

Imputed expenditure variables for Family Resources Survey 1995/6  
**USER DOCUMENTATION**

Contents

1. Description of the originating project	Page 1
2. Bibliography	Page 3
3. Data description	Page 4
3.1. Variable list	Page 5
3.2. Definition of expenditure categories	Page 6
3.3. Clusters	Page 16

Description of the originating project

The originating project was financed by the ESRC as part of the Analysis of Large and Complex Datasets (ALCD) Programme under the title “Data Enhancements to Improve the Scope and Reliability of Microsimulation Models” (H519255052). The project was completed in September 2000.

The aim of the project was to produce a data set containing detailed income and expenditure information for households in Great Britain. The data set was to be used as the source information for a tax-benefit microsimulation model to investigate the impact of changes to Government policies. A method was developed to enhance the 1995/6 Family Resources Survey (FRS) with expenditure data from the 1995/6 Family Expenditure Survey (FES). New data was not collected. All variables have been derived from existing data sources available from the Data Archive:

- 1995/6 Family Resources Survey = SN 3761
- 1995/6 Family Expenditure Survey = SN 3635.

The FRS contains detailed information on household income and the circumstances of each individual and was commissioned by the Department of Social Security (DSS). A general summary of the data, sample and survey procedures is provided in DSS (1997). The original purpose of the FES was to provide detailed household spending information for the Retail Prices Index through a two-week diary kept by the households. Other information is also recorded. A general summary is provided in ONS (1996). Both surveys were conducted annually; the year chosen covers the period of April 1995 – March 1996. In both cases a representative national sample of households is surveyed across England, Wales and Scotland (Northern Ireland households were included in the FES sample but are deleted so that both survey samples have the same national coverage).

The two surveys share many features such as the interviewing technique and the time period. However, there are some differences that may be influential, including:

- Response rates of FES and FRS are similar at 66% and 70% respectively. However, the pattern of differential non-response is thought to be different in the two surveys, due to different demands on respondents. The FRS is thought to better represent households in receipt of benefits (DSS, 1997).
- The criteria for inclusion in the sample as a responding household are more stringent in the FES than the FRS.
- The FES household is in fact a “spending unit” which is not quite the same as the FRS “household”. The former is a narrower unit, conditional on there being some common

housekeeping, including meal preparation. Thus some FRS households would be treated as separate units in the FES.

- The effective sample size of the FES is 6,663 households (excluding N. Ireland) and the FRS is 26,435.
- The methods of imputation for missing values in the two surveys are different.
- There are significant differences between the distributions of some key variables including income, (also found by another comparison of the FES and FRS data sets (DSS, 1996; Appendix 9).
- In the construction of common variables, variable definitions may differ in ways that are not evident from the documentation.

Our starting point is two separate household survey datasets. Expenditure information is brought from one household in the FES to a similarly defined household in the FRS. Households surveyed in the FES could not be identified in the FRS sample - the two surveys cover different household samples. Therefore IT IS NOT THE CASE that information from each household in the FES can be matched *directly* to a household in the FRS. Instead the independent data sources are statistically matched. Cohen (1991) and Baker, Harris and O'Brien (1989) provide reviews of the techniques and some practical applications of statistical matching.<sup>1</sup> The key problem for matching is classifying the samples into homogeneous groups, with the definition of similarity between households depending on the variables that are going to be imputed. In this case, these would be households with similar patterns of expenditures. The most commonly used method defines groups as cells in a cross-tabulation of common variables. In contrast, our groups are defined by variables that are common to both surveys but are not 'straight-edged' cells. The variables are used to identify clusters of households that have similar spending patterns. Expenditure variation within the clusters is expected to be less than the variation of the whole sample. Expenditure is statistically matched to the FRS within similarly defined clusters.

Two combinations of common variables have been used to produce two different ways of clustering the samples. The demographic characteristics were chosen for their ability to predict household spending. The procedure used to identify the clusters is Grade Correspondence Analysis. The procedure is complex and for that reason a full description is withheld. Basically the process draws out the relationships between the common variables and orders the sample of households to maximise these relationships. Clusters of households with similar spending patterns are selected from the reordered sample. See Taylor (2000) for a detailed description of how the clusters were chosen.

Within each cluster the households are ranked by income (disposable income minus housing costs) before sequentially matching households. The technique starts with the lowest income household from a donor cluster (FES) which is matched to the lowest income household in the equivalently defined cluster in the recipient file (FRS). Let this be called 'rank by income' matching. Households from any recipient cluster receive expenditures only from the corresponding donor cluster. The number of donor and recipient households in each cluster are not generally the same. In this case the appropriate ratio of multiple households is imputed from (or to).

---

<sup>1</sup> Our view is one of scepticism about the potential offered by the techniques applied in general, and we would always expect a synthetic database to be inferior to data collected within a single survey.

The 'rank by income' method can be augmented using common variables that have not been used to define clusters. Within the clusters, households can be split by the extra information. For example, households with children can be separated from households without children. Households are ranked by income within the split clusters and matching proceeds across the split clusters.

The project generated many alternative sets of imputed expenditures as various combinations of common variables and methods of identifying clusters were tested. Three are chosen for archiving based on their performance in evaluation exercises (Taylor et al, 2001). However, in each case the experiments used the same set of 27 expenditure categories. These 27 variables take different values depending on the particular imputation method.

