

Family Resources Survey

RELEASES 1997-98

RELEASE	CHANGES SINCE LAST RELEASE	RELEASE DATE
frs978a	FIRST RELEASE	17/03/99
frs978b	SECOND RELEASE. <ul style="list-style-type: none">• Adjustments to high SEPRO (profit amount) value (1 case).• Imputed TO1HR (hours worked) when EVROT (whether works overtime) had been imputed to 2. (59 cases)• Additional assets records included.• Imputed TTWFAR to 2 (distance travels to work varies) where it had previously been imputed to 4 (5 cases)• Added VAR variables including DC, DM, AA, JSTYP, MDUP, BO (benefit abbreviated), LT (benefit abbreviated), ST (benefit abbreviated).• Imputed TICKET to 2 (does not have a travel pass) where it had previously been imputed to 1 (4 cases).• Imputed HEALTH to 2 (does not have a long term illness) where it had previously been imputed to 1 (4 cases).• Imputed LIMIT to 2 (does not have an illness that limits activities) where it had previously been imputed to 2 (1 cases).• Imputed SMPRT to skipped (rate of statutory sickness pay) where it had previously been imputed to 1 and there was a value for SMPAM (amount of statutory sickness pay) (8 cases).• Correction to head of household person number to ensure that it is equal to head of benefit unit UPERSON.• Edited benefit amount causing negative income from benefits (1 case).• Deleted 1 case where 4 year old caring 6 hours a week• Government region now a DV as last year in order to group Merseyside with the North West. GVTREGN.• TOTCAPBU investigated and new assets data has increased the mean/median.• Correction (by creation of DV's) of ADULTH (number of adults in the household) and DEPCHLDH (number of children in the household).	19/05/99

	<ul style="list-style-type: none"> • Edited MRDSS2 to 2 where maintenance is received via DSS/CSA and amount received is given (MR2>0) • Addition of CHEARNS and HPERSON DV's on child table on flatfile 	
	<ul style="list-style-type: none"> • Rounded SEINCAMT in 47 cases • Deleted SMPAMT transactions in 7 cases from 82 to -9. • Deletion of wrongly created IS records in 3 cases • Created housing benefit record in 1 case 	
frs978c	<ul style="list-style-type: none"> • New grossing factors Changed grossing factors so that methodology for grossing consistent with that used in 1996-97. Previously, for 1997-98 an alternative methodology was used. 	20/7/99
frs978d	<ul style="list-style-type: none"> • Correction to INTRPAY. Conversion problem with INTRPAY – corrected by ONS. • Correction to NINEARNS. Error found with DV – corrected, minor affect on data. • Improvements to EMPSTATI. Improvements to this DV to ensure more ILO consistent. • Recoding of 14 values of SIC=89 to 0 and adding the category of 0=undefined to SIC (SIC removed – found to be unreliable) • Correction to DEPDEDs dv and put on flatfile 	2/11/99
frs978e	<ul style="list-style-type: none"> • Definitional changes to earnings derived variables to bring FRS earning dvs more in line with HBAI. This has lead to minor changes to various earnings and income related derived variables. For more details of the exact changes please contact the FRS team. 	23/5/00
frs978f	<ul style="list-style-type: none"> • Interim Grossing factor added (GROSSINT) 	25/11/02
frs978g	<p>Amendment of travel to work costs to use dual mileage rates (TTWCOST). See 2002-03 Changes documentation for full details.</p> <p>In 1996-97 the derived variable for specific household tenure types (TENTYPE) was changed from ten categories to eight. The format attached to this variable was not updated. This has now been corrected for</p>	24/11/03

	<p>all affected years.</p> <p>Family Type (FAMTHBAI) definition adjusted to be in line with HBAI definition introduced in 2001-02. See 2002-03 Changes documentation for full details.</p>	
Frs9798h	New Grossing regime (GROSS3) introduced - See paper for more details	22/11/04
Frs9798i	Revised weights issued for the new Grossing regime (GROSS3).	27/01/05
Frs9798j	Revised weights issued for the new Grossing regime (GROSS3) to correct for overestimation of the lone parent population control .	09/02/05

FAMILY RESOURCES SURVEY 1997/98:

SUMMARY OF EDITING AND IMPUTATION PROCEDURES CARRIED OUT BY DSS

For the 1997/98 data set, the following tasks were carried out by DSS.

1 Conversion of monetary amounts to weekly values

Many of the questions on the FRS ask for amounts received/paid and to what period they relate (eg benefit receipt, council tax payments). In these cases, amounts were converted to weekly equivalents. More information on which period code relates to which value is given in the Excel spreadsheet period34.xls.

