

A Short Introduction to the British (Quarterly) Labour Force Survey...

That is for those who actually want to start using it. If you have used this introduction and find it useful, in way of saying thank you may wish to send me a bottle of OMAS Amerigo Vespucci Red. (Yeah right, I'm only joking...)

Relatively important facts of LFS

1. Data are downloadable from the Data Archive (<http://www.data-archive.ac.uk>) in 2 Formats:
 - ✳ By year and quarter (with or without UACNTY and UALA) – Panel
 - ✳ By cohort – Longitudinal
 - ✳ Other summary data sets by area are also available.

2. User guides in PDFs are downloadable on ONS's website:
<http://www.statistics.gov.uk/statbase/Product.asp?vlnk=1537&More=N>
 - Vol 1: Background and methodology
 - Vol 2: LFS Questionnaire
 - Vol 2a: Transitional Questionnaire
 - Vol 3: Details of LFS variables
 - Vol 3: Details of LFS variables (1992-2002)
 - Vol 4: LFS Standard Derived Variables
 - Vol 5: LFS Classifications
 - Vol 6: LFS Local area data
 - Vol 7: LFS variables 1979-91
 - Vol 8: Household and family data
 - Vol 9: Eurostat and Eurostat derived variables

These User Guides are easiest to use in their original PDF formats as you can search through them inside the Acrobat Reader.

3. Sample drawn using postcodes (except for a small sample north of the Caledonian Canal where they use telephone numbers). Thus the sample is evenly distributed geographically over whole country. No clustering in cities.

4. Most people are interviewed 5 times (waves) – first time face to face, the rest via telephone.

5. Some people are interviewed *by proxy*, i.e. they weren't there at the time of interview and someone else answered the interviewer's questions *for* them. People who live alone might be under-represented. Sometimes the data are brought forward from previous quarter for various reasons. All these are indicated by IOUTCOME.

6. Quarters are defined as follow:
 - ✳ Spring – March to May
 - ✳ Summer – June to August
 - ✳ Autumn – September to November
 - ✳ Winter – December to February next year.

7. Anyone who works more than an hour a week would be prompted for questions on their income (ILO's definition for labour market participation). Thus covering more people who are on low pay than New Earnings Survey (NES) (but shouldn't be the case for the newly introduced Annual Hourly Earning Survey that replaces NES), although there are studies that show even LFS overestimated pay. See: <http://cemmap.ifs.org.uk/docs/cwp1203.pdf>
The names of the wage variables have changed in 1999.

8. Income data from 1992 to 1997 are only collected in the 5th wave, from 1997 it is collected in both 1st and 5th waves. Approximately **17000 observations** on hourly pay per quarter.
9. Individuals can be identified across quarters by a list of *System Variables*:
 - ✳ Interviewer's stint number (QUOTA)
 - ✳ Week they were interviewed (WEEK, also good for working out tenure, etc.)
 - ✳ Year that address first entered survey (W1YR)
 - ✳ Quarter that address first entered survey (QRTR)
 - ✳ Address number on interviewer's list (ADD)
 - ✳ Household number – there may be more than 1 household at the same address (HHLID)
 - ✳ Person ID within household (RECNO)
 - ✳ WAVFND

After you've appended the quarterly data into a panel, you might wish to generate a personal identifier (pid) and a household identifier (hid) using STATA command:

```
egen pid=group(quota week wlyr qrtr add hhld recno wavfnd)  
egen hid=group(quota week wlyr qrtr add hhld wavfnd)
```

10. You might want to weight your estimations. LFS comes with 2 sets of ready-made weights that were created using information gathered at Censuses: *Population* weights and *Income* weights. People of certain demographic groups are more likely to be left out than others and same can be said for people of certain income brackets. See User Guide Volume 1, Section 10 for more info.

Things that you might hate about LFS

1. Some questions are only asked in some quarters – usually in Spring and Autumn, or just Spring, or just Autumn. Gives you headaches when you're looping through the datasets, your loops will have to have built-in mechanisms that response to the differences across quarters.
2. Quite a few **new questions** were introduced in 1999 – it renders some estimation impossible or that you would only have 5 years (writing as at August 2005, roughly 19 cohorts) of data.
3. **Geographical variables:** LFS has redefined the geographical boundaries of Unitary Authorities and Local Authority Districts (UA/LA) in 1999 following the massive boundary change in 1996. County no longer exists as a geographical definition within LFS's geographically framework from 1999. Consequently if you're mapping UA/LA into counties, then the counties before and after 1999 aren't quite the same. Some of the common ways to chop UK into bits are:
 - ✳ Government Office Regions
 - ✳ Unitary Authorities + Counties (e.g. UACNTY, not available on general release)
 - ✳ Unitary Authorities + Local Authority Districts (e.g. UALA) With this you can map observations to, e.g., Health Authorities and Strategic Health Authorities using the *NHS Organisational Manual*, postcodes using *All Fields Postcodes Directory*. Unfortunately If more information on the geography of UK is needed visit:
http://www.statistics.gov.uk/geography/beginners_guide.asp
 - ✳ Local Education Authorities (LEA variable is not in general release!)