NOTE: There is no missing information. All of the households in the FES that were classified as having given a full interview by the FES data providers were used as expenditure donors. All of the households in the FRS that were classified as having given a full interview by the FRS data providers were recipients of the expenditure information.

### References

- Baker K, Harris P and O'Brien J (1989), 'Data fusion: an appraisal and experimental evaluation', *Journal of the market research society*, 31(2), pp153-212.
- Cohen M L (1991), 'Statistical Matching and Microsimulation Models', in C. F. Citro and E. A. Hanushek eds., *Improving Information for Social Policy Decisions. The Uses of Microsimulation Modeling* vol. II, National Academy Press, Washington D.C.
- Department of Social Security (1996), *Households Below Average Income: A statistical analysis 1979-1993/94*, London: The Stationary Office
- Department of Social Security (1997), *Family Resources Survey: Great Britain 1995-96*, London: The Stationary Office
- Office for National Statistics (1996), *Family Spending: A report on the 1995-96 Family Expenditure Survey*, London: The Stationary Office.
- Taylor R L (2000), 'Guidelines for Creating Clusters Using Grade Correspondence Analysis: Practical and Technical Issues', Microsimulation Unit Research Note MU/RN/39.
- Taylor R L, H Sutherland and J Gomulka (2001), 'Using POLIMOD to evaluate alternative methods of expenditure imputation', Microsimulation Unit Research Note MU/RN/38.

See also:

- Dayal N, J Gomulka, L Mitton, H Sutherland and R L Taylor (2000), 'Enhancing Family Resources Survey income data with expenditure data from the Family Expenditure Survey: data comparisons', Microsimulation Unit Research Note, MU/RN/40.
- Sutherland H, R L Taylor and J Gomulka (2001), 'Combining household income and expenditure data in policy simulations', Microsimulation Unit Discussion Paper MU0101.
- Taylor R L, J Gomulka and H Sutherland (2000), 'Creating Order out of Chaos? Identifying Homogeneous Groups of Households across Multiple Datasets' presented at the 26<sup>th</sup> General Conference of the *International Association for Research in Income and Wealth*, August/September.

All project papers are available at [www.econ.cam.ac.uk/dae/mu/alcd.htm](http://www.econ.cam.ac.uk/dae/mu/alcd.htm)

## Data Description

The data are contained in a single Excel file. The sample is the complete FRS sample i.e. 26,435. The households are identified by SERNUM, the existing FRS serial number. Consequently the information can be merged (case by case) into the existing FRS survey data. There is no information provided in the file that can be used to identify the households.

The file contains three sheets. Each sheet contains the identification number, the cluster set number and a full set of expenditure variables for the whole sample (30 variables). Each row in each data sheet contains the information for one household. Each column contains one variable. The variable name is in the first row. The expenditure variable names have suffix A, B or C depending on which method was used to create them. (Variable descriptions are provided below.) The methods used to create expenditure are:

- The expenditure data in the first sheet (impfrs2a) has been generated by Method A ('rank by income' matching within Test 42b.1 clusters).
- The expenditure data in the second sheet (impfrs2c) has been generated by Method B ('rank by income' matching within Test 41.1 clusters).
- The expenditure data in the third sheet (impfrs13) has been generated by Method C ('rank by income' matching within Test 42b.1 clusters that have been split into households with and without children).

See Taylor et al (2001) for a discussion of which set of expenditure variables is most appropriate for a particular application.

**WARNING:** It must be stressed that each sheet is self-contained. The expenditure variables are repeated in each sheet (although each is calculated in a different way), and should not be combined in any way across sheets. The user should not be tempted to average spending variables across the three methods. The three different estimates of expenditure give different values of expenditure for each household depending on the assumptions made throughout the construction of the data set. There is no restriction on the values that expenditure variables can take (e.g. such that total expenditure is less than income). Users who are unsure that the data is appropriate for the intended use are welcome to contact the data producers (not the Data Archive) for advice.

## VARIABLE LIST

Variable name	Description
SERNUM	FRS household identification number
Test42b1	Cluster set number (Test 42b.1)
Test411	Cluster set number (Test 41.1)
TOTAL_	Total household expenditure (£/week) = Sum of CAT1 to CAT27
CAT1_	Expenditure on housing and household services (£/week)
CAT2_	Expenditure on motoring expenditure (£/week)
CAT3_	Expenditure on food attracting VAT (£/week)
CAT4_	Expenditure on leisure goods and services (£/week)
CAT5_	Expenditure on adult clothing and footwear (£/week)
CAT6_	Expenditure on household goods, personal goods and services (£/week)
CAT7_	Expenditure on VAT-exempt goods (£/week)
CAT8_	Expenditure on VAT zero-rated food (£/week)
CAT9_	Expenditure on books and newspapers (£/week)
CAT10_	Expenditure on domestic fuel and power (£/week)
CAT11_	Expenditure on other zero-rated goods (£/week)
CAT12_	Expenditure on children's clothes (£/week)
CAT13_	Expenditure on insurance premia (£/week)
CAT14_	Expenditure on alcohol - beer (£/week)
CAT15_	Expenditure on alcohol - cider (£/week)
CAT16_	Expenditure on alcohol – fortified wine (£/week)
CAT17_	Expenditure on alcohol – wine (£/week)
CAT18_	Expenditure on alcohol - champagne (£/week)
CAT19_	Expenditure on alcohol - spirits (£/week)
CAT20_	Expenditure on cigarettes (£/week)
CAT21_	Expenditure on cigars (£/week)
CAT22_	Expenditure on pipe tobacco (£/week)
CAT23_	Expenditure on motor petrol and oils (£/week)
CAT24_	Expenditure on motoring diesel (£/week)
CAT25_	Expenditure on pools stakes (£/week)
CAT26_	Expenditure on other betting stakes (£/week)
CAT27_	Expenditure on lottery stakes (£/week)

