FEATURE

Kathryn Ashton and Katherine Kent
Office for National Statistics

Annual Population Survey household data sets

SUMMARY

The Office for National Statistics has developed new annual local area data sets called the Annual Population Survey (APS) household data sets. They allow for production of family and household labour market statistics at local areas and for small subgroups of the population across the UK.

The aim of this article is to raise awareness of the APS household data sets. The information is presented in two sections: the initial section describes the key features of the APS household data sets, while the second section presents some key results from the data to illustrate the type of analysis that can be produced.

The Annual Population Survey (APS) household data set comes from the Labour Force Survey (LFS) and the APS (person) data set. The former is a quarterly survey of households living at private addresses in the UK. The latter is created by combining individuals in waves one and five from four consecutive LFS quarters with the English, Welsh and Scottish Local Labour Force Surveys. The APS household data sets therefore contain results from four different sources. They cover the period January to December for individual years from 2004.

Sampling and data collection

The LFS consists of a rotating panel design where interviews for individuals in households aged 16 and over take place for five consecutive quarters. The first round of interviews are face to face, with the exception of north of the Caledonian Canal in Scotland, where households are interviewed by telephone. Thereafter, each interview takes place by telephone at quarterly intervals. The household address is the main sampling unit, rather than the people living at that address. Thus, if the occupants change, the address would remain in the sample, and any new occupant(s) interviewed.

The APS household sample is three times the size of the LFS sample. It contains information collected from a sample of around 160,000 households (300,000 people aged 16 or over), whereas the LFS data sets only have a sample size of around 53,000 households (100,000 people aged 16 or

over). The APS aims to get information on a minimum of 510 economically active persons for each local area authority and unitary authority. It is the recommended source for local area data because of the greater sample size. **Table 1** compares the sample sizes of the APS household and LFS household data sets.

However, users should note that, like the LFS, APS household data relate mainly to the population living in private households. The coverage excludes many groups living in communal establishments, except those in NHS housing and students in halls of residence, where inclusion takes place at their parents' address. Members of the armed forces are included if they live in private accommodation, and the economic status of under 16-year-olds is not collected. For further details about methodology, see ONS (2007).

Imputation

For some households taking part in the LFS, one or more household members may have unknown economic activity status, that is, it is not known whether they are employed, unemployed, inactive or a child. These individuals are known as 'non-responders'. This may be because they have either refused to take part in the survey, or were absent, and a proxy interview (that is, where an individual responds for someone else) was unavailable. Removal of non-responders takes place in the person data sets and weighting compensates for this. For the APS household data set, non-responders are necessary to identify

Table 1
Sample size of APS household and LFS household data sets

	APS			LFS ¹		
	Sample size	Number of households	People aged 16 and over	Sample size	Number of households	People aged 16 and over
2004	387,026	163,803	306,982	128,524	54,037	101,654
2005	383,771	162,771	305,393	126,587	53,275	100,389
2006	362,901	154,504	289,019	124,108	52,398	98,858
2007	365,926	156,141	292,997	123,715	52,202	98,327

Note:

1 Using April to June quarter data sets.

Source: APS household and LFS household data sets

relationships between household members, assign them to complete family units within the household, and derive family and household variables. To ensure that the economic status of all individuals within a household is known, a method of 'donor' imputation takes place for those with a missing economic activity status, allowing for analysis of the combined economic status of households.

Donor imputation on the APS household data sets uses the CANadian Edit and Imputation System (CANCEIS) software, developed by Statistics Canada. Its uses include imputing many variables into the 2001 Canadian Census and the imputation process of several Office for National Statistics (ONS) surveys. The procedure for imputing the APS household data set is based on the initial imputation of the variable for economic activity, which has the four categories described earlier. Non-responders under the age of 16 are coded as 'child'; the remaining non-responders are imputed using the donor imputation method. Donors for cases with no economic activity status are identified using several matching variables. These are age, sex, relationship to household reference person, household type and the economic status from the previous interview. Imputation takes place for non-responders using individuals with similar matching variables, likened to a 'nearest neighbour' form of imputation. The suitability of the imputation process is checked by comparing the distributions for key variables before and after imputation.

