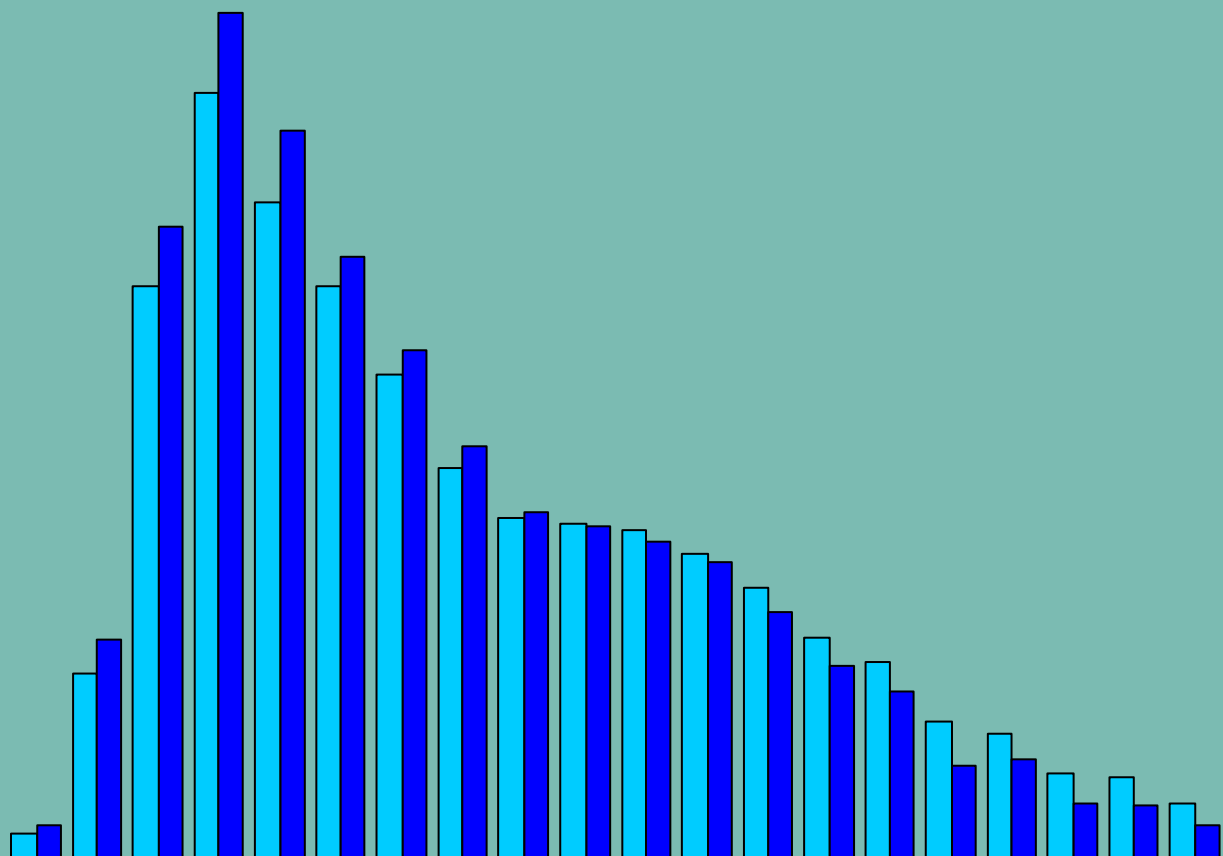


# Income imputation in the SHS and SHCS: examining the feasibility of using the FRS to broaden the measure of household income.

**Chris Martin,**  
Ipsos MORI Scotland

24 February 2009



# Contents

|  |           |
|--|-----------|
| <b>1. Introduction.....</b>  | <b>3</b>  |
| <b>2. Comparison of the FRS, SHS and SHCS surveys.....</b>   | <b>7</b>  |
| Sampling and timescales .....  | 7         |
| Definitions of Household and Household Reference Person .....  | 8         |
| Questionnaire coverage of income .....   | 9         |
| Questionnaire coverage of other factors.....   | 11        |
| Definitions of Household Income.....   | 14        |
| <b>3. Estimating the effect of definitional differences on measures of total household income .....</b>          | <b>17</b> |
| <b>4. Examining strategies to impute income for ‘other adults’.....</b>  | <b>23</b> |
| A short introduction to imputation.....  | 23        |
| Three potential strategies for modelling adult income data .....   | 25        |
| Modelling the determinants of income using variables common to the FRS, SHS and SHCS .....                       | 30        |
| Summary of proposed imputation methodology for other adults.....   | 35        |
| <b>5. Testing the proposed imputation strategy .....</b>   | <b>39</b> |
| Measuring the effect of the proposed imputation strategy.....  | 39        |
| Examining the robustness of using imputed data.....  | 40        |
| Limitations of using household income estimates in the SHS and SHCS .....  | 47        |
| <b>6. Other potential ways to improve the income data in the SHS and SHCS .....</b>                              | <b>51</b> |
| Data cleaning and processing routines .....  | 51        |
| Changes to the questionnaire .....   | 53        |
| Additional imputation for cases where the current measure of income is set to missing.....                       | 54        |
| Undertaking an annual re-imputation of the income data in the SHS together with the imputation of the SHCS ..... | 56        |

|   |           |
|---|-----------|
| <b>7. Summary and recommendations .....</b>                               | <b>61</b> |
| <b>Summary of recommended imputation strategy .....</b>                   | <b>62</b> |
| <b>Summary of potential changes to the SHS and SHCS</b>                   |           |
| <b>questionnaires.....</b>  | <b>62</b> |
| <b>Summary of data processing recommendations.....</b>                    | <b>63</b> |
| <b>Appendix 1: Detailed approach to the imputation of income of other</b> |           |
| <b>adults using hot-deck imputation .....</b>                             | <b>65</b> |
| <b>Retired adults.....</b>  | <b>65</b> |
| <b>Students .....</b>   | <b>67</b> |
| <b>Unemployed, looking after home, sick/disabled.....</b>                 | <b>69</b> |
| <b>Full-time workers .....</b>  | <b>72</b> |
| <b>Part-time workers .....</b>  | <b>75</b> |
| <b>Self-employed .....</b>  | <b>77</b> |
| <b>Appendix 2: Additional modelling tables for the SHS .....</b>          | <b>81</b> |



# 1. Introduction

- 1.1 In Scotland, the Scottish Household Survey (SHS) is the main general household survey, providing information on the composition, characteristics and behaviour of Scottish households. It covers a wide range of topics and allows links to be made between different policy areas at both a national and local level. Similarly, the Scottish House Condition Survey (SHCS) is the main national survey of the housing stock providing information on both households and dwellings. Neither of these surveys were, however, designed with the primary aim of collecting robust estimates of income.
- 1.2 The Family Resource Survey (FRS) is the currently accepted source of information on income across the UK, including Scotland. However, the FRS has a much smaller sample size in Scotland than the SHS (around 4,500 compared with around 15,500 a year) with limited scope for detailed analyses of sub-groups and smaller areas.
- 1.3 The definition of household income used by the FRS is different from that used by the SHS and SHCS. Most importantly, while the SHS and SHCS define household income as the income of the Highest Income Householder and their spouse<sup>1</sup>, the FRS definition includes income from all household members including children. There are various other differences between the SHS, SHCS and the FRS, which impact on the comparability of the household income measures these sources provide.
- 1.4 Information on income is of critical and growing importance, for example, to analyse the distribution of poverty and fuel poverty across Scotland. This has raised the question: are there practical ways in which the SHS and SHCS estimates of household income can be - and how far they should be - brought into line with those from the FRS through imputation or any other means?
- 1.5 This paper compares the SHS, SHCS and the FRS in order to examine the feasibility of imputing income data for all adults through modelling of FRS data. It compares the methodologies of the surveys, and examines whether there are significant differences in how information on income is collected and processed. It discusses the extent to which these differences, together with definitional differences, may affect estimates of household income. It explores potential strategies for imputation and makes recommendations on the best potential

model for imputation and the robustness of the resulting estimates. The proposed imputation methodology is summarised in paragraph 4.42 and detailed in Appendix 1. Finally, it discusses other potential modifications to improve the household income measure in the SHS and SHCS.

---

<sup>1</sup> It also includes the contribution of other household members to household finances by recording income received by the HH (and spouse) as ‘dig money’.



# Comparison of the FRS, SHS and SHCS surveys



## 2. Comparison of the FRS, SHS and SHCS surveys

### Sampling and timescales

- 2.1 The sample coverage and the sampling strategy for the SHS, SHCS and the FRS are broadly consistent. All three are designed to provide nationally representative samples of private households across the whole of Scotland<sup>2</sup>. All three surveys use the Postcode Address File (PAF) for Scotland as the sampling frame.
- 2.2 There are minor differences with regard to clustering and stratification between the surveys. The SHS uses a combination of clustered and unclustered sampling. In areas of high population density, a simple random sample of dwellings is selected. In areas of lower population density, datazones are selected before addresses within datazones are sampled. The SHCS is, however, completely unclustered.
- 2.3 The FRS sample in Scotland is clustered, using postcode sectors as the primary sampling unit. The SHS uses the urban rural classification and the Scottish Index of Multiple Deprivation for stratification<sup>3</sup>, while the FRS uses three variables: the proportion of economically active adults; the proportion of economically active men who are unemployed; and the proportion of households where the main householder is in NS-SEC categories 1 to 3. These differences however are minor and do not raise significant difficulties in using the FRS as a resource for designing imputation strategies for the SHS.
- 2.4 However, there are two factors that may influence the scope for using FRS data to impute income in the Scottish surveys, namely their respective sample sizes and the data delivery timescales.
- 2.5 Annually, the SHS collects around 15,500 surveys, while the SHCS collects around 3,800 surveys<sup>4</sup>. In contrast the FRS, although it boosts the number of surveys conducted in Scotland, collects around 4,400 interviews. Around 14-15% of households contain “other adults”, household members where income information is collected in the FRS but not in the SHS/SHCS.

---

<sup>2</sup> Before 2003-04, the FRS excluded addresses north of the Caledonian Canal with the exception of the Inverness post-code area.

<sup>3</sup> Previously, the SHS used Mosaic for stratification in place of the SIMD.

<sup>4</sup> Household surveys. Annually, the target for both household and dwelling surveys is around 3,000.

**Table 2-1: Number of Benefit Units in Scotland in the FRS, 2005-06.**

| No. of Benefits Units | Unweighted Count | Unweighted Distribution | Weighted Distribution |
|-----------------------|------------------|-------------------------|-----------------------|
| 1                     | 3809             | 85.8%                   | 83.7%                 |
| 2                     | 485              | 10.9%                   | 11.8%                 |
| 3                     | 124              | 2.8%                    | 3.7%                  |
| 4                     | 17               | 0.4%                    | 0.6%                  |
| 5                     | 2                | 0.1%                    | 0.1%                  |
| 6                     | 1                | 0.0%                    | 0.1%                  |
| Total                 | 4438             | 100%                    | 100%                  |

- 2.6 Table 2-1 shows the sample size by number of benefit units. All households with more than one Benefit Unit – this term is described in full in the following section - will contain “other adults”. The number of interviews collected across Scotland by the FRS in households with “other adults” is around 600-650. This limits the potential imputation strategies available<sup>5</sup>.
- 2.7 With regard to timescales, the FRS is structured around the financial year. Fieldwork periods run from April to March, with the first release of data around twelve months after the end of fieldwork. For the SHS and SHCS, fieldwork is organised around calendar years, with the data released around 7 and 9 months respectively after the end of the fieldwork period. **Any strategy that requires FRS data for the same fieldwork period as the SHS/SHCS will introduce a delay of at least a year to the delivery of the household income measure in the SHS/SHCS.**

## Definitions of Household and Household Reference Person

- 2.8 The definition of a household is consistent between the FRS and the SHS/SHCS; namely ‘a single person or group of people living at the same address who either share one meal a day or share the living accommodation, i.e. a living room’. For example, a group of students with a shared living room would be counted as a single household even if they did not eat together, but a group of bedsits at the same address would not.
- 2.9 The three surveys define a Household Reference Person (HRP<sup>6</sup>) in the same way. The HRP is classified as the highest income householder, without regard to sex. In a single adult household the HRP is the sole householder (i.e. the person

<sup>5</sup> This is discussed in greater depth in Section 3.

<sup>6</sup> The HRP in the SHS is commonly referred to as the HIH, the Highest Income Householder. As both definitions are the same, for ease, we use the term HRP.

in whose name the accommodation is owned or rented). If there are two or more householders, the Head is the householder with the highest personal income from all sources. If there are two or more householders who have the same income, the eldest householder will be the Head.

- 2.10 In the FRS, households are broken down into one or more benefit units. 'Benefit unit' is a standard DWP term that relates to the tighter family definition of 'a single adult or couple living as married and any dependent children'. A dependent child is aged under 16 or an unmarried 16 to 18-year-old in full time non-advanced education. For example, a husband and wife living with their young children and an elderly parent would be one household but two benefit units. While the SHS and SHCS do not routinely use the concept of Benefit Units, these can be created for most years of the surveys using the household grid data<sup>7</sup>.

### Questionnaire coverage of income

- 2.11 As the FRS is primarily concerned with estimating receipt of income, it covers receipt of all benefits in detail, both in the coverage of sources of income, and in the amount of information it collects about each source. Neither the SHS nor the SHCS can give the same time in the questionnaire to gathering information on income. It is also worth noting that the FRS is a longer questionnaire, with an average interview length of close to 90 minutes<sup>8</sup>, compared to around 45 minutes for the SHS and SHCS.
- 2.12 With regard to the coverage of sources of income, the main difference is that the FRS asks about the income of all members of the household, while the SHS and the SHCS only asks about the income of the Highest Income Householder and their spouse if necessary.
- 2.13 However, in terms of what components of income are covered by the two surveys, the differences between the two surveys are minor. Both ask about income from earnings, from a variety of benefits, and from various miscellaneous sources, such as private pensions, income from investments etc. Comparing the 2005/2006 questionnaires, the following components were asked about in the FRS but not specifically in the SHS:

---

<sup>7</sup> Since its inception in 1999, the SHS has collected standard demographic information on all household members. For some years however, only the relationship between the HRP and other members has been collected, rather than the relationships between all household members.

<sup>8</sup> One hour and 26 minutes – see [http://www.dwp.gov.uk/asd/frs/2005\\_06/chapter8.pdf](http://www.dwp.gov.uk/asd/frs/2005_06/chapter8.pdf)

- Lone Parent Benefit Run-On
- War widows Pension
- Maternity grant, funeral grant or community grant from Social fund.
- Guardian's allowance
- Work search premium
- Bereavement payment lump sum
- Winter fuel payment lump sum
- Back to Work Bonus
- Child Maintenance Bonus
- Government Training Allowances

2.14 The SHCS covers slightly more components of income than the SHS. The following components of income were asked about in the FRS but not specifically in the SHCS:

- Work search premium
- Bereavement payment lump sum
- Back to Work Bonus
- Government Training Allowances

2.15 Almost all of the benefits that are listed above are received by a small proportion of households, and therefore, do not comprise a major component of income. Winter fuel payments, however, are received by almost all pensioner households. Although they are not generally the main source of income, they do significantly contribute to pensioner households' income.

2.16 It should be noted that the SHS and SHCS, however, do include two 'catch-all' questions: one asking about receipt of income from "any other benefits" and one asking about receipt of income from "any other sources". It is possible that some income from the above benefits is captured in these questions.

2.17 In addition to the difference in coverage of sources of income, it is also worth noting that the FRS asks more detailed questions on particular types of income. For example, it routinely probes people for any income from investments such as interest payments on current accounts. Although these sources of income are widespread they normally comprise a small proportion of household income<sup>9</sup>. Although consideration of the effects of such differences are beyond the scope of this study, it is worth noting that the FRS may find higher levels of receipt for some sources of income than the SHS and SHCS because of such probes.

---

<sup>9</sup> Raab, G., MacDonald, C, and Macintyre, C. *Comparison of Income Data between Surveys of Scottish Households: Research commissioned by Communities Scotland*, 2004.

- 2.18 A further difference between the FRS and the SHS/SHCS that may influence estimates of household income is in the way they ask about benefit income. The SHS/SHCS asks, "Which of these are you (or your partner) receiving?" In contrast, for some benefits, the FRS makes explicit that benefits received on behalf of someone else should be included, asking, "...are you at present receiving any of the state benefits shown on this card - either in your own right or on behalf of someone else in your household?" This is particularly an issue in relation to Disability Living Allowance.
- 2.19 To compare the two surveys, we have to make an assumption of how SHS/SHCS respondents interpret this question. If they interpret these questions as asking about all benefit income they receive, irrespective of whom it is for, then the effect will be less than if they mention only the income they receive for themselves (or their partner). In the FRS, income from benefits at the adult level is assigned according to entitlement rather than receipt.
- 2.20 The FRS also asks a series of questions on receipt of various welfare entitlements such as free school meals, and free milk, and various benefits in kind from jobs. In the calculation of total household income, when a household member receives any of these entitlements, a monetary valued is added to their total household income<sup>10</sup>.

### Questionnaire coverage of other factors

- 2.21 It is as important to examine the questionnaire coverage relating to non-income questions as it is to examine the coverage of income sources. These questions shape the possible ways in which a wider definition of household income could be imputed in the SHS/SHCS. For example, while the SHS/SHCS ask about all household members' main activity, they don't specifically ask if adults, other than the HRP and their spouse, are in any form of employment. Therefore, it is impossible to distinguish between, for example, HE students who are working and those who are not working if they are not the HRP or spouse. So, while examining the difference in levels of income between students who have a job and those that don't in the FRS may be of interest, we would not be able to impute income separately for working and non-working students in the SHS/SHCS.

---

<sup>10</sup> This included payments relating to welfare milk, free school milk, free school meals, concessionary TV licence, company car, medical or dental insurance, workplace nursery, and childcare vouchers.

- 2.22 Only explanatory variables that are common to both the FRS and the SHS/SHCS can be used in any imputation strategy. The FRS provides data on the Standard Occupational Classification (SOC) and the Standard Industrial Classification (SIC) of all adults in the household. The SHS/SHCS only collect this for the HRP and spouse. They do not collect this information for other adults in the household, the group where income information is missing. So, while it is possible to use the FRS to model the difference in incomes level between, say “Managers and Senior Officials” and those in “Elementary Occupations,” this information is of little potential use to the imputation as we can’t distinguish which other adults in the SHS/SHCS are in which categories.
- 2.23 The main non-income variables that are common to both the FRS and the SHS/SHCS, and that are correlated to income levels are:
- Economic status/main activity of each person
  - Age of person
  - Sex of person
  - Relationship to Household Reference Person
  - Household composition
  - Tenure
  - Access to a car/van
  - Council Tax Banding (only SHCS)
- 2.24 There are a number of other household variables that are common across the different surveys but do not appear to prove to be useful in modelling income over and above the variables listed<sup>11</sup>.
- 2.25 While almost all of the variables are directly comparable between the surveys, economic status, a key indicator of income, is not. The FRS uses the ILO definition of employment status, derived from a number of different variables. In the SHS and the SHCS, the indicator of economic status for other adults comes from the ‘household grid’ series of questions, with the question simply asking, “And which of the items on this card would you say best describes [name]’s current status”.
- 2.26 Table 2-2 shows the distribution of response to those who are not in the first benefit unit – the “other adults” - in the SHS and the FRS. The main difference is in the classification of students. In the ILO definition of economic status, adults who are both working, and in Higher Education or Further Education are

---

<sup>11</sup> For example, number of bedrooms, and characteristics of the HRP. These are discussed more fully in the section on modelling full-time employees’ income.

classified as working. This helps to explain why in the SHS, 24% of other adults are in Further or Higher Education, while only 9% are classified as students in the FRS.

2.27 The ILO definition of employment status cannot be calculated in the SHS/SHCS. However, the FRS does record whether adults are studying at a university of college or training for a qualification in nursing or similar. By using this information, it is possible to recreate a recoded version of economic status in the FRS – using the assumption that all adults who are studying at a university or college, or training for a qualification in nursing or similar would consider themselves to be in further or higher education<sup>12</sup>. This helps to bring the distribution of economic status between the SHS and FRS into line.

**Table 2-2: Economic status of “other adults” FRS 2005-2006, SHS 2005 and 2006**

| Economic Status              | SHS  | FRS  | FRS Recoded |
|------------------------------|------|------|-------------|
| Full-time employees          | 44%  | 46%  | 42%         |
| Part-time employees          | 7%   | 14%  | 7%          |
| Self-employed                | 2%   | 4%   | 4%          |
| Looking after home or family | 2%   | 1%   | 1%          |
| Retired                      | 6%   | 7%   | 7%          |
| Unemployed                   | 8%   | 8%   | 8%          |
| In FE/HE                     | 24%  | 9%   | 20%         |
| Govt work/training           | 1%   | 0%   | 0%          |
| Sick or disabled             | 5%   | 6%   | 6%          |
| Short-term sick              | 1%   | 1%   | 1%          |
| Other/Other inactive         | 1%   | 4%   | 3%          |
|                              | 100% | 100% | 100%        |

2.28 Without this recode, the FRS definition of students would not include students who are working, and therefore, any imputation to the SHS/SHCS is likely to underestimate any income. This does raise a concern though. **As the measure for economic status is not directly comparable between the SHS/SHCS and the FRS, there is a possibility that even with this recode there are a small proportion of respondents who would be characterised differently between the two measures.** This needs to be borne in mind when making final decisions on the potential uses of imputed data.

<sup>12</sup> The variable edtyp, “Education Type” was used to recode those studying full-time or part-time at a college or university, or training for a qualification in nursing or similar as students in the FRS data.

