

## **LIS Policy on the Treatment of Shadow files (Wave V and beyond)**

### **Description**

In some surveys, individuals and/or households were interviewed even if they were not included in the representative cross-section. (These include, for example, an oversampling of new immigrants starting in a late wave in the German panel, and split-off households in Russia.) LIS refers to these as the wave observations that are not included in the cross-section. These observations have valid values for the survey and LIS variables, but have zero or missing cross-sectional household weights. Rather than discard this information, LIS has included these observations in a “shadow” file for the relevant country and year. (See country-specific documentation for shadow files for details of their use.) While they are not part of the representative cross-section, the information could be of interest to researchers concentrating on a specific sub-sample (e.g., recent immigrants or young families) where weights are not required in the analysis. In some cases, the data available in the shadow files are extensive.

In countries with cross-sectional surveys, zero-weighted households are usually excluded by the data provider and the remaining households are reweighted accordingly. In some cases, however, there may be many individuals with partially missing information who are zero-weighted, but included in the data provider’s file. Using the shadow file, therefore, could expand sample size and add insight into specific groups or types of information.

Shadow files are most likely to exist in countries where the original data LIS use are panel data. In panel data, individuals or households may be added to, or remain in, the current data set even if they do not belong in the representative cross-section for a given year. In those cases where the data provider interviews the individuals who are not members of the cross-section, LIS includes this information in the shadow files.

### **Shadow files in Wave V**

In Wave V, eight countries provide information that LIS includes in both the household- and person-level shadow files:

#### Austria (AT00)

The shadow files include the same information as the regular file, including dpi. These appear to be households that were erroneously interviewed in the 2001 wave of the panel.

There are 13 households that are not part of the cross-section for 2000, but which are part of the longitudinal sample. These households include 21 interviewed adults and 7 children.

### Belgium (BE00)

The shadow files include the same information as the regular file, including dpi.

There are 370 households that are not part of the cross-section for 2000, but which are part of the longitudinal sample. These households include 565 interviewed adults, 93 non-interviewed adults, and 212 children.

### Germany (DE00)

Zero-weighted households include West German households with new immigrants and East German households that relocated to West Germany after June 1990. Only households from GSOEP Subsample D (Immigrants) have zero weights.

There are 164 households that are not part of the cross-section for 2000, but which are part of the longitudinal sample. These households include 310 interviewed adults, 14 non-interviewed adults, and 138 children.

### Estonia (EE00)

Data for Estonia come from the Household Income and Expenditure Survey. This survey is conducted through diaries. The four modules are: Household Picture; Income; Non-food Expenditure; and Food Expenditure. The shadow files consist of households (and their members) that took part to the interview and filled the Household Picture, but did not complete the Income and/or Expenditures diaries.

There are 1,885 households that are not part of the weighted cross-section for 2000. These households include 1,736 that did not complete the Income diary; 15 that did not complete the Non-food Expenditure diary; and 134 that did not complete the Food Expenditure diary. At the person level, this corresponds to 3,859 adults and 757 children without an Income diary; 24 adults and 4 children with no Non-food Expenditure diary; and 263 adults and 55 children without a Food Expenditure diary.

### Luxembourg (LU00)

The shadow files include the same information as the regular file, including dpi.

There are 2 households that are not part of the cross-section for 2000, but which are part of the longitudinal sample. These households include 2 interviewed adults and 0 children.

### Poland (PL99)

Data for Poland come from the Household Budget Survey. These households do not report income and, therefore, dpi cannot be computed.

There are 90 households that are not part of the weighted cross-section for 1999. These households include 192 interviewed adults and 41 children.

### Russia (RU00)

The Russia Longitudinal Monitoring Survey is a longitudinal annual survey. Individuals and families that separate from a weighted household (split-off households) are followed and interviewed, but are not included in the weighted sample. The shadow files include the same information as the regular file, including dpi.

There are 687 households that are not part of the weighted cross-section for 2000. These households include 1,555 adults, 461 children, and 1 individual with no reported wage assumed to be an adult (based on household relationships).

### **Shadow files in Wave VI**

In Wave VI, no countries provide information that LIS includes in both the household- and person-level shadow files.