Weighting

Weighting of data in the APS household data sets ensures that the estimates are representative of the whole population. The weighting calibrates the survey estimates on five-year age groups by gender, and separately on unitary authority and local authority population totals. The method counts the number of individuals within each household in each of these groups,

and uses a least-squares method to derive household weights.

The weighting procedure used in producing the APS household data sets is generalised regression estimation, which uses the Generalised Estimation System computer software package. The starting point is the file of household-level records and the data are then put under the weighting procedure. This procedure assumes that the corresponding APS person data set, which is used to derive the population totals using the final calibrated weight, exists. The household-file design weights are rescaled so that they sum to the total UK population size.

Unlike in the person data sets, weights for each person in the same household are equal. This ensures that weighted estimates at the household level are consistent. For example, the number of married men living in a household type would equal the number of married women. Note that although the weighting method forces agreement with population figures at the person level, it does not fix the weighted number of families or households. In other words, the total number of households may not equal the actual number of households in the population. A different sample would produce different estimates of the total number of families and households.

Household/family-level variables

The APS household data sets include all the variables on the LFS and APS person data sets, except for the income variables. They also include key family and household-level derived variables. These variables allow for an analysis of the combined economic activity status of the family or household. A full list of the household and family variables available on the data set is given in the Appendix.

Uses

The APS household data set provides improved data on key family and

household-level variables that can inform against agreed Public Service Agreement targets. There is, for example, a growing demand for statistics on workless households and lone-parent employment rates for unitary and local authority districts and for parliamentary constituency areas. Besides better local authority-level data, the increased sample size available from the APS shown in Table 1 brings better national estimates.

ONS already publishes statistics on workless households, and on working-age people and children living in workless households, by region and ethnicity; it also publishes statistics on lone-parent employment rates for the UK as whole. These are produced from the LFS household data sets and are published in the annual 'Work and Worklessness among Households': First Release.

Equivalent figures for lower-level geographies are not published regularly because of the smaller LFS sample sizes, which result in wide margins of uncertainty. Sometimes it is not possible to publish estimates because the LFS sample is too small and the estimates are potentially disclosive (that is, it may be possible to identify an individual respondent).

Another reason affecting the quality of LFS-based family and household statistics is that figures for subgroups cannot be adjusted to compensate for people, families and households with unknown economic activity status. The LFS figures for regions, local areas and other subgroups may therefore underestimate the numbers of people, families and households in each economic activity status category.

Users should note that APS household data sets are only available from 2004 onwards and therefore do not permit long-run trend analysis. Also, because the APS is an annual data set which is published six months after the end of the annual period, the estimates are less timely than those for the quarterly LFS. However, by covering a whole year, the APS household data set overcomes seasonality.

Results

The following section gives an outline of the types of household and family analysis available, using the January to December 2007 APS household data set. The results are restricted to working-age households, where at least one person in the household is of working age (16 to 64 for men and 16 to 59 for women). This is comparable with the 'Work and Worklessness among Households': First Release, which uses LFS

household data for April to June of each year. These estimates will differ because of the imputation for unknown economic status on the APS household and the fact that it covers January to December of each year.

Sampling variability

Like all sample surveys, estimates from the APS household are subject to sampling variability. The smaller the group whose size is being estimated, the less precise that estimate is. One estimate of the variability among the estimates from all possible samples is the relative standard error (RSE). The RSE represents the standard error as a percentage of the quantity being estimated. As the RSE increases, the estimate itself becomes less reliable. Results are presented with the RSE associated with each estimate. Those estimates which have a relative standard error of 20 per cent or more are not considered reliable for practical purposes. In other words, if a different sample is taken from the same population, it is likely that the estimate may differ greatly from the estimate of the current sample.

Working and workless households

A 'working' household is one where all adults of working-age are in employment. A 'workless' household is one where none of the working-age adults are in employment. The latter includes households where all adults are unemployed, those where all adults are economically inactive, and a mixture of both economic statuses. There are also households containing a mixture of adults defined as working and workless, known as mixed households.

Table 2 shows that, of the 19.34 million working-age households in the UK in 2007, 57.2 per cent (11.06 million) were working households, 27.2 per cent (5.26 million) were mixed households and 15.6 per cent (3.02 million) were workless. Among the regions, Inner London had the highest percentage of workless households, at 21.8 per cent (0.24 million). The South East region had the lowest percentage of workless households, at 10.9 per cent (0.28 million), and the highest percentage of households in which all adults were working, at 61.9 per cent (1.55 million). Northern Ireland had the lowest percentage of working households, where just under half of households contained adults who were all working.

