Oversampling of the wealthy in the Spanish Survey of Household Finances (EFF)

Olympia Bover[1]

1. Introduction

One distinctive characteristic of the EFF (Encuesta Financiera de las Familias, or Survey of Household Finances), which follows the example of the SCF (Survey of Consumer Finances) in the US, is that it oversamples wealthy households. The distribution of wealth is heavily skewed; moreover, some types of assets are held by only a small fraction of the population. Therefore, it was considered important to have a sample that would be representative not only of the population but also of aggregate wealth, and that would facilitate the study of financial behaviour at the top of the wealth distribution.

This paper describes oversampling of the wealthy carried out in the Spanish Survey of Household Finances (EFF). This was achieved through a collaboration involving the National Statistics Office (INE), the Tax Authorities (TA), and the Banco de España (BE), employing an elaborate coordination mechanism that enabled the TA's strict confidentiality requirements to be observed at all times. A complex procedure for replacing non-responding households was incorporated in the sample design to ensure oversampling in the final result. Details are given, below, on the degree of oversampling in the final sample and on some practical problems encountered, along with examples of the benefits of oversampling.

2. Designing the sample

Basis for oversampling of the wealthy

Spain has a wealth tax ("Impuesto sobre el Patrimonio"), and the EFF oversampling is based on data from the individual wealth tax files. This is in contrast with the SCF, where a wealth index is constructed by drawing on information about asset income from individual income tax files, since there is no wealth tax in the US. In 1999 (the tax year used in selecting our sample), people subject to the wealth tax in Spain were those with taxable wealth of over 104,000 euros. In 1999, approximately 980,000 individuals (some 700,000 households) filed a wealth tax return.

The wealth strata were defined on the basis of households' percentile positions in the wealth tax distribution. We defined eight strata, which were oversampled at progressively higher rates. Strata 2 and 3 capture slightly less than half of the distribution of taxable wealth. Strata 4, 5 and 6 capture the third and fourth quartiles, except for approximately the last percentile and a half, which fall within the last two strata.

Finally, in Navarre and the Basque Country, there was no oversampling of the wealthy, since the National Tax Office does not maintain personal tax file information for those regions.

---

[1] Directorate General, Economics, Research, and Statistics, Banco de España.

Confidentiality guarantees

The Tax Office is subject to very stringent confidentiality requirements, and is not permitted to release any personal tax information (even in the form of intervals) to anyone, including the Statistics Office. To overcome this problem and allow for wealth tax oversampling while preserving confidentiality, the National Tax Office volunteered to select the random sample itself, following the sample design requirements specified by the Banco de España and the National Statistics Office.

As a result of the collaboration of the Statistics Office and the Tax Office, there is a unique population framework for the sampling: the mid-2001 Regular Municipal Census, in which the units consist of households, as defined by address. With this information from the Statistics Office, the Tax Office constructed three variables for each address, based on both wealth and income tax information. These data served as the starting point for the sampling.

The first variable, the wealth stratum indicator, is based on total household declared taxable wealth, which was obtained by adding up the returns of all relevant household members. The second variable, applied to those filing income tax but not wealth tax, indicates the quartile, in the national taxable income distribution, to which the household belongs. Finally, information on household per capita income was added. The income variables were helpful in the selection of sample replacements (as we shall see below) and in ensuring that households from all income levels were included in the sample. This inclusion was achieved by using random start systematic sampling in an appropriately ordered data frame. Furthermore, the income quartile indicator was used to correct for non-response in large cities. The tax information available at the time was dated 1999, which created a limited degree of mismatch between the two sources.

Sampling

First wave

The sampling design varied for each of the three following cases:

1. Municipalities with over 100,000 inhabitants. For large towns, the sampling was random within the eight wealth strata.
2. Municipalities with 100,000 or fewer inhabitants. For small municipalities, the sampling was based on a two-stage cluster design, with the primary sampling units (PSUs) being selected first, with probabilities proportional to their population. Within PSUs, the selection of households differed according to the number of wealth tax filers in the PSU.
3. Finally, in Navarre and the Basque Country, where no oversampling of the wealthy was possible because the National Tax Office does not maintain personal tax file information for those regions, the sample was selected according to a two-stage stratified cluster design, with six strata defined according to size of municipality.

