHOG
Size - floor space (sq. m) and/or number of bedrooms	CML ¹³⁶ CORE* ¹³⁷
Social sector lettings	HSSA CLG Housing (Rents, letting and tenancies) CORE* ¹³⁸
Transaction type e.g. private sale, estate agency, right-to-buy, family transaction, divorce settlement, re-mortgage	CLG ¹³⁹ SEH* ¹⁴⁰
Type of purchaser e.g. first-time buyer (never owned), 'first-time' buyer (previously owned or jointly owned), owner mover, second home, buy-to-let etc.	SEH* ¹⁴¹ CML ¹⁴² Nationwide House price index ¹⁴³ CORE ¹⁴⁴

¹²⁶ Shared ownership only

¹²⁷ Gives median percent advance and income multiple. Also data on arrears and repossession and on type of mortgage.

¹²⁸ Shared ownership only

¹²⁹ Can work out which households moved from renting into ownership in last 5 years, and which formed directly into ownership

¹³⁰ Can work out the previous tenure of those now in owner-occupation

¹³¹ Shared ownership only

¹³² Quarterly mix adjusted house price index by region

¹³³ Averages, data from the Regulated Mortgage survey

¹³⁴ Size of deposit, but no data on cash purchasers

¹³⁵ Cash purchases and purchases with a mortgage

¹³⁶ Number of rooms

¹³⁷ Shared ownership only, number of bedrooms

¹³⁸ RSLs and participating LAs only

¹³⁹ Right-to-Buy figures only

¹⁴⁰ Asks about details of why the previous seller sold the house (e.g. divorce, moving abroad, etc)

¹⁴¹ Can work out which households moved from renting into ownership in last 5 years, and which formed directly into ownership

¹⁴² Identifies first-time-buyers (i.e. those not moving directly from owner-occupation), home movers and Buy-to-let

Aggregate data:	
Lower and upper quartile house prices	Land Registry Hometrack CML
Median house price	Land Registry Hometrack CML
Mean house price	Land Registry CLG Housing (Housing market & House prices) ¹⁴⁵ Halifax House price index ¹⁴⁶ Nationwide House price index CML
Other measures of dispersion of house prices	Land Registry Hometrack CML
Affordability measures	Hometrack# RICS ONS Halifax House price index Nationwide House price index
House price indices and annual inflation	CLG Housing (Housing market & House prices) Halifax House Price index Nationwide House price index
Mortgage payment protection insurance	Nationwide website
Building insurance	Nationwide website

5. Land data	
For land with planning permission (past or current):	
Address/location	NLUD ¹⁴⁷ Emap Glenigan ¹⁴⁸
No of units or site area	Local planning offices
Status of planning permission, e.g. implemented, extant etc.	Local planning offices
Time between purchase and initial application	No known data source
Time between initial application and planning permission granted	CLG Development Control Statistics BERR ¹⁴⁹
Time between planning permission granted and start	Local planning offices
Time between start and completion ¹⁵⁰	Local planning offices
Type of land e.g. greenfield/ brownfield	NLUD ¹⁵¹
Type of planning permission, i.e. outline, full	Local planning offices
Land available for housing by:	
Brownfield/ greenfield site	NLUD ¹⁵² CLG Planning Previously-developed brownfield land*

¹⁴³ Identifies first-time-buyers (i.e. those not moving directly from owner-occupation) and home movers

¹⁴⁴ Shared ownership only

¹⁴⁵ Simple average house prices. Not mix adjusted

¹⁴⁶ Mix adjusted prices at regional or postal district level

¹⁴⁷ Brownfield only

¹⁴⁸ Emap Glenigan for a fee can track any site from permission stage to completion

¹⁴⁹ By Local Authority

¹⁵⁰ This is not always straightforward as some sites may have multiple applications