Within each region, there are wide variations in the percentage of workless households. **Map 1** shows the percentages

for each region of Great Britain in 2007. Wales contained local authorities with some of the highest percentages of workless households (in Merthyr Tydfil and Neath Port Talbot approximately one-quarter of households were workless), but also contained Monmouthshire, which had a low percentage (12.4 per cent).

However, when picking out individual local authorities, the data should be considered along with sampling errors since, as with any sample survey, estimates are subject to a margin of uncertainty. Of the 408 local authorities in the UK, 226 (55.4 per cent) had an RSE of greater than 20 per cent for the percentage of workless households. The full list is available on the ONS website.¹

Children in workless households

Table 3 shows the number and percentage of children living in working, mixed and workless households by Government Office Region. Of children living in the UK in 2007, 52.7 per cent (6.05 million) lived in working households, 32.0 per cent (3.67 million) in mixed households and the remaining 15.3 per cent (1.76 million) in workless households. Among the regions, Inner London had the highest percentage of children in workless households, at 28.7 per cent (0.17 million), and the South East the lowest, at 9.6 per cent (0.15 million).

Workless households by ethnic group

As well as being more suitable for local area analysis, the APS household data sets provide more robust estimates for small subgroups such as ethnicity. **Table 4** shows working, mixed and workless households by ethnic origin of the household reference person. The household reference person covers individuals such as those who own the accommodation or are legally responsible for the rent. In 2007, Indian households had the lowest percentage of workless households, at 10.6 per cent (0.04 million). 'Other Black' had the highest percentage of workless households, at 31.0 per cent (0.01 million). However, the 'Other Black' ethnic group had a large RSE and is an unreliable estimate. White households had the highest percentage of working households, at 58.8 per cent (10.26 million), and Pakistani/Bangladeshi households had the lowest, at 18.8 per cent (0.06 million).

Dissemination and disclosure

The APS household microdata will be available through the UK Data Archive at Essex University, where it can be accessed by academic institutions and members of

the public. More detail is available on the Data Archive website.² Access to the data is through a special licence file which, while limiting user access, allows for household variables to remain on the data set, and for lower geographical coding and no age banding.

Statistical disclosure control methodology is applied to the APS household data. This ensures that information attributable to an individual is not disclosed in any publication.³

A user guide will illustrate background to the data sets and the uses of the data, demonstrating basic household level analyses; this will be available from October 2008.⁴

Between October 2008 and February 2009, a consultation on the content of the 'Work and Worklessness among Households': First Release, which uses LFS household data, will take place. The consultation will consider moving the source of the annual release over to the APS household data sets and allow for a review of the content of the release.

Notes

- 1 See www.statistics.gov.uk/statbase/product.asp?vlnk=15150
- 2 See www.data-archive.ac.uk
- 3 For more information on ONS statistical disclosure control methodology, see www.statistics.gov.uk/about/data/methodology/general_methodology/sdc.asp
- 4 See www.statistics.gov.uk/statbase/product.asp?vlnk=1537

CONTACT

✉ elmr@ons.gsi.gov.uk

REFERENCES

- Office for National Statistics (2004) 'Statistical Disclosure Control' at www.statistics.gov.uk/about/data/methodology/general_methodology
- Office for National Statistics (2007) 'Labour Force Survey User Guide Volume 1: Background and Methodology' at www.statistics.gov.uk/downloads/theme_labour/lfsug_vol1_2007.pdf

Table 2

Working-age households:¹ by combined economic activity status of household and region, January to December 2007