## Definitions of Household Income<sup>13</sup>

2.29 In the SHS and the SHCS, total household income is currently defined as the income received from earnings, benefits and other miscellaneous sources, for the Highest Income Householder and their spouse (if applicable). The FRS, in contrast, collects income for all adults in the household. Therefore, for one person households, and two person household where the household members are married (or living as married), the definition of Household Income is consistent between the SHS/SHCS and the FRS with regard to coverage of adult household members. For all households with three or more adults, and for two person households where the household members are not married (or living as married) the definitions are not consistent.

2.30 The SHS and SHCS definitions of household income – income of the HRP (and their spouse if applicable) – can be reframed as the adult income of all adults in the first Benefit Unit (BU). In other words, with regard to adult income, coverage in the SHS/SHCS differs from the FRS only with regard to adults in a different BU to the HRP.

2.31 Unlike the SHS and SHCS, the FRS definition of household income also includes children's<sup>14</sup> income. The FRS collects information on a number of different components of children's incomes, principally from:

- children's earnings,
- Educational Maintenance Allowances,
- trust funds,
- and education grants.

2.32 Although the SHS and SHCS does not collect any information on children's income, it is possible that income from an educational grant for a child is captured in the 'catch-all' question about "other sources of income" (assuming that the income is received by the parent on behalf of the child).

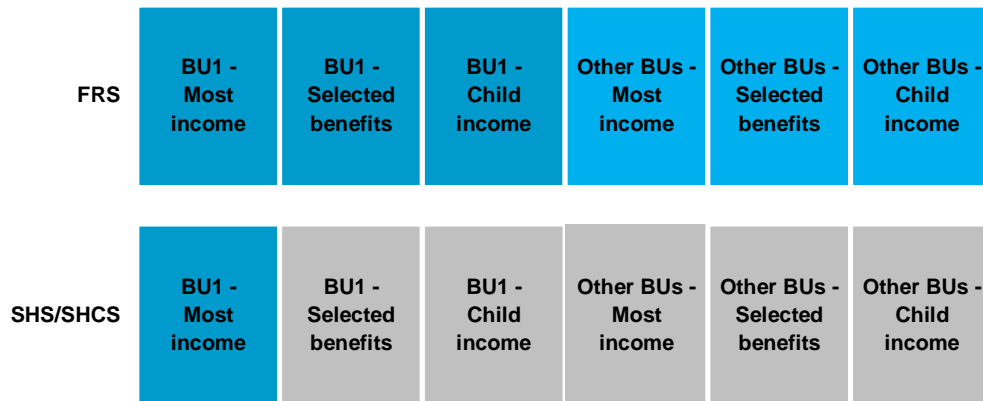
---

<sup>13</sup> The main analysis variable in the FRS is "Gross Household Income". In the SHS and SHCS, all amounts are reported for net household income. We have converted the gross Household Income data in the FRS to net amounts for ease of comparison.

<sup>14</sup> The definition of an adult and a child differs between the SHS and the FRS. In the SHS, all those who are aged 16 and over are defined as an adult. In the FRS, unmarried 16-18 year olds in full-time non-advanced education are not considered as adults.



**Figure 2.1: Comparison of the household measures of income between FRS and SHS, SHCS**



2.33 Overall, there are three main ways in which the FRS definition of household income is 'wider' than that of the SHS/SHCS, illustrated in Figure 2.1<sup>15</sup>:

- it covers adults in additional benefit units, i.e. adults who are not the HRP or their spouse;
- it covers children's incomes;
- it covers a number of additional sources of income not asked about in the SHS.

<sup>15</sup> Note that the diagram does not represent the relative size of the elements of income. These are detailed in Table 3-1 in the next chapter of the report.

Estimating the effect of  
definitional differences  
on measures  
of total household income

### 3. Estimating the effect of definitional differences on measures of total household income

- 3.1 This chapter examines the relative effects of the three differences between the SHS<sup>16</sup> and the FRS definitions of household income on the measures: children income, benefits not asked about in the SHS, and inclusion of income from adults who are not the HRP or spouse. It also examines how this differs across household types, and where imputation (or changes in definition) is likely to impact on the estimates of household income.
- 3.2 Table 3-1 shows a breakdown of total household income in the FRS by the components detailed in Figure 2.1. Across the UK as a whole, around 91% of the FRS measure of household income is made up from the income of the HRP and their spouse from sources asked about in the SHS. In other words, the SHS definition of household income would give, on average, an estimate of average household income that is 9% less than the FRS definition across the UK. Restricting the analysis to Scottish cases in the FRS only, the difference is around 8%.

**Table 3-1: Breakdown of Household income in the FRS across all households. UK and Scotland. FRS 2005-06<sup>17</sup>**

|   | <b>UK</b>     | <b>Scotland</b> |
|---|---------------|-----------------|
| <b><i>HRP and Spouse – SHS sources (SHS definition)</i></b> | <b>91.0%</b>  | <b>92.0%</b>    |
| HRP and Spouse – FRS only benefits                          | 0.5%          | 0.6%            |
| HRP and Spouse's children's income                          | 0.2%          | 0.2%            |
| <b><i>All income – HRP and Spouse</i></b>                   | <b>91.8%</b>  | <b>92.8%</b>    |
| Other adults – Income from SHS sources                      | 8.2%          | 7.2%            |
| Other adults – Income from benefits only asked in FRS       | 0.0%          | 0.0%            |
| Other adults' children's income                             | 0.0%          | 0.0%            |
| <b><i>All income – Other adults</i></b>                     | <b>8.2%</b>   | <b>7.2%</b>     |
| <b>Total household income</b>                               | <b>100.0%</b> | <b>100.0%</b>   |
| n   | 28,029        | 4,438           |

- 3.3 The effect of the inclusion of children's incomes is minimal. Across the UK, children's income account for around 0.2% of total household income in the FRS.

<sup>16</sup> To aid clarity, we have not included the SHCS in this analysis, but have indicated where the results are likely to be different from the SHS.

<sup>17</sup> All figures in this report have been rounded. Where percentages do not sum up exactly to 100%, this is due to rounding.

- 3.4 Benefits asked about in the FRS but not the SHS account for around half of one percent of total household income (0.5% in UK, 0.6% in Scotland)<sup>18</sup>. Given that the SHCS covers more benefits than the SHS including winter fuel payments – in other words that it is closer to the FRS in its coverage than the SHS - a corresponding analysis would show that less than half of one percentage is missed by the SHCS in this way.
- 3.5 Most of the difference between the FRS and the SHS estimates is accounted for by the income of other adult household members. Income from the other adults – adults who are not the HRP or their spouse – make up around 8% of household income across the UK overall. In Scotland, the corresponding figure is slightly lower at around 7% of household income. In other words, imputing the income of other adults in the household in the SHS/SHCS should increase the estimate of total household income by around 8-9% overall.
- 3.6 Clearly, however, the difference between the SHS and the FRS estimates of household income will differ between different types of household. Table 3-2 shows the breakdown of household income by household type across the UK overall<sup>19</sup>.
- 3.7 With regard to benefits asked about in the FRS but not the SHS, the largest effect is in single pensioner households where 3.3% of household income comes from these benefits. This is primarily due to the effect of the non-inclusion of winter fuel allowance in the SHS. In two pensioner households, 1.8% of household income is accounted for by these missing benefits.
- 3.8 With regard to children's income the largest effect on household income is seen in single parents households where between 1.2% and 1.3% of the total household income comes from this source.
- 3.9 It is again clear that the effects of benefits asked about in the FRS but not in the SHS and of children's income are small compared to the effect of income of other adults. This is particularly the case in households with three or more adults where between 21% (three or more adults and three or more children households) and 37% (three adults and no children households) of household income comes from other adults.

---

<sup>18</sup> These do not include income imputed from welfare benefits such as school meals, free prescription charges etc.

<sup>19</sup> Figures for Scotland are not shown due to the small sample size.

**Table 3-2: Breakdown of Household income in the FRS by household composition, FRS 2005-06 (Row percentages)**

| Household composition                           | HRP/ Spouse – SHS def. | HRP/ Spouse – FRS only bens | HRP/ Spouse – Child inc | Total – HRP/ Spouse | Others – SHS def. | Others – FRS only bens | Others – Child income | Total – Others |
|---|------------------------|-----------------------------|-------------------------|---------------------|-------------------|------------------------|-----------------------|----------------|
| One adult, no children over pension age         | 96.7%                  | 3.3%                        | -                       | <b>100.0%</b>       | -                 | -                      | -                     | -              |
| One adult, no children, under pension age       | 99.9%                  | 0.1%                        | -                       | <b>100.0%</b>       | -                 | -                      | -                     | -              |
| Two adults, no children, both over pension age  | 97.0%                  | 1.8%                        | -                       | <b>98.7%</b>        | 1.2%              | 0.0%                   | -                     | <b>1.3%</b>    |
| Two adults, no children, one over pension age   | 87.3%                  | 1.0%                        | -                       | <b>88.3%</b>        | 11.5%             | 0.1%                   | -                     | <b>11.7%</b>   |
| Two adults, no children, both under pension age | 95.9%                  | 0.1%                        | -                       | <b>95.9%</b>        | 4.1%              | 0.0%                   | -                     | <b>4.1%</b>    |
| Three or more adults, no children               | 62.7%                  | 0.2%                        | -                       | <b>62.8%</b>        | 37.1%             | 0.1%                   | -                     | <b>37.2%</b>   |
| One adult, one child                            | 98.1%                  | 0.7%                        | 1.2%                    | <b>100.0%</b>       | -                 | -                      | -                     | -              |
| One adult, two children                         | 98.1%                  | 0.7%                        | 1.2%                    | <b>100.0%</b>       | -                 | -                      | -                     | -              |
| One adult, three or more children               | 96.9%                  | 1.8%                        | 1.3%                    | <b>100.0%</b>       | -                 | -                      | -                     | -              |
| Two adults, one child                           | 97.5%                  | 0.1%                        | 0.4%                    | <b>98.0%</b>        | 2.0%              | 0.0%                   | 0.0%                  | <b>2.0%</b>    |
| Two adults, two children                        | 98.9%                  | 0.1%                        | 0.4%                    | <b>99.5%</b>        | 0.5%              | 0.0%                   | 0.0%                  | <b>0.5%</b>    |
| Two adults, three or more children              | 98.1%                  | 0.3%                        | 1.1%                    | <b>99.5%</b>        | 0.5%              | 0.0%                   | 0.0%                  | <b>0.5%</b>    |
| Three or more adults, one child                 | 73.0%                  | 0.1%                        | 0.9%                    | <b>74.1%</b>        | 25.9%             | 0.1%                   | 0.0%                  | <b>25.9%</b>   |
| Three or more adults, two children              | 77.2%                  | 0.2%                        | 1.2%                    | <b>78.5%</b>        | 21.4%             | 0.1%                   | 0.0%                  | <b>21.5%</b>   |
| Three or more adults, three or more children    | 70.5%                  | 0.6%                        | 1.1%                    | <b>72.2%</b>        | 27.4%             | 0.3%                   | 0.0%                  | <b>27.8%</b>   |
| Total   | 91.0%                  | 0.5%                        | 0.2%                    | <b>91.8%</b>        | 8.2%              | 0.0%                   | 0.0%                  | <b>8.2%</b>    |

3.10 In comparison, in all types of two adult households, less than 5% of total household income comes from other adults. This is due not to the level of income of other adults, but the likelihood that these types of household will contain an 'other adult'. Two adult types tend to be composed of an HRP and a spouse and no other adults.

3.11 When the analysis is limited to households with other adults – households where the SHS and SHCS will have missing income data - it is clear that there is likely to be a sizeable effect on the estimate of household income (see Table 3-3). For example, in households with two adults both under pension age and with no children and where they are not in the same benefit unit, over a third (37.2%) of income comes from the other adult. These types of household account for 2.6% of all households and 16% of households where there are other adults. In households with two adults both over pension age and not in the same benefit

unit, over 40% of income comes from other adults. This type of household only accounts for 0.2% of all households.

**Table 3-3: Households with more than one Benefit Unit - breakdown of income, distribution of household types, and sample sizes, FRS 2005-06**

| Household composition                           | Breakdown of household income |                         | Distribution of household types |  | Sample Sizes                  |                                       |
|---|-------------------------------|-------------------------|---------------------------------|--|-------------------------------|---------------------------------------|
|   | HRP and Spouse - SHS def.     | Other adults - SHS def. | Percentage of all households    | Percentage of all households with 2+ BUs | Sample size in FRS 05-06 - UK | Expected number of cases in SHS 05-06 |
| Two adults, no children, both over pension age  | <b>54.2%</b>                  | 43.3%                   | <b>0.2%</b>                     | 1%                                       | 74                            | 40                                    |
| Two adults, no children, one over pension age   | <b>56.0%</b>                  | 42.3%                   | <b>1.0%</b>                     | 6%                                       | 357                           | 150                                   |
| Two adults, no children, both under pension age | <b>62.7%</b>                  | 37.2%                   | <b>2.6%</b>                     | 16%                                      | 605                           | 400                                   |
| Three or more adults, no children               | <b>62.7%</b>                  | 37.1%                   | <b>8.4%</b>                     | 51%                                      | 1855                          | 1300                                  |
| Two adults, one child                           | <b>69.3%</b>                  | 28.9%                   | <b>0.6%</b>                     | 3%                                       | 190                           | 90                                    |
| Two adults, two children                        | <b>73.9%</b>                  | 25.0%                   | <b>0.2%</b>                     | 1%                                       | 67                            | 30                                    |
| Two adults, three or more children              | <b>77.9%</b>                  | 18.6%                   | <b>0.1%</b>                     | 1%                                       | 35                            | 20                                    |
| Three or more adults, one child                 | <b>73%</b>                    | 25.9%                   | <b>2.2%</b>                     | 13%                                      | 569                           | 340                                   |
| Three or more adults, two children              | <b>77.2%</b>                  | 21.4%                   | <b>0.8%</b>                     | 5%                                       | 233                           | 130                                   |
| Three or more adults, three or more children    | <b>70.5%</b>                  | 27.4%                   | <b>0.3%</b>                     | 2%                                       | 100                           | 45                                    |
| Total   | <b>65.3%</b>                  | 34.1%                   | <b>16.3%</b>                    | 100%                                     | 4084                          | 2550                                  |

3.12 In total, around 16% of all households contain other adults, adults whose income is not recorded in the SHS or the SHCS. This equates to around 2,550 cases per year in the SHS and around 4,000 cases per year in the FRS across the UK as a whole. Over half of these are households with three or more adults and no children.

3.13 It is interesting to compare the SHS estimate of household income and the FRS estimate of household income, adjusted to account for the difference in definition. In the financial year 2005-06, the median net household income according to the FRS was £367 a week. Excluding income from “other” adults, children’s income, and benefits not asked about in the SHS from the FRS estimate, the median net household income drops £36 to £331 a week. This compares with an estimate of

£317 in the SHS (See Table 3-4). Given the differing length and the focus of the two surveys, the difference in estimates is reassuringly small<sup>20</sup>.

**Table 3-4: Average weekly household income April 05 – Mar 06 Scotland: Comparison of FRS and SHS.**

|   | Median | Mean | N      |
|---|--------|------|--------|
| SHS data – SHS definition of household income             | £317   | £391 | 15,428 |
| FRS data – Adjusted to SHS definition of household income | £331   | £421 | 4,438  |
| FRS data – FRS definition of household income             | £367   | £458 | 4,438  |

3.14 In summary, the effect of the non-inclusion of children's income in the SHS and SHCS is likely to be small, less than a third of one percent<sup>21</sup> overall. The effect of the non-inclusion of particular benefits in the SHS is also small, around half of one percent overall. **Given that winter fuel payments account for a sizeable proportion of these benefits, consideration should be given to including receipt of winter fuel payments in the SHS questionnaire.** Alternatively, as there is near 100% take-up of this benefit, it would be relatively simple to impute receipt in eligible households<sup>22</sup>. **Consideration should also be given to including the other benefits asked about in the FRS but not the SHS and SHCS in the respective questionnaires.**

3.15 **Clearly, the largest difference between the FRS and the SHS and SHCS estimates of income is from other adults. Overall, this accounts for around 7-8% of household income, and will vary widely between different household types. The rest of this paper discusses potential imputation strategies to address this.**

<sup>20</sup> And even smaller had benefits in kind also been excluded from the adjusted FRS figure.

<sup>21</sup> It is arguable whether children's income should be considered part of household income for most types of analysis but this debate is beyond the scope of this paper.

<sup>22</sup> It could be argued that this approach is more accurate. Over-estimation due to not accounting for non take-up is likely to have a smaller effect than under-estimation due to under-reporting by respondents of Winter Fuel payments. These are a seasonal benefit that is paid directly to bank accounts of people over pensionable age.

# Examining strategies to impute income for 'other adults'



## 4. Examining strategies to impute income for 'other adults'

4.1 Methods of imputation almost always rely on modelling existing data. The more variance that these models explain, the better the resulting imputation strategies will be. This chapter provides a brief introduction to imputation methods, before examining how much of the variance in income of other adults overall can be accounted for by variables that could be used to design an imputation strategy. It concludes by presenting our recommended approach for imputing income of other adults.

### A short introduction to imputation

4.2 Imputation is a process whereby data from existing cases is used to compute values for cases with missing information.

4.3 There are a number of standard methods used to impute data. While all methods of imputation aim to get as close to the “true” values of the missing variables as possible, there are two partially conflicting ideals when imputing data<sup>23</sup>, namely:

- Imputed data should be as close to the real (unknown) value as possible.
- Imputed data should preserve the shape of the distribution of the values as far as possible.

4.4 In other words, imputation techniques tend to either try to ensure that the distribution of the imputed data is maintained or minimise the mismatch between imputed and “true” data. Two simplistic methods of imputation show this contrast: imputing from a randomly selected case and imputing the overall mean. Imputing from a randomly selected case maintains the distribution of the data, but may lead to a sizeable mismatch between the imputed and “true” value. Imputing the overall mean compresses the distribution of the data but minimises the possible mismatch between the imputed and the “true” value. This comparison of simple

---

<sup>23</sup> Lessler, J.T. & Kalsbeek, W.D. (1992) *Nonsampling Errors in Surveys*. New York: Wiley

methods highlights the fundamental tension between two different classes of imputation method<sup>24</sup> that are of potential use:

- Methods where imputation for a case will always produce the same answer. These methods are normally called deterministic methods.
- Random methods, where repeating the imputation for a case could produce different results. These methods are sometimes termed stochastic methods.

- 4.5 Deterministic methods always produce the same value for an imputation. These methods range from very simple techniques such as imputing the overall mean or imputing a class mean to Predictive Mean Matching. In Predictive Mean Matching, a regression model is calculated using existing data and imputed values are obtained from the regression equation. These are compared with values for complete cases and the case with the closest match provides the final imputed value.
- 4.6 A major weakness of deterministic methods of imputation is that they reduce the variance within the final distribution and result in fewer cases with values in the tails. In other words, the distribution is squeezed towards the mean. This can lead to problems in the data analysis. It makes estimating sampling errors for survey estimates more problematic with the possibility that tests of significance will be affected. It can also distort relationships between survey variables not used in any of the modelling, although the better the model of complete information, the smaller the disruption to the distribution.
- 4.7 With regards to imputation of income, there is an additional difficulty of using such methods. Distributions of income tend to be positively skewed, with the mean value tending to be higher than the median value. While regression can be used to examine patterns of income by examining the square root of income – this corrects for the skew to a large extent - the use of the parameter scores from the regression model to compute values for income remains problematic.
- 4.8 In comparison, stochastic methods are generally more suitable to imputing income. These methods build in an element of chance to the imputation process. As such, they maintain the distribution of imputed variables far better than deterministic methods, although there is a higher chance that in any individual

---

<sup>24</sup> GSS Task Force on Imputation (1996) *Report of the Task Force on Imputation*. GSS Methodology Series No. 3: HMSO. A third class of methods covered there, are where values for variables are calculated from other information supplied by the same cases, such as calculating amount of child benefit received from number of children and whether a lone parent. Given the limited amount of data we have on other adults common across the surveys, these methods are not practical.

case, the imputed value will be significantly different from the actual (and unknown) value. Therefore, we would recommend using stochastic methods of imputation where possible.

- 4.9 There is a wide range of stochastic methods ranging in complexity from simply imputing a randomly selected case to Hierarchical Hot Deck imputation. In Hot Deck imputation, respondents are sorted into imputation groups according to likely determinants. A simple example would be imputing earnings using imputation groups of self-employed men, self-employed women, male employees and female employees. Cases with missing data are donated values from cases with data which match on these characteristics.
- 4.10 Where several factors are correlated to the variable of interest, creating donor groups using all the factors is likely to result in some cells with no donor cases or very few donor cases. Hierarchical Hot Deck imputation can be used in these instances. This is an iterative technique. Imputation is first attempted using donor groups based on all significant factors. Where there are empty cells and the imputation fails, the factor with the least explanatory power is collapsed or excluded to create a model with fewer donor groups. This process is repeated until a match is found.

### Three potential strategies for modelling adult income data

- 4.11 **There are three broad strategies that could be used to model adult income data, and therefore provide the basis for imputing income of other adults in the SHS and the SHCS:**
- Model the FRS data to examine the differences between income of HRP's and other adults to inform an imputation strategy that uses existing SHS data of HRP's and spouses to impute income for other adults.
  - Use the income collected in the Random Adult section of the SHS to create an imputation strategy that uses existing SHS data of other adults to impute income for other adults in the SHS and SHCS.
  - Model FRS data of other adults to design an imputation strategy for the SHS and SHCS that uses only FRS data.
- 4.12 The first strategy, using the existing SHS data for householders to impute income for other adults is very appealing. It would preclude the need to use case-level FRS data in the imputation process and reduce the timescales for the production of the amended household income measures. This strategy would involve examining the difference between income levels of HRP's and the income levels

of other adults. Householders, obviously, are likely to have much higher levels of income than non-householders. However, this strategy has serious limitations, and is not recommended. These problems can be illustrated by examining pensioner incomes.