¹⁵¹ Brownfield only

	LUCS
Location	NLUD ¹⁵³ LUCS
Site area	NLUD ¹⁵⁴
Whether in flood plain	LUCS
Land ownership e.g. public, private	NLUD ¹⁵⁵
Percentage of all land available	LUCS
Land Banks	
Amount of land held	No known data source
Extent of land banks by developer and/or location	No known data source
Reason for land bank, e.g. operational, investment, planning issues etc.	No known data source
Land Prices	
Average land prices by region	Valuation Office Agency
6. Demolitions	
Location - district	HSSA RSR ¹⁵⁶
Number of housing units demolished	CLG Housing (stock) ¹⁵⁷ HSSA Regional AMR RSR ¹⁵⁸
Property characteristics	No known data source
Reason for demolition	CLG Housing (stock) ¹⁵⁹
Tenure category e.g. social rented, intermediate housing, private rented/owner-occupied etc.	HSSA ¹⁶⁰
7. Conversions	
Loss in housing units due to conversions over time (as far back as possible) by:	
• location	LAs/Regional Assemblies
• number of housing units	LAs/Regional Assemblies
• property characteristics	LAs/Regional Assemblies
• tenure category e.g. social rented, intermediate housing, private rented/owner-occupied etc.	LAs/Regional Assemblies
Gain in housing units due to conversions over time (as far back as possible) by:	
• location	LAs/Regional Assemblies
• number of housing units	LAs/Regional Assemblies
• property characteristics	LAs/Regional Assemblies
• tenure category e.g. social rented, intermediate housing, private rented/owner-occupied etc.	LAs/Regional Assemblies
Net gains to housing stock due to conversions	CLG Housing (stock) ¹⁶¹ Regional Assemblies

¹⁵² Brownfield only

¹⁵³ Brownfield only

¹⁵⁴ Brownfield only

¹⁵⁵ Brownfield only

¹⁵⁶ HA stock only

¹⁵⁷ National estimate only

¹⁵⁸ HA stock only

¹⁵⁹ National estimate only, distinguishes slum clearance and other

¹⁶⁰ Identifies social sector and private sector only.

¹⁶¹ National estimate only

8. Change of Use	
Number of housing units changed use	CLG Housing (stock)¹⁶²
Loss in housing units due to change of use over time (as far back as possible) by:	
Location	Regional AMR ¹⁶³
Property characteristics	No known data source
Tenure category e.g. social rented, intermediate housing, private rented/owner-occupied etc.	No known data source
	No known data source
Gain in housing units due to change of use over time (as far back as possible) by:	
Location	Regional AMR ¹⁶⁴
Property characteristics	No known data source
Tenure category e.g. social rented, intermediate housing, private rented/owner-occupied etc.	No known data source
9. Demographics	
Population over time (as far back as possible and also forecasts and projections) by:	
Location	ONS ¹⁶⁵ APS ¹⁶⁶
Person characteristics	ONS APS ¹⁶⁷ GAD Projections ¹⁶⁸

Households over time (as far back as possible and also forecasts and projections) by:	
Household type, e.g. one-person household, couple household with no children, couple household with dependent children, couple household with non-dependent children etc.	CLG Housing (Household estimates and projections) Census ¹⁶⁹
Location	Census ¹⁷⁰ CLG Housing (Household estimates and projections)
Household Formation	
Household formation rates	CLG Housing (Household estimates and projections)
Headship rates	Census CLG Housing (Household estimates and projections)
Household type of newly formed households, e.g. one-person household, couple household with no children, couple household with dependent children, couple household with non-dependent children etc.	SEH*

