ELSA Financial Derived Variables

Zoe Oldfield

Institute for Fiscal Studies

Overview

What are they?

What information do they contain?

How to use them?

The ELSA questionnaire

- Includes hundreds of questions regarding respondents' economic situation (earnings, benefits, pensions, assets, debts, etc...)

- ... spread over several different modules:
  - Work and Pensions (WP)
  - Income and Assets (IA)
  - Housing (HO)

- Lots of flexibility offered to respondent in the way that they answer the questions

What are the financial derived variables?

- We've combed through ELSA's many financial questions and
- transformed income flows into **uniform periods** (weekly)
- imputed amounts where we have incomplete or missing values
- converted amounts so they relate to a meaningful unit of observation
- added components together into economically meaningful **aggregates (**employment income, benefit income etc.)
- …so that you don't have to.

Which files are we talking about?

Data files	
Wave_X_financial_derived_variables	• Contains only information relating to financial variables • One observation per ELSA respondent
Wave_X_IFS_derived variables	• Contains other derived variables that are commonly used by IFS (and other) researchers • One observation per ELSA respondent

Which files are we talking about?

Documentation	
User_Guide_Financial_Derived Variables.pdf	Basic user guide
User_Guide_IFS_Derived_Variables.pdf	Basic user guide
Financial_Derived_Variables_Relationships.xlsx	Detailed information about the source variables for each financial variables (one tab for each wave)
IFS_derived_variables_description.xlsx	Detailed information on how each non-financial derived variables is created
Financial_Derived_variables_and Imputation_procedures.pdf	A more in-depth description of the way that we impute missing data

Which files are we talking about?

Documentation	
User_Guide_Financial_Derived Variables.pdf	Basic user guide
User_Guide_IFS_Derived_Variables.pdf	Basic user guide
Financial_Derived_Variables_Relationships.xlsx	Detailed information about the source variables for each financial variables (one tab for each wave)
IFS_derived_variables_description.xlsx	Detailed information on how each non-financial derived variables is created
Financial_Derived_variables_and Imputation_procedures.pdf	A more in-depth description of the way that we impute missing data

Which files are we talking about?

Documentation	
User_Guide_Financial_Derived Variables.pdf	Basic user guide
User_Guide_IFS_Derived_Variables.pdf	Basic user guide
Financial_Derived_Variables_Relationships.xlsx	Detailed information about the source variables for each financial variables (one tab for each wave)
IFS_derived_variables_description.xlsx	Detailed information on how each non-financial derived variables is created
Financial_Derived_variables_and Imputation_procedures.pdf	A more in-depth description of the way that we impute missing data

Q. What do we end up with?

A. Still a lot of variables to deal with!

But, the variables can be used as they are without further processing and once you know the naming conventions and the structure of the variables, it should (hopefully) be easier to work with!

Three categories of variable

- Single components of income/wealth
  - income support, incapacity benefit, ESA…
  - Take home pay, odd job income, private pension income, other annuity income…
  - ISAs, second homes, other assets…
- Aggregate measures of income/wealth
  - Total income, state benefit income, employment income, self employment income..
  - Total wealth, housing wealth, financial wealth..
- Auxiliary variables
  - Variables designed to give you extra information about each variable: why it's missing, what the underlying quality of the variable is

General naming convention

Every variable in the financial derived variables is named in the following way:

$$stem\_1^{st}suffix\_2^{nd}suffix$$

General naming convention

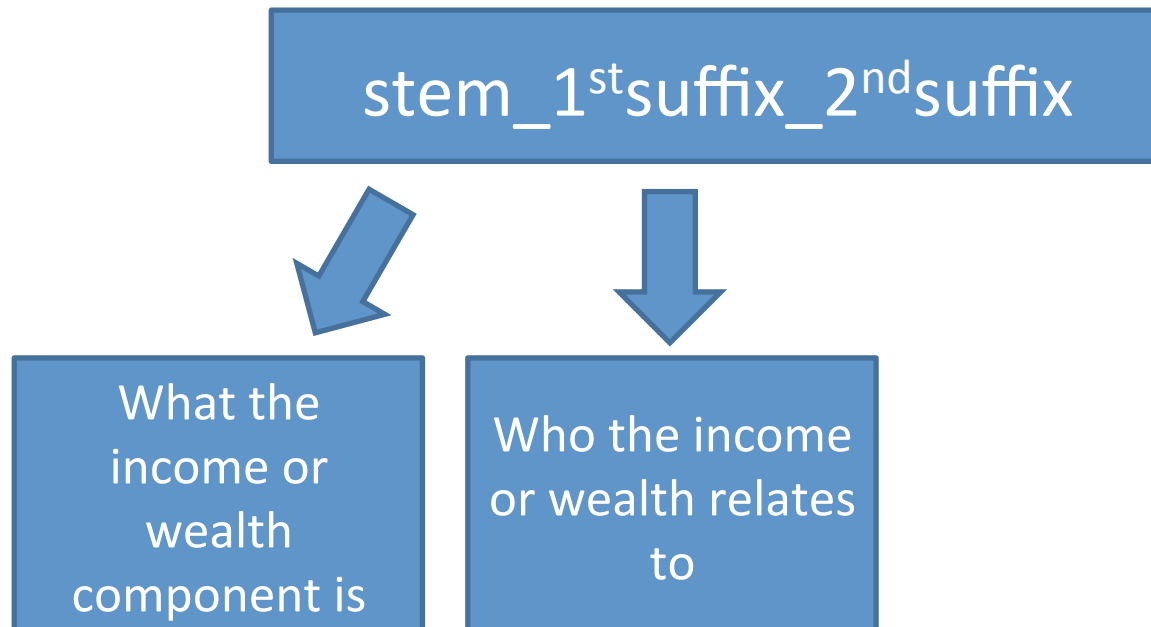Every variable in the financial derived variables is named in the following way:

