

lar payments; income from long-term care insurance (only for Austria and Germany).

The available data at the household level include: income from household members not interviewed; income from other payments, such as housing allowances, child benefits, poverty relief, etc.; income actually received from secondary homes, holiday homes or real estate, land or forestry; capital income (interest from bank accounts, transaction accounts or saving accounts; interest from government or corporate bonds; dividend from stocks or shares; interest or dividend from mutual funds or managed investment accounts). For homeowners, the data at the household level also include imputed rent, based on the self-assessed home value minus the net residual value of the debt (payments for mortgages or loans). The interest rate used for imputed rents is fixed at 4% for all countries.

The SHARE definition of income does not include home business and „other types of debts“: in the latter case we are not able to separate the amount of the debts on cars and other vehicles from the total amount of debts.

*Imputations:* Whenever a respondent did not know or refused to give the exact amount in a certain question, unfolding brackets (UB) questions were asked to recover that value (see above). Different cut-offs were used across countries.

As far as UB observations are concerned, we implemented a simple hot-deck procedure to impute values for those cases in which the exact amount are missing. At this stage, only the amount variable is imputed. Also, we imputed one variable at a time and did only one round of imputations for each variable. No stratification was made, except by country (due to the differences in the cut-offs).

In the event of a „refusal“ or „don't know“ answer to all UB questions, we stratify by country and age classes, except for financial assets, where income is computed on the basis of the stock values (whether exact records exist or just imputed).

In the event of “invalid” („refusal“, „don't know“, or missing) values on frequency variables (for instance the period covered by a payment and the number of months in which the respondent has received the payment in 2003), a linear regression technique was applied to impute such frequencies. In particular, we used the linear regression only for the frequencies of received pension. The regression conditions upon the following independent variables: age, sex and dummy indicators for whether the associated amount variable belongs to the intervals defined by the 1st, 2nd, and 3rd quartile.

We produce the estimated coefficients for each frequency variable within each country. In a few cases the hot-deck procedure may fail because there are no donors that can be used for that specific interval.

### 7.11 Wealth Imputation

*Dimitrios Christelis, Tullio Jappelli, and Mario Padula*

*The Definition of Wealth:* SHARE contains the following information on the ownership and value of the following assets.

- Real assets, i.e. the ownership and value of the primary residence, of other real estate, of the share owned of own businesses and of owned cars.
- Gross financial assets, i.e. the ownership and value of bank accounts, government and corporate bonds, stocks, mutual funds, individual retirement accounts, contractual savings for housing and life insurance policies.
- Mortgages and financial liabilities.

The values of these variables are summed over all household members in order to generate the corresponding household-level variables. As with income, whenever a respondent did not know or refused to give the exact amount in a certain question, unfolding brackets (UB) questions were asked to recover that value, where different entry points were used across countries.

*Imputations:* Imputation is performed using the hotdeck imputation package in STATA, which is based on the approximate Bayesian bootstrap described in Rubin and Schenker (1986). This procedure requires the classification (by some variables, e.g. unfolding bracket values, age, etc.) of the non-missing observations in cells, from which bootstrap samples are drawn and values from these samples are used to impute the missing observations in each.

We impute asset values in two steps. (1) If an individual gives a response of „don't know“ or refuses to answer the ownership question, then ownership is imputed. The imputation is done using country and age as classificatory variables for the hotdeck procedure. (2) The amount is imputed when ownership is imputed, when the individual gives a response of don't know/refusal and either does not start the unfolding brackets procedure, does not complete it, or completes it without giving a specific amount as an approximate answer, or when the original answer is deemed illegitimate for other reasons.

In the end we divided the variables into three groups according to the criteria by which the cell classification for imputation was made (all imputations were made separately for each country).

- Housing, bank accounts and cars: These variables contained numerous positive non-missing values, reflecting the wide ownership of the corresponding assets. In the case in which we did not know the bracket value we used age as an additional variable. When we knew the bracket value, we used it together with age.
- Mortgage: We needed to link the value of the mortgage to the value of the house, in order to avoid as much as possible the case where the imputed value of the mortgage was greater than the value of the house. Thus, when we did not know the bracket value of the mortgage, we used the bracket value of the house as a classificatory variable; when we knew the bracket value of the mortgage we used it for the imputation and we excluded the bracket value of the house because its inclusion would have made the cells too thin.
- Other real estate, bonds, stocks, mutual funds, individual retirement accounts, contractual savings for housing, life insurance, own business and owned share thereof and financial liabilities: These variables exhibited relatively few positive non-missing values. We used age to define the imputation cells when we did not know the bracket value, while we used the bracket value for their definition when we knew it.

## 7.12 Methodological Issues in the Elicitation of Subjective Probabilities

*Luigi Guiso, Andrea Tiseno, and Joachim Winter*

Non-response rates for the subjective expectations questions are generally low. For the “sunny day” question, the non-response rate is 3.2% and for the subjective survival question it is 7.9%. There is only minor variation of non-response rates across countries—the smallest non-response rates (below 5%) are observed in Austria, Switzerland, and Germany; the largest non-response rate to the subjective survival question of about 15% in Spain.