Understanding Society User Support - Support #257

Residential mobility and LSOA codes

05/02/2014 11:03 AM - Rory Coulter

Status:	Closed		**Start date:**	05/02/2014
Priority:	Normal		**% Done:**	100%
Assignee:				
Category:	Special license			

Description

Hi Jakob,

Apologies for the long post. I'm using UKHLS studies SN6614 and SN7248 (Special License Census 2011 LSOA codes) and have a couple of queries about residential mobility and the geocodings in the data.

1. First I was wondering if you could advise as to whether there is any best practice as to how to derive an indicator of individual residential mobility for waves 2 and 3 of UKHLS. This was quite easy in BHPS as there was a 'movest' variable coding whether the person changed address since the last wave. BHPS also had the 'plnew' question and this could be compared with movest to easily derive a mobility dummy. Neither of these options seems to be available in UKHLS as plnew is no longer asked and there is no movest equivalent. From previous forum posts I can see that there are several ways mobility could be defined, although each has problems:

a) Using xwhist on xwavedat - although it is not entirely clear what a change in postcode actually means. Is this a move over any distance or could the person have moved within their postcode area without this being recorded?
b) By comparing LSOA/datazones at t-1 with t (more on this below)
c) Using the origadd and reasons for moving variables derived from the questionnaires-although the origadd indicator seems to be derived at the household not individual level. Am I correct in assuming that this means that people are only routed towards the reasons for moving questions in the annual event module if their whole household had moved since t-1? This is quite different from BHPS and it seems odd given that households are not consistent units through time.

My current approach is to try and do all of these and somehow triangulate them into a new variable. Does this seem appropriate or is there something really obvious I am missing? In any case it would be very useful to have a simple 'residential move since last wave' dummy in UKHLS (like movest in BHPS). Are there any plans to produce such a variable?

2. I have some concerns about some of the LSOA/datazone information in SN7248. The study documentation tells us that at wave 2 37500 households had postcodes matched to their records and this was 35128 at wave 3. However in the SN7248 there are lots of hidp records with blank micro-geographic codes (5315 at wave 1; 3719 at wave 2; 3320 at wave 3). This isn't mentioned in the study documentation. Is there any reason for this discrepancy with the number of households with valid postcodes? I merged the LSOA identifiers onto the hhsamp files and most of these cases seem to come from Scottish households, many of whom completed full or proxy interviews.

Thanks in advance for any help you can provide.

Best wishes,
Rory

History

#1 - 05/07/2014 02:14 PM - Redmine Admin

- Target version set to X M

- % Done changed from 0 to 50

> a) Using xwhist on xwavedat - although it is not entirely clear what a change in postcode actually means. Is this a move over any distance or could the person have moved within their postcode area without this being recorded?

Moves within the same unit postcode are not detected in this way.

> b) By comparing LSOA/datazones at t-1 with t (more on this below)

Could be used if only moves between different LSOAs are of interest.

c) Using the origadd and reasons for moving variables derived from the questionnaires-although the origadd indicator seems to be derived at the household not individual level. Am I correct in assuming that this means that people are only routed towards the reasons for moving questions in the annual event module if their whole household had moved since t-1? This is quite different from BHPS and it seems odd given that households are not consistent units through time.

The universe for e.g. movy1 (reason for move) can be found in the questionnaire specification;

```
If (ff_ivlolw = 1 | Ff_everint = 1) // interviewed at prior wave or has been interviewed previously
And If ((HHGrid.OrigAdd = 1 & (HHGrid.NewPer = 2 | AdCts = 2)) | (HHGrid.OrigAdd = 2 & AdCts <> 1)) // HH inte
rviewed at current address previously and
respondent is a rejoiner or respondent has not lived at current address continuously since previous interview,
 or HH interviewed at a different address previously and respondent has not lived at the current
address continuously since previous interview
```

So, not solely dependent on origadd.

My current approach is to try and do all of these and somehow triangulate them into a new variable. Does this seem appropriate or is there something really obvious I am missing? In any case it would be very useful to have a simple 'residential move since last wave' dummy in UKHLS (like movest in BHPS). Are there any plans to produce such a variable?

Thanks for the suggestion. We will take it into account when we next revise the added-value content. Please note that we have not (yet) tried to resolve any potential conflicts between approach a and c to date.