¹⁶² National estimate only

¹⁶³ Regional level data

¹⁶⁴ Regional level data

¹⁶⁵ Down to district level

¹⁶⁶ Back in time, not projections

¹⁶⁷ Age and sex

¹⁶⁸ Now ONS

¹⁶⁹ Over time in the past

¹⁷⁰ Over time in the past

Individual income of newly formed households	SEH*
Person characteristics of newly formed households including current and previous tenure	SEH*
Internal Migration: In- and Out-flows over time and also forecasts and projections by	
Current tenure	SEH* ¹⁷¹
Location of origin	Census* ¹⁷² ONS ¹⁷³ HMRC* ¹⁷⁴
Location of destination	Census* ¹⁷⁵ ONS
Person characteristics	Census* ¹⁷⁶ ONS ¹⁷⁷
Reason for move e.g. to seek employment, to care for family etc.	SEH* ¹⁷⁸
International Migration: In- and Out-flows over time and also forecasts and projections by:	
Current tenure	GLA ^{179,180}
Location/country of origin	ONS ¹⁸¹ Census ¹⁸² GLA ¹⁸³
Location of destination	ONS ¹⁸⁴ GLA ¹⁸⁵
Occupational classification (SIC) prior to move	International passenger survey
Person characteristics	ONS ¹⁸⁶ International passenger survey GLA ¹⁸⁷
Previous tenure e.g. social rented, private rented, private rented (with housing benefit), owned outright, owned with a mortgage, living with parents/friends but not responsible (or jointly responsible) for payment of mortgage/rent	No known data source
Qualifications	International passenger survey
Reason for move e.g. to seek employment, to care for family etc.	International passenger survey
Total numbers	ONS ¹⁸⁸

¹⁷¹ Not projections

¹⁷² Not projections

¹⁷³ To MSOA level, total flows only

¹⁷⁴ Estimates

¹⁷⁵ Not projections

¹⁷⁶ Not projections

¹⁷⁷ By age and sex

¹⁷⁸ Not projections

¹⁷⁹ London only

¹⁸⁰ In-migration only

¹⁸¹ Total international migration figures at regional level

¹⁸² In-migration only

¹⁸³ London only

¹⁸⁴ Total international migration figures at regional level

¹⁸⁵ London only

¹⁸⁶ Total international migration figures at regional level by age and sex

¹⁸⁷ London only

¹⁸⁸ Compiled from the International Passenger Survey and Home Office data

Annex 3: Glossary

Affordability

A measure of whether households can access and sustain the costs of private sector housing. Frequently used in the context of market housing affordability.

Affordable housing

Affordable housing is defined according to Planning Policy Statement No. 3: Housing as either social rented housing or low cost home ownership (see Shared ownership and Shared equity HomeBuy below).

Affordability ratio

The ratio of house prices to incomes.

Average (mean) income

Average (mean) income is calculated by adding together all incomes and dividing the sum by the number of the responses or, when income bands are used, by multiplying the total number of responses in each income range by the mid-point of the range and dividing by the total population. The average income is often higher than most people's incomes, because very high income outliers skew the result higher (in contrast, the median income resists such skew).

Bedroom Standard

'Bedroom standard' is used as an indicator of occupation density. The standard number of bedrooms allocated to each household so that no one has to share a bedroom unless they are a) a couple, b) both aged under 10 or c) both aged under 21 and of the same sex. No more than two people may share a bedroom.

Brownfield land

Land that has been previously developed.

Concealed household requiring self-contained accommodation

A person or a family unit living within another household wanting to move to their own accommodation and form a separate household (e.g. adult children living with their parents). A concealed household is one that does not include the household reference person.

Communal establishment

An establishment providing managed residential accommodation (for e.g. prisons, large hospitals, hotels, nursing homes, halls of residence). Small hotels and guesthouses are treated as communal establishments if they have the capacity to have 10 or more guests. Sheltered housing is treated as communal establishment if less than half the residents possess their own cooking facilities.

Decent Homes standard

A decent home can be defined as one that meets the following criteria: (1) meets the current statutory minimum standard for housing; (2) is in a reasonable state of repair; (3) has reasonably modern facilities and services; (4) provides a reasonable degree of thermal comfort. In terms of fitness, each home must achieve **all** of the following: (1) be structurally stable; (2) be free from disrepair; (3) be free from damp levels that could affect the health of the tenant; (4) have lighting, heating and ventilation; (5) have a piped supply of wholesome water; (6) have facilities for preparation and cooking of food; (7) have a suitably located toilet; (8) have a bath **or** shower with hot

and cold water; (9) have suitable drainage. The decent homes standard applies only to social housing, not to the private sector.

Dependent child

A dependent child is defined in the Census as a person in a household aged 0 - 15 (whether or not in a family) or a person aged 16-18 who is a full-time student in a family with parent(s). Definitions of 'dependent child' for migration or benefit purposes may be different.