## Definition of expenditure categories

These are derived variables created from the FES diary expenditure variables. A detailed description of the expenditures contained in each category is provided with summary statistics of each variable. The exact FES variable names are given alongside the FES description of the expenditure. Negative values are permitted to represent refunds or items that have been sold.

### CAT1\_

#### Housing expenditure:

b102	/* central heating repairs - second dwelling	*/
+ b104	/* central heating repairs - main dwelling	*/
+ b107	/* house maintenance, etc., main dwelling	*/
+ b108	/* house maintenance, etc., second dwelling	*/
+ d010401t	/* central heating installation (DIY)	*/
+ d010402t	/* double glazing, kitchen units, sheds (DIY)	*/
+ d010501t	/* doors, baths and other fittings	*/
+ d010502t	/* tools	*/
+ d010503t	/* paint, wallpaper, timber	*/
+ d010504t	/* other materials, including hire of equipment	*/

#### Household services:

b166	/* telephone - household share of account	*/
+ b272	/* legal fees: main dwelling	*/
+ b273	/* moving house costs	*/
+ b274	/* legal fees: second dwelling	*/
+ b280	/* property transaction: purchase and sale	*/
+ b281	/* property transaction: sale only	*/
+ b282	/* property transaction: purchase only	*/
+ b283	/* property transaction: other payments	*/
+ d010606t	/* second dwelling - telephone account payments	*/
+ d080104t	/* cleaning and dyeing	*/
+ d080105t	/* laundry, laundrette	*/
+ d080106t	/* repairs to footwear	*/
+ d080107t	/* repairs to personal goods	*/
+ d080202t	/* telephone purchase	*/
+ d080204t	/* telephone coin and other payment	*/
+ d080205t	/* mobile phone purchase	*/
+ d080206t	/* mobile phone account payments	*/
+ d080207t	/* answering and fax machines, modems	*/
+ d080405t	/* legal fees paid to banks	*/
+ d080406t	/* legal fees paid to solicitors	*/
+ d080407t	/* court fines, architect, other professional fees	*/
+ d080501t	/* contract catering for weddings, etc.	*/
+ d080502t	/* rental/hire of electrical or household equipment	*/

#### Other household expenditure

b101	/* central heating installed - second dwelling	*/
+ b103	/* central heating installed - main dwelling	*/
+ b105	/* house extensions, etc., main dwelling	*/
+ b106	/* house extensions, etc., second dwelling	*/

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	23.28	0	976.41
Method B	23.18	0	976.41
Method C	23.28	-2.2x10 <sup>-14</sup>	976.14

**CAT2\_**

## Motoring expenditure

b183	/* credit purchase - net cost of new vehicles	*/
+ b244	/* cost of new car/van bought outright	*/
+ b247	/* cost of motorcycle bought outright	*/
+ b248	/* expenditure on car leasing	*/
+ b249	/* amount paid for car/van servicing	*/
+ b250	/* amount paid for other works and repairs on car/van	*/
+ b251	/* amount paid for spare parts for car/van	*/
+ b252	/* amount paid for motorcycle servicing, repairs, spare parts	*/
+ d100105t	/* new car/van: loan or hp	*/
+ d100107t	/* new of secondhand motorcycle: loan or hp	*/
+ d100203t	/* car/van accessories and fittings	*/
+ d100205t	/* motorcycle accessories	*/
+ d100403t	/* AA and RAC subscriptions	*/
+ d100404t	/* driving lessons	*/
+ d100405t	/* antifreeze, battery water, cleaning materials	*/
+ d100406t	/* parking fees, tolls and permits	*/
+ d100408t	/* motoring fees and penalties	*/
+ d110302t	/* hire of self-drive cars	*/

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	12.04	0	2318.15
Method B	12.02	$-1.6 \times 10^{-13}$	2318.15
Method C	12.08	$-1.9 \times 10^{-14}$	2318.15

**CAT3\_**

## Food

d031005t	/* carbonated drinks	*/
+ d031101t	/* ice cream and sorbets	*/
+ d031201t	/* hot take away meals eaten at home	*/
+ d031301t	/* meals bought and eaten at workplace	*/
+ d031401t	/* hot food eaten on premises	*/
+ d031402t	/* cold food eaten on premises	*/
+ d031403t	/* hot food eaten off premises	*/
+ d031405t	/* confectionery eaten off premises	*/
+ d031406t	/* ice cream eaten off premises	*/
+ d031407t	/* soft drinks eaten off premises	*/
+ d031501t	/* hot food from other outlets	*/
+ d031503t	/* confectionery from other outlets	*/
+ d031504t	/* ice cream from other outlets	*/
+ d031505t	/* soft drinks from other outlets	*/
+ b260t	/* school meals - total paid last week	*/