2. I have some concerns about some of the LSOA/datazone information in SN7248. The study documentation tells us that at wave 2 37500 households had postcodes matched to their records and this was 35128 at wave 3. However in the SN7248 there are lots of hidp records with blank micro-geographic codes (5315 at wave 1; 3719 at wave 2; 3320 at wave 3). This isn't mentioned in the study documentation. Is there any reason for this discrepancy with the number of households with valid postcodes?

The Scottish 2011 Census identifiers were not available before the deadline for this deposit. This is documented in the source, ONSPD Nov 2013, but unfortunately not in our documentation.

Let us know if you have got any further questions.

Jakob

#2 - 05/16/2014 10:05 AM - Rory Coulter

Hi Jakob,

Thanks for your very informative post, that clears almost everything up. It would be great to have a 'mover' dummy in future releases of the data, if only so that work could be more easily replicated.

One further clarification would be helpful. On the xwhist variable there are quite a lot of cases where a postcode is present at wave 1, then not at wave 2 but the person is back at wave 3. Examples could be 'FMS' or 'FON'. If a person has a value of 'O' or 'M' in wave 2, then am I correct to assume that any 'S' or 'N' values at wave 3 refer to the individual's mobility behaviour between wave 1 and wave 3? So basically if an 'S' or 'N' follows a non-contact, then all we know is that that individual moved at some point between the first and last observation. If we were interested in mobility between pairs of waves then we couldn't make use of these cases as it is not possible to know when the move occurred (unless this is recorded on plnowm and plnowy4).

Thanks again for your help.

Redmine Admin wrote:

a) Using xwhist on xwavedat - although it is not entirely clear what a change in postcode actually means. Is this a move over any distance or could the person have moved within their postcode area without this being recorded?

Moves within the same unit postcode are not detected in this way.

b) By comparing LSOA/datazones at t-1 with t (more on this below)

Could be used if only moves between different LSOAs are of interest.

c) Using the origadd and reasons for moving variables derived from the questionnaires-although the origadd indicator seems to be derived at the household not individual level. Am I correct in assuming that this means that people are only routed towards the reasons for moving questions in the annual event module if their whole household had moved since t-1? This is quite different from BHPS and it seems odd given that households are not consistent units through time.

The universe for e.g. movy1 (reason for move) can be found in the questionnaire specification;

[...]

So, not solely dependent on origadd.

My current approach is to try and do all of these and somehow triangulate them into a new variable. Does this seem appropriate or is there something really obvious I am missing? In any case it would be very useful to have a simple 'residential move since last wave' dummy in UKHLS (like movest in BHPS). Are there any plans to produce such a variable?

Thanks for the suggestion. We will take it into account when we next revise the added-value content. Please note that we have not (yet) tried to resolve any potential conflicts between approach a and c to date.

2. I have some concerns about some of the LSOA/datazone information in SN7248. The study documentation tells us that at wave 2 37500 households had postcodes matched to their records and this was 35128 at wave 3. However in the SN7248 there are lots of hidp records with blank micro-geographic codes (5315 at wave 1; 3719 at wave 2; 3320 at wave 3). This isn't mentioned in the study documentation. Is there any reason for this discrepancy with the number of households with valid postcodes?

The Scottish 2011 Census identifiers were not available before the deadline for this deposit. This is documented in the source, ONSPD Nov 2013, but unfortunately not in our documentation.

Let us know if you have got any further questions.

Jakob

#3 - 05/21/2014 04:06 PM - Redmine Admin

- % Done changed from 50 to 80

On the xwhist variable there are quite a lot of cases where a postcode is present at wave 1, then not at wave 2 but the person is back at wave 3. Examples could be 'FMS' or 'FON'. If a person has a value of 'O' or 'M' in wave 2, then am I correct to assume that any 'S' or 'N' values at wave 3 refer to the individual's mobility behaviour between wave 1 and wave 3? So basically if an 'S' or 'N' follows a non-contact, then all we know is that that individual moved at some point between the first and last observation

That is correct.

If we were interested in mobility between pairs of waves then we couldn't make use of these cases as it is not possible to know when the move occurred (unless this is recorded on plnowm and plnowy4).

As mentioned above, we have yet to try and tie admin and response data together and resolve any potential conflicts between these different data sources. I agree that some assumptions might be needed.
Should also mention that in terms of data access, it is possible to apply for secure data service access to postcode grid references.

Jakob

#4 - 05/30/2014 08:49 AM - Redmine Admin

- Status changed from New to Closed

- % Done changed from 80 to 100