Deprivation indices

Indices that are used to measure poverty and deprivation in a given area/locality. Deprivation indices commonly combine a series of measures of 'deprivation'. All Census based deprivation indices are generally composed of surrogate or proxy measures of deprivation rather than direct measures.

Equivalised income

Equivalised median income is defined as a household's total disposable income divided by its equivalent size to take account of the size and composition of the household and is attributed to each household member (including children). The equivalised income of large households is therefore less than their actual income.

Greenfield land

Land that has not been previously developed.

Homeless household (Statutory)

A household is accepted as statutorily homeless by the authority if it meets the criteria set out in the Housing Act 1996.

Household characteristics

This term is commonly used to refer to different characteristics that can be used to classify households into categories according to the type of the household (see below for a definition), the household size, household income and the ethnicity, age, and gender of the head of the household/household reference person.

Household expenditure

Any spending done by a person living alone or by a group of people living together in shared accommodation and with common domestic expenses.

Household type / household composition

A household comprises either one person living alone or a group of people (not necessarily related) living at the same address with common housekeeping - that is, sharing at least one meal a day or sharing a living room or sitting room. A household can consist of one or more benefit units. Households are commonly classified according to their composition / type (such as single, couple, couple with children, lone parent, multi-adult, multi-adult with children, single pensioner, pensioner couple). The categories used to classify households by type/composition vary between different data sources.

Household in need

A household that is currently living in unsuitable housing.

Housing register

A register of people waiting for affordable housing. It may have two components: a list for those not currently occupying affordable housing (more properly known as the

Housing Register) and a Transfer List for those tenants who wish to move to another affordable home within the same district.

Inadequate housing

Housing which is inadequate or unsuitable to the needs of the household comprising a range of criteria on house condition, size, cost and security of tenure.

Intermediate housing

Housing at prices or rents above those of social rented but below market prices or rents. This can include sub-market renting or low cost homeownership.

Lower quartile

This refers to the bottom 25% of a distribution. The upper 25% (or the bottom 75%) is the upper quartile. Distributions may also be divided into deciles, such that the lowest decile is the bottom 10% of the distribution.

Mean

A term used to refer to an average calculated by dividing the total by the number of cases in the sample. Mean income, for example, would refer to the total income of area residents, divided by the number of residents.

Median income

The median is the middle number present in a set of data when the incomes of all household are arranged in an order of highest to lowest (if number of values in a set is even, the average between the two middle values are used). This is considered by many statisticians to be a better indicator of what is a “typical” income than the average household income as it is not dramatically affected by unusually high or low values.

Neighbourhood Statistics

ONS provided data on a range of topics down to ward and parish level, available on-line.

Opportunity cost of a deposit

The money a home-owner could be earning on the deposit that was used to buy their home if they had instead invested it in a savings account or ISA etc.

Overcrowding

Overcrowding occurs when, using the **bedroom standard** (see above) there are insufficient bedrooms in a home for the people living there. Overcrowding can also be measured by using a ‘persons per room’ measure.

Owner occupier user cost

The cost of owner occupation that takes into account mortgage repayments (often based on purchasing a lower quartile house prices) and other housing related costs (e.g. property insurance, mortgage payment protection insurance and opportunity cost of deposit).

Persons per room

The number of persons per room is a way of determining whether a home is overcrowded. Rooms that are counted include living rooms, bedrooms and large kitchens. For the space and floor area calculations:

- children under one year old are ignored
- children under ten years old and over one count as a half
- rooms under 50 square feet are ignored.

As a general rule:

- 1 room = 2 people
- 2 rooms = 3 people
- 3 rooms = 5 people
- 4 rooms = 7.5 people
- 5 or more rooms = 2 people per room.