stem_1stsuffix_2ndsuffix

What the income or wealth component is

General naming convention

Every variable in the financial derived variables is named in the following way:

$$stem\_1^{st}suffix\_2^{nd}suffix$$

| What the income or wealth component is | Who the income or wealth relates to |

General naming convention

Every variable in the financial derived variables is named in the following way:

stem_1stsuffix_2ndsuffix

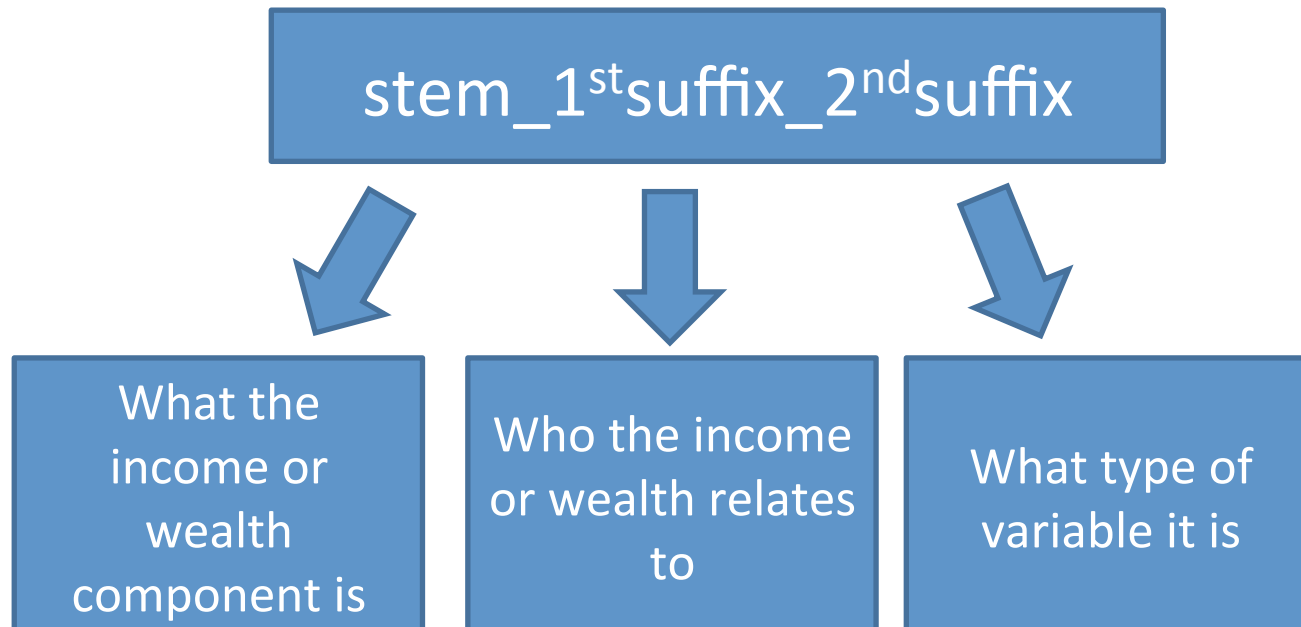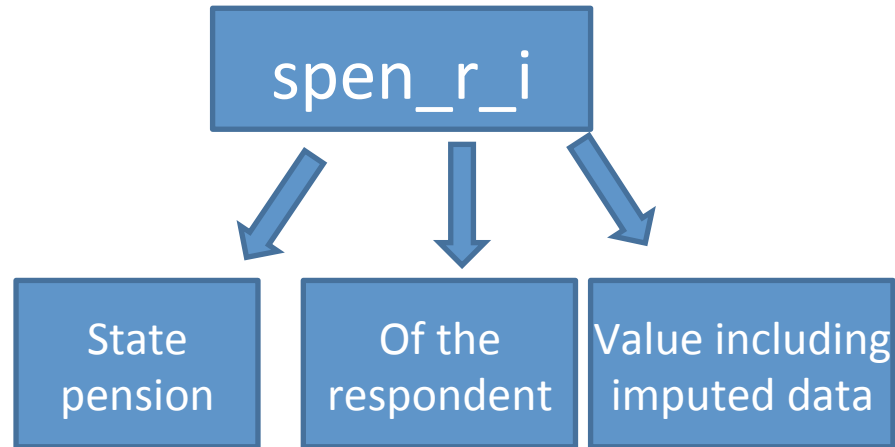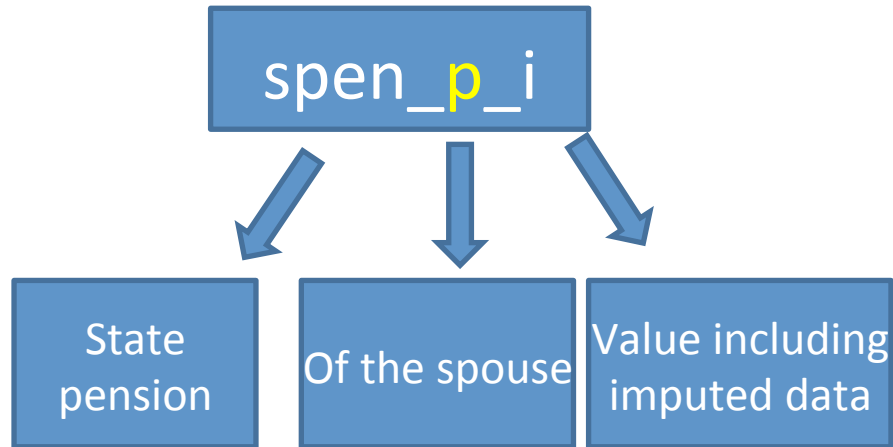What the income or wealth component is

