



Understanding Society: the UK Household Longitudinal Study



<http://www.understandingsociety.org.uk/>

UK HLS Background

- *Understanding Society* is a longitudinal study based on a household panel design
- Basic design similar to BHPS
- Target sample size of 40,000 households – largest Household Panel Survey
- Main fieldwork due to start in January 2009

Importance of the Household focus

- Strength of the HPS model shown by range of studies internationally (e.g. PSID, SOEP, HILDA)
- Important for research on households and individuals
 - consumption and income, where within-household sharing of resources is important
 - demographic change, where the household itself is often the object of study

Importance of the Household focus

- The household itself is often the object of study
 - demographic change
 - consumption
 - income
 - within-household sharing of resources

Importance of the Household focus

- Can investigate family factors in decision making
- Observing multiple generations allows examination of long-term transmission processes
- Comparative analysis of sibling outcomes
- Opportunities to explore linkages outside the household

Some key features of Understanding Society

- Very large sample size proposed (40K households)
- Large sub samples(4,000 Scottish households)
- Representative sample of whole population
(all ages)
- Multi-purpose multi-topic design to meet a wide range
of disciplinary and inter-disciplinary research needs

Some key features of Understanding Society

- Ethnic minority research
- Research linking social and biomedical sciences
- Innovation in data collection methods

Understanding Society Sample

- Approx. 27,000 households - The fieldwork for this sample will commence in January 2009
- A boost ethnic minority sample, focussed on five main ethnic minority groups, comprising 4,000 households
- Incorporating the BHPS sample of approximately 8,400 households
- An Innovation Panel of 1500 households to enable methodological research (panel began in January 2008)

Understanding Society Sample

- 40K households gives an opportunity to explore issues where other longitudinal surveys are too small
- Small subgroups, such as teenage parents or disabled people

Understanding Society Sample

- Analysis at regional and sub-regional levels, allowing examination of the effects of geographical variation
- Large sample size allows high-resolution analysis of events in time, for example focussing on single-year age cohorts

UK HLS Opportunities

- Starting again, compared with BHPS, an opportunity to review activities and see which are worthwhile to continue, which not
- Focus on new research issues
- Opportunities for mixed methods:
 - Data linkage admin, organisation, spatial
 - Bio-markers and health indicators
 - Qualitative data
 - Other non-standard data: diaries, visual, audio

UK HLS Opportunities

- Use of different modes
 - e.g. web to collect data with higher frequency
 - Experiment with new technology as it is introduced
- Overall aim - to build a robust survey structure within which can experiment and innovate while minimising risk



Understanding Society: the UK Household Longitudinal Study



<http://www.understandingsociety.org.uk/>

A now for something related...

Weights – A very brief introduction

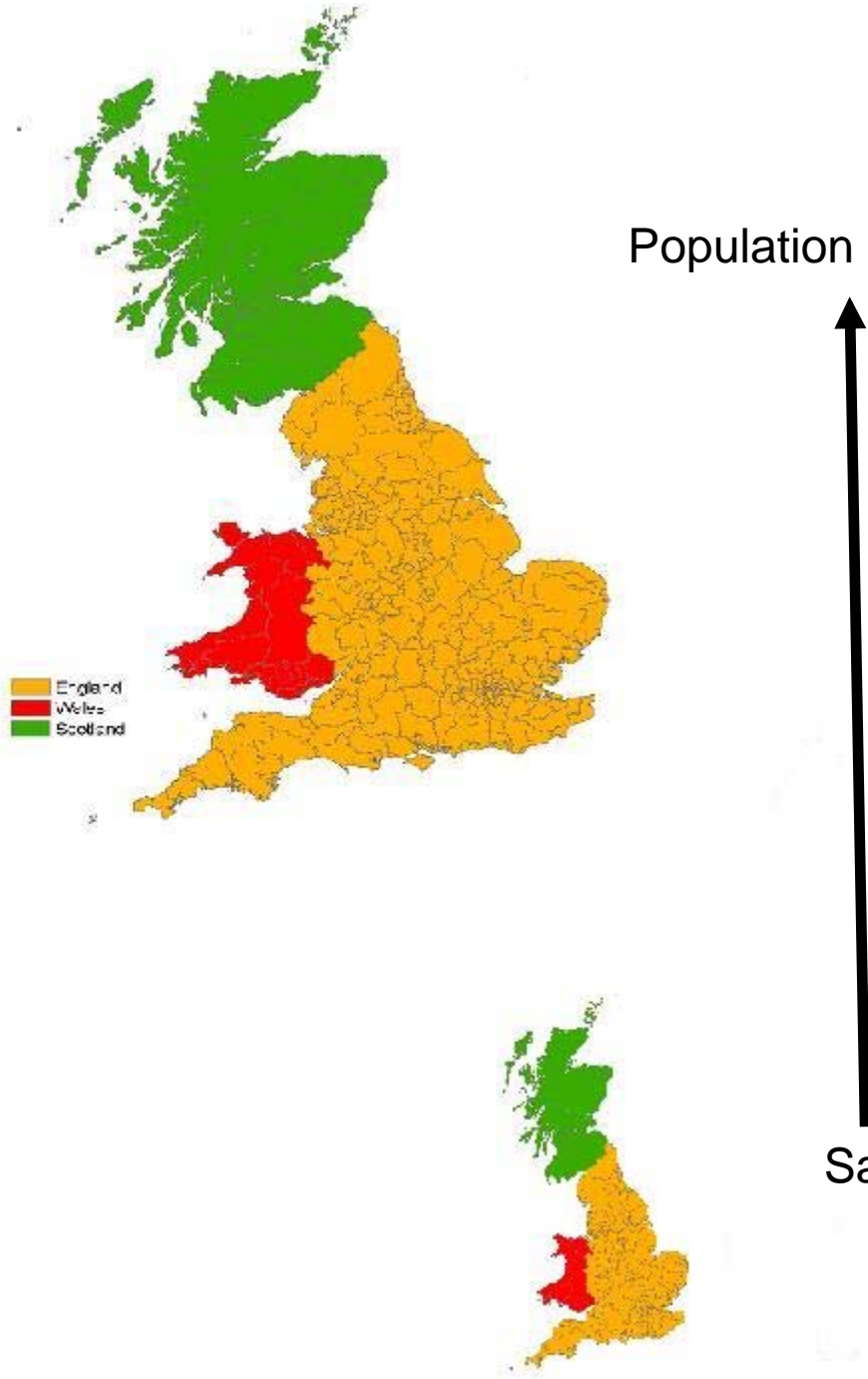
- Weighting data is a good practice in social survey data analysis
- The appropriate weight to use will depend on your specific analysis