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	14.35	$-2.4 \times 10^{-14}$	243.45
Method B	14.40	$-6.2 \times 10^{-14}$	243.45
Method C	14.37	0	243.45

**CAT4\_**

## Leisure goods and services

b194	/* rent for TV/cable/satellite/vcr weekly amount	*/
+ d110101t	/* purchase of boats, wheelchairs etc (assume new)	*/
+ d110102t	/* accessories and repairs of bicycles, etc.	*/
+ d110103t	/* bicycle purchase	*/
+ d120101t	/* television purchase	*/
+ d120102t	/* satellite dish purchase	*/
+ d120103t	/* satellite dish installation	*/

- + d120104t /\* video recorder purchase \*/
- + d120105t /\* audio equipment, cd player purchase \*/
- + d120106t /\* personal computers, printers, calculators \*/
- + d120107t /\* musical instruments \*/
- + d120108t /\* records, cds, cassettes, software, disks \*/
- + d120109t /\* cassette cases, record tokens, racks, etc. \*/
- + d120110t /\* blank and pre-recorded video cassettes \*/
- + d120111t /\* repair & maintenance of TVs, audio and video equipment \*/
- + d120112t /\* computer software and games cartridges \*/
- + d120113t /\* console computer games \*/
- + d120114t /\* spare parts for TV, video \*/
- + d120201t /\* sports goods and equipment \*/
- + d120304t /\* personal stationery and writing material \*/
- + d120401t /\* toys, hobbies and games (inc electronic) \*/
- + d120402t /\* photo equipment, developing, etc. \*/
- + d130101t /\* cinemas \*/
- + d130102t /\* live entertainment, theatres, concerts \*/
- + d130103t /\* admission to clubs, discos, bingo, museums \*/
- + d130105t /\* participant sports (excluding subscriptions) \*/
- + d130106t /\* spectator sport admission charges \*/
- + d130202t /\* TV rental: slot meter payments \*/
- + d130204t /\* satellite TV subscription to channels \*/
- + d130205t /\* cable TV connection and subscription \*/
- + d130304t /\* payments for school trips, etc \*/
- + b480 /\* holiday package UK \*/
- + b482 /\* holiday hotel UK \*/
- + b484 /\* holiday self-catering UK \*/
- + d110207t /\* airfares within UK \*/
- + d110208t /\* airfares outside of UK \*/

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	17.17	-2.1x10 <sup>-13</sup>	1200.78
Method B	17.24	-7.8x10 <sup>-15</sup>	1200.78
Method C	17.14	-9.3x10 <sup>-15</sup>	1200.78

### CAT5\_

#### Adult clothing and footwear

- d060101t /\* men's outerwear \*/
- + d060102t /\* women's outerwear \*/
- + d060201t /\* men's underwear \*/
- + d060202t /\* women's underwear \*/
- + d060301t /\* men's accessories - belts, headgear ties \*/
- + d060302t /\* women's accessories - belts, headgear ties \*/
- + d060401t /\* men's footwear \*/
- + d060402t /\* women's footwear \*/
- + d060404t /\* footwear - adult/child not specified \*/
- + d060501t /\* haberdashery, including buttons, needles, zips \*/
- + d060502t /\* clothing materials and charges \*/

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	13.39	-1.1x10 <sup>-14</sup>	579.63
Method B	13.46	-2.0x10 <sup>-12</sup>	579.63
Method C	13.43	-6.2x10 <sup>-13</sup>	579.63

### CAT6\_

#### Household goods

- b270 /\* furniture \*/
- + d070101c /\* furniture - purchases by children \*/



+ b271	/* soft floor coverings	*/
+ d070102c	/* soft floor coverings - purchases by children	*/
+ d070103t	/* hard floor coverings	*/
+ d070104t	/* bedding excluding beds and mattresses	*/
+ d070105t	/* curtains, cushions and other soft furnishings	*/
+ d070201t	/* electric and combined cookers	*/
+ d070202t	/* electric washing machines and spin driers	*/
+ d070203t	/* electric refrigerators and freezers	*/
+ d070204t	/* dishwashers, microwaves and other appliances	*/
+ d070205t	/* electrical tools	*/
+ d070206t	/* minor electrical equipment	*/
+ d070207t	/* gas cookers	*/
+ d070208t	/* other gas appliances	*/
+ d070209t	/* electric consumables, batteries, fuses, etc.	*/
+ d070210t	/* repairs to gas and electric appliances	*/
+ d070211t	/* gas/electric appliances: spare parts	*/
+ d070301t	/* kitchen utensils and equipment	*/
+ d070302t	/* kitchen disposables - kitchen towels, foil, etc.	*/
+ d070303t	/* china, glass, pottery, cutlery, clocks	*/
+ d070304t	/* decorative goods, eg., vases	*/
+ d070305t	/* other household hardware and appliances	*/
+ d070306t	/* matches	*/
+ d070401t	/* detergents, washing-up liquid, washing powder	*/
+ d070402t	/* disinfectants, polishes, etc	*/
+ d070501t	/* toilet paper	*/
+ d070601t	/* pet food	*/
+ d070602t	/* pet purchase, accessories, etc	*/
+ d070701t	/* garden equipment, eg, barbecues, etc	*/
+ d070702t	/* garden tools and accessories	*/
+ d070703t	/* plants, seeds, fertilisers, insecticides	*/
+ d070801t	/* unspecified household goods	*/