Who the income or wealth relates to

What type of variable it is

Example – state pension (respondent)

Example – state pension (partner)

1st suffix can be:

suffix	Relates to...
_r	The respondent
_p	The partner of the respondent
_bu	The benefit unit.
_hh	The household

1st suffix can be:

suffix	Relates to...
_r	The respondent
_p	The partner of the respondent
_bu	The benefit unit.
_hh	The household

1ˢᵗ suffix can be:

suffix	Relates to...
_r	The respondent
_p	The partner of the respondent
_bu	The benefit unit.
_hh	The household

What is a "benefit unit"?

- Technically we are referring to a unit that is used for the assessment of state benefits

- It is a couple or a single person (plus any dependent children they may have)

- A "household" (all individuals living together) can have multiple benefit unit

What if they don't know the amount?

- Respondent is asked a series of follow up questions designed to extract as much information as possible
- Of the form "is it more or less than £X?"
- Respondent can say "don't know" at any stage
- We end up with:
  - Closed band (between £X and £Y)
  - Open band (more than £Y)
  - "Missing positive"
  - "Missing completely"

2ⁿᵈ suffix can be:

suffix	What it means
_i	These variables contain the (imputed if necessary) value of the income or wealth
_t	These variables indicate what information the respondent gave us: 0 – zero (no income/wealth) 1 – continuous answer 2 – closed band 3 – open band 4 – missing positive 5 – missing completely *Used only for components of income or wealth*
_o	The number of observations that the imputation was based on
_f	Validation flag – gives the reason why a variable is missing (i.e. not imputed)
_s	Indicates that the variable is an aggregate (summary) measure of income or wealth
_ni2 _ni3 _ni4	Number of components of a summary measure of income/wealth that were of type 2 (closed band), 3 (open band) or 4 (missing positive) or more *Used only for aggregate (summary) measures of income or wealth*

2nd suffix can be:

suffix	What it means
_i	These variables contain the (imputed if necessary) value of the income or wealth
_t	These variables indicate what information the respondent gave us: 0 – zero (no income/wealth) 1 – continuous answer 2 – closed band 3 – open band 4 – missing positive 5 – missing completely *Used only for components of income or wealth*
_o	The number of observations that the imputation was based on
_f	Validation flag – gives the reason why a variable is missing (i.e. not imputed)
_s	Indicates that the variable is an aggregate (summary) measure of income or wealth
_ni2 _ni3 _ni4	Number of components of a summary measure of income/wealth that were of type 2 (closed band), 3 (open band) or 4 (missing positive) or more *Used only for aggregate (summary) measures of income or wealth*

2ⁿᵈ suffix can be:

suffix	What it means
_i	These variables contain the (imputed if necessary) value of the income or wealth
_t	These variables indicate what information the respondent gave us: 0 – zero (no income/wealth) 1 – continuous answer 2 – closed band 3 – open band 4 – missing positive 5 – missing completely *Used only for components of income or wealth*
_o	The number of observations that the imputation was based on
_f	Validation flag – gives the reason why a variable is missing (i.e. not imputed)
_s	Indicates that the variable is an aggregate (summary) measure of income or wealth
_ni2 _ni3 _ni4	Number of components of a summary measure of income/wealth that were of type 2 (closed band), 3 (open band) or 4 (missing positive) or more *Used only for aggregate (summary) measures of income or wealth*

2nd suffix can be:

suffix	What it means
_i	These variables contain the (imputed if necessary) value of the income or wealth
_t	These variables indicate what information the respondent gave us: 0 – zero (no income/wealth) 1 – continuous answer 2 – closed band 3 – open band 4 – missing positive 5 – missing completely *Used only for components of income or wealth*
_o	The number of observations that the imputation was based on
_f	Validation flag – gives the reason why a variable is missing (i.e. not imputed)
_s	Indicates that the variable is an aggregate (summary) measure of income or wealth
_ni2 _ni3 _ni4	Number of components of a summary measure of income/wealth that were of type 2 (closed band), 3 (open band) or 4 (missing positive) or more *Used only for aggregate (summary) measures of income or wealth*

What measures of income and wealth are available?