Weights – A very brief introduction

- There are separate weights for each BHPS Wave
 - Respondent individuals
 - Enumerated individuals
 - Households

The basic principles

- Cross-sectional weights
- Regional sample weights
- Longitudinal weights



Adjust for unequal selection probability and non-response

(regional samples are examples of unequal selection probabilities)

Often $wXRWGHT_{18}$

Regional Samples

- The regional samples (boosts) facilitate
 - Inter-regional or comparative analyses (e.g. England / Scotland)
 - Regional analyses (e.g. Scotland only)

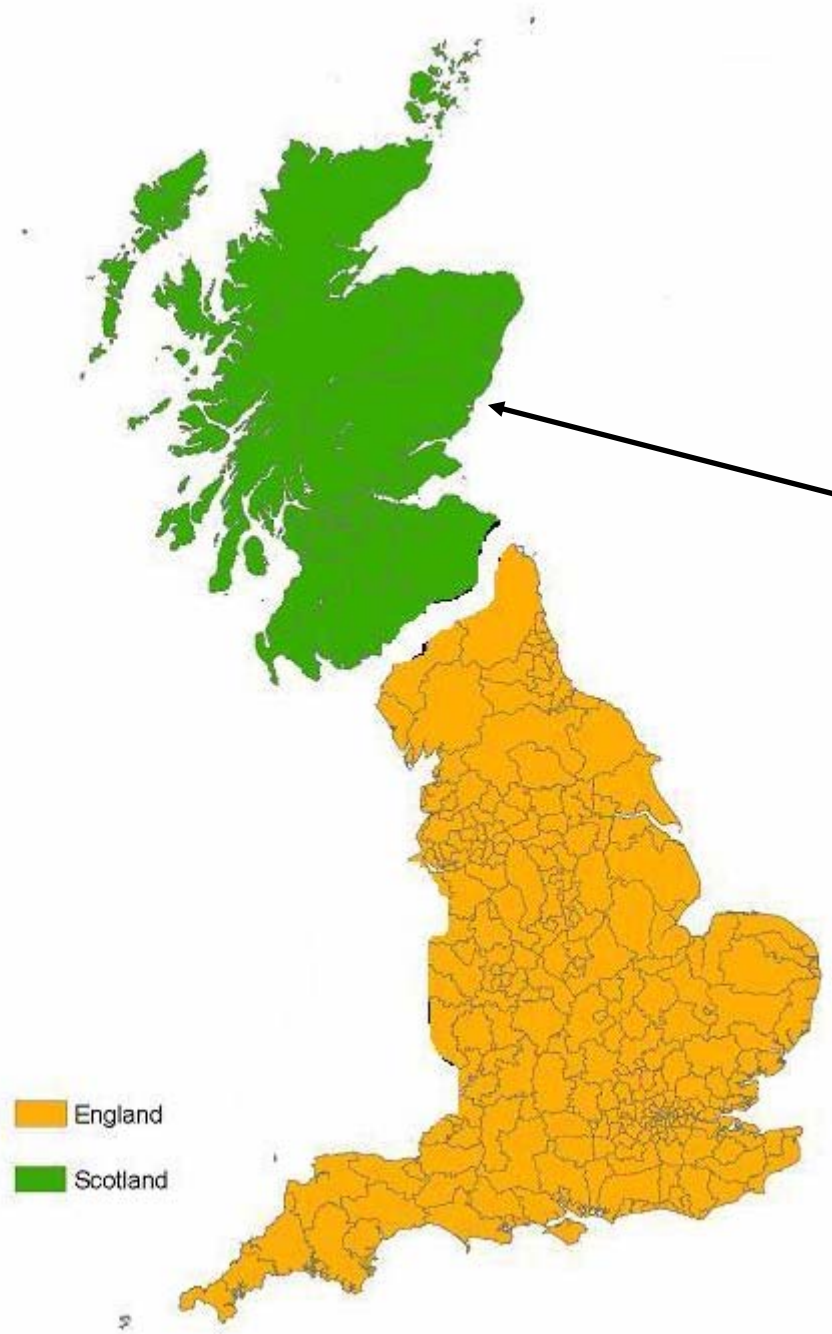
Wave i (1999)