**Table 4-1: Mean weekly adult income of pensioners by age and whether HRP or other adults<sup>25</sup>. FRS 2005- 06 UK**

|                 | Other adults |            | HRP  |            | Ratio of Means |
|-----------------|--------------|------------|------|------------|----------------|
|                 | Mean         | SE of Mean | Mean | SE of Mean |                |
| <b>Up to 59</b> | £77          | £17        | £296 | £23        | 26%            |
| <b>60-64</b>    | £108         | £18        | £237 | £7         | 45%            |
| <b>65-69</b>    | £132         | £12        | £241 | £9         | 55%            |
| <b>70-74</b>    | £144         | £12        | £233 | £5         | 62%            |
| <b>75-79</b>    | £146         | £9         | £223 | £3         | 66%            |
| <b>80-84</b>    | £147         | £8         | £206 | £3         | 71%            |
| <b>85up</b>     | £151         | £11        | £202 | £3         | 75%            |
| <b>Total</b>    | £141         | £4         | £227 | £2         | 62%            |

4.13 Table 4-1 shows the mean weekly adult income of pensioners by age and whether they are the HRP or not. Overall, the income of pensioners who are other adults is around 62% of pensioners who are the HRP (£141 compared to £227). This could be used to create a simple hot deck imputation strategy: extract all values of pensioner HRP income in the SHS, multiply them by a weighting factor of 0.62, and use these values to donate to non-householder pensioners. However, it is clear from Table 4-1 that pensioner incomes tends to decrease with age among the HRP group and increase with age among the other adults group. By not taking into account age, the imputed levels of income for younger pensioners would be overstated and the imputed income for older pensioners would be understated.

4.14 This could be overcome by imputing income separately for different age groups of pensioners, using the different weights – based on the ratio of the means - to adjust the data. However, as the cells sizes become smaller and the standard errors of the means used to create the ratios become larger, the weights would become less robust. Introducing a further explanatory variables - for example, distinguishing between other adults who are related to the HRP and those not related to the HRP - would further decrease the robustness of these weights. There is a further danger because this strategy would build-in systematic error if

<sup>25</sup> Income of pensioner spouses is not shown.

any factors influencing the difference between the HRP and other adults where missed.

- 4.15 The second strategy would be to use the income collected in the Random Adult section of the SHS. Between 1999 and 2004, the Random Adult section of the questionnaire included a set of questions on income that corresponded to those asked of the HRP and their spouse. Using this data for imputation is also very appealing in that it would again preclude the need to use FRS data.
- 4.16 However, this strategy is not recommended due to the limited number of cases with full data. In imputation, the size of the donor group compared to the size of the group where imputation is required is critical. It is generally accepted that the ratio of donors to missing cases should be maximised, and that where possible, there should be more donor cases than missing cases. Otherwise, there is a danger that the random variations in income levels are likely to be magnified<sup>26</sup>.
- 4.17 Overall, the random adult income section will be asked of around a third of “other adults”. For example, in three-person households, where there is one HRP, a spouse of the HRP and another adult, the random adult will be the ‘other’ adult in a third of the cases. In four person households, with two other adults, the random adult will be an “other adult” in half of the cases. In such households, there will be either 1 or 2 other adults whose income needs to be imputed. Therefore, this strategy would involve using, on average, one case of full data to impute two missing cases<sup>27</sup>, and given natural variations across groups, a much poorer ratio for some groupings.
- 4.18 We would, however, suggest that where this data does exist that it should be used as an alternative to imputation. Should it be decided to retrospectively create a variable indicating household income including all adults in the SHS, we would suggest using this information to reduce the number of cases requiring imputation between 1999 and 2004<sup>28</sup>.
- 4.19 An alternative, preferable strategy would be to use the FRS to model the income of other adults in Scotland without reference to income from HRP or using the Random Adult data in the imputation strategy. This has the advantage over the

---

<sup>26</sup> But potentially, indicators of variance, such as Standard Deviation (SD) calculations, will be deflated. Consider, for example, if six cases are imputed from 2 donor cases. There will be two peaks in the data, potentially near the population mean/median, but potentially also close to the tails. But more importantly, the difference between the values of the imputed data will be artificially small.

<sup>27</sup> Indeed, there will be a slightly poorer ratio of donor to missing cases as 1 in every 10 households is missing a random adult interview.

first strategy in that it would not be affected by the variation in levels of income of HRP, and would only have to account for differences among other adults.

4.20 As noted in Chapter 2, the sample size of the FRS in Scotland is considerably smaller than that of the SHS. This influences the choice of whether to use the FRS data for Scotland only or to use data from across the UK. Table 4-2 shows the sample sizes in the FRS when analysed by relationship to HRP, economic status and whether in Scotland. In the 2005-06 FRS data, there was a sample size of 5,410 other adults (adults who are not the HRP or their spouse) across the UK. In Scotland, the total sample size of other adults was 803 including only 21 who are self-employed and 204 who are students. Given that the annual sample size of the FRS in Scotland is around a quarter of that of the SHS, only using Scottish data would mean using, on average, one case from the FRS to impute four cases in the SHS. For example, using the 21 cases where the other adult is self-employed to impute around 85 corresponding cases in the SHS.

**Table 4-2: Unweighted sample size in FRS 2005-06 by relation to HRP, Economic Status and area. Scotland versus rest of UK.**

|            |                   | Full-time work | Part-time work | Self-employed | Unemp, carer, etc. | Retired       | Student      | Total         |
|------------|-------------------|----------------|----------------|---------------|--------------------|---------------|--------------|---------------|
| Rest of UK | Is HRP            | 9,625          | 1,514          | 1,995         | 3265               | 6,489         | 703          | 23,591        |
|            | Spouse, cohabitee | 3,683          | 2,656          | 904           | 2786               | 3,022         | 436          | 13,487        |
|            | Child, step etc.  | 1,607          | 298            | 125           | 738                | 16            | 648          | 3,432         |
|            | Parent, step etc. | 14             | 7              | 5             | 29                 | 229           | 2            | 286           |
|            | Other relative    | 96             | 12             | 9             | 77                 | 41            | 39           | 274           |
|            | Not related       | 236            | 28             | 35            | 73                 | 39            | 204          | 615           |
|            | <b>Total</b>      | <b>15,261</b>  | <b>4,515</b>   | <b>3,073</b>  | <b>6968</b>        | <b>9,836</b>  | <b>2,032</b> | <b>41,685</b> |
| Scotland   | Is HRP            | 1,869          | 271            | 280           | 581                | 1,301         | 136          | 4,438         |
|            | Spouse, cohabitee | 723            | 467            | 141           | 437                | 551           | 75           | 2,394         |
|            | Child, step etc.  | 269            | 51             | 16            | 122                | 2             | 131          | 591           |
|            | Parent, step etc. | 5              | 2              | 1             | 3                  | 39            | 0            | 50            |
|            | Other relative    | 18             | 2              | 1             | 10                 | 17            | 10           | 58            |
|            | Not related       | 20             | 3              | 3             | 13                 | 2             | 63           | 104           |
|            | <b>Total</b>      | <b>2,904</b>   | <b>796</b>     | <b>442</b>    | <b>1166</b>        | <b>1,912</b>  | <b>415</b>   | <b>7,635</b>  |
| Total      | Is HRP            | 11,494         | 1,785          | 2,275         | 3846               | 7,790         | 839          | 28,029        |
|            | Spouse, cohabitee | 4,406          | 3,123          | 1,045         | 3223               | 3,573         | 511          | 15,881        |
|            | Child, step etc.  | 1,876          | 349            | 141           | 860                | 18            | 779          | 4,023         |
|            | Parent, step etc. | 19             | 9              | 6             | 32                 | 268           | 2            | 336           |
|            | Other relative    | 114            | 14             | 10            | 87                 | 58            | 49           | 332           |
|            | Not related       | 256            | 31             | 38            | 86                 | 41            | 267          | 719           |
|            | <b>Total</b>      | <b>18,165</b>  | <b>5,311</b>   | <b>3,515</b>  | <b>8134</b>        | <b>11,748</b> | <b>2,447</b> | <b>49,320</b> |

<sup>28</sup> The random adult income section was simplified to a single question from 2005 onwards.

4.21 A modification to this strategy would be to use data for a number of years from the FRS. Using Scottish data from the FRS from four different years would give a rough ratio of 1 donor case for every 1 case to impute. This is not recommended for three main reasons. First, it would be difficult to correct for any trends in income levels over time. Second, it would still provide a significantly smaller donor pool than using a single year of FRS data for the UK as a whole. Third, if this strategy was employed, it would still not be possible to design an imputation method that took into account regional variations within Scotland in income levels in any useful way<sup>29</sup>. Additionally, this method would also involve linking donor data from four different datasets, adding to the practical issues in carrying out the imputation.

4.22 Limiting the analysis to FRS cases in Scotland will therefore significantly constrain the ability to model the effect of various factors due to small sample sizes, and would risk magnifying random variability. However, using data from the FRS across the UK as a whole risks biasing the distribution of income in Scotland if there are significant differences between Scotland and the rest of the UK.

**Table 4-3: Mean weekly adult income by region and whether HRP/Spouse or other adults. FRS 2005- 06 UK**

| FRS Region                   | HRP/Spouses | Other adults | Total       |
|------------------------------|-------------|--------------|-------------|
| North East                   | £246        | £143         | £231        |
| North West and Merseyside    | £266        | £164         | £253        |
| Yorks and Humberside         | £266        | £145         | £250        |
| East Midlands                | £272        | £149         | £257        |
| West Midlands                | £259        | £147         | £242        |
| Eastern                      | £310        | £167         | £293        |
| London                       | £362        | £201         | £334        |
| South East                   | £335        | £173         | £313        |
| South West                   | £277        | £166         | £265        |
| Wales                        | £251        | £146         | £237        |
| Northern Ireland             | £250        | £158         | £234        |
| <b>UK Excluding Scotland</b> | <b>£292</b> | <b>£165</b>  | <b>£275</b> |
| <b>Scotland</b>              | <b>£274</b> | <b>£149</b>  | <b>£258</b> |
| Total                        | £290        | £164         | £273        |

<sup>29</sup> The FRS does not as standard provide details of the location of interviews in Scotland down to LA level. Even if it did, given that income levels are more correlated to other variables in the modelling than region, by only using cases in one LA to impute to a missing cases in another LA is likely to lead to less accurate rather than better estimates. Consider the following example. We want to impute a full-time worker, living in Glasgow, who is living with his parents, in a household with a car. We have a two potential donor cases – one that is exactly the same except for the location, and one that is exactly the same, except that it is a parent who is living with grown up children. As income is more closely correlated to the relationship to the HRP than it is to geography, we would use the first potential donor case.

- 4.23 Table 4-3 shows the mean weekly adult income by region. Overall, the mean income in Scotland among HRPs and spouses is £274, £18 less than the average across the rest of the UK. Among other adults, the average weekly income in Scotland is £149, £16 less than the rest of the UK.
- 4.24 It is likely that differences by region in income levels are related to other factors and in particular, to whether income primarily comes from benefits or from earnings. While earnings levels are likely to vary across the UK, benefit levels are likely to vary less, if at all<sup>30</sup>. Therefore, any simple weighting strategy that doesn't account for such differences may have an adverse (and hidden) effect on imputed values<sup>31</sup>.
- 4.25 In order to explore as fully as possible the determinants of adult income, we have used data from the whole of the UK. However, in order to guard against biasing the imputed data because of the geographical differences in levels of income, we have also examined whether there are significant differences by region.

### **Modelling the determinants of income using variables common to the FRS, SHS and SHCS**

- 4.26 Table 4-4 shows average income levels by selected factors collected in both the FRS and the SHS/SHCS. As discussed previously, only factors that are common between surveys can be used in the imputation processes.
- 4.27 Overall, income levels differ markedly by relationship to the HRP and economic status, and to a less extent, by tenure and council tax banding. Among other adults, those who are not related to the HRP have the highest income levels on average (mean of £191). In contrast, parents of the HRP have the lowest average income (mean of £140).
- 4.28 Economic status is closely related to income among other adults in the household. Other adults who are full-time employees have a mean net income of £239, those in part-time employment have a mean income of £129, while students have a mean income of £123 and those who are unemployed, looking after home or family, or sick and disabled have a mean income of £52. Other adults who are retired have an average weekly net income of £141.

<sup>30</sup> This is explored further in the detailed imputation strategy outlined in Appendix 1.

<sup>31</sup> For example, weighting values imputed from non-Scottish data by 0.9 (£149/£165) is likely to artificially depress pensioner incomes.



**Table 4-4: Net adult household income by selected factors collected in both the FRS and SHS. FRS 2005-06. UK**

| <b>Relationship to HRP</b>             | <b>Mean</b> | <b>Median</b> | <b>SD</b> | <b>N</b> |
|--|-------------|---------------|-----------|----------|
| Is HRP                                 | £346        | £276          | £360      | 28,015   |
| Spouse, cohabite                       | £193        | £154          | £230      | 15,864   |
| Child, step etc.                       | £161        | £155          | £128      | 4,018    |
| Parent, step etc.                      | £140        | £127          | £99       | 336      |
| Other relative                         | £162        | £140          | £135      | 332      |
| Not related                            | £191        | £165          | £134      | 719      |
| Total                                  | £273        | £219          | £311      | 49,284   |
| <b>Economic status (Other adults)</b>  |             |               |           |          |
| <b>Economic status (Other adults)</b>  | <b>Mean</b> | <b>Median</b> | <b>SD</b> | <b>N</b> |
| Full-time work                         | £239        | £220          | £118      | 2,265    |
| Part-time work                         | £129        | £118          | £72       | 403      |
| Self-employed                          | £219        | £189          | £188      | 190      |
| Unemployed, looking after family etc.  | £52         | £44           | £60       | 1,065    |
| Retired                                | £141        | £134          | £88       | 385      |
| Student                                | £123        | £106          | £102      | 1097     |
| Total                                  | £164        | £152          | £128      | 5,405    |
| <b>Tenure (Other adults)</b>           |             |               |           |          |
| <b>Tenure (Other adults)</b>           | <b>Mean</b> | <b>Median</b> | <b>SD</b> | <b>N</b> |
| Owns outright                          | £186        | £169          | £151      | 1,521    |
| Buying with Mortgage                   | £162        | £159          | £114      | 2,133    |
| Shared ownership                       | £144        | £103          | £122      | 11       |
| Rents                                  | £147        | £132          | £118      | 1,686    |
| Rent-free                              | £191        | £182          | £158      | 54       |
| Total                                  | £164        | £152          | £128      | 5,405    |
| <b>Council Tax band (Other adults)</b> |             |               |           |          |
| <b>Council Tax band (Other adults)</b> | <b>Mean</b> | <b>Median</b> | <b>SD</b> | <b>N</b> |
| A                                      | £133        | £122          | £101      | 938      |
| B                                      | £147        | £137          | £111      | 989      |
| C                                      | £164        | £154          | £131      | 1,043    |
| D                                      | £185        | £180          | £131      | 916      |
| E                                      | £181        | £174          | £124      | 566      |
| F                                      | £193        | £175          | £152      | 275      |
| G                                      | £199        | £164          | £193      | 172      |
| H                                      | £285        | £206          | £269      | 23       |
| Household not valued separately        | £191        | £173          | £134      | 63       |
| Total                                  | £164        | £153          | £128      | 4,985    |

4.29 Relationship to the HRP and economic status are clearly two key variables in modelling adult income. In order to examine how far we can model income levels from the variables that are collected in the FRS, the SHS and the SHCS, and what factors best account for differences in income, a regression model of income was computed.

4.30 As mentioned previously, income distributions are almost always positively skewed, with a long right tail. In other words, there is a much wider range of

values above the median than below it. As regression models assume that the distribution of the dependent variable is normal, the model was run on the square root of the income values. This significantly reduces the skew. Most of the variables that are common to all three surveys were included as dependent factors<sup>32</sup>, namely:

- Relationship to the HRP
- Economic Status
- Sex
- Age (entered as set of dummy variables as the relationship between age and income is likely to be non-linear)
- Tenure
- Whether the household has a car or a van
- Whether in Scotland or not
- Council Tax banding (in the SHCS, not in the SHS).

4.31 The results from this regression model are encouraging (shown in Table 4-5 overleaf) for two primary reasons.

4.32 Firstly, the adjusted  $R^2$  of the model is 0.48<sup>33</sup>. In other words, around 48% of the variance can be explained by these factors. **Given the limited number of explanatory variables common to the FRS and the SHS/SHCS, this is a sizeable amount. It follows that an imputation strategy that uses these variables should capture around half of the variation in income of other adults.**

4.33 Secondly, the p-value of the 'whether in Scotland or not' co-efficient suggests that this factor is not significant. Indeed, at 0.75, it is not close to the commonly accepted level of 0.05 significance. This suggests that, once the effects of the other factors have been accounted for, this has little impact.

4.34 The model shows that most of the factors included do help to account for variance in adult income levels. Almost all of the co-efficients relating to age, relationship to HRP, economic status, and car ownership are significant. Sex and tenure appear to be less important in terms of modelling income overall.

---

<sup>32</sup> Number of bedrooms, for example, is collected in all three surveys. However in none of the models run did it prove significant in accounting for any of the variance in income after the other variables had been included.

<sup>33</sup> Analysis of a second year of FRS data 2006-2007 suggests that this is relatively stable

4.35 Removing Council Tax Banding from the model – this variable is not currently in the SHS data - does not significantly reduce the amount of variance explained, with the adjusted  $R^2$  remaining at 0.48 (see Table A2.1 for details).

**Table 4-5: Regression model of square root of net adult income<sup>34</sup> of other adults by factors collected in both the FRS and SHCS. FRS 2005-06. UK**

|  | Unstandardized Coefficients |     | Sig.    |
|--|-----------------------------|-----|---------|
|  | B                           | SE  | P-Value |
| (Constant)   | 12.1                        | 0.3 | 0.00    |
| Parent   | -1.0                        | 0.4 | 0.01    |
| Other relative   | 0.0                         | 0.2 | 0.86    |
| Not related  | 1.2                         | 0.2 | 0.00    |
| PT worker  | -3.4                        | 0.2 | 0.00    |
| Self-employed  | -2.3                        | 0.3 | 0.00    |
| Unemployed etc   | -9.1                        | 0.1 | 0.00    |
| Retired  | -7.0                        | 0.6 | 0.00    |
| Student  | -4.3                        | 0.1 | 0.00    |
| Sex (female)   | 0.0                         | 0.1 | 0.77    |
| Aged 20-24   | 2.1                         | 0.1 | 0.00    |
| Aged 25-29   | 3.2                         | 0.2 | 0.00    |
| Aged 30-34   | 3.7                         | 0.2 | 0.00    |
| Aged 35-39   | 4.5                         | 0.3 | 0.00    |
| Aged 40-44   | 3.6                         | 0.3 | 0.00    |
| Aged 45-49   | 4.3                         | 0.4 | 0.00    |
| Aged 50-54   | 4.8                         | 0.5 | 0.00    |
| Aged 55-59   | 4.3                         | 0.5 | 0.00    |
| Aged 60-64   | 4.8                         | 0.6 | 0.00    |
| Aged 65-69   | 5.4                         | 0.7 | 0.00    |
| Aged 70-74   | 6.1                         | 0.8 | 0.00    |
| Aged 75-79   | 6.5                         | 0.8 | 0.00    |
| Aged 80-84   | 6.8                         | 0.8 | 0.00    |
| Aged 85+   | 6.6                         | 0.8 | 0.00    |
| Scottish survey  | -0.1                        | 0.2 | 0.75    |
| CT Band B  | 0.0                         | 0.2 | 0.96    |
| CT Band C  | 0.4                         | 0.2 | 0.00    |
| CT Band D  | 0.6                         | 0.2 | 0.00    |
| CT Band E  | 0.4                         | 0.2 | 0.05    |
| CT Band F  | 0.8                         | 0.2 | 0.00    |
| CT Band G  | 0.9                         | 0.3 | 0.00    |
| CT Band H  | 3.1                         | 0.7 | 0.00    |
| Mortgage   | -0.2                        | 0.1 | 0.21    |
| Shared owner   | 0.3                         | 1.0 | 0.78    |
| Rent   | -0.3                        | 0.2 | 0.09    |
| Rent free  | -0.3                        | 0.5 | 0.61    |
| Household has car/van?   | 0.4                         | 0.1 | 0.01    |
| Adjusted $R^2$ of Model = 0.48   |                             |     |         |
| Constant – Child, 16-19, in FT work, owning property outright, CT band A |                             |     |         |

<sup>34</sup> From sources collected in the SHS. For the remainder of this report, net adult household income measures exclude the sources of income that are collected as part of the FRS but not the SHS and detailed on page 4.

- 4.36 Although almost half of the variance in income levels overall can be explained, this is driven primarily by economic status. **Appendix 1 provides our suggested best approach to imputing income for other adults separately based on economic status, and details the amount of variance explained by variables common to the FRS and the SHCS within each of the different economic statuses**<sup>35</sup>.
- 4.37 Within economic statuses, far less variance in income levels can be accounted for by variables that are common to both the FRS and the SHS/SHCS. Among full-time employees, around 15-16% of variance in income levels can be explained. Among students, the figure is around 9-10%, while among those who are retired between 6-12% of the variance can be explained<sup>36</sup>.
- 4.38 The other key consideration, as noted above, is whether there is a significant difference between Scotland and the rest of the UK. Among students, pensioners, and other economically inactive, once other factors have been controlled for, there is no evidence of significant differences between Scotland and the rest of the UK. This is reassuring. However, among workers, the analysis does show that there are regional effects, and that the split between Scotland and the rest of the UK does prove significant. This suggests that an imputation strategy would lead to systematically biasing the imputed data if this was not taken into account.
- 4.39 Full details of regional variations of income levels of earners are provided in Appendix 1. Tables A1.8, A1.11 and A1.14 show that the income levels in London are out of line with the rest of the UK among full-time, part-time and self-employed workers. Although there is regional variation in income levels, once data from London is excluded from the analysis, the difference between Scotland and the rest of the UK is not significant<sup>37</sup>. **It follows that, imputation strategies for full-time, part-time and self-employed workers, once data from London are excluded, would not lead to artificially high estimates of income in Scotland.**

---

<sup>35</sup> And Appendix 2 provides details of the amount of variance explained by variables common to both the FRS and the SHS. In other words, repeating the analysis with CT band excluded.