Government Office Region	Working households ²		Households containing both working and workless members		Workless households ³		Total	
	Level	Relative standard error ⁴	Level	Relative standard error ⁴	Level	Relative standard error ⁴	Level	Relative standard error ⁴
Thousands								
England	9,298	0.4	4,427	0.6	2,454	0.9	16,179	0.2
North East	435	1.5	227	2.1	153	2.9	815	0.7
North West	1,216	1.1	572	1.5	404	2.1	2,192	0.5
Yorkshire and The Humber	930	1.3	447	1.8	273	2.8	1,650	0.6
East Midlands	816	1.4	380	2.2	193	3.5	1,389	0.7
West Midlands	913	1.3	481	1.8	266	2.9	1,659	0.6
East of England	1,067	1.3	481	2.0	218	3.5	1,766	0.6
London	1,358	1.2	728	1.5	460	2.2	2,546	0.5
Inner London	597	1.8	271	2.4	242	3.0	1,109	0.8
Outer London	762	1.6	457	1.9	218	3.4	1,437	0.7
South East	1,594	1.0	702	1.6	280	2.9	2,576	0.5
South West	968	1.3	408	2.0	209	3.3	1,585	0.6
Wales	479	1.2	266	1.5	169	2.2	914	0.5
Scotland	1,018	1.0	399	1.6	287	2.3	1,703	0.5
Great Britain	10,794	0.3	5,092	0.5	2,910	0.8	18,796	0.1
Northern Ireland	267	2.5	169	2.9	106	4.7	542	1.1
United Kingdom	11,061	0.3	5,260	0.5	3,016	0.8	19,337	0.1
Percentages								
England	57.5	0.3	27.4	0.6	15.2	0.9	100.0	
North East	53.4	1.3	27.9	2.2	18.7	2.7	100.0	
North West	55.5	0.9	26.1	1.6	18.4	2.0	100.0	
Yorkshire and The Humber	56.4	1.0	27.1	1.9	16.5	2.6	100.0	
East Midlands	58.8	1.2	27.3	2.3	13.9	3.3	100.0	
West Midlands	55.0	1.1	29.0	1.9	16.0	2.7	100.0	
East of England	60.4	1.0	27.2	2.1	12.4	3.4	100.0	
London	53.3	0.9	28.6	1.6	18.1	2.1	100.0	
Inner London	53.8	1.4	24.4	2.6	21.8	2.8	100.0	
Outer London	53.0	1.3	31.8	2.0	15.2	3.2	100.0	
South East	61.9	0.8	27.3	1.6	10.9	2.8	100.0	
South West	61.1	1.0	25.8	2.1	13.2	3.2	100.0	
Wales	52.4	0.9	29.1	1.6	18.5	2.1	100.0	
Scotland	59.8	0.8	23.4	1.7	16.8	2.1	100.0	
Great Britain	57.4	0.3	27.1	0.6	15.5	0.8	100.0	
Northern Ireland	49.3	2.2	31.1	3.2	19.6	4.3	100.0	
United Kingdom	57.2	0.3	27.2	0.5	15.6	0.8	100.0	

Notes:

Source: APS household data sets

1 A household that includes at least one person of working age (man aged 16 to 64 or woman aged 16 to 59).

2 A working-age household where all members aged 16 or over are in employment.

3 A working-age household where no one aged 16 or over is in employment.

4 0≤RSE<5: estimates are considered precise; 5≤RSE<10: estimates are considered reasonably precise; 10≤RSE<20: estimates are considered acceptable; RSE≥20: estimates are not considered reliable for practical purposes.

Map 1

Percentage of workless households¹ in Great Britain: by local authority, 2007

1 A workless household is a working-age household where no one aged 16 or over is in employment.

2 Sample size is too small to provide an estimate for City of London, Isles of Scilly, Epsom and Ewell and Oadby and Wigston.

Table 3

Children¹ living in working-age households: by combined economic activity status of household and region, January to December 2007