- In total there are:
  - 45 components of income
  - 31 components of wealth
  - 7 aggregate (summary) measures of income

- Best place to find out what those are is Table 5 in the User_guide_Financial_derived_variables_.pdf
  - Tells you whether available individual/benefit unit/household level
  - Tells you what category of income or wealth it is
  - Tells you whether it's a component or summary measure of income or wealth

What's included in the FDVs?
Summary measures of income

- **eqtotinc_bu_s** = Total income adjusted for family size ('equivalised')

- **totinc_bu_s =** Unadjusted ('unequivalised') total income

- Income quintiles and deciles also available:
  - **yq5_bu_s** = Equivalised income quintiles
  - **yq10_bu_s** = Equivalised income deciles

- Users may wish to recalculate quintiles/deciles when working with subgroups
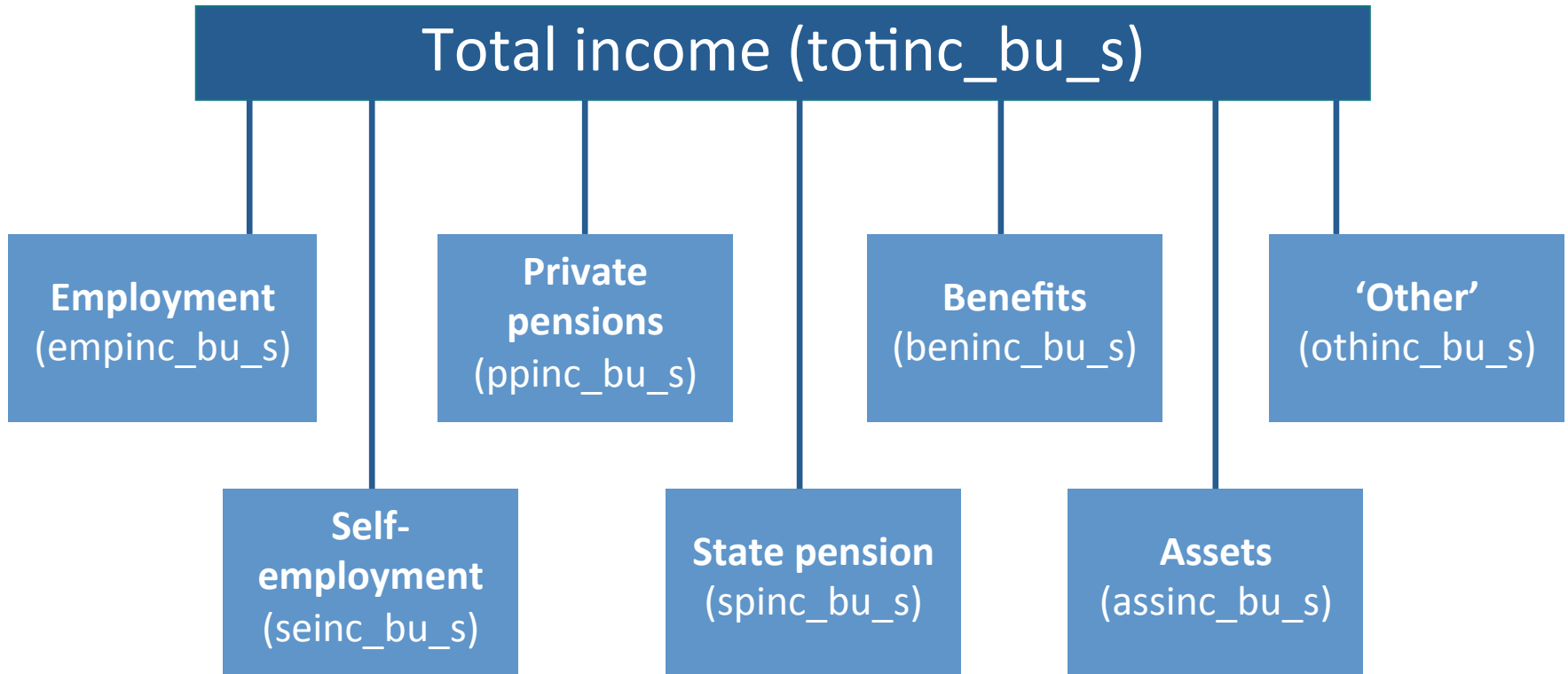 - To do this you rank individuals according to their income (within subgroup) and split into equal groups

What's included in the FDVs?
Summary measures of income

Total income (totinc_bu_s)

What's included in the FDVs?
Summary measures of income

Total income (totinc_bu_s)

Employment (empinc_bu_s)

Self-employment (seinc_bu_s)

Private pensions (ppinc_bu_s)

State pension (spinc_bu_s)

Benefits (beninc_bu_s)

Assets (assinc_bu_s)