<sup>36</sup> These figures are for both the SHS and the SHCS models. As the SHS does not contain information on CT band, slightly less of the variance can be explained by factors that could be used in the imputation of income for this survey.

<sup>37</sup> See Tables A1.9, A1.12 and A1.15.

4.40 Analysis of the 2006-07 FRS dataset suggest that **the level of variance explained overall and within each of the economic status groups is relatively consistent year on year. While there are minor differences in the significance and explanatory power of different factors, they do not call into question the overall suggested approach.**

4.41 As economic status is clearly – both intuitively and statistically – linked to income, we suggest structuring the detailed imputation strategies around this variable. However, given the obvious danger of using data covering the whole of the UK to impute income in Scotland, we suggest excluding data from London for the imputation of income from earnings.

### Summary of proposed imputation methodology for other adults

4.42 We would suggest that the best possible model for imputation would be to use a series of hierarchical hot deck imputation routines to impute income using the following approach:

- Adjust the *net* income of all other adults in the FRS across the UK to take account of the different coverage of benefits in the SHS and the SHCS. This is to ensure that the definition of adult income is consistent within the SHS and SHCS (in other words, that HRP/Spouses' income is not based on a different definition from other adults).
- Use the FRS to model the income levels of other adults in order to design a series of donor group typologies and to confirm whether there is evidence of any significant differences between Scotland the rest of the UK. This should be undertaken separately by economic status. Several donor group typologies should be designed for each to ensure that there are no cells where a match can't be made between the FRS and SHS/SHCS. Factors that have the smallest explanatory power should be collapsed or excluded first. Donor groups should be amended each year using FRS data.
- Discard extreme values of income<sup>38</sup> to minimise the potential mismatch between the actual (unknown) value of income and the imputed value, while maintaining the shape of the distribution.
- Create a 'donor data' dataset from the FRS that only includes the income and the various typology variables.
- Recreate the donor typologies in the SHS and the SHCS datasets and link these datasets to the FRS 'donor dataset'.

---

<sup>38</sup> This would be done by excluding the lowest and highest 5% of values.

- Impute all other adults' income in the SHS/SHCS using hierarchical hot decking.
- Create an additional measure of household income in the SHS by summing the existing income measure with the imputed income of other adults in the household.



# Testing the proposed imputation strategy



## 5. Testing the proposed imputation strategy

5.1 Imputation of income for other adults was carried out for the SHS to test the suggested imputation methodology using the methods detailed in Appendix 1 and summarised in the previous chapter. This chapter details the effect that the proposed income imputation strategy would have on overall estimates for household income. It examines the robustness of using these estimates. It also discusses the limitations of using the proposed estimates of household income.

### Measuring the effect of the proposed imputation strategy

5.2 Table 5-1 shows the average weekly household income before and after imputation. Overall, the mean weekly household income increases from £391 to £427 while the median income in the SHS increases from £317 to £346 a week. This represents a 9% increase overall and is almost exactly in line with the proportion of household income from other adults found in the FRS.

**Table 5-1: Average weekly household income pre and post imputation of other adults. SHS, April 05 – Mar 06.**

|                            | <b>Median</b> | <b>Mean</b> |
|----------------------------|---------------|-------------|
| SHS data – Pre-imputation  | £317          | £391        |
| SHS data – Post-imputation | £346          | £427        |

5.3 Table 5-2 compares the proportion of household income from “other adults” in the FRS to the SHS following imputation across the different types of household composition discussed in Chapter 3. Among all households where there are other adults in the FRS, the income of these other adults accounts for 34% of household income. Following imputation in the SHS, the corresponding proportion of income is 32%. That the proportion is slightly lower for the SHS is likely to be due at least partially to the effect of excluding the highest 5% and lowest 5% of values in each economic status from the donor groups (as the mean will be more effected by the exclusion of the highest 5% of values).

5.4 When broken down by household type, the increase in the household income estimates in the SHS due to imputation is also in line with what we would expect from the FRS. In the most common household composition with other adults – three adults and no children – the respective figures for the SHS and the FRS is

37% and 34%. These results are very positive, especially given that household composition was not used explicitly as a controlling factor in the imputation models

**Table 5-2: Proportion of income from ‘other adults’ in households with more than one Benefit Unit, and distribution of household types with total sample, FRS, SHS after imputation, 2005-06.**

| Household composition                           | Proportion of household income from other adults (where other adults present in household) |            | Percentage of all households |              |
|---|--|------------|------------------------------|--------------|
|   | FRS  | SHS        | FRS                          | SHS          |
| Two adults, no children, both over pension age  | 43%  | 37%        | 0.2%                         | 0.3%         |
| Two adults, no children, one over pension age   | 42%  | 37%        | 1.0%                         | 1.5%         |
| Two adults, no children, both under pension age | 37%  | 36%        | 2.6%                         | 2.8%         |
| Three or more adults, no children               | 37%  | 34%        | 8.4%                         | 8.3%         |
| Two adults, one child                           | 29%  | 32%        | 0.6%                         | 0.7%         |
| Two adults, two children                        | 25%  | 23%        | 0.2%                         | 0.2%         |
| Two adults, three or more children              | 19%  | 19%        | 0.1%                         | 0.0%         |
| Three or more adults, one child                 | 26%  | 25%        | 2.2%                         | 2.7%         |
| Three or more adults, two children              | 21%  | 23%        | 0.8%                         | 0.9%         |
| Three or more adults, three or more children    | 27%  | 23%        | 0.3%                         | 0.3%         |
| <b>Total</b>                                    | <b>34%</b>   | <b>32%</b> | <b>16.3%</b>                 | <b>17.8%</b> |

5.5 The proportion of households where there is more than one benefit unit differs marginally between the FRS and the SHS. From Table 5-2, it can be seen that the FRS estimates that around 16.3% of households in Scotland have more than one benefit unit according to the FRS, but the equivalent analysis of the SHS gives a figure of 17.8%. This difference is within sample error and is not likely to compromise the suggested imputation strategy<sup>39</sup>.

### Examining the robustness of using imputed data

5.6 The imputation of income for the HRP and spouse has a greater impact on household income than the imputation of the income of other adults. In order to examine the robustness of using imputed data, we analyse the impact of the existing imputation procedures before discussing the effect of the additional proposed imputation strategy for other adults. As way of an example, we also

<sup>39</sup> This difference may be due to non-response bias and the respective weighting strategies used to correct for this.

examine the effect that the proposed imputation strategy would have on estimates of relative poverty in Scotland.

- 5.7 As an example of a possible analysis that might be undertaken using imputed income data, Table 5-3 illustrates net total household income in the SHS by local authority and year. It shows the difference in estimates between 2005 and 2006 for all households in each local authority, and for household where imputation of earnings – for the HRP or Spouse - was not necessary<sup>40</sup>. Such an analysis should, of course, be limited to the larger local authorities where the sample can be considered representative on an annual basis (these are highlighted in bold in Table 5.3), but all local authorities are included for illustration.
- 5.8 Overall, there are 8 local authorities (highlighted in yellow in Table 5-3) where excluding cases with imputed earnings leads to a different interpretation of change in income levels over time. In two of these eight local authorities, Glasgow and Edinburgh, a single year of SHS data is considered robust enough for analysis.
- 5.9 Two interpretations of the data could be made. First that change across years between the estimates before and after imputation is due to noise created by the imputation process. Variability in earnings is only partially accounted for by the imputation process. By maintaining the unexplained variance in the donor data in the imputed data, the imputation process is, in effect, distributing this unexplained variance across the imputed values.
- 5.10 Secondly, that the imputation process is helping to correct for under-reporting of earnings. Differences in income levels (or lack of differences) may reflect differences in the proportion of earners who do not give their income between years. In other words, the imputation process is correcting for differences in non-response to the earnings questions.
- 5.11 Comparing the figures of a local authority where the impact of imputation makes a difference to the imputation to one where imputation does not – Glasgow and Fife for example - provides illumination. There is a significant difference between years in the mean level of household income after imputation of earnings but not before imputation in Glasgow. In contrast, there is no significant difference before or after imputation between 2005 and 2006 in Fife.

---

<sup>40</sup> This shows the effect of the imputation of earnings of the HRP and spouse, and not of other adults. This imputation has a far greater effect on the estimates of income and allows a comparison across years.

**Table 5-3: Income by local authority area and year, before and after imputation of earnings of HRP and Spouse (before imputation of income from other adults), SHS 2005 and 06<sup>41</sup>**

|                          | All households including those with imputed earnings |                |               |            | Excluding cases where earnings have been imputed |                |               |            |
|--------------------------|--|----------------|---------------|------------|--|----------------|---------------|------------|
|                          | 2005   | 2006           | Difference    | Sig?       | 2005   | 2006           | Difference    | Sig?       |
| South Ayrshire           | £18,773  | £18,048        | £-725         | NS         | £17,914  | £16,402        | £-1,512       | NS         |
| <b>South Lanarkshire</b> | <b>£21,764</b>                                       | <b>£21,231</b> | <b>£-534</b>  | <b>NS</b>  | <b>£20,393</b>                                   | <b>£20,146</b> | <b>£-247</b>  | <b>NS</b>  |
| Stirling                 | £25,240  | £23,512        | £-1,728       | NS         | £21,408  | £21,397        | £-11          | NS         |
| West Dunbartonshire      | £18,159  | £18,970        | £811          | NS         | £16,636  | £17,876        | £1,240        | NS         |
| West Lothian             | £22,033  | £22,998        | £965          | NS         | £21,122  | £21,841        | £719          | NS         |
| Eilean Siar              | £17,958  | £18,813        | £855          | NS         | £15,743  | £16,389        | £646          | NS         |
| Aberdeen City            | £22,457  | £22,628        | £171          | NS         | £21,743  | £22,001        | £257          | NS         |
| Aberdeenshire            | <b>£22,852</b>                                       | <b>£24,779</b> | <b>£1,926</b> | <b>SIG</b> | <b>£21,671</b>                                   | <b>£22,820</b> | <b>£1,149</b> | <b>NS</b>  |
| Angus                    | <b>£20,286</b>                                       | <b>£23,126</b> | <b>£2,840</b> | <b>SIG</b> | <b>£18,279</b>                                   | <b>£20,794</b> | <b>£2,515</b> | <b>NS</b>  |
| Argyll and Bute          | £20,809  | £20,719        | £-90          | NS         | £19,445  | £19,517        | £72           | NS         |
| Scottish Borders         | £20,755  | £19,212        | £-1,544       | NS         | £18,994  | £16,541        | £-2,453       | NS         |
| Clackmannanshire         | £18,564  | £21,852        | £3,288        | SIG        | £17,632  | £20,421        | £2,789        | SIG        |
| Dumfries and Galloway    | £17,166  | £18,514        | £1,348        | NS         | £15,359  | £16,666        | £1,307        | NS         |
| Dundee City              | <b>£16,716</b>                                       | <b>£17,971</b> | <b>£1,254</b> | <b>NS</b>  | <b>£14,939</b>                                   | <b>£17,106</b> | <b>£2,167</b> | <b>SIG</b> |
| East Ayrshire            | £18,598  | £20,110        | £1,513        | NS         | £17,840  | £18,305        | £464          | NS         |
| East Dunbartonshire      | £22,909  | £24,010        | £1,101        | NS         | £21,841  | £23,053        | £1,212        | NS         |
| East Lothian             | £21,374  | £19,983        | £-1,391       | NS         | £20,325  | £17,663        | £-2,663       | NS         |
| East Renfrewshire        | <b>£24,803</b>                                       | <b>£27,652</b> | <b>£2,849</b> | <b>NS</b>  | <b>£23,135</b>                                   | <b>£26,925</b> | <b>£3,790</b> | <b>SIG</b> |
| <b>Edinburgh City</b>    | <b>£20,376</b>                                       | <b>£21,916</b> | <b>£1,540</b> | <b>SIG</b> | <b>£19,500</b>                                   | <b>£20,890</b> | <b>£1,390</b> | <b>NS</b>  |
| Falkirk                  | £19,136  | £19,949        | £813          | NS         | £17,875  | £18,172        | £296          | NS         |
| <b>Fife</b>              | <b>£20,045</b>                                       | <b>£20,051</b> | <b>£6</b>     | <b>NS</b>  | <b>£18,758</b>                                   | <b>£17,975</b> | <b>£-783</b>  | <b>NS</b>  |
| <b>Glasgow City</b>      | <b>£16,568</b>                                       | <b>£17,474</b> | <b>£906</b>   | <b>SIG</b> | <b>£15,630</b>                                   | <b>£16,077</b> | <b>£447</b>   | <b>NS</b>  |
| Highland                 | <b>£20,162</b>                                       | <b>£22,256</b> | <b>£2,094</b> | <b>SIG</b> | <b>£18,335</b>                                   | <b>£19,652</b> | <b>£1,318</b> | <b>NS</b>  |
| Inverclyde               | £18,651  | £19,155        | £504          | NS         | £16,970  | £17,844        | £874          | NS         |
| Midlothian               | £21,074  | £21,240        | £166          | NS         | £19,795  | £19,750        | £-45          | NS         |
| Moray                    | £21,804  | £21,436        | £-368         | NS         | £19,862  | £19,521        | £-340         | NS         |
| North Ayrshire           | £17,000  | £22,641        | £5,641        | NS         | £15,815  | £20,923        | £5,107        | NS         |
| <b>North Lanarkshire</b> | <b>£19,868</b>                                       | <b>£19,680</b> | <b>£-188</b>  | <b>NS</b>  | <b>£18,743</b>                                   | <b>£18,212</b> | <b>£-531</b>  | <b>NS</b>  |
| <b>Orkney</b>            | <b>£18,017</b>                                       | <b>£19,943</b> | <b>£1,926</b> | <b>SIG</b> | <b>£15,951</b>                                   | <b>£17,588</b> | <b>£1,637</b> | <b>NS</b>  |
| Perth and Kinross        | £23,360  | £23,181        | £-179         | NS         | £21,649  | £20,609        | £-1,040       | NS         |
| Renfrewshire             | £20,638  | £21,358        | £720          | NS         | £19,181  | £20,205        | £1,025        | NS         |
| Shetland                 | £19,653  | £23,380        | £3,727        | SIG        | £17,683  | £22,080        | £4,397        | SIG        |
| <b>Overall</b>           | <b>£20,028</b>                                       | <b>£20,894</b> | <b>£866</b>   | <b>SIG</b> | <b>£18,631</b>                                   | <b>£19,282</b> | <b>£651</b>   | <b>SIG</b> |

5.12 The largest effect on the imputed values will come from earnings of the HRPs main job. Table 5-4 shows the proportion of households where earnings was imputed in 2005 and 2006 for these two local authorities. In Glasgow, the proportion of HRPs whose earnings was imputed was 9% in 2005, and increased

<sup>41</sup> The five local authorities in bold are those where one year of SHS data is considered robust enough for analysis. The eight local authorities highlighted in yellow are where excluding cases with imputed earnings leads to a different interpretation of change in income levels over time.

to 14% in 2006. As a higher proportion of income from earnings was missing in the pre-imputation estimates in 2006 than in 2005 in Glasgow, any real increase in household income is likely to have been masked by these changes in non-response. In Fife, although there is small increase in the proportion of households where earnings for the HRP's main job was imputed between 2005 and 2006 (11% compared to 13%), there is also a small increase in the proportion of households where the HRP was not in paid employment (44% to 46%).

**Table 5-4: Summary of imputation of earnings of the HRP in Glasgow and Fife. SHS 2005 and 2006**

| Year | Summary of earnings of HRPs main job | Fife | Glasgow City |
|------|--------------------------------------|------|--------------|
| 2005 | Not in paid employment               | 44%  | 56%          |
|      | In employment, income given          | 46%  | 35%          |
|      | In employment, income imputed        | 11%  | 9%           |
|      | All                                  | 100% | 100%         |
| 2006 | Not in paid employment               | 46%  | 53%          |
|      | In employment, income given          | 40%  | 33%          |
|      | In employment, income imputed        | 13%  | 14%          |
|      | All                                  | 100% | 100%         |

**5.13 This analysis suggests that the imputation of earnings of the HRP is helping to correct for differences in non-response to the earnings questions, and therefore, that the imputed estimates (and differences across years) are more robust and useful than the estimates with the imputed data excluded.**

5.14 Table 5-5 shows the effect of imputing income of other adults using the strategy outlined previously on estimates of household income in different local authorities. Again, all local authorities are shown but only the figures of the local authorities highlighted in bold should be considered robust given that the estimates are based on a single year of data.

5.15 Of the five local authorities where annual results can be considered robust, the smallest increase in the median income from the additional imputation is seen in Fife (5%) while the largest increase is seen in North Lanarkshire (15%). This again raises the question, is this variation a result of the noise built into the imputation process, or is the imputation process reflecting differences in these local authorities?

**Table 5-5: Net household income before and after imputation of other adult's income by local authority, SHS 2005-2006**

| Local authority          | Mean              |                  |               | Median            |                  |               |
|--------------------------|-------------------|------------------|---------------|-------------------|------------------|---------------|
|                          | Before imputation | After imputation | %age increase | Before imputation | After imputation | %age increase |
| Aberdeen City            | £434              | £466             | 7%            | £343              | £362             | 5%            |
| Aberdeenshire            | £456              | £489             | 7%            | £395              | £420             | 6%            |
| Angus                    | £400              | £430             | 7%            | £335              | £371             | 11%           |
| Argyll and Bute          | £402              | £426             | 6%            | £334              | £353             | 6%            |
| Clackmannanshire         | £381              | £412             | 8%            | £330              | £355             | 8%            |
| Dumfries and Galloway    | £333              | £366             | 10%           | £274              | £290             | 6%            |
| Dundee City              | £332              | £363             | 10%           | £253              | £275             | 9%            |
| East Ayrshire            | £346              | £385             | 11%           | £303              | £326             | 8%            |
| East Dunbartonshire      | £474              | £505             | 7%            | £406              | £431             | 6%            |
| East Lothian             | £420              | £443             | 6%            | £323              | £340             | 5%            |
| East Renfrewshire        | £472              | £500             | 6%            | £411              | £438             | 7%            |
| <b>Edinburgh City</b>    | <b>£402</b>       | <b>£437</b>      | <b>9%</b>     | <b>£323</b>       | <b>£346</b>      | <b>7%</b>     |
| Eilean Siar              | £372              | £423             | 14%           | £317              | £361             | 14%           |
| Falkirk                  | £365              | £391             | 7%            | £300              | £319             | 6%            |
| <b>Fife</b>              | <b>£395</b>       | <b>£423</b>      | <b>7%</b>     | <b>£330</b>       | <b>£348</b>      | <b>5%</b>     |
| <b>Glasgow City</b>      | <b>£326</b>       | <b>£369</b>      | <b>13%</b>    | <b>£259</b>       | <b>£290</b>      | <b>12%</b>    |
| Highland                 | £393              | £428             | 9%            | £338              | £369             | 9%            |
| Inverclyde               | £361              | £405             | 12%           | £307              | £333             | 9%            |
| Midlothian               | £411              | £456             | 11%           | £366              | £407             | 11%           |
| Moray                    | £418              | £452             | 8%            | £346              | £384             | 11%           |
| North Ayrshire           | £400              | £428             | 7%            | £284              | £304             | 7%            |
| <b>North Lanarkshire</b> | <b>£380</b>       | <b>£428</b>      | <b>13%</b>    | <b>£307</b>       | <b>£352</b>      | <b>15%</b>    |
| Orkney                   | £358              | £397             | 11%           | £325              | £341             | 5%            |
| Perth and Kinross        | £458              | £485             | 6%            | £361              | £402             | 11%           |
| Renfrewshire             | £407              | £442             | 9%            | £337              | £367             | 9%            |
| Scottish Borders         | £373              | £394             | 6%            | £295              | £310             | 5%            |
| Shetland                 | £398              | £446             | 12%           | £380              | £392             | 3%            |
| South Ayrshire           | £358              | £398             | 11%           | £275              | £316             | 15%           |
| <b>South Lanarkshire</b> | <b>£417</b>       | <b>£459</b>      | <b>10%</b>    | <b>£358</b>       | <b>£398</b>      | <b>11%</b>    |
| Stirling                 | £495              | £534             | 8%            | £370              | £393             | 6%            |
| West Dunbartonshire      | £357              | £402             | 13%           | £278              | £319             | 15%           |
| West Lothian             | £426              | £463             | 9%            | £368              | £392             | 7%            |
| <b>Total</b>             | <b>£391</b>       | <b>£427</b>      | <b>9%</b>     | <b>£317</b>       | <b>£346</b>      | <b>9%</b>     |

5.16 While in Fife, 15% of households contain 'other adults', in North Lanarkshire, 24% of households contain 'other adults'. Households with three or more adults and no children – the household type where the inclusion of other adult income will have the most effect on household income – are also more prevalent in North Lanarkshire (10%) than in Fife (7%). Therefore, the difference in the impact of the additional imputation on North Lanarkshire and Fife is primarily due to differences in the distribution of household types. **This again suggests that the imputation strategy is successfully reflecting differences between the local authorities.**

5.17 One of the potential uses of using the broader definition of household income is to produce estimates of the proportion of households in relative poverty. The aim of reducing relative poverty is given and defined in National Indicator 14, namely to “decrease the proportion of individuals living in private households with an equivalised income of less than 60% of the UK median before housing costs”<sup>42</sup>.