1.1 During the conversion process amounts were not converted where:

- 1.1.1 payments were one off or lump sum payments (period code 95)
- 1.1.2 "none of the above" (period code 97)
- 1.1.3 period code missing
- 1.1.4 payments were less than 1 week (period code 90)

1.2 However, for those items of income and expenditure which feed in to derived variables used by the DSS, missing, 90, 95 and 97 period code payments were scrutinised and edited to a weekly value. Remaining 90, 95 and 97 period codes will appear in analyses as outliers. Users will need to consider whether to edit or delete these cases. The easiest way to identify such variables is to consult minmaxan.xls and search on maximum values of 95 or 97. The link between period codes and monetary amounts is given in period34.xls.

2 Validation, editing and imputation

Information about procedures carried out by DSS are contained in the file methodology chapter of the latest FRS publication.

2 Anonymisation

2.1 OPCS/SCPR have their own procedures to ensure the confidentiality of respondents. Names and addresses are kept separately from the data and are not supplied to the DSS.

2.2 Additional steps have been taken by the DSS prior to release of the data outside the department. These are:

2.2.1 Local Authority Code has been removed.

2.2.2 Monetary amounts relating to council tax variables have been rounded to whole pounds. Variables affected are:

Variable	Description	Table
ctamt	last CT payment	househol
ctrebamt	amount of CT rebate	househol
ctredamt	amount of transitional reduction	househol
cwatamt	amount included in rent for CT water charge	househol
indinc	Derived Variable (DV) for adult income	adult
inrpinc	DV for adult RP/IS income	adult
indisben	DV for adult disability benefit income	adult
inirben	DV for adult income related benefit income	adult
innirben	DV for adult non-income related benefit income	adult
inothben	DV for adult other benefits	adult
buinc	DV for benefit unit income	benunit
burpinc	DV for benefit unit RP/IS income	benunit
budisben	DV for benefit unit disability benefit income	benunit
buirben	DV for benefit unit income related benefit income	benunit
bunirben	DV for BU non-income related benefit income	benunit
buothben	DV for BU other benefit income	benunit
hhinc	DV for household income	househol
hhrpinc	DV for HH RP/IS income	househol
hhdisben	DV for HH disability benefit income	househol
hhirben	DV for HH income related benefit income	househol
hhnirben	DV for HH non-income related benefit income	househol
hhothben	DV for HH other benefit income	househol
hbeninc	DV for HH benefit income	househol
cwatamtd	DV for council tax water charge	househol
burent	DV for BU rent	benunit
hhrent	DV for HH rent	househol
hscosthh	DV for HH housing costs	househol

2.3 However, assurances given to interviewees allow DSS to provide unanonymised data in very restricted circumstances. For more information, please contact Mehdi Hussain at the address given below.

ASD3E
Analytical Services Division
Department of Social Security
4th Floor
The Adelphi
1-11 John Adam Street
London WC2N 6HT

3 For further information write to Mehdi Hussain at the same address or telephone 0171-962-8092

IMPUTATION OF THE 1997/98 FAMILY RESOURCES SURVEY

1. INTRODUCTION

A combination of methods were used in to impute values for missing data in the 1997/8 FRS dataset. The sections below provide an outline of the imputation process as well as more detailed descriptions of each of the different methods used. This document, intended for user information, applies to the imputation of missing data in all the FRS tables except BENEFITS, for which a different approach is used.

2. METHODOLOGY

The process of imputation for 1997-98 was:

1. **Bulk Edits.** For some missing values the imputed value is the same in every case e.g. closing down a route.
2. **Algorithms.** The algorithm may simply just set all missing values to a set value, e.g. the mean, or it may be more complex and involve formulae relating to a number of other variables, where a relationship exists.
3. **Hot Decking.** Hot deck imputation is fundamentally a process of assigning a non-missing value, taken from a 'donor' case, to a 'target' case, which had a missing value for the variable of interest. The donor case has to fit the same 'factor' values as that of the target case. Factor values are variables, which will have an impact on the variable to be imputed.
4. **Neural.** Neural Networks can impute data by finding patterns in the dataset and then applying those patterns to the data that require imputation.
5. **Mop-up.** Any remaining variables will be imputed at this stage in the imputation process.

In general the process was performed in the order as given above (with the aim of providing as complete a dataset as possible for the neural imputation stage).