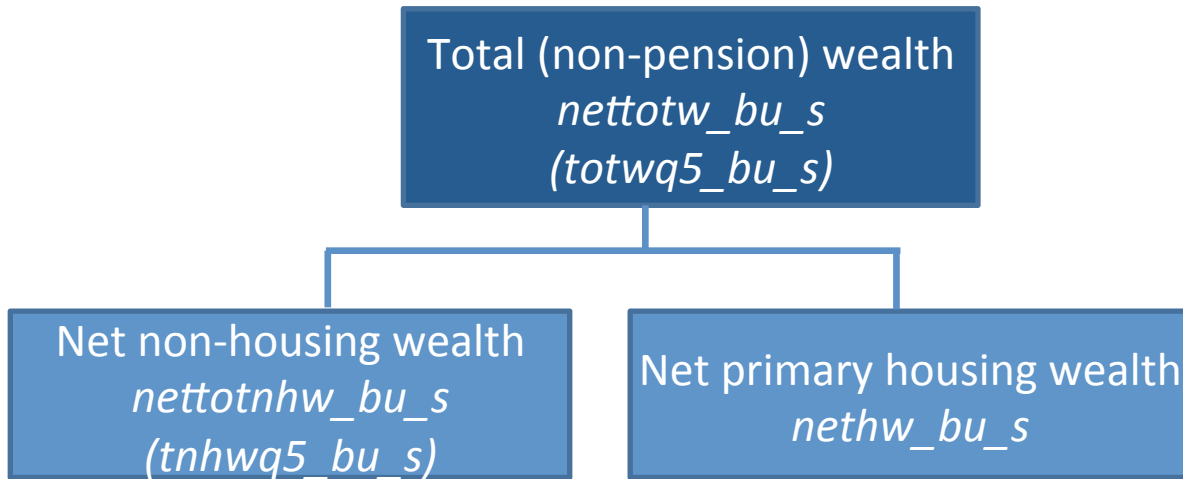**'Other' (othinc_bu_s)**

What's included in the FDVs?
Summary measures of wealth

Total (non-pension) wealth
nettotw_bu_s
(totwq5_bu_s)

What's included in the FDVs?
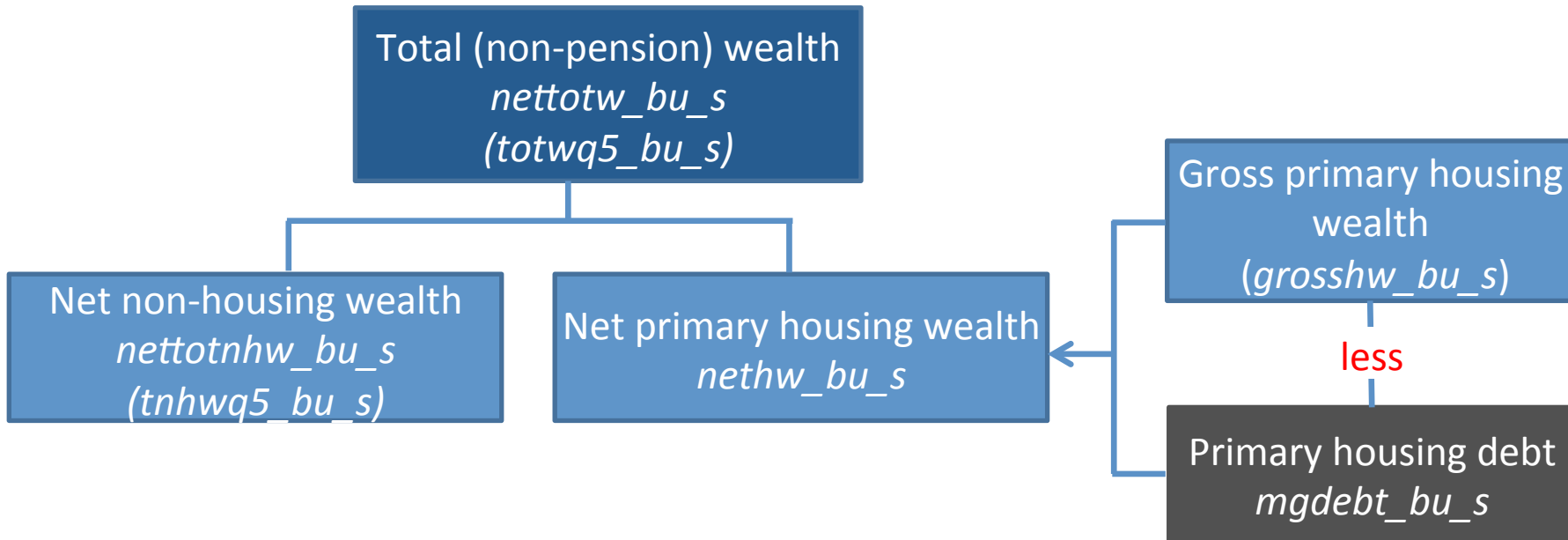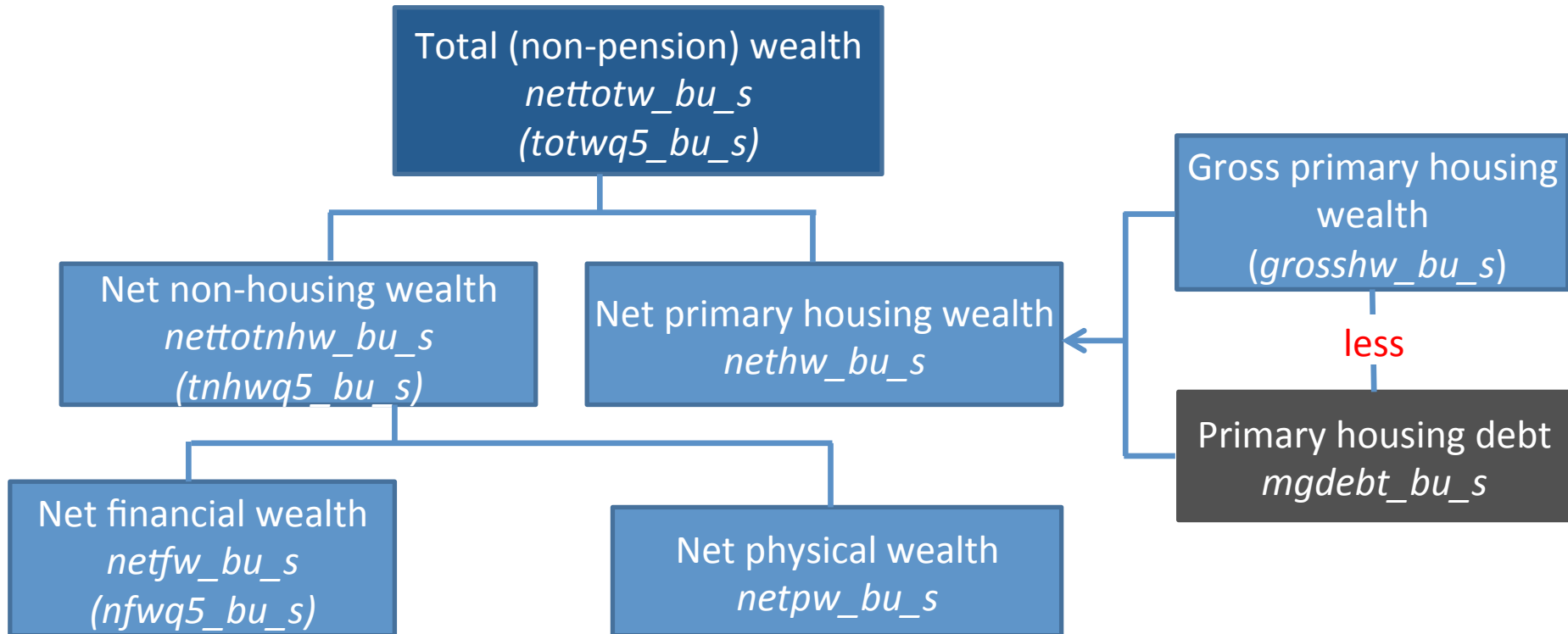## Summary measures of wealth