5.18 Table 5-6 shows the median net household income across Scotland using the old and proposed new definition, for both equivalised and non-equivalised income estimates. After imputation of income from other adults, the median equivalised income in Scotland is £355 a week. This compares to a figure of £362 across the UK for the same period according to the FRS, a difference of £7 per week.

**Table 5-6: Median net household income before and after imputation of other adult’s income, SHS, FRS 2005-2006**

|   | <b>Median</b> |
|---|---------------|
| <b>SHS Non-equivalised old (narrow) definition</b>  | £317          |
| <b>SHS Equivalised old (narrow) definition</b>      | £323          |
| <b>SHS Non-equivalised new (broader) definition</b> | £346          |
| <b>SHS Equivalised new (broader) definition</b>     | <b>£355</b>   |
| <b>FRS Equivalised FRS for UK</b>                   | <b>£362</b>   |

5.19 What effect does this have on households classified as being in relative poverty? As noted previously, imputation of income for other adults increases the estimates of household income overall in the SHS. Therefore, basing estimates of the proportion of households in relative poverty (less than <60% of the median income) on the FRS estimates of the median equivalised income across the UK would mean that fewer households are classed as being in relative poverty using the new broader definition of household income. Overall, 23% of households in Scotland had an equivalised household income of less than £217 per week (60% of £362) before imputation of income of other adults. After imputation the proportion drops to 19%.

5.20 More importantly, the imputation of other adults’ income impacts on the types of household that would be classified as in relative poverty. Table 5-7 shows the proportion of different types of household that would be considered to be in relative poverty using the narrow and broad definitions of household income. For ease of comparison, the threshold was based on the figures for 60% of the median income before and after imputation (£194 and £217).

<sup>42</sup> As part of the Scottish Government’s National Performance Framework, in the Scottish Budget Spending Review 2007. See [www.scotland.gov.uk/Publications/2007/11/13092240/0](http://www.scotland.gov.uk/Publications/2007/11/13092240/0)

5.21 There are clear differences by household type. In households with a single person over pensionable age, the use of the broader estimates of household income increases the estimate of those in relative poverty from 23% to 30%. This is because the threshold figure for 60% of median income increases, but the imputation of other adults' income does not affect these households. Relative poverty rates also increase significantly in single parents households (for example, from 22% to 30% in households with one adult and one child). In comparison, the proportion of households with three adults and no children in relative poverty drops from 24% to just 4%.

5.22 Imputation of other adults' income will clearly have a significant impact on estimates of relative poverty, both in the level of relative poverty, and where it is found.

**Table 5-7: Proportion of households in relative poverty (below 60% of the median net household income) before and after imputation of other adult's income by household type. SHS, 2005-2006**

| Household type                                  | SHS new (broad) definition<br>%age <60% median<br>equivalised income<br>(£217) | SHS old (narrow) definition<br>%age <60% median<br>equivalised income<br>(£194) |
|---|--|---|
| One adult, no children over pension age         | 30%  | 23%   |
| One adult, no children, under pension age       | 21%  | 17%   |
| Two adults, no children, both over pension age  | 35%  | 29%   |
| Two adults, no children, one over pension age   | 17%  | 23%   |
| Two adults, no children, both under pension age | 9%   | 12%   |
| Three or more adults, no children               | 4%   | 24%   |
| One adult, one child                            | 30%  | 22%   |
| One adult, two children                         | 42%  | 30%   |
| One adult, three or more children               | 52%  | 36%   |
| Two adults, one child                           | 8%   | 9%  |
| Two adults, two children                        | 9%   | 8%  |
| Two adults, three or more children              | 21%  | 16%   |
| Three or more adults, one child                 | 6%   | 18%   |
| Three or more adults, two children              | 9%   | 19%   |
| Three or more adults, three or more children    | 14%  | 31%   |
| <b>All</b>                                      | <b>18.7%</b>   | <b>18.4%</b>  |

5.23 The results of the trial imputation are positive. The increase in the mean estimate for household income is almost exactly in line with the proportion of household income from other adults found in the FRS. The estimates are what we would expect when broken down by household type. Equivalised median income after



imputation given by the SHS for Scotland is £7 a week lower than the respective measure given by the FRS for the UK. The analysis of the imputation of earnings data for the HRPs' main job suggests that this part of the imputation process helps to correct for differing levels of non-responses, making analysis across years more rather than less robust. The analysis of the effect of imputing other adults' income shows that this successfully reflects differences in household composition across local authorities.

**5.24 At the national and local authority level, the estimates of household income will be robust.** As the imputation strategy employed used stochastic methods to maintain the variability in the data, the effect on the calculation of standard errors of estimates and confidence intervals will be minimised, allowing the robustness of estimates to be estimated through standard measures.

### **Limitations of using household income estimates in the SHS and SHCS**

**5.25 These findings suggest that it is feasible to impute income for other adults in the SHS and SHCS using the FRS, in order to create a wider definition of household income.** It is, however, important to note the limitations of the resulting data. Clearly, the imputed income of other adults in the SHS should not be examined separately from the rest of household income. The income sections of the SHS and the SHCS are not designed to capture individual adult income, but rather, to produce a measure of household income. **We would recommend that for external use of the data, information on income from other adults is not included separately in the dataset<sup>43</sup>.**

**5.26 Particular care should be taken when undertaking sub-group analysis, where the use of imputed data could be misleading.** For most sub-group analyses of income, the effect of the additional imputation should be minimal. However, the smaller the sub-group, the more caution should be taken when using imputed data. This will be particularly important when the sub-group is likely to have an income distribution that is not accounted for by the imputation strategy.

**5.27 As the imputation of other adult's income impacts most on households with more than 2 adults, the risk of misinterpreting data are particularly high when examining differences among this sub-group.** For example,

consider a comparison of income levels of student households in Glasgow with student households in Edinburgh. The primary factor limiting the robustness of this analysis will be the small sample size. The imputation strategy will account for the fact that students living with their parents will have, on average, a lower income than those living on their own. However, any difference in earning levels for student jobs in Glasgow with those in Edinburgh will not be accounted for in the imputation strategy.

- 5.28 It is worth emphasizing that neither the current procedures for imputing income of the HRP or spouse in the SHS and SHCS, nor the proposed procedures for imputing the income of other adults are able to incorporate regional variations, especially at the local authority level. This is primarily because differences in income by region are small once other factors have been controlled for, or because such differences in income levels are limited to particular sub-groups where a small sample size make such differences difficult to estimate. The overall effect of the imputation will be to flatten the differences across variables that are not included in the imputation process: high 'true' values will be slightly underestimated and low 'true' amounts will be overestimated. Given that the variation will be small, or limited to a small number of cases, the effect on the estimates using imputed income will also be small. However, **any policy initiatives that impact on income levels of non-householders, perhaps in particular localities within Scotland, will not be reflected in the estimates of household income.**
- 5.29 Good interpretation of data needs to take into consideration the potential impact of the imputation procedures, together with the nature and potential uses of the output. **In order to aid interpretation, we would recommend that an imputation flag variable is created.** This would indicate cases where other adults' income has been imputed. It would allow imputed data to be excluded from any analyses that might be sensitive to the imputation procedures. Results with and without imputed data could be compared in order to assess the effects of imputation on the findings. In this way, the data would benefit from the advantages of imputation while reducing the dangers of misinterpretation.
- 5.30 Finally, although it is beyond the scope of this paper to consider what should constitute the definition of a household income, it is worth noting that the creation of a broader definition of household income would give two different estimates of

---

<sup>43</sup> It would be, of course, easy to calculate this by subtracting the old narrower definition of household income from the new broader definition of household income.

a household's income in the SHS. Which estimate is more preferable will depend partly on the type of analysis undertaken and its purpose. However, such decisions will also, inevitably, depend on the view taken on a long-running question, how far do other adults' income constitute a household resource?: should an adult child's income be considered in its entirety to be contributing to the household income, or should only any dig money they paid be considered part of the household's resources?

Other potential ways  
to improve the income data  
in the SHS and SHCS

## 6. Other potential ways to improve the income data in the SHS and SHCS

- 6.1 In this section, we explore other areas that could bring potential benefits to estimating income in the SHS and the SHCS, namely:
- the processes used to check and clean the data,
  - potential changes to the questionnaire,
  - additional imputation routines to impute income in the small proportion of households where the measure of household income is currently set to missing,
  - and annual re-imputation of income data in the SHS combined with the imputation of the SHCS.

### Data cleaning and processing routines

- 6.2 Staff in the FRS team at the DWP undertake a considerable amount of data checking and cleaning of the data. This involves a high level of staff input, with the timescales for data production about twice as long as for the SHS.
- 6.3 However, most importantly, this relies on the detailed data that is collected as part of the FRS interview. There is less scope for such detailed data checking in the SHS and the SHCS, particularly regarding level of benefit receipt. Unlike the FRS, eligibility cannot be calculated for all but a few types of benefit.
- 6.4 There is a particular danger of introducing bias when cleaning the income data in that it is much easier to identify misreported high levels of income than it is to identify low levels of income. For example, receipt of £1,000 a fortnight in Income Support is an obvious mistake, while receipt of £10 a fortnight in Income Support could be a correct response or it could be understated by a factor of 10.
- 6.5 There is a second potential issue. Some respondents may report the correct amount of benefit income they receive but allocate it to a single benefit rather than a number of benefits. Therefore, limiting the amount received to an upper threshold may improve the estimate for income received for that benefit, but underestimate the estimate of household income.
- 6.6 Currently in the SHS and the SHCS, outliers are all individually checked, and any values that are clear mistakes are corrected. However, any value which is over the maximum threshold but could be a valid amount for total benefit income

received is not amended (though the case would be excluded as a potential donor case for imputation).

- 6.7 However, it would be possible to improve these procedures by making them more systematic. This would involve the following. **Each benefit would be checked against the minimum and maximum thresholds and amended to be within these limits, but with the difference given to a new computed variable.** This new variable could then be checked against various indicators (such as benefits received, household type, economic status) to decide whether to include this amount in the total income from benefits. All negative values (where it is likely that a respondent will have under-reported levels of a benefit received) would be excluded (as the adjustment will have been made to the individual benefit.) For positive values, the decision would be less straightforward, and the decision would involve considering whether it is more likely that there is over-reporting the amount of income received, or under-reporting the benefits that they received. Therefore, a figure of an extra £100 a week might be dropped if there was no evidence that the household received any means tested benefits or disability benefits, but included if they did receive a single means tested benefit, and had a low overall income.
- 6.8 While there is scope with the FRS data to calculate likely eligibility for benefits, and potentially to impute receipt, the SHS and the SHCS are more limited with this respect. Only the amount received and not whether a benefit is received has been amended previously. **Given the difference between eligibility and take-up of benefits, and the limited availability of data to calculate eligibility, we would not recommend imputing receipt of any benefits, with the possible exception of Child Benefit and Winter Fuel Payment.** These two benefits have very high take-up rates, and eligibility can be calculated from the existing data. We do recommend that there may be amendments that could be made to the questionnaires to ensure that receipt is correctly identified.
- 6.9 Currently, all households with a net total household income is set to missing if the computed figure is less than £25 a week. Although a small proportion of households will have a lower income than this – and be living off savings or loans – it is likely that some households will have either under-reported receipt of benefits or the imputation process has resulted in a low value being given. In the SHS and SHCS, households where income is set to missing are not currently inspected at the end of the process. **We would recommend that these are visually inspected on a case by case basis, and where necessary, either re-**

**impute components of income or assess whether it is likely that they are receiving other benefits.**

- 6.10 Finally, high levels of given earnings are currently excluded from the imputation routines, but low levels of earnings are not. **We would suggest amending the current imputation procedures to exclude both the highest 5% and the lowest 5% of values for earnings from use as donor cases.**

## Changes to the questionnaire

- 6.11 There are a number of changes that could be made to the SHS and the SHCS questionnaires that would help improve the estimation of income levels, short of repeating the income section for all adults in the household. Clearly, any amendments that add to the interview length, or require additional budget for processing, need to be balanced against other competing demands. Five potential changes are suggested below for consideration:
- 6.12 **Supplement the main activity question with the questions required to calculate the ILO definition of economic status for ‘other adults’.** This would ensure that the definitions of economic status are consistent between the FRS and the SHS/SHCS. It would require several additional questions asked about each member of household, such as whether undertaken paid employment in last 7 days, whether currently absent from a job, and whether waiting to start a new job. These questions are currently asked of the HRP in the SHS (in section HG of the questionnaire) and SHCS but not of all adults in the household.
- 6.13 **Add questions to calculate NS-SEC<sup>44</sup> for all workers in the household.** The greatest variation in income levels is related to earnings. Computing NS-SEC for workers would allow a much greater amount of the variance in income levels to be explained and incorporated into the imputation routines. For example, it would allow us to use the income of someone in a higher managerial position to impute a missing income of a person in a similar occupation. Currently, it is not possible to distinguish such a person from, for example, someone in routine agricultural work. This would involve additional questions for all workers who are in the household (they are already asked of the HRP). Three of these are open questions that would also require to be coded<sup>45</sup>.

<sup>44</sup> National Statistics – Social Economic Classification

<sup>45</sup> Namely, name/title of job, description of job, description of firm/organisation.

#### 6.14 **Add additional checks to the CAPI script related to the receipt of benefits.**

As noted above, as the SHS and SHCS do not – and cannot feasibly - collect such comprehensive information on households to allow benefit eligibility to be calculated. However, consideration should be given to adding the following soft checks to the CAPI scripts to ask for clarification in the following instances:

- Children in household and no Child Benefit received.
- No children in household and Child Benefit received.
- Main activity is retired, age is greater than retirement age, and does not receive state pension or occupational pension.
- Main activity is not retired, age is less than retirement age and does receive state pension.
- Main activity is unemployed (and no-one in household in employment) and no receipt of means-tested benefits.
- Main activity is long-term sick/disabled or short-term sick disabled and not receiving any likely benefits (Incapacity Benefit, DLA, Statutory sick pay etc.)

6.15 It should be noted that soft checks such as these can add to the length of the interview, the flow of the interview, and the time taken to script the questionnaire. Additionally, where checks are using information from different parts of the questionnaire, such as the household grid section and the income section, it can be difficult to return to the earlier section of the questionnaire to amend.

6.16 **Amend the questionnaire to ask about income from investments as a separate question rather than as part of “all other sources”.** As noted by Raab et al (2004), the SHS and the SHCS currently under-estimate income from investments. While it is unlikely that these surveys can afford the space to ask a similar barrage of questions as is currently asked in the FRS to cover investment income, restructuring the questionnaire to ask about investment income separately is likely to lead to increased reporting of such income.

6.17 **Add an additional question to ask if households are living off savings or loan.** We discuss this option further in the next section.

### **Additional imputation for cases where the current measure of income is set to missing**

6.18 As mentioned previously, all cases where net household income is less than £25 a week is currently set to missing in the SHS and the SHCS. This results in around 2-4% of cases per year in the SHS and the SHCS with missing income. In comparison, analysis of the FRS suggests that in 2005-2006 around 2.5% of



households received less than £100 a week. It is therefore likely that a proportion of the cases with missing incomes in the SHS and SHCS are genuinely receiving less than £25 a week in income, but that for some, the income level is under-reported.

6.19 Table 6-1 details the summary of the income variable in the SHS by the economic status of the HRP. Clearly, those with missing incomes are more likely to be economically inactive. While households where the HRP is retired account for 27% of all households where no imputation has been carried out, they make up 57% of all cases with missing income. A similar pattern is seen for students.

**Table 6-1: Summary of income by economic status of HRP. SHS 2005-2006**

|   | <b>Income - none imputed</b> | <b>Income - some imputed</b> | <b>Missing</b> |
|---|------------------------------|------------------------------|----------------|
| Self employed                               | 5%                           | 9%                           | 1%             |
| Full time employment                        | 55%                          | 34%                          | 4%             |
| Part time employment                        | 6%                           | 6%                           | 1%             |
| Looking after home/family                   | 2%                           | 5%                           | 5%             |
| Permanently retired from work               | 27%                          | 31%                          | 57%            |
| Unemployed and seeking work                 | 1%                           | 4%                           | 8%             |
| Higher/further education                    | 1%                           | 2%                           | 15%            |
| Government work/training scheme             | 0%                           | 0%                           | -              |
| Permanently sick or disabled                | 3%                           | 8%                           | 7%             |
| Unable to work due to short term ill-health | 0%                           | 1%                           | 1%             |
| Other                                       | 0%                           | 0%                           | 2%             |
| Total                                       | 100%                         | 100%                         | 100%           |
| N   | 15,291                       | 14,566                       | 1,156          |

6.20 The changes recommended to the data cleaning procedures, the introduction of additional checks to the questionnaire, and the imputation of “other adults” income are likely to reduce the proportion of cases where income is set to missing. There will however, still be a proportion where income is under-reported and neither the data cleaning processes nor the imputation help to correct for this. Imputing household income for all households where income is currently set to missing may artificially inflate the income of households that are relying on savings or loans. However, setting these households’ income to missing leads to over-estimation of income levels overall.

6.21 **Rather than imputing income of these cases, consideration should be given to asking an additional question in the SHS and SHCS to those who are not in employment or receiving any benefits, about whether they are living off savings, loans or any other source of income not asked about.** However, this is likely to be a sensitive question, and could potentially have an adverse

affect on the proportion of respondents agreeing to be recontacted. An alternative strategy would be to assume these incomes are genuine and should not be set to missing.

## Undertaking an annual re-imputation of the income data in the SHS together with the imputation of the SHCS

- 6.22 As mentioned previously, imputation procedures are more robust where there are a large number of cases, where the ratio of donors to missing cases is high, and when variation in the donor cases can be explained by other factors in the data. Ideally, there should be more donor cases than missing cases. Currently the income imputation in the SHS is done on a quarterly basis. **Undertaking an annual re-imputation of income data in the SHS, together with the imputation of the SHCS data could potentially improve the robustness of the imputation of the income of the HRP and their spouse in both the SHS and the SHCS.**
- 6.23 Table 5.2 provides details of the level of receipt of selected components of household income in the SHS and SHCS and the ratio of donor cases to missing cases. It shows all components of income that are received by more than 10% of the sample and selected other components<sup>46</sup>. With a couple of significant exceptions, the components that are received by a high proportion of households have a high ratio of donors to missing cases.
- 6.24 For example, income from the HRP's earnings from their main job is received by the majority of all households (56%-57%). There is a relatively high ratio of donor cases to missing cases, with around 3 donor cases for every missing case. State retirement pension is received by just under a third of all households (31-32%) and also has a high ratio of donor cases to missing cases (around 2:1).
- 6.25 Of all the components received by more than 10% of households, only Council Tax benefit and Housing Benefit have a poor ratio of donor to missing cases. With regards to income from Housing Benefit, around a quarter of cases are imputed where possible by using the data relating to rent levels before and after housing benefit. This is more accurate than imputing using donor data, and leads to a slightly higher ratio of donor to missing cases than given in Table 6-2.

---

<sup>46</sup> It is reassuring to note that the SHS and the SHCS figures for level of receipt of almost all of these benefits are very similar. The only sizeable difference is in receipt of Council Tax Benefit. This is likely to be due to the fact that Council Tax benefit is asked about in different ways in each of the surveys.

- 6.26 Council Tax Benefit has the lowest ratio of donor cases to missing cases. In the 2007 SHS data, there was 3 missing cases for every donor case. In the SHCS, the ratio was even lower, with 5 missing cases for every donor case. As part of the data processing undertaken by the SHCS team at the Scottish Government, Council Tax banding is appended to the SHCS data and used in the imputation procedures. This is not undertaken in the SHS. **Given the power of Council Tax band in modelling Council Tax benefit, combining the datasets would not improve the imputation of this benefit as much as appending Council Tax band to the dataset.**
- 6.27 Combining the imputation of the SHCS data with an annual imputation of the SHS would not provide a higher ratio of donor cases to missing cases overall, but would give a higher ratio of all donor cases to missing SHCS cases. It would, more importantly, provide a larger pool of donor cases overall, with a combined sample roughly five times the size of the quarterly SHS sample. This would allow for more hot-deck classes to be used for imputation, enabling more factors to be built into the imputation process.
- 6.28 Using more detailed hot-deck classes would help improve the robustness of income from earnings more than it would improve the imputation of income from benefits. Household characteristics such as whether the household own a car, whether the property is a 5-room detached house or a 1-bedroom flat, and whether they own their property outright, will be correlated to earnings. These tend not to be correlated to level of benefits received. An annual re-imputation of the SHCS and the SHS combined would therefore improve the imputation of non-benefit data more than benefits data.
- 6.29 Components of income that are received by a small proportion of households are either benefits – such as Attendance Allowance, Severe Disablement Allowance, Statutory Paternity Pay, Industrial Injury Disablement Benefit – or miscellaneous sources of income – such as maintenance payments and mileage allowances. Currently, imputation for components of income that are received by a very small number of cases is carried out by either imputing the median amount or imputing a random cases for all donors. For example, Attendance Allowance is imputed using a single hot-deck group of all donor data. Undertaking an annual combined imputation of the SHS and SHCS would not enable a change of approach to the imputation of income for this benefit. No variables in the SHS or the SHCS help to distinguish between those receiving Attendance Allowance at the higher rate from those receiving it at the lower rate.