But the floor area of a room also determines how many people can sleep in it:

- floor area 110 sq feet (10.2 sq metres approx) = 2 people
- floor area 90 - 109 sq ft (8.4 - 10.2 sq m approx) = 1.5 people
- floor area 70 - 89 sq ft (6.5 - 8.4 sq m approx) = 1 person
- floor area 50 - 69 sq ft (4.6 - 6.5 sq m approx) = 0.5 people

Planning application

In order to be allowed to build on land, or change the use of land or buildings, a planning permission or planning consent is required. Since the Town and Country Planning Act entered into force in 1947, any new "development" has required planning permission, which can be obtained by submitting a planning application.

Quartile

In descriptive statistics, a quartile is any of the three values which divide the sorted data set into four equal parts, so that each part represents 1/4th of the sampled population. Each quartile value is the median of that quarter of the sample, for example, the lower quartile value is the median of the lower half of the data.

Regional dummy

A "dummy" is a way of adding a variable into a regression model. It is used for categorical variables: those which represent a set of mutually exclusive states (like male/female or north/midlands/south) rather than a quantity (like age, income or price). A regional dummy will capture the influence of all differences between regions that are not recorded in other variables within the model.

Re-let

Local Authority or RSL rented accommodation that becomes vacant due to the departure or death of a previous tenant; therefore the accommodation can be re-let to another tenant or new applicant on the Housing Register.

Residual income

Net income less all housing costs (such as mortgage outgoings or rent).

Shared ownership

A form of tenure whereby households partly own and partly rent their homes. Shared ownership properties are usually offered for sale by housing associations, who remain the owner of the second part of the property, on which they charge a sub-market rent. The borrower is usually entitled to buy subsequent shares in the property, a process known as 'staircasing'.

Shared equity HomeBuy

Shared Equity is the basis for the Government's Open Market HomeBuy scheme which is aimed primarily at Social Housing tenants and Key Workers – particularly in London and the South East and others who are considered to be in housing need. With the Open Market HomeBuy scheme, the successful applicant does not need a deposit and for the first five years of property ownership need not pay any interest on the equity loan. Shared equity requires a first time buyer to take out a mortgage for between 60 and 75% of the property and then also take out a lower cost or free top up loan or equity loan to fund the majority of the rest. In return for borrowing both the

mortgage and equity loan to afford that property the lenders will require a proportion of any increase in equity/value of that property when the occupier comes to sell it.

Sheltered accommodation

Sheltered Housing provides accommodation to meet the needs of elderly people or others with special needs to enable them to live an independent life within the community. It commonly takes the form of a group of self-contained properties (flats, bungalows, bedsits), which are specifically designed to meet the needs of the residents. They are usually served by a resident warden and connected to the warden by means of an intercom system.

Supported housing

Housing specially designed, managed or adapted to meet the needs of particular client groups requiring additional management support (such as people with learning difficulties, young people at risk, the frail elderly, the blind, mentally ill or mentally handicapped people and women with dependents requiring refuge from harassment and violence). Accommodation can be grouped individual self-contained flats, or may be shared housing, with or without shared communal facilities, or can make use of the ordinary housing stock via "floating support". Support staff may be based on-site or be visiting staff. People can live in supported housing for varying lengths of time, dependent on their individual needs and the type of service.

Temporary accommodation

Temporary accommodation is a short-term basis accommodation often provided by the Council or a Housing Association to homeless people and people who are at risk of homelessness. Temporary accommodation can consist of bed and breakfast, a room in a hostel or an assured shorthold tenancy on a flat or house.

Tenancy

The renting agreement whereby tenants rent the accommodation.

Tenure

A way in which a property is occupied. Commonly division is made between owner-occupation, social renting and private renting, although some data sources provide more detailed tenure categories (such as owned with mortgage, owned outright, and living rent-free).

Travel to work area

A Travel to Work Area (TTWA) is a statistical tool used by UK Government agencies and local authorities, especially by the Department for Work and Pensions and Job Centres, to indicate an area where the population would generally commute to a larger town, city or conurbation for the purposes of employment. The zones must be defined so that the bulk of their resident population also work within the same area. The Office for National Statistics (ONS) has worked with Newcastle University to apply a complex allocation process to define a set of Travel to Work Areas (TTWAs). The fundamental criterion is that, of the resident economically active population, at least 75 per cent actually work in the area, and also, that of everyone working in the area, at least 75 per cent actually live in the area. The resulting pattern is that, although the definitive minimum working population in a TTWA is 3,500, many are much larger - indeed, the whole of London and surrounding area forms one TTWA. The 243 current TTWAs were defined in 2007 using 2001 Census information on home and work addresses.