#### Personal goods and services

d090101t	/* toilet soap	*/
+ d090102t	/* toiletries, including cotton wool, toothpaste	*/
+ d090103t	/* toilet requisites, including flannel, nail brush	*/
+ d090104t	/* hair products, including shampoo, colourings	*/
+ d090105t	/* cosmetics and related accessories	*/
+ d090201t	/* jewellery, watches other personal effects	*/
+ d090202t	/* leather and travel goods	*/
+ d090301t	/* baby toiletries and disposables	*/
+ d090302t	/* baby equipment: prams, carry cots, bedding	*/
+ d090501t	/* glasses, lenses, prescription sunglasses	*/
+ d090502t	/* optical accessories, including contact lens fluid	*/
+ d090601t	/* hairdressing, beauty treatments, health clubs	*/
+ d090701t	/* personal goods unspecified	*/
+ d140301t	/* pocket money to child not keeping diary (assume on VA Tted goods) */	*/

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	31.98	0	1335.12
Method B	31.97	0	1335.12
Method C	31.97	-1.3x10 <sup>-14</sup>	1335.12

#### CAT7\_

Items which cannot attract VAT (no value added)

d010201t	/* main dwelling purchase/deposit	*/
+ d010601t	/* second dwelling purchase	*/
+ d010602t	/* second dwelling rent	*/
+ d100407t	/* garage rent, other costs	*/

+ b182 /\* credit purchase - net cost of secondhand vehicles \*/  
 + b245 /\* cost of secondhand car/van outright \*/  
 + d100106t /\* secondhand car/van: loan or hp \*/  
 + b265 /\* expenditure on maintenance allowance \*/  
 + hhb334 /\* amount of money sent abroad (sum over household of individual level variable b334) \*/  
 + d130405t /\* duty free goods bought in UK \*/  
 + d140302t /\* cash gifts to those outside the household \*/  
 + hhp186 /\* cash gifts deducted straight from earnings (sum over household of individual level variable p186) \*/  
 + d080101t /\* domestic help, eg, gardener, window cleaner \*/  
 + d080102t /\* child care payments \*/  
 + d080103t /\* nursery, creche, playschools \*/  
 - b187 /\* WE road tax refund over last year \*/  
 + b179 /\* WE road tax paid over last year \*/

#### Licences and taxes

d080403t /\* stamp duty, certificates and licences (excl. TV licence) \*/  
 + b181 /\* TV licence - amount paid last year \*/  
 + d010607t /\* second dwelling - TV licences \*/

#### Life insurance

b196 /\* life insurance before April 1984 - amount premium\*/  
 + b197 /\* life insurance after April 1984 – amount premium \*/

#### Postal services

d080201t /\* postage and poundage\*/

#### Finance

b180 /\* bank charges - net amount last 3 months \*/  
 + b238 /\* annual standing charge for credit cards – weekly amount \*/  
 + d080401t /\* bank and post office counter charges \*/  
 + d130407t /\* commission on travellers' cheques/foreign currency \*/

#### Education

b160 /\* education - amount paid last quarter \*/  
 + b162 /\* leisure class fees paid - amount \*/  
 + b164 /\* children outside household - education fees last quarter \*/

#### Health

d090403t /\* nhs medical, dental and optical fees \*/  
 + d090404t /\* private medical, dental and optical fees \*/

#### Burial and cremation

d080408t /\* funeral expenses \*/

#### Subscriptions to trade unions, professional and leisure organisations

d080301t /\* subscriptions - trade union / professional organisations \*/  
 + d080302t /\* subscriptions - leisure activities, eg., national trust \*/  
 + d080303t /\* subscriptions - sports and social clubs \*/  
 + d080304t /\* other subscriptions, eg political parties, library, etc \*/

#### Miscellaneous leisure activities and expenditures that are unlikely to attract VAT

d130104t /\* social events including car boot sales, etc \*/  
 + d130404t /\* money spent abroad excluding duty-free goods \*/  
 + b481 /\* holiday package - abroad \*/  
 + b483 /\* holiday hotel - abroad \*/  
 + b485 /\* holiday self-catering - abroad \*/

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	41.90	-22.58	4078.38
Method B	41.77	-22.58	4078.38
Method C	42.00	-22.58	4078.38

### CAT8\_

Zero rated food

p518t /\* total expenditure on food \*/  
- vatexp4 /\* VATted expenditure on food \*/

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	37.92	-7.1x10 <sup>-15</sup>	226.99
Method B	37.97	-7.1x10 <sup>-15</sup>	226.99
Method C	37.87	-7.1x10 <sup>-15</sup>	226.99

### CAT9\_

Books and newspapers

d120301t /\* newspapers \*/  
+ d120302t /\* magazines and periodicals \*/  
+ d120303t /\* books \*/

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	4.01	-1.3x10 <sup>-14</sup>	82.64
Method B	4.02	-9.9x10 <sup>-15</sup>	82.64
Method C	4.02	0	82.64

### CAT10\_

Domestic fuel and power

p517t /\* fuel, light and power \*/

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	12.75	-3.46	130.22
Method B	12.68	-3.46	130.22
Method C	12.74	-3.46	130.22