- 6.30 **There are no major practical barriers to re-imputing the SHS income together with the SHCS on an annual basis.** Differences in the structure of the data<sup>47</sup> are minimal and would at worst limit the advantages of combining the data.
- 6.31 **The main practical considerations relate to data delivery timings and any guidelines related to re-releasing revised estimates of SHS data.** With regard to timings, currently the SHS data is provided three months after the end of fieldwork on a quarterly basis, while the SHCS is provided six months after the end of fieldwork on an annual basis. Re-imputing the SHS income data on an annual basis, together with the SHCS would extend the timescales for the final annual SHS data by around 3 months. It would not affect the SHCS data delivery timescales. Imputing income for other adults in the household, using FRS data would add a further 12 months to the timescales.
- 6.32 Releasing SHS quarterly estimates of income, revised annual estimates of income, and also further estimates of income using the broader, all adults in the household, definition may not be a preferred option of the Scottish Government. Using SHS data to impute SHCS data, but not undertaking an annual re-imputation of SHS data would not impact the current data delivery timescales for the SHCS.

---

<sup>47</sup> For example, the SHS collected information on whether the household had a computer while the SHCS has never collected this information. This has been used in the imputation of earnings in the SHS.

**Table 6-2: Level of receipt of ratio of donor cases to missing cases in selected components of income. SHS and SHCS 2007. Unweighted.**

| Component of Income           | SHS 2007     |             |         |       | SHS 2007 Q1  |             |         |       | SHCS 2007    |             |         |       |
|-------------------------------|--------------|-------------|---------|-------|--------------|-------------|---------|-------|--------------|-------------|---------|-------|
|                               | %age receive | Donor cases | Missing | Ratio | %age receive | Donor cases | Missing | Ratio | %age receive | Donor cases | Missing | Ratio |
| HiH Income for main job       | 57%          | 7,364       | 2,317   | 3.2   | 56%          | 1,757       | 541     | 3.2   | 56%          | 1,611       | 563     | 2.9   |
| Spouse's income for main jobs | 31%          | 3,883       | 1,363   | 2.8   | 30%          | 920         | 308     | 3.0   | 32%          | 925         | 314     | 2.9   |
| State Retirement Pension      | 30%          | 2,913       | 1,382   | 2.1   | 30%          | 715         | 329     | 2.2   | 31%          | 743         | 383     | 1.9   |
| Child Benefit                 | 26%          | 3,450       | 878     | 3.9   | 26%          | 835         | 206     | 4.1   | 27%          | 818         | 196     | 4.2   |
| Non-state pension             | 20%          | 2,271       | 763     | 3.0   | 20%          | 605         | 97      | 6.2   | 20%          | 487         | 197     | 2.5   |
| Council Tax Benefit           | 19%          | 868         | 2,332   | 0.4   | 19%          | 222         | 550     | 0.4   | 13%          | 84          | 391     | 0.2   |
| Housing Benefit               | 14%          | 1,204       | 1,103   | 1.1   | 14%          | 286         | 275     | 1.0   | 14%          | 288         | 268     | 1.1   |
| Child Tax Credit              | 13%          | 1,331       | 637     | 2.1   | 13%          | 327         | 157     | 2.1   | 13%          | 332         | 161     | 2.1   |
| Investments                   | 9%           | 713         | 757     | 0.9   | 9%           | 219         | 123     | 1.8   | 8%           | 94          | 183     | 0.5   |
| Income Support                | 7%           | 767         | 460     | 1.7   | 7%           | 181         | 100     | 1.8   | 8%           | 194         | 125     | 1.6   |
| DLA                           | 7%           | 220         | 610     | 0.4   | 6%           | 59          | 142     | 0.4   | 6%           | 114         | 126     | 0.9   |
| Working Tax Credit            | 6%           | 358         | 358     | 1.0   | 6%           | 84          | 77      | 1.1   | 6%           | 142         | 83      | 1.7   |
| Pension Credit                | 5%           | 476         | 392     | 1.2   | 5%           | 122         | 88      | 1.4   | 6%           | 111         | 96      | 1.2   |
| Attendance Allowance          | 3%           | 206         | 219     | 0.9   | 3%           | 53          | 50      | 1.1   | 2%           | 47          | 44      | 1.1   |
| Severe Disablement Allowance  | 1%           | 38          | 60      | 0.6   | 1%           | 9           | 14      | 0.6   | 0%           | 8           | 8       | 1.0   |

# Summary and recommendations

## 7. Summary and recommendations

- 7.1 Differences in the estimates of household income between the FRS and the SHS/SHCS are driven primarily by the inclusion of income of 'other adult' – adults who are not the HRP or their spouse – in the FRS. Other methodological and definitional differences have a much smaller effect on the estimate of household income.
- 7.2 Overall, even though the FRS questionnaire is almost twice the length of the SHS, once this difference has been accounted for, the estimates for income levels in Scotland are very similar.
- 7.3 Around half of the variation in the income of other adults in the FRS can be accounted for by variables that are common to both the FRS and the SHS/SHCS. The main drivers of income that could be used in an imputation strategy are economic status and relationship to the HRP. Less variation in income levels within different economic statuses can be explained by variables common to all three surveys.
- 7.4 **It is feasible to impute reasonable approximations of income for other adults in the SHS and SHCS to produce estimates of household income that will be robust at the national and local authority level. Imputation of income of other adults will increase the mean estimate of household income across Scotland by around 9%.**
- 7.5 **The strategy employed to impute income of other adults should be based around economic status.** The economic status variable in the FRS is not consistent with that used in the SHS and SHCS, but an alternative version can be created for use in an imputation strategy.
- 7.6 The imputation strategy should use stochastic methods, namely hot deck hierarchical imputation, in order to maintain the distribution of income levels.
- 7.7 **FRS data for the whole of the UK – with the exception of earnings data in London – should be used as the donor data for the imputation of other adults.** FRS data for Scotland only and the Random Adult income data collected until 2004 in the SHS do not provide enough cases to develop a robust imputation strategy.
- 7.8 There are no major practical barriers to imputing income of other adults in the SHS and the SHCS. However, the need to use case level data from the FRS

would mean that the estimates of the broader definition of household income, based on the imputed data, would only be available one year after the existing estimates are released.

7.9 A number of other recommendations are made to improve the estimates of household income in the SHS and the SHCS.

### Summary of recommended imputation strategy

- Adjust the *net* income of all other adults in the FRS across the UK to take account of the different coverage of benefits in the SHS and the SHCS. This is to ensure that the definition of adult income is consistent within the SHS and SHCS (in other words that HRP/Spouses' income is not based on a different definition from other adults).
- Use the FRS to model the income levels of other adults in order to design a series of donor group typologies and to confirm whether there is evidence of any significant differences between Scotland the rest of the UK. This should be undertaken separately by economic status. Several donor group typologies should be designed for each to ensure that there are no cells where a match can't be made between the FRS and SHS/SHCS. Factors that have the smallest explanatory power should be collapsed or excluded first. Donor groups should be amended each year using FRS data.
- Discard extreme values of income – the highest and lowest 5% - to minimise the potential mismatch between the actual (unknown) value of income and the imputed value, while maintaining the shape of the distribution.
- Create a 'donor data' dataset from the FRS that only includes the income and the various typology variables.
- Recreate the donor typologies in the SHS and the SHCS datasets and link these datasets to the FRS 'donor dataset'.
- Impute all other adults' income in the SHS/SHCS using hierarchical hot decking.
- Create an additional measure of household income in the SHS by summing the existing income measure with the imputed income of other adults in the household. The documentation and naming of the two household income measures would require careful consideration to avoid confusion.

### Summary of potential changes to the SHS and SHCS questionnaires.

7.10 Any questionnaire amendments need to be considered alongside other competing demands for space:



- Supplement the main activity question with the questions required to calculate the ILO definition of economic status.
- Add questions to calculate NS-SEC for all workers in the household.
- Additional checks related to the receipt of benefits
- Amend the questionnaire to ask about income from investments as a separate question rather than as part of the “all other sources” question.
- Add an additional question to ask if households are living off savings or loans.

### **Summary of data processing recommendations**

- Improve the routines used to check benefit levels received by systematically examining over-reported and under-reported estimates against minimum and maximum thresholds.
- Consider imputing receipt of Child Benefit and Winter Fuel Payments for households who are eligible but not reporting receiving these benefits.
- Exclude both high and low outliers of earnings from the imputation of income from earnings of HRPs and spouses.
- Add an additional step to the processing procedures to examine cases where income is less than £25 a week following completion of the imputation processes.
- Consider undertaking an annual re-imputation of the income from the Highest Income Householder and their spouse for the SHS together with the imputation of the SHCS income.

# Appendix 1: Detailed approach to the imputation of income of other adults using hot-deck imputation

# Appendix 1: Detailed approach to the imputation of income of other adults using hot-deck imputation

*A1.1* This appendix provides details of the models of income of other adults by economic status and the suggested donor group typologies for hot-deck imputation, for both the SHS and SHCS. It presents the analysis of the FRS for factors collected in both the FRS and the SHCS. Appendix 2 provides the corresponding analysis for all factors collected in both the FRS and the SHS.

## Retired adults

*A1.2* There are 385 adults in the FRS in 2005-06 who are retired and could potentially be used to impute income in the SHS. There are 226 cases to impute in the SHS for 2005-2006. We would expect around 230 cases in the SHS and around 60 cases in the SHCS to be imputed each year.

*A1.3* There is a relatively low adjusted  $R^2$  of 0.12 for this model. There is not a significant difference between Scotland and the rest of the UK. The three factors that have any significance in modelling pensioner income are age, relationship to the HRP and council tax banding. For imputation in the SHCS, we would suggest the donor groupings structure as follows:

- Level 1<sup>48</sup>: Age (7 bandings shown above), Relationship to HRP (3 bandings: Parent/Child/Other relative or unrelated), Council Tax band (2 bandings A-E, F-H).
- Level 2: Age (4 bandings, collapsing 70-74, 75-79, 80-84 and 85+ into one band), Relationship to HRP (2 bandings: Parent of HRP/not parent of HRP).
- Level 3 Age (4 bandings).

*A1.4* For imputation in the SHS<sup>49</sup>, we would suggest the donor groupings structure as follows:

- Level 1: Age (7 bandings shown above), and Relationship to HRP (3 bandings: Parent/Child/Other relative or unrelated).
- Level 2: Age (4 bandings, collapsing 70-74, 75-79, 80-84 and 85+ into one band), Relationship to HRP (2 bandings: Parent of HRP/not parent of HRP).

<sup>48</sup> The different levels show how the donor groups would be sequentially collapsed.

<sup>49</sup> See Table A2.

- Level 3 Age (4 bandings).

**Table A1.1: Regression model of square root of net adult income of retired adults who are not HRP or their spouse by selected factors collected in both the FRS and SHCS. FRS 2005-06. UK**

|   | Unstandardised co-efficients |     | Sig.    |
|---|------------------------------|-----|---------|
|   | B                            | SE  | P-value |
| (Constant)  | 10.4                         | 1.1 | 0.00    |
| Child   | 2.9                          | 1.2 | 0.02    |
| Other relative                                      | 0.7                          | 0.6 | 0.28    |
| Not related   | 1.7                          | 0.7 | 0.01    |
| Sex (female)  | -0.3                         | 0.4 | 0.53    |
| Aged up to 59                                       | -4.6                         | 1.2 | 0.00    |
| Aged 60-64  | -1.3                         | 1.1 | 0.23    |
| Aged 70-74  | 1.3                          | 0.8 | 0.09    |
| Aged 75-79  | 1.4                          | 0.7 | 0.05    |
| Aged 80-84  | 1.6                          | 0.7 | 0.02    |
| Aged 85+  | 1.3                          | 0.7 | 0.07    |
| In Scotland   | 0.7                          | 0.7 | 0.33    |
| CT Band B   | -0.2                         | 0.6 | 0.80    |
| CT Band C   | 0.1                          | 0.6 | 0.93    |
| CT Band D   | 0.8                          | 0.6 | 0.17    |
| CT Band E   | -0.6                         | 0.7 | 0.42    |
| CT Band F   | -0.0                         | 0.8 | 0.97    |
| CT Band G   | 0.3                          | 1.0 | 0.75    |
| CT Band H   | 9.0                          | 1.7 | 0.00    |
| Mortgage  | -0.5                         | 0.4 | 0.33    |
| Rent  | -0.1                         | 0.6 | 0.78    |
| Household has car/van?                              | -0.1                         | 0.5 | 0.78    |
| Adjusted R <sup>2</sup> = 0.12                      |                              |     |         |
| Constant – Parent of HRP, Owned outright, Age 65-69 |                              |     |         |

A1.5 Table A1.2 shows the pattern of average adult income among retired other adults by age and relationship to HRP for illustration.

**Table A1.2: Mean net adult income of pensioners who are not HRP or their spouse by age and relationship to HRP. FRS 2005-06. UK**

|                      | Parent | Child | Other Relation | Not related | All  | Row Ns |
|----------------------|--------|-------|----------------|-------------|------|--------|
| <b>Aged up to 59</b> | £106   | £82   | £57            | £12         | £77  | 18     |
| <b>Aged 60-64</b>    | £58    | £187  | £129           | £97         | £108 | 18     |
| <b>Aged 65-69</b>    | £112   | £136  | £138           | £171        | £132 | 51     |
| <b>Aged 70-74</b>    | £127   | -     | £162           | £211        | £144 | 54     |
| <b>Aged 75-79</b>    | £139   | -     | £189           | £134        | £146 | 69     |
| <b>Aged 80-84</b>    | £145   | -     | £155           | £161        | £147 | 89     |
| <b>Aged 85+</b>      | £152   | -     | £147           | £153        | £151 | 86     |
| <b>All</b>           | £137   | £121  | £156           | £156        | £141 |        |
| <b>Col Ns</b>        | 248    | 38    | 58             | 41          |      | 385    |

## Students

A1.6 There is data for 1,087 adults in the FRS in 2005-06 who are students and could potentially be used to impute income in the SHS. There are 870 cases in the SHS for 2005-6 that are students and we would expect around 220 cases in the SHCS and around 900 cases in the SHS to be imputed each year.

**Table A1.3: Regression model of square root of net adult income of students who are not HRP or their spouse by selected factors collected in both the FRS and SHCS. FRS 2005-06. UK**

|  | Unstandardised co-efficients |      | Sig.                           |
|--|------------------------------|------|--------------------------------|
|  | B                            | SE   | P-Value                        |
| (Constant)   | 5.8                          | 0.6  | 0.00                           |
| Parent   | -14.6                        | 3.2  | 0.00                           |
| Other relative                                     | 0.2                          | 0.6  | 0.81                           |
| Not related  | 2.3                          | 0.4  | 0.00                           |
| Sex  | 0.2                          | 0.3  | 0.37                           |
| Aged 20-24   | 2.0                          | 0.3  | 0.00                           |
| Aged 25-29   | 2.1                          | 0.5  | 0.00                           |
| Aged 30-34   | 4.4                          | 1.1  | 0.00                           |
| Aged 35-39   | 1.4                          | 1.5  | 0.35                           |
| Aged 40-44   | 5.7                          | 1.8  | 0.00                           |
| Aged 45-49   | 6.0                          | 3.0  | 0.05                           |
| Aged 50+   | 5.5                          | 2.9  | 0.06                           |
| In Scotland  | -0.1                         | 0.4  | 0.87                           |
| CT Band B  | 0.7                          | 0.4  | 0.11                           |
| CT Band C  | 1.1                          | 0.4  | 0.00                           |
| CT Band D  | 1.0                          | 0.4  | 0.02                           |
| CT Band E  | 1.0                          | 0.5  | 0.03                           |
| CT Band F  | 0.9                          | 0.7  | 0.19                           |
| CT Band G  | 0.6                          | 0.7  | 0.40                           |
| CT Band H  | 2.4                          | 1.4  | 0.07                           |
| Mortgage   | 0.7                          | 0.4  | 0.04                           |
| Shared owner                                       | 1.4                          | 1.6  | 0.39                           |
| Rent   | -0.6                         | 0.4  | 0.18                           |
| Rent free  | 0.4                          | 1.1  | 0.69                           |
| Household has car/van?                             | 1.30                         | 0.36 | 0.00                           |
|  |                              |      | Adjusted R <sup>2</sup> = 0.10 |
| Constant – Child of HRP, Owned outright, Age 16-19 |                              |      |                                |

A1.7 There is a low adjusted R<sup>2</sup> of 0.10 among this grouping. There is not significant difference between Scotland and the rest of the UK. The three factors that are significant in modelling student's income are age, relationship to the HRP and council tax banding. For the SHCS imputation, we would suggest structuring the donor grouping as follows:

- Level 1: Age (5 bandings: 16-19, 20-24, 25-29, 30-39, 40+), Relationship to HRP (3 bandings: Child or other relative/Parent/Unrelated), Council Tax band (3 bandings A-C, D-F, G-H) and Access to a car (Yes, No).
- Level 2: Age (5 bandings), Relationship to HRP (3 bandings) and Access to a car (2 bandings)
- Level 3 Age (5 bandings).

A1.8 For the SHS imputation, we would suggest structuring the donor grouping as follows:

- Level 1: Age (5 bandings: 16-19, 20-24, 25-29, 30-39, 40+), and Relationship to HRP (3 bandings: Child or other relative/Parent/Unrelated) and Access to a car (2 bandings: yes, no)
- Level 2: Age (5 bandings: 16-19, 20-24, 25-29, 30+), and Relationship to HRP (3 bandings)
- Level 3 Age (5 bandings).

A1.9 Table A1.4 displays net income of students by these bandings.

**Table 1.4: Mean net adult income of students who are not HRP or their spouse by age, CT band and relationship to HRP. FRS 2005-06. UK**

| Age   | Relationship to HRP   | CT Band |      |       | Total |
|-------|-----------------------|---------|------|-------|-------|
|       |                       | A-C     | D-F  | G & H |       |
| 16-19 | Child, other relative | £88     | £109 | £93   | £95   |
|       | Unrelated             | £103    | £127 | £111  | £108  |
|       | Total                 | £91     | £112 | £93   | £97   |
| 20-24 | Child, other relative | £126    | £169 | £105  | £143  |
|       | Unrelated             | £123    | £142 | £281  | £131  |
|       | Total                 | £125    | £160 | £111  | £139  |
| 25-29 | Child, other relative | £149    | £95  | £90   | £123  |
|       | Unrelated             | £130    | £121 | -     | £127  |
|       | Total                 | £142    | £103 | £90   | £124  |
| 30-34 | Child, other relative | £253    | £159 | -     | £207  |
|       | Unrelated             | £204    | -    | -     | £204  |
|       | Total                 | £230    | £159 | -     | £206  |
| 35-39 | Child, other relative | £228    | £243 | -     | £237  |
|       | Unrelated             | £268    | £66  | -     | £132  |
|       | Total                 | £249    | £139 | -     | £178  |
| 40+   | Child, other relative | £98     | £264 | -     | £220  |
|       | Parent                | £0      | -    | -     | £0    |
|       | Unrelated             | £393    | -    | -     | 393   |
|       | Total                 | £85     | £264 | -     | £189  |
| Total | Child, other relative | £111    | £143 | £96   | £121  |
|       | Parent                | £0      | -    | -     | £0    |
|       | Unrelated             | £122    | £136 | £155  | £127  |
|       | Total                 | £114    | £141 | £98   | £123  |

## Unemployed, looking after home, sick/disabled

A1.10 There is data for 1,065 adults in the FRS in 2005-06 who are unemployed, looking after home or family, long or short term sick or disabled or otherwise economically inactive that could potentially be used to impute income in the SHS. There are 577 cases to impute in the SHS and we would expect around 150 cases in the SHCS and around 600 cases in the SHS to be imputed each year.

A1.11 Table A1.5 shows the mean income level across these various groupings. Income levels are highest among the long term sick and lowest among the unemployed and 'other inactive' groupings. It should be noted that the categories used for main activity in the SHS and SHCS are slightly different from the FRS, in that they do not capture those who are short term sick or have an 'other inactive' category. It is arguable whether some respondents who are classified as 'other inactive' in the FRS would classify themselves as 'looking after home/family' in the SHS/SHCS.

**Table A1.5: Net adult income of other adults who are unemployed, looking after home and sick/disabled who are not HRP or spouse by detailed economic status. FRS 2005-06, UK.**

|                                  |        | Other parts of UK | Scotland | Total | Total N |
|----------------------------------|--------|-------------------|----------|-------|---------|
| <b>Unemployed</b>                | Mean   | £33               | £36      | £33   | 458     |
|                                  | SD     | £50               | £41      | £49   |         |
|                                  | Median | £1                | £33      | £2    |         |
| <b>Looking after home/family</b> | Mean   | £70               | £127     | £71   | 65      |
|                                  | SD     | £58               | £95      | £59   |         |
|                                  | Median | £77               | £145     | £78   |         |
| <b>Long term sick</b>            | Mean   | £97               | £105     | £98   | 324     |
|                                  | SD     | £61               | £56      | £61   |         |
|                                  | Median | £93               | £93      | £93   |         |
| <b>Short term sick</b>           | Mean   | £40               | £70      | £41   | 29      |
|                                  | SD     | £41               | £93      | £41   |         |
|                                  | Median | £44               | £91      | £44   |         |
| <b>Other inactive</b>            | Mean   | £27               | £40      | £27   | 189     |
|                                  | SD     | £47               | £61      | £48   |         |
|                                  | Median | £0                | £0       | £0    |         |
| <b>Total</b>                     | Mean   | £51               | £58      | £52   | 1065    |
|                                  | SD     | £61               | £58      | £60   |         |
|                                  | Median | £44               | £45      | £44   |         |

A1.12 Modelling the income of the unemployed, long-term sick and disabled and those looking after home or family, the potential imputation factors give an adjusted R<sup>2</sup> of 0.39 overall (Table A1.6 overleaf). There is not a significant difference

between Scotland and the rest of the UK, although the p-value for this component suggests that it is close to being significant. The three factors that are most significant are age, whether related to the HRP, and whether unemployed, sick, looking after home or family, or otherwise inactive.