Government Office Region	Working households ²		Households containing both working and workless members		Workless households ³		Total	
	Level	Relative standard error ⁴	Level	Relative standard error ⁴	Level	Relative standard error ⁴	Level	Relative standard error ⁴
Thousands								
England	5,050	0.7	3,145	1.0	1,479	1.7	9,674	0.2
North East	238	2.9	129	4.2	81	5.7	448	1.8
North West	694	1.9	368	3.1	232	3.9	1,294	1.2
Yorkshire and The Humber	514	2.4	315	3.5	156	5.5	985	1.5
East Midlands	466	2.7	257	4.2	101	6.7	825	1.7
West Midlands	532	2.4	338	3.5	186	5.3	1,056	1.4
East of England	574	2.6	361	3.7	122	6.9	1,057	1.6
London	586	2.5	580	2.6	350	3.7	1,516	1.2
Inner London	183	4.4	232	4.1	167	5.2	582	2.0
Outer London	403	3.0	348	3.4	183	5.4	934	1.6
South East	902	1.9	535	2.8	153	5.8	1,590	1.2
South West	543	2.4	262	4.0	98	6.9	903	1.6
Wales	288	2.1	163	3.1	101	4.3	552	1.3
Scotland	520	1.8	232	3.2	132	4.7	884	1.2
Great Britain	5,859	0.6	3,540	0.9	1,712	1.5	11,110	0.1
Northern Ireland	186	4.4	135	5.5	47	9.9	368	2.5
United Kingdom	6,045	0.6	3,674	0.9	1,759	1.5	11,478	0.0
Percentages								
England	52.2	0.6	32.5	1.0	15.3	1.7	100.0	
North East	53.1	2.3	28.8	4.2	18.1	5.3	100.0	
North West	53.6	1.6	28.4	3.1	17.9	3.7	100.0	
Yorkshire and The Humber	52.2	2.0	31.9	3.5	15.9	5.1	100.0	
East Midlands	56.6	2.2	31.2	4.2	12.3	6.4	100.0	
West Midlands	50.4	2.1	32.0	3.5	17.6	4.9	100.0	
East of England	54.3	2.1	34.1	3.7	11.6	6.6	100.0	
London	38.6	2.2	38.3	2.6	23.1	3.4	100.0	
Inner London	31.5	4.1	39.9	4.1	28.7	4.6	100.0	
Outer London	43.1	2.7	37.3	3.4	19.6	5.0	100.0	
South East	56.7	1.5	33.6	2.8	9.6	5.6	100.0	
South West	60.1	1.9	29.0	4.0	10.8	6.6	100.0	
Wales	52.2	1.7	29.5	3.1	18.3	3.9	100.0	
Scotland	58.9	1.4	26.2	3.2	14.9	4.4	100.0	
Great Britain	52.7	0.6	31.9	0.9	15.4	1.5	100.0	
Northern Ireland	50.6	3.6	36.6	5.5	12.8	9.5	100.0	
United Kingdom	52.7	0.6	32.0	0.9	15.3	1.5	100.0	

Notes:

Source: APS household data sets

¹ Children refers to children under 16.² A working-age household where all members aged 16 or over are in employment.³ A working-age household where no one aged 16 or over is in employment.⁴ 0≤RSE<5: estimates are considered precise; 5≤RSE<10: estimates are considered reasonably precise; 10≤RSE<20: estimates are considered acceptable; RSE≥20: estimates are not considered reliable for practical purposes.

Table 4

Working-age households:¹ by ethnic origin of household reference person and combined economic activity status of household, January to December 2007

Ethnic origin of household reference person	Working households ²		Households containing both working and workless members		Workless households ³		Total	
	Level	Relative standard error ⁴	Level	Relative standard error ⁴	Level	Relative standard error ⁴	Level	Relative standard error ⁴
Thousands								
White	10,264	0.3	4,593	0.6	2,612	0.9	17,469	0.2
All other ethnic groups	793	1.8	664	1.8	402	2.4	1,859	1.0
Mixed	73	6.0	24	10.1	34	8.6	130	4.4
Asian	280	3.0	389	2.3	141	4.1	809	1.6
Indian	164	3.9	161	3.7	39	7.8	363	2.5
Pakistani/Bangladeshi	57	6.7	174	3.5	74	5.5	306	2.6
Other Asian	59	6.6	54	6.6	28	9.8	140	4.2
Black	265	3.1	130	4.3	130	4.3	524	2.1
Black Caribbean	126	4.5	58	6.6	57	6.6	241	3.2
Black African	131	4.4	66	5.9	66	6.0	263	3.0
Other Black	8	18.1	6	22.3	6	19.8	20	11.5
Chinese	47	7.6	25	9.8	18	12.1	91	5.4
Other	129	4.6	96	5.0	80	5.4	305	2.8
Total ⁵	11,061	0.3	5,260	0.5	3,016	0.8	19,337	0.1
Percentages								
White	58.8	0.3	26.3	0.6	15.0	0.6	100.0	
All other ethnic groups	42.7	1.3	35.7	1.5	21.6	1.5	100.0	
Mixed	56.0	3.9	18.1	9.2	25.9	9.2	100.0	
Asian	34.6	2.4	48.0	1.8	17.4	1.8	100.0	
Indian	45.0	2.9	44.4	2.9	10.6	2.9	100.0	
Pakistani/Bangladeshi	18.8	6.0	56.9	2.4	24.3	2.4	100.0	
Other Asian	41.9	5.0	38.3	5.3	19.8	5.3	100.0	
Black	50.5	2.2	24.8	3.8	24.7	3.8	100.0	
Black Caribbean	52.2	3.1	24.2	5.8	23.6	5.8	100.0	
Black African	49.7	3.1	25.1	5.2	25.2	5.2	100.0	
Other Black	40.9	13.8	28.1	18.7	31.0	18.7	100.0	
Chinese	51.6	5.2	28.1	8.4	20.2	8.4	100.0	
Other	42.3	3.4	31.6	4.2	26.1	4.2	100.0	
Total ⁵	57.2	0.3	27.2	0.5	15.6	0.8	100.0	