### CAT11\_

Fares and other travel costs

b158t /\* school travel - amount paid last week \*/  
+ b216 /\* bus/tube/rail season ticket \*/  
+ b217 /\* bus/coach season ticket \*/  
+ b218 /\* rail/tube season ticket \*/  
+ d110202t /\* railway and tube fares other than season \*/  
+ d110204t /\* bus and coach fares other than season \*/  
+ d110206t /\* combined fares other than season \*/  
+ d110209t /\* water travel \*/  
+ d110301t /\* taxis and hired cars with drivers \*/  
+ d110303t /\* other personal travel including coach trips \*/

Caravans and houseboats

d010202t /\* caravan/mobile home purchase/decoration \*/

Drugs and medicines

d090401t /\* NHS prescription charges and payments \*/  
+ d090402t /\* medicines and medical goods (not NHS) \*/

## Charitable gifts

z16=d140303t /\* charitable gifts and donations \*/

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	8.21	$-1.1 \times 10^{-14}$	6538.50
Method B	8.22	$-2.0 \times 10^{-14}$	6538.50
Method C	8.20	$-5.0 \times 10^{-14}$	6538.50

**CAT12\_**

## Childrens clothes

d060103t /\* boys' outerwear \*/  
 + d060104t /\* girls' outerwear \*/  
 + d060105t /\* infants' outerwear \*/  
 + d060203t /\* children's underwear \*/  
 + d060303t /\* children's accessories, eg., scarves, headgear \*/  
 + d060403t /\* children's footwear \*/

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	3.41	$-2.6 \times 10^{-13}$	229.74
Method B	3.38	$-3.0 \times 10^{-14}$	229.74
Method C	3.39	$-1.1 \times 10^{-13}$	229.74

**CAT13\_**

## All insurance except life insurance

b110 /\* structure insurance - last payment \*/  
 + b168 /\* contents insurance - amount of last premium \*/  
 + b188 /\* vehicle insurance - amount paid last year \*/  
 + b205 /\* deductions for friendly societies from earnings \*/  
 + b206 /\* other insurance - total amount premium \*/  
 + b208 /\* mortgage protection - pre April 1984 – amount premium \*/  
 + b213 /\* mortgage protection - post April 1984 – amount premium \*/  
 + b229 /\* medical insurance - total amount premium \*/  
 + b486 /\* expenditure on holiday insurance \*/  
 + d130406t /\* non-package holiday / other travel insurance \*/

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	9.93	$-1.0 \times 10^{-13}$	178.62
Method B	9.90	0	178.62
Method C	9.93	$-2.6 \times 10^{-14}$	178.62

**CAT14\_**

## Beer, lager

d040101t /\* beer and lager (off-licence) \*/  
 +d040201t /\* beer and lager (licenced premises) \*/

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	6.69	$-4.7 \times 10^{-13}$	170.21
Method B	6.74	$-8.3 \times 10^{-13}$	170.21
Method C	6.71	$-4.7 \times 10^{-14}$	170.21

**CAT15\_**

## Cider

d040102t /\* cider (off-licence) \*/  
 +d040202t /\* cider (licenced premises) \*/

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	.29	0	46.29
Method B	.29	-6.7x10 <sup>-16</sup>	46.29
Method C	.29	-2.5x10 <sup>-13</sup>	46.29

#### CAT16\_

Fortified wine

d040105t /\* fortified wine (off-licence) \*/  
+d040205t /\* fortified wine (licenced premises) \*/

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	.36	0	29.66
Method B	.37	-3.1x10 <sup>-15</sup>	29.66
Method C	.37	-1.9x10 <sup>-14</sup>	29.66

#### CAT17\_

Unfortified wine

d040103t /\* unfortified wine (off-licence) \*/  
+d040203t /\* unfortified wine (licenced premises) \*/

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	2.02	-5.1x10 <sup>-14</sup>	112.09
Method B	2.03	-1.7x10 <sup>-13</sup>	112.09
Method C	2.03	-5.4x10 <sup>-13</sup>	112.09

#### CAT18\_

Champagne and sparkling wine

d040104t /\* champagne and sparkling wine (off-licence) \*/  
+d040204t /\* champagne and sparkling wine (licenced premises)\*/

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	.11	0	54.49
Method B	.11	-9.3x10 <sup>-17</sup>	54.49
Method C	.11	-1.7x10 <sup>-16</sup>	54.49

#### CAT19\_

Spirits and liquers

d040106t /\* spirits and liquers (off-licence) \*/  
+d040206t /\* spirits and liquers (licenced premises) \*/

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	1.59	-1.3x10 <sup>-15</sup>	99.33
Method B	1.60	-3.3x10 <sup>-13</sup>	99.33
Method C	1.60	-2.5x10 <sup>-14</sup>	99.33

Undefined alcohol purchases are allocated according to average alcohol purchases across categories 14 to 19.