**Table A1.6: Regression model of square root of net adult income of other adults who are unemployed, looking after home and sick/disabled who are not HRP or spouse by factors common to both FRS and SHCS and detailed economic status. FRS 2005-06. UK**

|   | Unstandardised Coefficients |     | Sig.    |
|---|-----------------------------|-----|---------|
|   | B                           | SE  | P-value |
| (Constant)  | 2.4                         | 0.5 | 0.00    |
| Parent  | 0.3                         | 0.8 | 0.76    |
| Other relative  | -0.4                        | 0.4 | 0.36    |
| Not related   | 1.8                         | 0.4 | 0.00    |
| Sex   | 0.2                         | 0.2 | 0.44    |
| Aged 20-24  | 1.9                         | 0.3 | 0.00    |
| Aged 25-29  | 2.9                         | 0.4 | 0.00    |
| Aged 30-34  | 3.2                         | 0.5 | 0.00    |
| Aged 35-39  | 5.2                         | 0.6 | 0.00    |
| Aged 40-44  | 3.4                         | 0.6 | 0.00    |
| Aged 45-49  | 4.0                         | 0.7 | 0.00    |
| Aged 50-54  | 5.9                         | 0.7 | 0.00    |
| Aged 55-59  | 4.7                         | 0.7 | 0.00    |
| Aged 60+  | 5.3                         | 0.8 | 0.00    |
| Scottish survey   | 0.8                         | 0.4 | 0.05    |
| CT Band B   | -0.9                        | 0.3 | 0.00    |
| CT Band C   | -0.4                        | 0.3 | 0.14    |
| CT Band D   | -0.2                        | 0.4 | 0.64    |
| CT Band E   | -0.7                        | 0.5 | 0.12    |
| CT Band F   | -0.8                        | 0.6 | 0.17    |
| CT Band G   | -2.6                        | 0.8 | 0.00    |
| Mortgage  | -0.1                        | 0.3 | 0.84    |
| Shared owner  | -2.2                        | 3.2 | 0.48    |
| Rent  | 0.5                         | 0.3 | 0.08    |
| Rent free   | 3.0                         | 2.1 | 0.16    |
| Looking after home/family                                     | 2.7                         | 0.5 | 0.00    |
| Long term sick  | 3.4                         | 0.3 | 0.00    |
| Short term sick   | 0.8                         | 0.7 | 0.23    |
| Other inactive  | -1.2                        | 0.3 | 0.00    |
| Household has car/van?  | -0.2                        | 0.3 | 0.39    |
| Adjusted R <sup>2</sup> = 0.39                                |                             |     |         |
| Constant: Child, 16-19, Unemployed, CT band A, Owner occupied |                             |     |         |

A1.13 We would suggest the structuring the donor grouping for the SHS and the SHCS imputation as follows:



- Level 1: Detailed economic status (5 bandings: unemployed, long-term sick, sick-term sick, looking after home/family and other inactive) Age (6 bandings: 16-19, 20-29, 30-39, 40-49,50-59,60+), Relationship to HRP (2 bandings: Child/Parent/Other relative against unrelated)
- Level 2: Age (4 bandings: 16-24, 25-39, 40-59, 60+), and detailed economic status (3 bandings).
- Level 3 Age (2 bandings: 16-24<sup>50</sup>, 25+) and detailed economic status (3 bandings).

A1.14 Table A1.7 shows the mean income of other adults by these bandings.

**Table A1.7: Mean net adult income of adults who are unemployed, looking after home and long-term sick/disabled who are not HRP or spouse by relationship to HRP and age. FRS 2005-06, UK.**

| Economic Status                | Age          | Relationship to HRP           |             |
|--------------------------------|--------------|-------------------------------|-------------|
|                                |              | Child, Parent, other relative | Unrelated   |
| Unemployed                     | 16-19        | £16                           | £75         |
|                                | 20-24        | £30                           | £70         |
|                                | 25-29        | £57                           | £101        |
|                                | 30-34        | £36                           | £44         |
|                                | 35-39        | £39                           | £346        |
|                                | 40-44        | £44                           | £80         |
|                                | 45-49        | £49                           | -           |
|                                | 50-54        | £53                           | -           |
|                                | 55-59        | £25                           | -           |
|                                | 60+          | £193                          | -           |
|                                | <b>Total</b> | <b>£27</b>                    | <b>£84</b>  |
| Looking after home/<br>family  | 16-19        | £66                           | -           |
|                                | 20-24        | £70                           | £66         |
|                                | 25-29        | £99                           | £4          |
|                                | 30-34        | £60                           | -           |
|                                | 35-39        | £173                          | -           |
|                                | 40-44        | £39                           | -           |
|                                | 55-59        | £80                           | -           |
|                                | 60+          | £65                           | -           |
|                                | <b>Total</b> | <b>£73</b>                    | <b>£53</b>  |
| Long term sick and<br>Disabled | 16-19        | £48                           | £64         |
|                                | 20-24        | £80                           | £162        |
|                                | 25-29        | £85                           | £131        |
|                                | 30-34        | £112                          | £150        |
|                                | 35-39        | £108                          | £245        |
|                                | 40-44        | £106                          | £72         |
|                                | 45-49        | £89                           | -           |
|                                | 50-54        | £133                          | £142        |
|                                | 55-59        | £119                          | £86         |
|                                | 60+          | £115                          | £138        |
|                                | <b>Total</b> | <b>£95</b>                    | <b>£132</b> |

<sup>50</sup> The majority of these cases are aged 16-24.

## Full-time workers

A1.15 There is data for 2,265 adults in the FRS in 2005-06 who are employed full-time and could potentially be used to impute income in 1,594 cases in the SHS. We would expect around 400 cases in the SHCS and 1,600 cases in the SHS to be imputed each year.

A1.16 Unlike the previous models there is a significant difference between income levels in Scotland and in the rest of the UK among other adults who are full time workers.

**Table A1.8: Regression model of square root of net adult income of adults in full-time work who are not HRP or spouse by region. FRS 2005-06. UK**

| Region                              | Mean        | SD          | SE of Mean  | Median      | Count        |
|-------------------------------------|-------------|-------------|-------------|-------------|--------------|
| North East                          | £222        | £113        | £11         | £211        | 82           |
| North West and Merseyside           | £234        | £167        | £9          | £218        | 227          |
| Yorks and Humberside                | £207        | £84         | £6          | £198        | 170          |
| East Midlands                       | £221        | £90         | £7          | £210        | 130          |
| West Midlands                       | £213        | £96         | £6          | £202        | 194          |
| Eastern                             | £241        | £113        | £7          | £210        | 172          |
| <b>London</b>                       | <b>£300</b> | <b>£126</b> | <b>£6</b>   | <b>£274</b> | <b>290</b>   |
| South East                          | £249        | £113        | £222        | £113        | 283          |
| South West                          | £236        | £95         | £234        | £167        | 127          |
| Wales                               | £207        | £97         | £207        | £84         | 92           |
| Northern Ireland                    | £222        | £98         | £213        | £96         | 186          |
| <b>UK (exc Scotland)</b>            | <b>£241</b> | <b>£120</b> | <b>£2</b>   | <b>£221</b> | <b>1,953</b> |
| <b>UK (exc Scotland and London)</b> | <b>£229</b> | <b>£114</b> | <b>£2</b>   | <b>£210</b> | <b>1,663</b> |
| <b>Scotland</b>                     | <b>£215</b> | <b>£100</b> | <b>£221</b> | <b>£90</b>  | <b>312</b>   |
| Total                               | £239        | £118        | £241        | £113        | 2265         |

A1.17 Table A1.8 shows the difference by region. While the average income in Scotland among full-time workers was £215, across the rest of the UK the average income was £241. Earnings in London are out of line with the rest of the UK. Excluding London, the average income in the rest of the UK is £229. When data from London is excluded from the model, the difference between Scotland and the rest of the UK, after controlling for the other factors, is not significant (Table A1.9). This suggests that earnings data from London should be excluded from the imputation. This would reduce the number of potential donor cases to 1,975.

**Table A1.9: Regression model of square root of net adult income of adults in full-time work who are not HRP or spouse by factors common to FRS and SHCS. FRS 2005-06, UK excluding London.**

|                        | Unstandardized Coefficients |     | Sig.   |
|------------------------|-----------------------------|-----|--|
|                        | B                           | SE  | P-value                                      |
| (Constant)             | 11.9                        | 0.4 | 0.00   |
| Parent                 | 0.1                         | 1.0 | 0.92   |
| Other relation         | 0.4                         | 0.3 | 0.26   |
| Not related            | 1.2                         | 0.3 | 0.00   |
| Sex (Female vrs Male)  | -0.3                        | 0.1 | 0.02   |
| Aged 20-24             | 1.9                         | 0.2 | 0.00   |
| Aged 25-29             | 3.0                         | 0.2 | 0.00   |
| Aged 30-34             | 3.5                         | 0.3 | 0.00   |
| Aged 35-39             | 3.5                         | 0.4 | 0.00   |
| Aged 40-44             | 2.7                         | 0.4 | 0.00   |
| Aged 45-49             | 2.8                         | 0.5 | 0.00   |
| Aged 50-54             | 4.3                         | 0.6 | 0.00   |
| Aged 55-59             | 2.6                         | 0.8 | 0.00   |
| Aged 60-up             | 1.7                         | 1.0 | 0.10   |
| Scottish survey        | -0.3                        | 0.2 | 0.24   |
| CT Band B              | 0.3                         | 0.2 | 0.17   |
| CT Band C              | 0.7                         | 0.2 | 0.00   |
| CT Band D              | 0.8                         | 0.2 | 0.00   |
| CT Band E              | 0.5                         | 0.3 | 0.06   |
| CT Band F              | 1.5                         | 0.3 | 0.00   |
| CT Band G              | 0.6                         | 0.5 | 0.24   |
| CT Band H              | 5.5                         | 1.2 | 0.00   |
| Mortgage               | -0.2                        | 0.2 | 0.35   |
| Rent                   | -0.6                        | 0.2 | 0.00   |
| Rent free              | 0.5                         | 0.7 | 0.49   |
| Household has car/van? | 1.0                         | 0.2 | 0.00   |
|                        |                             |     | Adjusted R <sup>2</sup> = 0.16               |
|                        |                             |     | Constant = Child, Aged 16-19, Owned Outright |

A1.18 The adjusted R<sup>2</sup> of the model of income of full-time workers shown in Table A1.9 is 0.16. Given that full-time workers account for nearly half of all other adults, it is important to maximise the power of the model as far as possible. A number of other variables were included in the model to try to improve the model of income of full-time workers: number of bedrooms, whether the HRP is of retirement age, sick or unemployed, and the English urban/rural classification. None of these factors proved to be significant. Indeed, when included as the only variable, the urban/rural classification resulted in a model with an adjusted R<sup>2</sup> of 0.008 – in other words it helps to explain less than one percent of the variance.

A1.19 Given that the average income of full-time workers in Scotland is close to that in the rest of the UK once London is excluded, the suggested model of imputation

will not introduce systematic bias to the imputed data due to regional variations. However, neither will it account for such variation.

A1.20 Age, sex, whether related to the HRP, Council Tax band, tenure and whether have access to a car all help to model income of full-time workers. For the SHCS imputation, we would propose structuring the donor grouping as follows:

- Level 1: Age (6 bandings: 16-19, 20-29, 30-39, 40-49, 50-59, 60+), Sex (2 bandings), Relationship to HRP (2 bandings: Child/Parent/Other relative against unrelated), Council Tax band (3 bandings: A-C, D-F, G-H), Tenure (2 bandings: Rent vrs other) and access to a car (2 bandings)
- Level 2: Age (4 bandings: 16-24, 25-39, 40-59, 60+), Sex (2 bandings), Relationship to HRP (2 bandings) and Council Tax band (3 bandings)
- Level 3: Age (4 bandings) and Sex (2 bandings).

A1.21 For the SHS imputation, we would propose structuring the donor grouping as follows:

- Level 1: Age (6 bandings: 16-19, 20-29, 30-39, 40-49, 50-59, 60+), Sex (2 bandings), Relationship to HRP (2 bandings: Child/Parent/Other relative against unrelated), Tenure (2 bandings: Rent vrs other) and access to a car (2 bandings)
- Level 2: Age (4 bandings: 16-24, 25-39, 40-59, 60+), Sex (2 bandings), Relationship to HRP (2 bandings) and access to a car (2 bandings)
- Level 3: Age (4 bandings) and Sex (2 bandings)

A1.22 Table A1.10 illustrates the pattern of mean incomes of adults employed full-time by age, sex and council tax bands.

**Table A1.10: Mean net adult income of adults employed full-time who are not HRP or spouse by age, Council Tax band and sex. FRS 2005-06. UK excluding London**

|              | Male                |             |             |             | Female              |             |             |             | Total               |             |             |             |
|--------------|---------------------|-------------|-------------|-------------|---------------------|-------------|-------------|-------------|---------------------|-------------|-------------|-------------|
|              | Council Tax Banding |             |             |             | Council Tax Banding |             |             |             | Council Tax Banding |             |             |             |
|              | A-C                 | D-F         | G-H         | Total       | A-C                 | D-F         | G-H         | Total       | A-C                 | D-F         | G-H         | Total       |
| 16-19        | £171                | £167        | £146        | <b>£169</b> | £145                | £172        | £206        | <b>£154</b> | £162                | £169        | £179        | <b>£164</b> |
| 20-24        | £231                | £230        | £213        | <b>£230</b> | £197                | £234        | £213        | <b>£211</b> | £218                | £232        | £213        | <b>£223</b> |
| 25-29        | £238                | £284        | £321        | <b>£256</b> | £255                | £285        | £208        | <b>£266</b> | £243                | £284        | £272        | <b>£259</b> |
| 30-34        | £272                | £279        | £468        | <b>£281</b> | £264                | £275        | £447        | <b>£272</b> | £270                | £279        | £463        | <b>£279</b> |
| 35-39        | £256                | £334        | £467        | <b>£282</b> | £242                | £277        | -           | <b>£254</b> | £253                | £317        | £467        | <b>£276</b> |
| 40-44        | £251                | £268        | -           | <b>£254</b> | £201                | £209        | -           | <b>£203</b> | £244                | £258        | -           | <b>£247</b> |
| 45-49        | £239                | £262        | -           | <b>£249</b> | £233                | £450        | £1,056      | <b>£348</b> | £237                | £275        | £1,056      | <b>£277</b> |
| 50-54        | £299                | £315        | -           | <b>£306</b> | £225                | £348        | £556        | <b>£315</b> | £269                | £333        | £556        | <b>£311</b> |
| 55-59        | £265                | £308        | -           | <b>£274</b> | £226                | £255        | -           | <b>£239</b> | £251                | £274        | -           | <b>£258</b> |
| 60+          | £212                | £295        | -           | <b>£265</b> | £169                | £220        | -           | <b>£211</b> | £203                | £265        | -           | <b>£247</b> |
| <b>Total</b> | <b>£226</b>         | <b>£242</b> | <b>£281</b> | <b>£232</b> | <b>£200</b>         | <b>£243</b> | <b>£268</b> | <b>£218</b> | <b>£217</b>         | <b>£242</b> | <b>£275</b> | <b>£227</b> |

## Part-time workers

A1.23 There is data for 403 other adults in the FRS in 2005-06 who are employed part-time, that could be used to impute 239 cases in the SHS. We would expect around 60 cases in the SHCS and around 240 cases in the SHS to be imputed each year.

A1.24 While the mean level of income in Scotland among part-time workers does not differ significantly from the rest of the UK, the level of income in London remains an obvious outlier (Table A1.11). We would again suggest that data from London is excluded from the imputation processes, reducing the number of potential donor cases to 366.

**Table A1.11: Net adult income of all part-time employed adults who are not HRP or Spouse by region. FRS 2005-2006, UK.**

| Region                              | Mean        | SD         | SE of mean | Median      | Count      |
|-------------------------------------|-------------|------------|------------|-------------|------------|
| North East                          | £116        | £61        | £11        | £122        | 24         |
| North West and Merseyside           | £130        | £78        | £10        | £119        | 49         |
| Yorks and Humberside                | £136        | £76        | £13        | £115        | 27         |
| East Midlands                       | £119        | £56        | £12        | £112        | 18         |
| West Midlands                       | £99         | £45        | £6         | £95         | 36         |
| Eastern                             | £124        | £46        | £9         | £118        | 22         |
| <b>London</b>                       | <b>£159</b> | <b>£98</b> | <b>£12</b> | <b>£142</b> | <b>37</b>  |
| South East                          | £132        | £70        | £8         | £122        | 53         |
| South West                          | £129        | £61        | £8         | £126        | 34         |
| Wales                               | £122        | £68        | £12        | £111        | 23         |
| Northern Ireland                    | £122        | £54        | £16        | £113        | 22         |
| <b>UK (exc Scotland)</b>            | <b>£129</b> | <b>£71</b> | <b>£3</b>  | <b>£119</b> | <b>345</b> |
| <b>UK (exc Scotland and London)</b> | <b>£124</b> | <b>£65</b> | <b>£3</b>  | <b>£115</b> | <b>308</b> |
| <b>Scotland</b>                     | <b>£128</b> | <b>£81</b> | <b>£13</b> | <b>£106</b> | <b>58</b>  |
| Total                               | £129        | £72        | £3         | £118        | 403        |

A1.25 The model of income levels of part-time workers explains a relatively small amount of the variation within this group, having a low adjusted  $R^2$  of 0.14. The three factors that have any significance in modelling part-time workers income are age, tenure and council tax banding. For the SHCS imputation, we would suggest the donor groupings structure as follows:

- Level 1 Age (6 bandings: 16-19, 20-29, 30-39, 40-49, 50-59, 60+), Tenure (2 bandings: Renting versus not renting), and Council Tax band (3 bandings: A-C, D-F, G-H).
- Level 2: Age (6 bandings) and Tenure (2 bandings).
- Level 3 Age (5 bandings, collapsing 16-29).

**Table A1.12: Regression model of square root of net adult income of adults in part-time work who are not HRP or spouse by factors common to FRS and SHCS. FRS 2005-06, UK excluding London.**

|  | Unstandardized Coefficients |     | Sig.                           |
|--|-----------------------------|-----|--------------------------------|
|  | B                           | SE  | P-value                        |
| (Constant)                                   | 9.6                         | 0.6 | 0.00                           |
| Parent                                       | -1.9                        | 1.3 | 0.14                           |
| Other relation                               | 0.6                         | 1.0 | 0.53                           |
| Not related                                  | 0.4                         | 0.6 | 0.51                           |
| Sex (Female vrs Male)                        | 0.0                         | 0.3 | 0.92                           |
| Aged 20-24                                   | 0.9                         | 0.3 | 0.01                           |
| Aged 25-29                                   | 2.4                         | 0.6 | 0.00                           |
| Aged 30-34                                   | 2.1                         | 0.7 | 0.00                           |
| Aged 35-39                                   | 1.9                         | 0.9 | 0.03                           |
| Aged 40-44                                   | 1.6                         | 1.1 | 0.14                           |
| Aged 45-49                                   | 1.9                         | 1.5 | 0.20                           |
| Aged 50-54                                   | 1.6                         | 1.3 | 0.21                           |
| Aged 55-59                                   | 2.9                         | 1.2 | 0.02                           |
| Aged 60-up                                   | 3.5                         | 1.0 | 0.00                           |
| Scottish survey                              | -0.1                        | 0.5 | 0.88                           |
| CT Band B                                    | 0.6                         | 0.4 | 0.12                           |
| CT Band C                                    | 1.4                         | 0.4 | 0.00                           |
| CT Band D                                    | 1.4                         | 0.5 | 0.01                           |
| CT Band E                                    | 0.3                         | 0.5 | 0.60                           |
| CT Band F                                    | 1.7                         | 0.7 | 0.01                           |
| CT Band G                                    | 0.6                         | 1.0 | 0.50                           |
| CT Band H                                    | -8.2                        | 2.4 | 0.00                           |
| Mortgage                                     | -0.4                        | 0.4 | 0.24                           |
| Shared owners                                | -3.2                        | 2.1 | 0.12                           |
| Rent   | -1.1                        | 0.4 | 0.02                           |
| Rent free                                    | 0.4                         | 1.3 | 0.78                           |
| Household has car/van?                       | 0.0                         | 0.4 | 0.97                           |
|  |                             |     | Adjusted R <sup>2</sup> = 0.14 |
| Constant = Child, Aged 16-19, Owned Outright |                             |     |                                |

A1.26 For the SHS, we would suggest the following donor groupings structure:

- Level 1 Age (6 bandings: 16-19, 20-29, 30-39, 40-49, 50-59, 60+), Tenure (2 bandings: Renting versus not renting) and access to a car (2 bandings: Yes/No)
- Level 2: Age (6 bandings) and Tenure (2 bandings).
- Level 3 Age (5 bandings, collapsing 16-29).

**Table A1.13: Mean net adult household income of adults employed part-time who are not HRP or spouse by age and Council Tax band. FRS 2005-06, UK excluding London.**

| Age band     | Council Tax Banding |             |             |             |
|--------------|---------------------|-------------|-------------|-------------|
|              | A-C                 | D-F         | G & H       | Total       |
| 16-19        | £105                | £100        | £54         | £102        |
| 20-24        | £119                | £147        | £117        | £126        |
| 25-29        | £145                | £183        | -           | £159        |
| 30-34        | £156                | £229        | -           | £159        |
| 35-39        | £123                | £191        | £199        | £148        |
| 40-44        | £149                | £158        | -           | £154        |
| 45-49        | £93                 | £304        | -           | £122        |
| 50-54        | £137                | -           | -           | £137        |
| 55-59        | £116                | £244        | -           | £140        |
| 60+          | £142                | £235        | -           | £180        |
| <b>Total</b> | <b>£119</b>         | <b>£140</b> | <b>£103</b> | <b>£124</b> |

## Self-employed

A1.27 There is data for 190 other adults in the FRS data for 2005-06 who are self-employed and could be used to impute the 69 cases in the SHS. We would expect less than 20 cases in the SHCS and around 70 cases in the SHS to be imputed each year.