Under-occupancy

A household is under-occupying if it has more bedrooms than it “needs” according to the **bedroom standard** (see above). It is households with two or more bedrooms more than the bedroom standard that are most often of concern.

Annex 4: List of Abbreviations

ABI	Annual Business Inquiry
AMR	Regional Annual Monitoring Reports
APS	Annual Population Survey
ARLA	Association of Residential Letting Agents
ASHE	Annual Survey of Hours and Earnings
BERR	Department for Business Enterprise and Regulatory Reform
BHPS	British Household Panel Survey
BME	Black and Minority Ethnic
BPSA	Business Plan Statistical Appendix
BRES	Business Register and Employment Survey
CACI	Company name
CLG	Communities and Local Government
CML	Council of Mortgage Lenders
CORE	Continuous Recording System
DHS	Decent Homes Standard
DWP	Department of Work and Pensions
EEA	European Environment Agency
EERA	East of England Regional Assembly
EFS	Expenditure and Food Survey
EHCS	English House Condition Survey
EHS	English Housing Survey
EMRA	East Midlands Regional Assembly
FRS	Family Resources Survey
GAD	Government Actuary's Department
GHS	General Household Survey

GLA	Greater London Authority
HBF	Home Builders Federation
HC	Housing Corporation
HMRC	Her Majesty's Revenue and Customs (formerly Inland Revenue)
HRP	Household Reference Person
HRA	Housing Revenue Account
HSSA	Housing Strategy Statistical Appendix
IHS	Integrated Household Survey
IR	Inland Revenue
LA	Local authority
LFS	Labour Force Survey
LHOG	London Home Ownership Group
LPA	Local Planning Authority
LUCS	Land Use Change Statistics
NeSS	Neighbourhood Statistics
NHS	National Health Service
NLA	National Landlords' Association
NLUD	National Land Use Database
NOMIS	(formerly National Online Manpower Information Service)
NPD	National Pupil Dataset
NROSH	National Register of Social Housing
ONS	Office for National Statistics
RICS	Royal Institution of Chartered Surveyors
RMS	Regulated Mortgage Survey
RSL	Registered Social Landlord
RSR	Regulatory and Statistical Return

RTB	Right to Buy
RTPI	Royal Town Planning Institute
SEERA	South East England Regional Assembly
SEH	Survey of English Housing
SHMA	Strategic Housing Market Assessments
SML	Survey of Mortgage Lenders
TEDS	Tenancy Deposit Scheme
VOA	Valuation Office Agency
WMRA	West Midlands Regional Assembly

Annex 5: Private rented sector dwelling stock and buy-to-let loans

This annex relates to section 6.2: new supply to the private rented sector.

Official estimates of the private rented dwelling stock in the UK in 1996, 2001 and 2006 are in Table A.1. 2006 is the most recent year; 2001 is the last year of a sequence of years in which the number of dwellings in the sector changed by only small amounts from year to year; and 1996 was the first year of that sequence.

Table A.1 Estimates of the private rented dwelling stock in 1996, 2001 and 2006

	1996	2001	2006	Increase 2001-2006
England	2,191	2,152	2,611	459
Wales	104	127	137	10
Scotland	155	156	178	22
Northern Ireland	22	37	69	32
United Kingdom	2,472	2,471	2,995	523

Source: Wilcox (ed) UK Housing Review 2007-2008, Tables 16a and 17c

Numbers of buy-to-let loans are shown in Table A.2. They are published in half year amounts. The interval between the end of the first half of 2001 and the end of the first half of 2006 correspond most closely to the interval between the figures for 2001 and 2006 in Table A.1.