Total (non-pension) wealth
nettotw_bu_s
(totwq5_bu_s)

Net non-housing wealth
nettotnhw_bu_s
(tnhwq5_bu_s)

Net primary housing wealth
nethw_bu_s

What's included in the FDVs?
Summary measures of wealth

Total (non-pension) wealth
nettotw_bu_s
(totwq5_bu_s)

Net non-housing wealth
nettotnhw_bu_s
(tnhwq5_bu_s)

Net primary housing wealth
nethw_bu_s

Gross primary housing wealth
(grosshw_bu_s)

less

Primary housing debt
mgdebt_bu_s

What's included in the FDVs?
Summary measures of wealth

Total (non-pension) wealth
nettotw_bu_s
(totwq5_bu_s)

Gross primary housing wealth
(grosshw_bu_s)

Net non-housing wealth
nettotnhw_bu_s
(tnhwq5_bu_s)

Net primary housing wealth
nethw_bu_s

less

Primary housing debt
mgdebt_bu_s

Net financial wealth
netfw_bu_s
(nfwq5_bu_s)

Net physical wealth
netpw_bu_s

What's included in the FDVs?
Summary measures of wealth

Total (non-pension) wealth
nettotw_bu_s
(totwq5_bu_s)

Net non-housing wealth
nettotnhw_bu_s
(tnhwq5_bu_s)

Net primary housing wealth
nethw_bu_s

Gross primary housing wealth
(grosshw_bu_s)

less

Primary housing debt
mgdebt_bu_s

Net financial wealth
netfw_bu_s
(nfwq5_bu_s)

Net physical wealth
netpw_bu_s

Gross financial wealth
grossfw_bu_s

less

Financial debt
debt_bu_s

© Institute for Fiscal Studies

What's included in the FDVs?
Summary measures of wealth

Total (non-pension) wealth
nettotw_bu_s
(totwq5_bu_s)

Net non-housing wealth
nettotnhw_bu_s
(tnhwq5_bu_s)

Net primary housing wealth
nethw_bu_s

Gross primary housing wealth
(grosshw_bu_s)

less

Primary housing debt
mgdebt_bu_s

Net financial wealth
netfw_bu_s
(nfwq5_bu_s)

Net physical wealth
netpw_bu_s

Gross financial wealth
grossfw_bu_s

less

Financial debt
debt_bu_s

© Institute for Fiscal Studies

Which measure of income or wealth should I use?

- This really depends on the research question
- Good place to start is the summary variables
  - Remember these end in "_s"
- Also look at:
  - Table 5 in the FDV user guide
  - Figures 1 and 2 in the FDV user guide
  - Financial_derived_Variable_Relationships.xls (and filter by category of income)

Which variables in the questionnaire do the derived variables correspond to?

- This is documented in:

  Financial_Derived_Variable_Relationships.xls

- It's also noted in the variable description

trusts_bu_i: "BU Trusts (iauit) - value (incl. imputed values)"

This variable is based on iauit

Is there an individual measure of income?

- The short answer is no
  - ELSA is designed to measure income at the benefit unit level
  - All summary measures relate to the benefit unit level

But…

- Some components of income are collected at the individual level
  - These are measures which are easily assignable to an individual (state benefits, employment income etc.)