Undefined alcohol = d040107t /\* alcohol from off-licence not specified \*/  
d040207t /\* alcohol from licenced premises not specified \*/

**CAT20\_**

Cigarettes

d050101t /\* cigarettes \*/

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	5.14	$-1.2 \times 10^{-13}$	116.42
Method B	5.14	$-1.5 \times 10^{-13}$	116.42
Method C	5.14	$-3.1 \times 10^{-13}$	116.42

**CAT21\_**

Cigars

d050103t /\* cigars \*/

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	.14	0	42.78
Method B	.14	0	42.78
Method C	.14	0	42.78

**CAT22\_**

Pipe tobacco

d050102t /\* pipe tobacco \*/

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	.39	0	28.48
Method B	.39	0	28.48
Method C	.39	$-1.6 \times 10^{-13}$	28.48

**CAT23\_**

Petrol

d100301t /\* petrol \*/

Motoring fuels other than petrol and diesel

d100303t

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	9.39	$-2.0 \times 10^{-13}$	129.96
Method B	9.46	$-3.5 \times 10^{-13}$	129.96
Method C	9.36	$-1.4 \times 10^{-13}$	129.96

**CAT24\_**

Diesel oil

d100302t /\* diesel oil \*/

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	.74	0	90.20
Method B	.75	$-2.2 \times 10^{-13}$	90.20
Method C	.74	0	90.20

**CAT25\_**

Gambling on the pools

d130501t /\* football pools stakes \*/

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	.37	$-1.8 \times 10^{-14}$	41.60
Method B	.37	$-8.9 \times 10^{-14}$	41.60
Method C	.38	$-1.2 \times 10^{-13}$	41.60

**CAT26\_**

Other betting except National Lottery

d130502t /\* bingo stakes \*/  
 + d130503t /\* lottery stakes (except national lottery) \*/  
 + d130504t /\* bookmaker, tote, other betting stakes \*/  
 + d130506t /\* Irish lottery stakes \*/

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	1.16	$-1.8 \times 10^{-14}$	92.20
Method B	1.19	$-8.9 \times 10^{-14}$	92.20
Method C	1.17	$-2.4 \times 10^{-14}$	92.20

**CAT27\_**

National Lottery and other lottery stakes

d130505t /\* National Lottery stakes \*/  
 + d130507t

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	2.23	$-2.3 \times 10^{-14}$	89.00
Method B	2.24	$-1.8 \times 10^{-14}$	89.00
Method C	2.24	$-3.6 \times 10^{-14}$	89.00

**TOTAL\_ = CAT1\_ + CAT2\_ + CAT3\_ + ... + CAT27\_**

Imputation Method	Mean (£/week)	Minimum (£/week)	Maximum (£/week)
Method A	260.97	10.29	6822.91
Method B	261.05	10.29	6822.91
Method C	261.05	10.29	6822.91

## CLUSTERS

There are 7 household characteristics used to define the clusters. The description of the clusters takes the characteristics in order, stating which characteristics the households have. If there is no statement about a particular characteristic then it is allowed to take on any value i.e. no restriction, with the exception of the number of cars used by the household which is always specified. The characteristics are:

- Number of cars used by the household (not just owned) – 0, 1, 2+.
- Number of adults in the household – over 16 years or over 18 years if in full-time education, at least one in every household.
- Number of children in the household – under 17 years or 17/18 years in full-time education  
OR children present in the household – none, 1+.
- Number of adults of each gender in the household.
- Housing tenure – ownall (own outright), ownsome (have mortgage or partly own), rent (rentf = furnished, rentunf1 = unfurnished from Local Authority or Housing Association, rentunf2 = other unfurnished), and rfree (rent free).
- Employment status of adults in the household – work (employed or self-employed), retired, other (unemployed, sick or unoccupied).
- Expenditure region – high (Greater London, South East, East Midlands), mid (North West, Yorkshire and Humberside, West Midlands, East Anglia, South West, Scotland), low (North, Wales).

The numbering of the clusters and the order in which the clusters are listed is significant. The numbering relates to the tool used to define the clusters (see Taylor (2000)). The listed order of the clusters relates to the description offered for each cluster. The description is a translation of the programme written to create the clusters which should be read as “all of the households included in the definition minus households who fall into clusters listed above”.



## Test 42b.1 – associated with Methods A and C

The clusters have been defined using the variables: number of cars, number of adults, disaggregated housing tenure and employment status.