Table A.2 Buy-to-let mortgages outstanding, issued, and terminated 2001-2006

Outstanding, end of first half 2001	232,900
Outstanding, end of first half 2006	767,300
Net increase 2001-2006	534,400
New mortgages issued, second half of 2001 to first half of 2006 (inclusive)	953,100
Mortgages terminated	418,700

Source: UK Housing Review 2007/2008, Table 55

References

- Anderson B (2007) *Creating Small-area Income Estimates: Spatial Microsimulation Modelling*, CLG, London
- Barker K (2003) *Review of Housing Supply: Securing Our Future Housing Needs Interim Report: Analysis*, HM Treasury, London
- Barker K (2004) *Review of Housing Supply: Delivering Stability: Securing Our Future Housing Needs Final Report*, HM Treasury, London
- Barker K (2006) *Barker Review of Land Use Planning*, HM Treasury, London.
- Bramley G (1993a) 'The impact of land use planning and tax subsidies on the supply and price of housing in Britain', *Urban Studies*, Vol 30, no. 1, 5-30
- Bramley G (1993b) 'Land use planning and the housing market in Britain – the impact on house building and house prices', *Environment and Planning A*, Vol 25, no. 7, 1021-1052
- Bramley G and Smart G (1996) 'Modelling local income distributions in Britain', *Regional Studies* Vol 30 no 3, 239-255
- Bramley G and Lancaster S (1998) 'Modelling local and small area income distributions in Scotland', *Environment & Planning C: Government and Policy*, 16, 681-706
- Cheshire P (2005) 'Unpriced regulatory risk and the competition of rules: unconsidered implications of land use planning', *Journal of Property Research*, Vol, 22, no. 2/3, 225-244
- Communities and Local Government (2007) *Strategic Housing Market Assessment Practice Guidance*, CLG, London
- Gibbons S and Machin S (2003) *Rail Access and House Prices: An Evaluation of the Wider Benefits of Transport Improvements*, Department for Transport, London.
- Holmans A (1995) *Housing Demand and Need in England 1991 - 2011*, Joseph Rowntree Foundation, York.
- Holmans A (2000) *Housing Demand and Need in England 1996-2016*, Town and Country Planning Association, London
- HM Treasury (2007) *Comprehensive Spending Review 2008-2011*, HM Treasury, London
- Meen G, Allmendinger P, Andrew M, Ball M, Cameron G, Evans A, Gibb K, Goody J, Holmans A, Kasparova D, Monk S, Muellbauer J, Murphy A, Whitehead C and Wilson A (2005) *Affordability Targets: Implications for Housing Supply*, CLG, London
- Meen G (2007) *A Blueprint for Modelling Affordability*, Reading University, Reading
- Meen G, Andrew M, Ball M, Goody J, Kasparova D, Pryce G and Whitehead C (2008) *Recent Developments in the Communities and Local Government Affordability Model*, CLG, London

Monk S, Crook T, Lister D, Lovatt R, Ni Luanaigh A, Rowley S and Whitehead C (2006) *Delivering Affordable Housing through the Planning System: Outputs and Outcomes*, JRF, York

Monk S, Pearce B and Whitehead C (1996) 'Land-use planning, land supply and house prices', *Environment and Planning A*, Vol 28, no. 3, 495-511

NHPAU (2008a) *Meeting the Housing Requirements of an Aspiring and Growing Nation*, NHPAU, Titchfield

NHPAU (2008b) *Buy-to-Let Mortgage Lending and the Impact on UK House Prices*, NHPAU Research Findings Number 1, NHPAU, Titchfield

Office for Population, Censuses and Surveys (1994) *Routes into Local Authority Housing*, HMSO, London (1991 survey, now out of print)

Office for Population, Censuses and Surveys (1992) *Shared Accommodation in England 1990*, HMSO, London

Tang, C (2008, forthcoming) *Rental Affordability of New Housing Association Tenants in England: The Mediating Role of Housing Benefit*, Housing Corporation, London.

Taylor M (2008) *Taylor Review of Rural Economy and Affordable Housing*, CLG, London

Wilcox S (2007) *Can Work, Can't Buy*, Joseph Rowntree Foundation, York