Notes:

Source: APS household data sets

1 A household that includes at least one person of working age (man aged 16 to 64 or woman aged 16 to 59).

2 A working-age household where all members aged 16 or over are in employment.

3 A working-age household where no one aged 16 or over is in employment.

4 0≤RSE<5: estimates are considered precise; 5≤RSE<10: estimates are considered reasonably precise; 10≤RSE<20: estimates are considered acceptable; RSE≥20: estimates are not considered reliable for practical purposes.

5 Total includes people with unknown ethnic origin.

Appendix

APS household and family-derived variables

Variable name	Description
AOFL16	Age of oldest dependent child in family unit aged under 16
AOFL19	Age of oldest dependent child in family unit aged under 19
AOHL16	Age of oldest dependent child in household aged under 16
AOHL19	Age of oldest dependent child in household aged under 19
AYFL19	Age of youngest dependent child in family unit aged under 19
AYHL19	Age of youngest dependent child in household aged under 19
CAIND	Child/adult indicator
EXTFU	Extended family unit
FDPCH15	Number of dependent children in family aged between 10 and 15
FDPCH16	Number of dependent children in family aged under 16
FDPCH19	Number of dependent children in family aged under 19
FDPCH2	Number of dependent children in family aged under 2
FDPCH4	Number of dependent children in family aged between 2 and 4
FDPCH9	Number of dependent children in family aged between 5 and 9
FMDP	Number of family units in household with dependent children only
FMNDP	Number of family units in household with non-dependent children only
FMPUS	Total number of family units in household with more than one person
FUTYPE6	Type of family unit
HDC515	Number of children in household aged between 5 and 15 years
HDPCH4	Number of children in household aged 4 years or less
HDPCH18	Number of children in household aged between 16 and 18 years
HDPCH19	Number of dependent children in household aged under 19
HEACOMB	Household economic activity
HEAHEAD	Economic activity of head of family unit
HEAWIFE	Economic activity of wife family unit
HHTYPE6	Type of household
HLDCMP6	Composition of household
HNDK	Number of people in household where basic economic activity is not known
HNEMP	Number of people in household who are unemployed
HNFTIME	Number of people in household who are working full-time
HNFTSTUD	Number of people in household who are full-time students
HNINAC05	Number of people in household who are inactive
HNWDSC	Number of people in household who are inactive and would like work but discouraged from seeking work
HNWFAM	Number of people in household who are inactive, would like work but looking after family/home
HNWSKD	Number of people in household who are inactive, would like work but currently sick/injured/disabled
HNWSTU	Number of people in household who are inactive, would like work but currently students
HNWK05	Number of people in household who are inactive and do not want work
HNSTSTUD	Number of people in household who are not full-time students
HNPEN	Number of people in household who are of pensionable age
HNPTIME	Number of people in household who are working part-time
HNUNEMP	Number of people in household who are unemployed
HNWKAGE	Number of people in household who are of working age
HNWOTH05	Number of people in household who are inactive for other reasons but would like to work
HOHID	Head of household identification
RELH06	Relationship to head of household
RELHFU	Relationship to head of family
RELHRP6	Relationship to household reference person
SMSXFU	Same sex family unit
TOTFU	Total number of family units in household
TOTNUM	Total number of eligible people in household
TOTXFU	Total number of extended families in household
XFMDC	Total number of extended family units in household with dependent children only
XFMNDC	Total number of extended family units in household with non-dependent children only