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.

An issue that has received some attention in the literature on probabilistic expectations is rounding to certain "focal values" (in particular, to 0%, 50%, and 100%, and to a lesser degree to other multiples of 10%). Even more striking than rounding is the excessive use of 50% responses. Some authors argue that a 50% response reflects "epistemic" uncertainty about the event in question (e.g., Bruine de Bruin, et al. 2002). In this case, 50% responses would be similar to a "don't know" response, and they would have to be dealt with differently than other multiples of 10% generated by rounding. While a deeper analysis of this issue is beyond the scope of this paper, it is nevertheless interesting to see whether the phenomenon of rounding and excessive 50% responses is present in the SHARE data as well, and even more importantly, whether there are any striking differences in response behaviour to probabilistic expectations questions across participating countries.

For instance, an analysis of the responses to the "sunny day" question confirms findings of other surveys such as HRS: Most of the responses are at focal values, in particular multiples of 10%, with a peak at 50% that cannot easily be explained by rounding. Overall, however, only about one fifth of all responses are at 50%, which is less than what has been found in other surveys. In the SHARE data, the prevalence of 50% responses is similar in all questions—between 20% and 30% of all responses. Second, there is some variation across countries. The question with the largest degree of cross-country variation in the use of 50% responses is the "sunny day" question, and it seems likely that the observed differences are due to actual differences in weather conditions and not in response behaviour—the Mediterranean countries simply have better weather, so the entire response distribution should be shifted to the right, reducing the number of 50% responses. For the other questions, the variation is rather small.

Future research will have to test whether these differences correctly reflect differences in the underlying expectations across countries or whether there are country-specific response styles for probabilistic expectations questions. Another methodological issue related to probabilistic expectations questions is whether there is a general tendency by respondents to be optimistic (i.e., to report high probabilities for positive and low probabilities for negative events) in hypothetical choice questions. A first impression of whether this effect exists can be obtained by correlating responses to a question that likely reflects an individual's overall optimism (in the case of SHARE, we use the "sunny day" question for this purpose) with the responses to substantive probabilistic questions.

Table 3 shows the correlation of responses to the substantive expectations questions with responses to the "sunny day" question. While all correlations are statistically signifi-

Table 3 Correlation of Responses to the Substantive Expectations Questions with Responses to the "Sunny Day" Question

Question		Correlation	p-value	N
ex007	Decrease in pensions	-0.0755	0.000	6240
ex008	Increase in retirement age	-0.0563	0.000	6268
ex009	Survival to target age	0.0979	0.000	15108
ex010	Better standard of living	0.1154	0.000	15618
ex011	Worse standard of living	0.0262	0.001	15531

Notes: Reported correlation are Spearman's rank correlation coefficients. The p-value is for the null hypothesis that the row variable is independent of the response to the "sunny day" question.

cant at any conventional confidence level (due to the large sample size), the absolute size of the correlation coefficients is small, which can be taken as evidence against a general tendency to be optimistic or pessimistic.

References

Bruine de Bruin, W., P. S. Fischbeck, N. A. Stiber, and B. Fischhoff. 2002.: What number is Fifty-Fifty? Redistributing excessive 50% responses in elicited probabilities. *Risk Analysis* 22:713-23.

Cutler, D.M. and E. Richardson. 1997: Measuring the Health of the U.S. population. *Brookings Papers on Economic Activity: Microeconomics 1997*, 217-71.

Hoynes, Hilary, M. Hurd and H. Chand. 1998: Household Wealth of the Elderly under Alternative Imputation Procedures. In *Inquiries of Economics of Aging*, ed. David Wise, 229-57. Chicago: The University of Chicago Press.

Rubin, D.B and N. Schenker. 1986: Multiple imputation for interval estimation from simple random samples with ignorable non-response. *Journal of the American Statistical Association* 81:366-74.