CHILDCARE AND EARLY YEARS PROVIDERS SURVEY 2011 – SPSS NAMING & WEIGHTING SUMMARY

VARIABLE NAMES

The following general naming conventions are used for variable within the SPSS datasets:

- If the variable name starts with a 'q' then it is just a standard question in the survey where respondents were only able to give one answer.
- For the childcare dataset, if the variable name starts with an '@' symbol then it is a standard question in the survey where respondents could give more than one answer and each iteration of the variable represents one answer code. In the early years and childminders datasets there is no '@' symbol beforehand, but you will see each variable has multiple iterations when more than one response is allowed.
- If a variable starts with a 'd' it is a derived variable
- If a variable starts with a 'b' it is a break variable (a variable we use as standard in the tables for subgroup analysis).
- If a variable starts with an 'n' then it is the unbanded version of numeric question.
- The staff level variables start with a 't' or 'nt' depending on whether it is a banded or unbanded derived variable.

TYPES OF DATABASE

There are two separate types of database – hierarchical and non-hierarchical.

The hierarchical datasets are only used when looking at data relating to 'all staff' (i.e. adding together the total number of staff at different levels).

It therefore follows, that all provider level questions, and also staff questions which only look at one staff level, should be analysed in the non-hierarchical file.

EARLY YEARS DATABASES

For the Early Years data set the different types of provider can be selected using the '**dsample**' variable.

Non-hierarchical:

The weights to be used in the non-hierarchical Early Years file are called:

wtotreg (this is the general weight including grossing and regional and exclusion weighting)

The following three weights are the staff weighting that should be applied when looking at one of the three different levels of staff in isolation – if you want to look at the total number of staff across several levels, you will need to use the hierarchical database:

wtotqstaf (for other qualified teachers)
wtotqnurs (for nursery nurses)
wtotpaid (for other paid early years support staff)

Hierarchical:

In the hierarchical dataset, all analysis will be run at staff level rather than provider level.

The one and only weight that should be used in the early years hierarchical dataset is called **hierwgt**.

There is a variable called '**Stnum1**' which tells you which type of staff each case is. The difference between stnum and stnum1 is that stnum1 bands all staff of the same category together.

The Early Years hierarchical weight was calculated / derived as follows: If stnum1=1, use wtotreg; if stnum1=2 or 3, use wtotqstaf; if stnum1=4 or 5, use wtotqnurse; and if stnum1=6 or 7, use wtotqpaid.

Because hierarchical data is at staff level, only the following variables will add up to the figures in the tables/report when aggregated (all other variables would only make sense if they were run for one particular member of staff – for example to find out what the characteristics of their provider was):

Tothour
Ntothour
Totage
Totqual
Totqts
Totqualw
Ntotannpay
Ntothrpay
Totannpay
Tothrpay

CHILDCARE DATABASES

For Childcare '@tabset1-6' allow you to select each of the different childcare groups.

Non-hierarchical:

There are three non-hierarchical childcare weights in total. In some earlier waves of the survey, different weights were used for each type of provider. This is no longer the case.

The weights are named in the non-hierarchical SPSS as follows (SPSS names shown in **BOLD)**:

wt_reg	Weights un-modulated variables relating to individual settings and the senior manager (region and grossing weight)
wt_supv	Weights un-modulated variables relating to supervisors
wt_paid	Weights un-modulated variables relating to other paid members of staff

The **supv** and **paid** weights should be applied to variables relating to these staff types (e.g. the paid staff weight applies to any questions relating to the qualifications and pay of 'other paid staff', as distinct from 'all paid staff', which are covered by the hierarchical weights in the other database).

The **reg** weights should be applied to any variables relating to either the setting or to the senior manager.

Hierarchical:

In the hierarchical dataset, all analysis will be run at staff level rather than provider level. There is a variable called 'Stnum1' which tells you which type of staff each case is. The difference between stnum and stnum1 is that stnum1 bands all staff of the same category together.

There is only one weight to use in the Childcare hierarchical dataset:

hierwgt (applies to all hierarchical variables and can be used for all care types)

Because hierarchical data is at staff level, only the following variables will add up to the tables when aggregated (all other variables would only make sense if they were ran for one particular member of staff – for example to find out what the characteristics of their provider was):

Tothour
ntothour
totqts
totqual
totqualw
totpay
ntotpay
totfipay
ntotfipay
toteyps
totinter
toteyqts