Cluster Number	Test 42b.1	FRS 1995/6	
		Frequency	Percent
1	Single retired (no other, work), rentunf1, no car	1596	6.0
2	Single retired (no other, work), ownall, no car	1235	4.7
3	Single other (no retired, work), rentunf1, no car	1182	4.5
4	Retired >1 (no other) OR retired&other (no work), rentunf1, no car	478	1.8
5	1/2 retired, rent OR 2 retired, ownall, (no other, work), no car	671	2.5
6	1/2 retired (no other, work), rentunf1, 1 car OR (retired&other OR 1/2/3 other (no work)), ownall OR 1 retired&>0 other, (rentunf2 OR rentf OR rfree), no car	524	2.0
7	Other (no retired, work), rentunf1 OR rent, no car	888	3.4
8	1 work&0/1/2 retired (no other), rentunf1, no car	449	1.7
9	Single retired (no other, work), ownall, 1 car	650	2.5
10	2 retired (no other, work), ownall, 1 car	1211	4.6
11	1 work&1 retired (no other) OR 1 work&1 other (no retired), no car	527	2.0
12	Retired and/or other (no work), not ownsome, 1 car	1213	4.6
14	Single worker (no retired, other), not rentunf1, no car	708	2.7
20	Workers (no retired, other), rentunf2 OR rentf OR rfree, 1 car	539	2.0
21	1 work&1 other (no retired), ownsome, 1 car	911	3.4
22	1 work&1 other (no retired), ownsome, 1 car OR (1 retired&1 other (no work) OR 2 other (no retired, work)), ownsome, 2+ cars OR 1 retired&1 other&work >0, ownsome OR ownall, 2+ cars OR 1 retired&work >0, rentunf2 OR rentf OR rfree, 2+ cars OR (1 retired&work >0 OR 1 other&work >0), rentunf1 OR ownall, 2+ cars	687	2.6
24	(1other&work >0, rentunf2 OR rentf OR rfree) OR (1 retired&1/2 work (no other) OR 2 retired&1/2 work (no other) OR 1 other&work >1 (no retired) OR 2 other&work >0 (no retired)), ownsome, 2+ cars	421	1.6
23	Workers (no retired, other), not ownsome, 2+ cars OR work >2 (no retired, other), ownsome, 1 car	813	3.1
25	2 workers (no retired, other), ownsome, 1 car	2520	9.5
26	Single other (no retired, work) OR 1 other&1 work (no retired), ownsome, 2+ cars	523	2.0
27	Single worker (no retired, other), ownsome, 1 car	1146	4.3
28	Work >2 (no retired, other), ownsome, 2+ cars	413	1.6
29	1/2 workers (no retired, other), ownsome, 2+ cars	2276	8.6
18	Retired and/or other (no work), 1+ cars	853	3.2
13	Retired and/or other, no car	568	2.1
15	Work >1 (no retired, other), no car	510	1.9
19	Workers (no retired, other), rentunf1 OR ownall, 1 car	1207	4.6
16	1 work&1 other (no retired), not ownsome, 1 car	608	2.3
17	Workers, retired and/or other, 1+ cars	1108	4.2
Total		26435	100.0

## Test 41.1 – associated with Method B

The clusters have been defined using the variables: number of cars, expenditure region, aggregated housing tenure, number of children, number of adult females, number of adult males and employment status.

Cluster Number	Test 41.1	FRS 1995/6	
		Frequency	Percent
1	Workers (no retired, other), ownsome/ownall, no children, 2+ cars	1528	5.8
2	Workers (no retired, other), ownsome/ownall, children, 2+ cars	1570	5.9
3	Workers (no retired, other), ownsome, no children, high, 1 car	890	3.4
4	Workers (no retired, other), ownsome, children, high, 1 car	620	2.3
5	Workers (no retired, other), ownsome, no children, mid/low, 1 car	1222	4.6
6	Work >0&(retired and/or other), ownsome, 2+ cars	926	3.5
7	Workers (no retired, other) ownsome, children, mid/low, 1 car	1106	4.2
8	Workers (no retired, other), ownall, 1 car	762	2.9
9	(Work >0 (no retired, other) OR work >0&(retired and/or other)), rent OR work >0&(retired and/or other), ownall OR rfree, 2+ cars	883	3.3
10	Work >0&1 other (no retired), ownsome, 1 car	1062	4.0
11	Retired and/or other (no work), 2+ cars	346	1.3
12	Work >0 (no retired, other), rent OR work >0&(retired and/or other), ownsome, children, 1 car	439	1.7
13	Retired and/or other (no work), ownsome, 1 car	542	2.1
14	Workers (no retired, other), rent, no children, 1 car	594	2.3
15	Workers (no retired, other) OR work >0&(retired and/or other), ownsome, no car	895	3.4
16	Work >0&(retired and/or other), not ownsome, children, 1 car	432	1.6
17	Retired and/or other (no work), not ownsome, children, 1 car	417	1.6
18	Work >0&(retired and/or other), no children, 1 car	1103	4.2
19	Workers (no retired, other) OR work >0&(retired and/or other), not ownsome, children, no car	519	2.0
20	Workers (no retired, other) OR work >0&(retired and/or other), not ownsome, no children, no car	991	3.7
21	Retired and/or other (no work), children, no car	1237	4.7
22	Retired and/or other (no work), ownsome, no children, no car OR other >0 (no retired, work), not ownsome, no children, 1 car	664	2.5
25	Retired (1female low/mid OR 2female mid OR 1male&1female low/mid) (no other, work), ownall, no children, 1 car	1030	3.9
23	(Retired >0 (no other, work) OR retired >0&other >0 (no work)), ownall OR rfree, no children, 1 car	1056	4.0
24	Retired >0 (no other, work) OR retired >0&other >0 (no work), rent, no children, 1 car	467	1.8
26	Other >0 (no retired, work), not ownsome, no children, no car	1095	4.1
29	Retired (1male low/mid OR 2female mid, 1male&1female low/mid) (no other, work), rent, no children, no car	511	1.9
30	Single retired female (no other, work), rent, no children, low/mid, no car	904	3.4
31	Single retired female (no other, work), ownall OR rfree, no children, low/mid, no car	670	2.5
27	Retired >0 (no other, work) OR retired >0&other >0 (no work), rent, no children, no car	842	3.2
28	Retired >0 (no other, work) OR retired >0&other >0 (no work), ownall OR rfree, no children, no car	1112	4.2
Total		26435	100.0