A1.28 As Table A1.14 shows income in London is atypical and should be excluded, reducing the number of potential donor cases to 160.

**Table A1.14: Net adult income of all self-employed adults who are not HRP or Spouse by region. FRS 2005-2006 UK**

| Region                              | Mean        | SD          | SE of Mean | Median      | Count      |
|-------------------------------------|-------------|-------------|------------|-------------|------------|
| North East                          | £156        | £128        | £77        | £172        | 2          |
| North West and Merseyside           | £246        | £146        | £34        | £250        | 12         |
| Yorks and Humberside                | £199        | £127        | £29        | £248        | 14         |
| East Midlands                       | £148        | £118        | £29        | £123        | 13         |
| West Midlands                       | £192        | £127        | £23        | £180        | 21         |
| Eastern                             | £215        | £235        | £57        | £124        | 13         |
| <b>London</b>                       | <b>£326</b> | <b>£265</b> | <b>£37</b> | <b>£294</b> | <b>30</b>  |
| South East                          | £176        | £123        | £23        | £175        | 19         |
| South West                          | £157        | £138        | £31        | £108        | 13         |
| Wales                               | £198        | £123        | £29        | £200        | 11         |
| Northern Ireland                    | £211        | £258        | £79        | £181        | 21         |
| <b>UK (exc Scotland)</b>            | <b>£221</b> | <b>£190</b> | <b>£13</b> | <b>£190</b> | <b>169</b> |
| <b>UK (exc Scotland and London)</b> | <b>£191</b> | <b>£151</b> | <b>£11</b> | <b>£176</b> | <b>139</b> |
| <b>Scotland</b>                     | <b>£182</b> | <b>£165</b> | <b>£44</b> | <b>£119</b> | <b>21</b>  |
| <b>All UK</b>                       | <b>£219</b> | <b>£188</b> | <b>£12</b> | <b>£189</b> | <b>190</b> |

A1.29 The model of income levels of self-employed workers explains a small amount of the variation within this group, having a low adjusted  $R^2$  of 0.12 (Table A1.15). The four factors that are significant in modelling self-employed income are age, sex, tenure and access to a car/van. We would suggest the donor groupings structure as follows in both the SHS and the SHCS:

- Level 1 Age (5 bandings: 16-29, 30-39, 40-49, 50-59, 60+), Sex (2 bandings), Access to a car (2 bandings) and Tenure (2 bandings: Owning outright against other tenures).
- Level 2: Age (5 bandings), Access to a car (2 bandings) and Tenure (2 bandings).
- Level 3 Age (5 bandings) and Access to a car (2 bandings).

**Table A1.15: Regression model of square root of net adult income of self-employed adults who are not HRP or spouse by factors common to FRS and SHCS. FRS 2005-06, UK excluding London.**

|   | Unstandardized Coefficients |     | Sig.    |
|---|-----------------------------|-----|---------|
|   | B                           | SE  | P-Value |
| (Constant)  | 11.9                        | 2.7 | 0.00    |
| Parent  | 2.6                         | 3.6 | 0.47    |
| Other relation  | 0.6                         | 1.9 | 0.74    |
| Not related   | -0.9                        | 1.4 | 0.53    |
| Sex (Female vrs Male)   | -2.3                        | 1.3 | 0.07    |
| Aged 20-24  | 2.5                         | 1.6 | 0.13    |
| Aged 25-29  | 3.1                         | 1.7 | 0.08    |
| Aged 30-34  | 0.6                         | 1.8 | 0.75    |
| Aged 35-39  | 5.9                         | 2.6 | 0.03    |
| Aged 40-44  | 0.2                         | 2.5 | 0.93    |
| Aged 45-49  | 5.2                         | 3.1 | 0.09    |
| Aged 50-54  | -2.1                        | 2.8 | 0.45    |
| Aged 55-59  | 1.7                         | 3.1 | 0.59    |
| Aged 60-up  | 0.7                         | 3.6 | 0.83    |
| Scottish survey   | -0.7                        | 1.6 | 0.64    |
| CT Band B   | 0.6                         | 1.4 | 0.66    |
| CT Band C   | -0.9                        | 1.2 | 0.45    |
| CT Band D   | -2.5                        | 1.3 | 0.07    |
| CT Band E   | -2.9                        | 1.7 | 0.09    |
| CT Band F   | -1.2                        | 2.7 | 0.65    |
| CT Band G   | 3.7                         | 5.4 | 0.49    |
| CT Band H   | -0.3                        | 3.0 | 0.93    |
| Mortgage  | -2.2                        | 1.0 | 0.04    |
| Rent  | -1.8                        | 1.4 | 0.20    |
| Rent free   | -10.6                       | 4.6 | 0.02    |
| Household has car/van?  | 4.1                         | 1.8 | 0.02    |
| Constant = Child, Aged 16-19, Owned Outright<br>Adjusted $R^2$ = 0.12 |                             |     |         |



**Table A1.16: Mean net adult household income of self-employed adults who are not HRP or spouse by sex, age and household access to a car. FRS 2005-06. UK excluding London**

| Age band     | Male                   |             |             | Female                 |             |             | Total                  |             |             |
|--------------|------------------------|-------------|-------------|------------------------|-------------|-------------|------------------------|-------------|-------------|
|              | Household has car/van? |             |             | Household has car/van? |             |             | Household has car/van? |             |             |
|              | No                     | Yes         | Total       | No                     | Yes         | Total       | No                     | Yes         | Total       |
| <b>16-19</b> | £3                     | £170        | <b>£142</b> | -                      | £117        | <b>£117</b> | £3                     | £154        | <b>£135</b> |
| <b>20-24</b> | £98                    | £209        | <b>£189</b> | £61                    | £176        | <b>£158</b> | £91                    | £201        | <b>£182</b> |
| <b>25-29</b> | £148                   | £206        | <b>£206</b> | -                      | £145        | <b>£145</b> | £148                   | £200        | <b>£199</b> |
| <b>30-34</b> | £330                   | £173        | <b>£175</b> | £120                   | £52         | <b>£60</b>  | £191                   | £154        | <b>£155</b> |
| <b>35-39</b> | -                      | £338        | <b>£338</b> | -                      | -           | <b>-</b>    | -                      | £338        | <b>£338</b> |
| <b>40-44</b> | £56                    | £204        | <b>£195</b> | -                      | £67         | <b>£67</b>  | £56                    | £187        | <b>£180</b> |
| <b>45-49</b> | £189                   | £359        | <b>£309</b> | £23                    | -           | <b>£23</b>  | £158                   | £359        | <b>£291</b> |
| <b>50-54</b> | -                      | £119        | <b>£119</b> | -                      | £0          | <b>£0</b>   | -                      | £98         | <b>£98</b>  |
| <b>55-59</b> | £82                    | £382        | <b>£279</b> | -                      | £154        | <b>£154</b> | £82                    | £267        | <b>£229</b> |
| <b>60+</b>   | -                      | £186        | <b>£186</b> | -                      | £207        | <b>£207</b> | -                      | £197        | <b>£197</b> |
| <b>Total</b> | <b>£91</b>             | <b>£208</b> | <b>£196</b> | <b>£68</b>             | <b>£139</b> | <b>£133</b> | <b>£87</b>             | <b>£195</b> | <b>£184</b> |
| <b>N</b>     | 13                     | 116         | <b>129</b>  | 2                      | 29          | <b>31</b>   | 15                     | 145         | <b>160</b>  |

# Appendix 2: Additional modelling tables for the SHS

## Appendix 2: Additional modelling tables for the SHS

**Table A2.1 (corresponds to Table 4-5): Regression model of square root of net adult income of other adults by factors collected in both the FRS and SHS. FRS 2005-06. UK**

|   | Unstandardized Coefficients |     | Sig.    |
|---|-----------------------------|-----|---------|
|   | B                           | SE  | P-Value |
| (Constant)  | 12.4                        | 0.2 | 0.00    |
| Parent  | -1.1                        | 0.4 | 0.01    |
| Other relative  | -0.1                        | 0.2 | 0.77    |
| Not related   | 1.2                         | 0.2 | 0.00    |
| PT worker   | -3.4                        | 0.2 | 0.00    |
| Self-employed   | -2.3                        | 0.3 | 0.00    |
| Unemployed etc  | -9.2                        | 0.1 | 0.00    |
| Retired   | -7.1                        | 0.6 | 0.00    |
| Student   | -4.3                        | 0.1 | 0.00    |
| Sex (female)  | 0.0                         | 0.1 | 0.89    |
| Aged 20-24  | 2.1                         | 0.1 | 0.00    |
| Aged 25-29  | 3.2                         | 0.2 | 0.00    |
| Aged 30-34  | 3.7                         | 0.2 | 0.00    |
| Aged 35-39  | 4.5                         | 0.3 | 0.00    |
| Aged 40-44  | 3.6                         | 0.3 | 0.00    |
| Aged 45-49  | 4.3                         | 0.4 | 0.00    |
| Aged 50-54  | 4.8                         | 0.5 | 0.00    |
| Aged 55-59  | 4.3                         | 0.5 | 0.00    |
| Aged 60-64  | 4.8                         | 0.6 | 0.00    |
| Aged 65-69  | 5.4                         | 0.7 | 0.00    |
| Aged 70-74  | 6.2                         | 0.8 | 0.00    |
| Aged 75-79  | 6.7                         | 0.8 | 0.00    |
| Aged 80-84  | 6.8                         | 0.8 | 0.00    |
| Aged 85+  | 6.9                         | 0.8 | 0.00    |
| Scottish survey   | 0.0                         | 0.2 | 0.93    |
| Mortgage  | -0.2                        | 0.1 | 0.09    |
| Shared owner  | 0.2                         | 1.0 | 0.81    |
| Rent  | -0.4                        | 0.2 | 0.00    |
| Rent free   | -0.3                        | 0.5 | 0.57    |
| Household has car/van?  | 0.4                         | 0.1 | 0.00    |
| Adjusted R <sup>2</sup> of Model = 0.48                       |                             |     |         |
| Constant – Child, 16-19, in FT work, owning property outright |                             |     |         |

**Table A2.2 (corresponds to Table A1.1): Regression model of square root of net adult income of retired adults who are not HRP or their spouse by selected factors collected in both the FRS and SHS. FRS 2005-06. UK**

|   | Unstandardised co-efficients |     | Sig.                           |
|---|------------------------------|-----|--------------------------------|
|   | B                            | SE  | P-value                        |
| (Constant)  | 10.4                         | 1.1 | 0.00                           |
| Child   | 3.0                          | 1.2 | 0.02                           |
| Other relative                                      | 0.7                          | 0.6 | 0.26                           |
| Not related   | 1.5                          | 0.7 | 0.03                           |
| Sex (female)  | -0.1                         | 0.4 | 0.77                           |
| Aged up to 59                                       | -4.4                         | 1.3 | 0.00                           |
| Aged 60-64  | -1.3                         | 1.1 | 0.23                           |
| Aged 70-74  | 1.2                          | 0.8 | 0.11                           |
| Aged 75-79  | 1.5                          | 0.7 | 0.04                           |
| Aged 80-84  | 1.6                          | 0.7 | 0.02                           |
| Aged 85+  | 1.6                          | 0.7 | 0.02                           |
| In Scotland   | 0.5                          | 0.7 | 0.46                           |
| Mortgage  | -0.4                         | 0.5 | 0.40                           |
| Rent  | -0.3                         | 0.6 | 0.58                           |
| Household has car/van?                              | -0.3                         | 0.5 | 0.55                           |
|   |                              |     | Adjusted R <sup>2</sup> = 0.06 |
| Constant – Parent of HRP, Owned outright, Age 65-69 |                              |     |                                |

**Table A2.3 (corresponds to Table A1.3): Regression model of square root of net adult income of students who are not HRP or their spouse by selected factors collected in both the FRS and SHS. FRS 2005-06, UK.**

|  | Unstandardised co-efficients |     | Sig.                           |
|--|------------------------------|-----|--------------------------------|
|  | B                            | SE  | P-Value                        |
| (Constant)   | 6.6                          | 0.6 | 0.00                           |
| Parent   | -14.2                        | 3.2 | 0.00                           |
| Other relative                                     | 0.1                          | 0.6 | 0.88                           |
| Not related  | 2.3                          | 0.4 | 0.00                           |
| Sex  | 0.2                          | 0.3 | 0.39                           |
| Aged 20-24   | 2.0                          | 0.3 | 0.00                           |
| Aged 25-29   | 2.3                          | 0.5 | 0.00                           |
| Aged 30-34   | 4.5                          | 1.1 | 0.00                           |
| Aged 35-39   | 1.7                          | 1.5 | 0.27                           |
| Aged 40-44   | 5.7                          | 1.8 | 0.00                           |
| Aged 45-49   | 5.9                          | 3.0 | 0.05                           |
| Aged 50+   | 5.7                          | 2.9 | 0.05                           |
| In Scotland  | 0.0                          | 0.4 | 0.94                           |
| Mortgage   | 0.7                          | 0.4 | 0.05                           |
| Shared owner                                       | 1.5                          | 1.6 | 0.36                           |
| Rent   | -0.8                         | 0.4 | 0.09                           |
| Rent free  | 0.3                          | 1.0 | 0.79                           |
| Household has car/van?                             | 1.4                          | 0.4 | 0.00                           |
|  |                              |     | Adjusted R <sup>2</sup> = 0.09 |
| Constant – Child of HRP, Owned outright, Age 16-19 |                              |     |                                |

**Table A2.4 (corresponds to Table A1.6): Regression model of square root of net adult income of other adults who are unemployed, looking after home and sick/disabled who are not HRP or spouse by factors common to both FRS and SHS and detailed economic status. FRS 2005-06, UK.**

|   | Unstandardized Coefficients |     | Sig.    |
|---|-----------------------------|-----|---------|
|   | B                           | SE  | P-value |
| (Constant)  | 2.1                         | 0.5 | 0.00    |
| Parent  | 0.5                         | 0.8 | 0.59    |
| Other relative  | -0.4                        | 0.4 | 0.37    |
| Not related   | 1.8                         | 0.4 | 0.00    |
| Sex   | 0.2                         | 0.2 | 0.49    |
| Aged 20-24  | 1.8                         | 0.3 | 0.00    |
| Aged 25-29  | 2.8                         | 0.4 | 0.00    |
| Aged 30-34  | 3.2                         | 0.5 | 0.00    |
| Aged 35-39  | 5.3                         | 0.6 | 0.00    |
| Aged 40-44  | 3.4                         | 0.6 | 0.00    |
| Aged 45-49  | 4.1                         | 0.7 | 0.00    |
| Aged 50-54  | 5.8                         | 0.7 | 0.00    |
| Aged 55-59  | 4.6                         | 0.7 | 0.00    |
| Aged 60+  | 5.2                         | 0.8 | 0.00    |
| Scottish survey   | 0.6                         | 0.4 | 0.15    |
| Mortgage  | 0.0                         | 0.3 | 0.96    |
| Shared owner  | -1.9                        | 3.2 | 0.55    |
| Rent  | 0.7                         | 0.3 | 0.02    |
| Rent free   | 2.3                         | 2.1 | 0.28    |
| Looking after home/family                                     | 2.7                         | 0.5 | 0.00    |
| Long term sick  | 3.4                         | 0.3 | 0.00    |
| Short term sick   | 0.5                         | 0.7 | 0.42    |
| Other inactive  | -1.3                        | 0.3 | 0.00    |
| Household has car/van?  | -0.4                        | 0.3 | 0.14    |
| Adjusted R <sup>2</sup> = 0.38                                |                             |     |         |
| Constant: Child, 16-19, Unemployed, CT band A, Owner occupied |                             |     |         |

**Table A2.5 (corresponds to Table A1.9): Regression model of square root of net adult income of adults in full-time work who are not HRP or spouse by factors common to FRS and SHS . FRS 2005-06, UK excluding London.**

|  | Unstandardised Coefficients |     | Sig.                           |
|--|-----------------------------|-----|--------------------------------|
|  | B                           | SE  | P-value                        |
| (Constant)                                   | 12.2                        | 0.4 | 0.00                           |
| Parent                                       | 0.2                         | 1.1 | 0.86                           |
| Other relation                               | 0.3                         | 0.3 | 0.44                           |
| Not related                                  | 1.2                         | 0.3 | 0.00                           |
| Sex (Female vrs Male)                        | -0.3                        | 0.1 | 0.03                           |
| Aged 20-24                                   | 1.9                         | 0.2 | 0.00                           |
| Aged 25-29                                   | 3.1                         | 0.2 | 0.00                           |
| Aged 30-34                                   | 3.5                         | 0.3 | 0.00                           |
| Aged 35-39                                   | 3.5                         | 0.4 | 0.00                           |
| Aged 40-44                                   | 2.7                         | 0.4 | 0.00                           |
| Aged 45-49                                   | 3.1                         | 0.5 | 0.00                           |
| Aged 50-54                                   | 4.3                         | 0.6 | 0.00                           |
| Aged 55-59                                   | 2.6                         | 0.8 | 0.00                           |
| Aged 60-up                                   | 1.8                         | 1.0 | 0.08                           |
| Scottish survey                              | -0.2                        | 0.2 | 0.43                           |
| Mortgage                                     | -0.2                        | 0.2 | 0.16                           |
| Rent   | -0.9                        | 0.2 | 0.00                           |
| Rent free                                    | 0.4                         | 0.7 | 0.55                           |
| Household has car/van?                       | 1.2                         | 0.2 | 0.00                           |
|  |                             |     | Adjusted R <sup>2</sup> = 0.15 |
| Constant = Child, Aged 16-19, Owned Outright |                             |     |                                |

**Table A2.6 (corresponds to Table A1.12): Regression model of square root of net adult income of adults in part-time work who are not HRP or spouse by factors common to FRS and SHS . FRS 2005-06, UK excluding London.**

|  | Unstandardised Coefficients |     | Sig.                           |
|--|-----------------------------|-----|--------------------------------|
|  | B                           | SE  | P-value                        |
| (Constant)                                   | 9.7                         | 0.6 | 0.00                           |
| Parent                                       | -1.7                        | 1.3 | 0.21                           |
| Other relation                               | 0.8                         | 1.0 | 0.46                           |
| Not related                                  | 0.5                         | 0.6 | 0.34                           |
| Sex (Female vrs Male)                        | 0.2                         | 0.3 | 0.55                           |
| Aged 20-24                                   | 1.0                         | 0.3 | 0.00                           |
| Aged 25-29                                   | 2.3                         | 0.6 | 0.00                           |
| Aged 30-34                                   | 2.4                         | 0.7 | 0.00                           |
| Aged 35-39                                   | 1.9                         | 0.9 | 0.04                           |
| Aged 40-44                                   | 1.9                         | 1.1 | 0.10                           |
| Aged 45-49                                   | 1.8                         | 1.5 | 0.23                           |
| Aged 50-54                                   | 1.3                         | 1.3 | 0.35                           |
| Aged 55-59                                   | 2.6                         | 1.3 | 0.04                           |
| Aged 60-up                                   | 3.6                         | 1.1 | 0.00                           |
| Scottish survey                              | 0.2                         | 0.5 | 0.68                           |
| Mortgage                                     | -0.4                        | 0.4 | 0.25                           |
| Shared owners                                | -3.1                        | 2.1 | 0.15                           |
| Rent   | -1.0                        | 0.4 | 0.02                           |
| Rent free                                    | 0.8                         | 1.3 | 0.56                           |
| Household has car/van?                       | 0.3                         | 0.4 | 0.49                           |
|  |                             |     | Adjusted R <sup>2</sup> = 0.09 |
| Constant = Child, Aged 16-19, Owned Outright |                             |     |                                |

**Table A2.7 (Corresponds to Table A1.15): Regression model of square root of net adult income of self-employed adults who are not HRP or spouse by factors common to FRS and SHS . FRS 2005-06, UK excluding London.**

|  | Unstandardised Coefficients |     | Sig.    |
|--|-----------------------------|-----|---------|
|  | B                           | SE  | P-Value |
| (Constant)   | 11.0                        | 2.5 | 0.00    |
| Parent   | 2.7                         | 3.3 | 0.41    |
| Other relation   | 0.7                         | 1.9 | 0.73    |
| Not related  | -0.4                        | 1.3 | 0.79    |
| Sex (Female vrs Male)  | -2.4                        | 1.2 | 0.05    |
| Aged 20-24   | 1.7                         | 1.6 | 0.28    |
| Aged 25-29   | 2.3                         | 1.7 | 0.18    |
| Aged 30-34   | -0.2                        | 1.8 | 0.93    |
| Aged 35-39   | 5.9                         | 2.6 | 0.03    |
| Aged 40-44   | -0.3                        | 2.3 | 0.88    |
| Aged 45-49   | 5.0                         | 3.1 | 0.11    |
| Aged 50-54   | -2.6                        | 2.8 | 0.35    |
| Aged 55-59   | 1.0                         | 2.9 | 0.73    |
| Aged 60-up   | 0.2                         | 3.2 | 0.96    |
| Scottish survey  | -0.8                        | 1.6 | 0.61    |
| Mortgage   | -1.9                        | 1.0 | 0.05    |
| Rent   | -1.2                        | 1.4 | 0.37    |
| Rent free  | -12.3                       | 4.5 | 0.01    |
| Household has car/van?   | 4.6                         | 1.8 | 0.01    |
| Constant = Child, Aged 16-19, Owned Outright<br>Adjusted R <sup>2</sup> = 0.11 |                             |     |         |