Important things to remember

- All income flows are measured weekly
- There are equivalised and unequivalised measures of aggregate income..
- ..but not wealth
- The unit of observation is the individual but measures of income/wealth (generally) relate to the benefit unit.
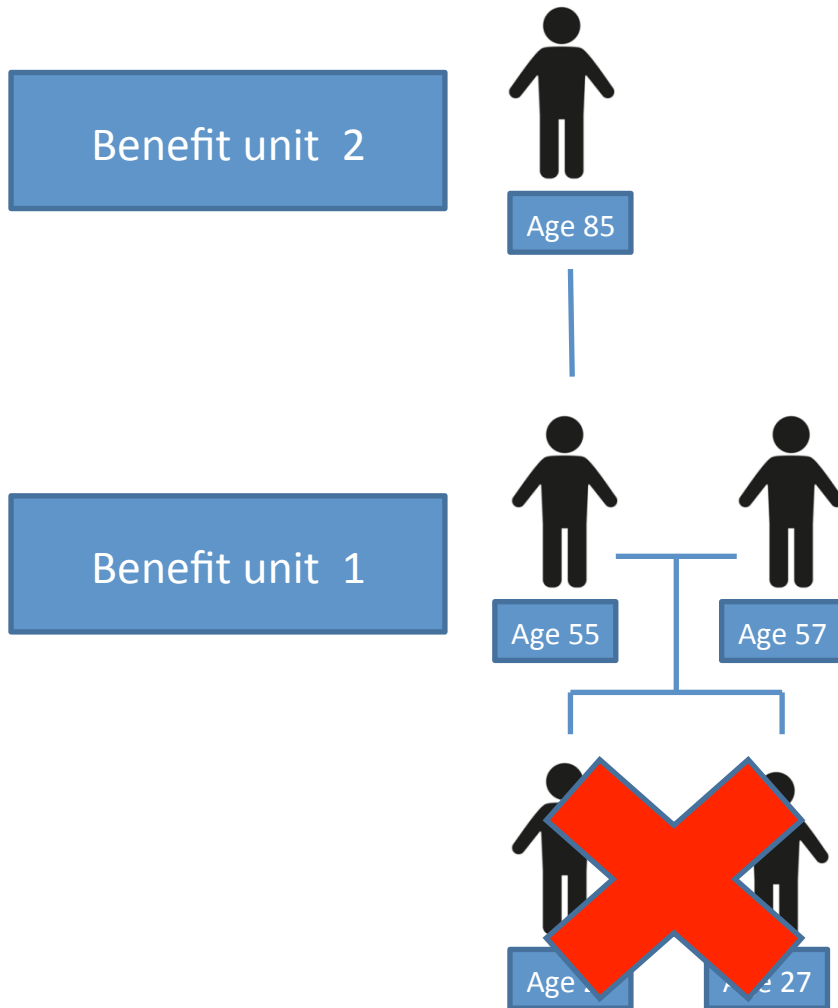
Example household
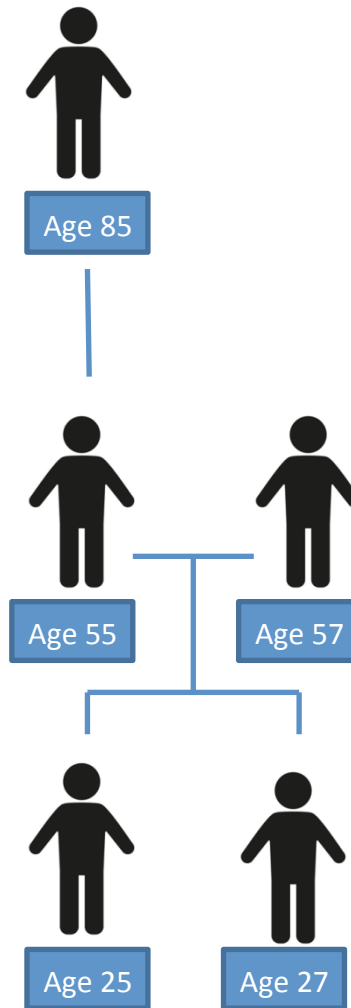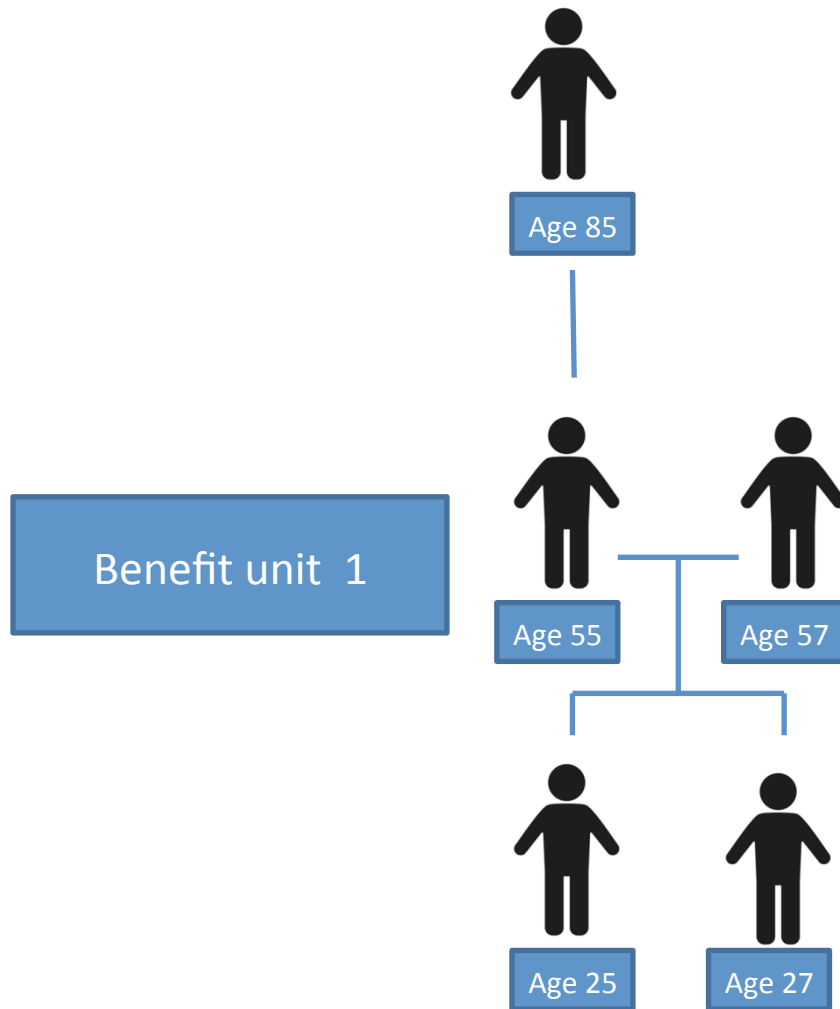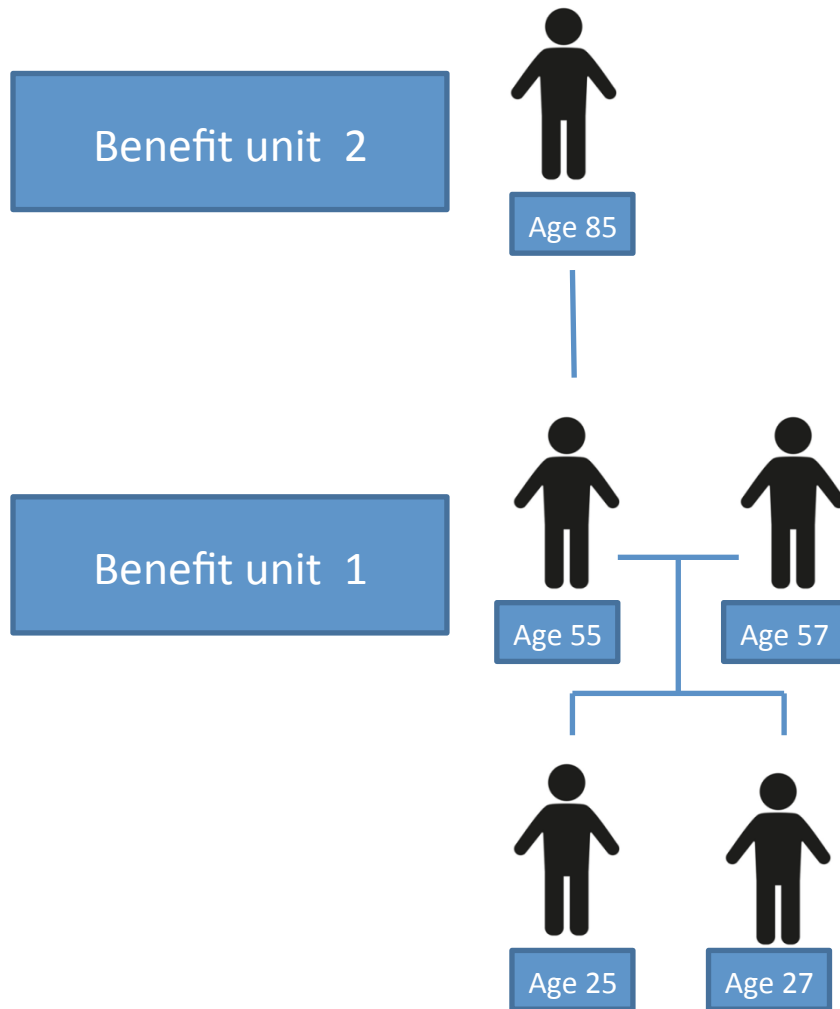
Example household

Example household

Example household

Example household

totinc_bu_s (total income)		
£300		
£300		
£150		

Benefit unit 1 — Age 57

Benefit unit 1 — Age 55

Benefit unit 2 — Age 85

Example household

Totinc_bu_s (total income)	thp_r_i (take home pay)	
£300	£220	
£300	£50	
£150	£0	

Benefit unit 1   Age 57

Benefit unit 1   Age 55

Benefit unit 2   Age 85

Example household

	Totinc_bu_s (total income)	thp_r_i (take home pay)	thp_bu_i (take home pay)
Benefit unit 1 — Age 57	£300	£220	£270
Benefit unit 1 — Age 55	£300	£50	£270
Benefit unit 2 — Age 85	£150	£0	£0

Non-financial derived variables

- Purpose: centralise the derivation of some of the more complex variables and reduce duplication of effort across the research community

- Includes additional identifiers, children variables, family type, labour market participation, summary pension variables, summary health, expectations, expenditure

Much more in the documentation

- More detail on how the variables were imputed
- More detail on how the variables were derived
- How to use the auxiliary variables
- More detail on the way the data are collected
- Information about the pension wealth variables
- Please come and ask me if you have questions

Thank you!