3. VARIATION IN MISSING VALUES

Approximately 60% of all FRS variables had some missing values. Of these 63% were missing less than 1% of expected completed values. Table 1a lists the variables where 10% or more of expected completed values were missing, along with the method of imputation used for each. Table 1b lists the top thirty variables with the highest number of missing values.

4. SUMMARY

Table 2 attached provides an overall summary of imputation outlining the number of missing values initially and how many were imputed by each method. It also provides a comparison with previous years, where possible. Users will note a small fall in the total percentage of missing values. This is most likely due to more efficient closing down of routes in the questionnaire where respondents are unlikely to know the answer to a particular question, for example type of council tax rebate, which is only asked where a statement has been consulted.

Table 1a. Variable with 10% or more of expected values missing.

TABLE	VARIABLE	DESCRIPTION FROM LABELS34	No. MISSING	% MISSING	METHOD OF IMPUTATION
ADULT	LN2RPINT	2nd loan repayment incl interest/capital	2	67%	ALGORITHM
HOUSEHOL	FRNINS2	Premium: Insured value of the furniture	176	64%	HOTDECK
RENTER	SERVAMT	Amount included in rent for services	250	62%	HOTDECK
JOB	SENIAMT	Amount of National insurance lump sum	257	58%	RSL SYSTEM
JOB	SMPRATE	Higher/lower rate of SMP	8	57%	ALGORITHM
HOUSEHOL	FRNINS1	Repayment : Insured value of furniture	201	54%	HOTDECK
ADULT	LN1RPINT	1st loan repayment incl interest/capital	9	45%	ALGORITHM
RENTER	WSINCAMT	Amount incl in rent for water/sewerage	731	45%	HOTDECK
HOUSEHOL	STRINS2	Insured value of the structure	342	42%	HOTDECK
JOB	TAXDAMT	How much income tax was deducted last time	78	41%	HOTDECK
ADULT	APAMT	Amount received from absent partner	15	39%	RSL SYSTEM
JOB	PRBEFORE	Profit before or after tax	85	36%	HOTDECK
ADULT	APDAMT	Amount from absent partner paid directly	20	34%	RSL SYSTEM
HOUSEHOL	STRAMT1	Amount: Insurance part of repayment	1006	27%	HOTDECK
ASSETS	HOWMUCHE	Value of asset (office edit)	1059	26%	LEAVE AS MISSING
HOUSEHOL	STRINS1	Repayment: Insured value of the structure	250	26%	HOTDECK
JOB	PROFIT1	Amount of net profit or loss	562	25%	ALGORITHM THEN HOTDECK
JOB	SETAXAMT	Amount of tax in last 12mths (s-empl)	320	24%	ALGORITHM
JOB	SMPAMT	Amount included in last wage for SMP	14	21%	ALGORITHM
ADULT	ROYR2	Amount of income as sleeping partner	8	20%	RSL SYSTEM
ADULT	OTAMT	Amt of extra income tax in last 12 mths	183	18%	RSL SYSTEM
ADULT	TTWREC	Travel costs: passengers' contributions	26	17%	RSL SYSTEM
ADULT	ED2SUM	Amount originally borrowed: 2nd loan	1	17%	IMPUTE SEPARATELY
JOB	SSPAMT	Amount included in last wage for SSP	17	16%	ALGORITHM
JOB	SEINCAMT	What is your income from this business	314	16%	HOTDECK
HOUSEHOL	CTCCDIS	Whether CTB includes childcare discount	211	15%	ALGORITHM
MORTGAGE	INCMSTY3	Year MPP taken out	1	14%	LEAVE AS MISSING
JOB	NIDAMT	How much National Insurance was deducted	10	14%	HOTDECK
JOB	OWNOTAMT	Amount: Other drawings from business	9	13%	HOTDECK
ASSETS	HOWMANY	Number of shares/bonds/units held	540	13%	HOTDECK
PENSION	TRIGHTS	Applying to trustees to take money out	9	13%	RSL SYSTEM
ACCOUNTS	ACCINT	Interest received	10210	12%	HOTDECK
ASSETS	ISSVAL	Value of NSC	14	12%	HOTDECK
ADULT	ALLPAY1	Amount of allowance: Friend/Relative	15	12%	RSL SYSTEM
ADULT	GRTVAL1	Current value of 1st grant (incl fees)	10	12%	ALGORITHM
JOB	GRSOFAR	Taxable gross earnings so far this year	996	11%	LEAVE AS MISSING
MORTCONT	OUTSAMT	Amount paid towards mortgage last time	40	11%	ALGORITHM
CHILD	CHAMTST	Income received from a Trust	4	11%	HOTDECK
ADULT	SHAREPAY	Amount of contribution towards housing etc	59	10%	RSL SYSTEM
HOUSEHOL	CHRGAMT8	Amount paid for other regular charges	7	10%	RSL SYSTEM

Table 2a. Top thirty variables with the highest number of expected values missing.