Original sample	n=9,101	(58%)
Scottish booster	n=2,446	(16%)

The Scots are over sampled -

Put simply in 1999 there were 49 million English people and 5 million Scots

Weights for regional samples (booster samples)



Scotland is 'over' sampled

Examples from the data

PID	iSEX	iMEMORIG	iREGION	iXRWTW1
10830766	male	original	O.London	.89
65434463	male	original	Scotland	.24
97568023	male	Scotland	Scotland	.24

wXRWTUK1 from wave 11 (2001) onwards

Weights for the analysis of regional samples (booster samples)



Examples from the data

PID	iSEX	iMEMORIG	iREGION	iXRWTW2
10830766	male	original	O.London	.602
65434463	male	original	Scotland	.597
97568023	male	Scotland	Scotland	.604

wXRWTUK2 from wave 11 (2001) onwards

Longitudinal Weights (wLRWGHT)

For many panel analyses only the cases who give a full interview are of interest

PID	aIVFIO	bIVFIO	cIVFIO	cLRWGHT
10004521	full int.	full int.	full int.	1.691
10002251	full int.	.	.	.
10042571	full int.	.	full int.	0
10087451	.	.	full int.	0

Comments

- Most statistical software packages do not calculate standard errors for weighted data correctly
- If you require weighted analyses then move from SPSS to Stata
 - *the svy suite is specially designed for survey data*

Comparing Stata & SPSS

```
use "D:\home\vgayle\BHPS_weights\kind2.dta", clear  
numlabel _all, add
```

```
svyset kpsu [pweight= kxrwtuk2], strata(kstrata) ///  
    singleunit(scaled)
```

Without the the 'singleunit (scaled)' stata will not estimate the standard errors and reports the following note

"Note: missing standard errors because of stratum with single

```
. svy:mean kpaygu, over(kregion)
(running mean on estimation sample)
```

Survey: Mean estimation

```
Number of strata = 121      Number of obs = 8698
Number of PSUs = 399      Population size = 8063.42
Design df = 278
```

```
_subpop_1: kregion = 17. wales
_subpop_2: kregion = 18. scotland
_subpop_3: kregion = 19. northern ireland
_subpop_4: kregion = 20. england
```

Over	Linearized			
	Mean	Std. Err.	[95% Conf. Interval]	
kpaygu				
_subpop_1	1242.017	29.55327	1183.841	1300.194
_subpop_2	1378.118	33.89045	1311.404	1444.833
_subpop_3	1249.98	.	.	.
_subpop_4	1446.1	20.67361	1405.403	1486.797

Note: variance scaled to handle strata with a single sampling unit.

end of do-file

* SPSS Comparison

	mean	s.e.	lower	upper
Wales	1242.017472	29.43089575	1184.081753	1299.95319
Scotland	1378.118036	33.75012108	1311.679776	1444.556296
N.I.	1249.980248	0	1249.980248	1249.980248
England	1446.099938	20.58800068	1405.571759	1486.628117

Comments

- Good practice to use weights
- Appropriate weight depends on your specific analysis
- For certain analyses weights may not be available
 - Think about sub-optimal weights
 - Think about constructing your own weights
- In practice many analysts use weights for descriptive statistics but do not use them in multivariate analyses
 - *BUT ALL GOOD RESEARCHERS CONSIDER THE IMPLICATIONS OF NON-RESPONSE FOR THEIR ANALYSIS*

Comments

- Important to consult Section V of the BHPS User Manual Volume A
- Table 25 is extremely helpful
- Further reading...

Lynn, P.J. (2005) 'Weighting', in the *Encyclopedia of Social Measurement*, Vol. 3, Elsevier Inc.