TABLE	VARIABLE	DESCRIPTION FROM LABELS34	No. MISSING	% MISSING	METHOD OF IMPUTATION
ACCOUNTS	ACCINT	Interest received	10210	12%	HOTDECK
ASSETS	HOWMUCH	Value of asset (respondent's estimate)	1936	9%	HOTDECK
BENEFITS	BENAMT	Benefit amount	1416	5%	BENEFIT EDITS
ASSETS	HOWMUCHE	Value of asset (office edit)	1059	26%	LEAVE AS MISSING
JOB	PAYAMT	Amount of last take home pay	1008	5%	HOTDECK
HOUSEHOL	STRAMT1	Amount: Insurance part of repayment	1006	27%	HOTDECK
JOB	GRSOFAR	Taxable gross earnings so far this year	996	11%	LEAVE AS MISSING
JOB	PAYE	Amount deducted for PAYE	981	5%	HOTDECK
JOB	NATINS	Amount deducted for NI	969	5%	HOTDECK
HOUSEHOL	STRAMT2	Amount: Insurance premium	845	7%	HOTDECK
RENTER	WSINCAMT	Amount incl in rent for water/sewerage	731	45%	HOTDECK
PENSION	PENPAY	Amount of last payment from pension	611	8%	HOTDECK
HOUSEHOL	WSEWAMT	Combined water/sewer rates: amount paid	608	4%	HOTDECK
ADULT	TTWMOD1	Travels to work by walking/bicycle	597	3%	RSL SYSTEM
ADULT	TTWMOD2	Travels to work by car/van	597	3%	RSL SYSTEM
ADULT	TTWMOD3	Travels to work by motorcycle	597	3%	RSL SYSTEM
ADULT	TTWMOD4	Bus/train/tube	597	3%	RSL SYSTEM
ADULT	TTWMOD5	Works bus/company transport	597	3%	RSL SYSTEM
ADULT	TTWMOD6	Travels to work by other means	597	3%	RSL SYSTEM
JOB	PROFIT1	Amount of net profit or loss	562	25%	ALGORITHM THEN HOTDECK
ASSETS	HOWMANY	Number of shares/bonds/units held	540	13%	HOTDECK
MORTGAGE	MORTLEFT	Amount outstanding on mortgage	531	6%	HOTDECK
ENDOWMNT	MENPOLAM	Last premium on endowment policy	503	6%	HOTDECK
INSURANC	POLAMT	Premium: Amount Paid on this Policy	418	7%	ALGORITHM
HOUSEHOL	STRINS2	Insured value of the structure	342	42%	HOTDECK
JOB	SETAXAMT	Amount of tax in last 12mths (s-empl)	320	24%	ALGORITHM
JOB	SEINCAMT	What is your income from this business	314	16%	HOTDECK
PENSION	PTAMT	Amount of tax deducted at source	305	6%	HOTDECK
JOB	DEDUC1	Amount deducted:pensions/superannuation	284	4%	HOTDECK
PENSION	PENTAX	Whether tax deducted at source on PENPAY	270	3%	RSL SYSTEM

Table 2. Summary of imputation

	1994/5 (v31)		1995/6 (v32)		1996/7 (v33)		1997/8 (v34)	
	Number	% required values	Number	% required values	Number	% required values	Number	% required values
Required values	10,409,900		11,279,900		11,642,424		12,083,694	
Missing values	55,300	0.53	57,100	0.51	51,917	0.45	47,523	0.39
Type of	1994/5 (v31)		1995/6 (v32)		1996/7 (v33)		1997/8 (v34)	
Imputation	Number	% missing values	Number	% missing values	Number	% missing values	Number	% missing values
Hot decks			32,100	56.2	28,189	54.3	25225	53.1
Gatekeeper/Batc h					1,883	3.6	1954	4.1
Algorithm					7,275	14.0	1412	3.0
Neural - mean			4,300	7.5	3,557	6.9	2828	6.0
Neural - neural	6,900	12.5	5,900	10.3	3,933	7.6	8397	17.7
Neural total	6,900	12.5	10,200	17.9	7490	14.4	11225	23.6
Total	6900.00	12.5	42300.00	74.1	44837.00	86.4	39816.00	83.8

Julie Stanborough
ASD3E

p:\frs\shared\frs34\impute\documentation\impute.doc