For reasons of confidentiality, stratum and cluster indicators cannot be provided. To calculate appropriate variance formulas, however, replicate weights are provided instead.

Second wave

The second wave of the EFF was designed as a panel with refreshment sample by wealth strata. The general principles used in the first wave of the survey were followed in the second. Specifically, the aim was to re-interview the 5,143 households that participated in the EFF2002 and to complement them in such a way as to obtain a sample representative of the 2005 population, while preserving oversampling of wealth.

For large municipalities, the refreshment sample required to obtain the desired 2005 sample was selected by wealth strata. For small municipalities there was no oversampling of the wealthy for the refreshment sample. In these cases a two-stage sampling design was performed for the same PSUs used for the EFF2002.

Replacements

Another important aspect of the EFF sample design was the replacement scheme. To preserve the oversampling scheme as much as possible, tightly controlled replacements were chosen. The use of controlled replacements is similar to post-stratification and weight adjustments done within cells when data collection is completed. In our case, having controlled replacements was an important advantage in that, since we had no indication of the wealth stratum to which the sample households belonged, no "directed" effort could be made during the field work, if we discovered particularly low response rates for certain strata.

Specifically, up to four replacements were provided for each household originally in the sample, and these were to serve as replacements for that household only. The replacements selected were the two households immediately above and the two immediately below the household's ranking based on income quartile (for non-wealth tax filers), wealth stratum, and per capita household income. Replacements had to belong to the same income quartile (for those not paying a wealth tax) or wealth stratum as the sample household. This was carried out within municipalities in the case of large cities, and within PSUs in the case of small ones, to ensure that replacements would not be too distant geographically from the original sample household. In some cases this meant that fewer than four replacements (in a few instances, none at all) were available. In Navarre and the Basque country, a more standard scheme was applied, in which a pool of eight replacement households was used, providing potential substitutes for eight sample households (within the same PSU).

3. Non-response

One of the characteristics of wealth and income surveys, due to the nature or difficulty of the questions asked, is a high unit non-response rate.

Could not establish contact (never at home)

The number of households for which the interviewer was unable to find anyone at home (having confirmed with neighbours etc that the household address was correct), despite at least five attempted visits, was very high. The number of these failed contacts as a proportion of the total number of attempted contacts varies with wealth stratum in a way that is not entirely random. Use of multiple residences was considered a possible reason for failure to establish contact with high wealth people during the fieldwork.

Refusal

There is a clear non-random component in cooperation rates – defined as completed/ (completed + refused) – which decrease as we move up the wealth strata, falling from 53.6% to 29.4%. It is clear from this pattern that overall cooperation or response rates are not very informative in the case of oversampling, since they are dependent on the degree of oversampling. To establish some meaningful comparison, we constructed cooperation rates by strata for the 1992 SCF. These cooperation rates for the list sample ranged from 52.6% for stratum 1 to 20.1% for stratum 7.

Adjusting sample weights to correct for unit non-response

To compensate for differential unit non-response, sample weights are adjusted within the cells defined by the various sampling frame variables, including, in particular, wealth strata and income quartiles. Given the above-mentioned confidentiality restrictions, sample and non-response weights are calculated by the Tax Office pursuant to detailed instructions from the Statistics Office.

4. Oversampling, and its usefulness, in the final sample

Degree of oversampling in the final sample

Before concluding, let us examine some figures indicating the degree of oversampling in our final sample.

Overall, slightly over 40% of the households that completed the interview are wealth tax filers (a figure obtained, due to confidentiality constraints, with the assistance of the Tax Office). Furthermore, aggregate tax return information indicates that 0.4% of the population of households hold 40% of total taxable wealth. We would therefore expect to have at most 20 such households in a random sample of 5,000 – an upper bound, since it assumes a non-differential rate of response. In contrast, our sample contains over 500.

In terms of actual net worth in the 2002 EFF data, we calculate oversampling rates at various points in the distribution. The oversampling rate is defined as the ratio of the number of observations actually in the sample for a specific percentile range of the distribution to the number of observations one would expect if the sample were randomly drawn from the population. A progressive oversampling of the wealthy was achieved. Specifically, while for the bottom 50% the rate is less than 1% (0.73%, to be precise), for the wealthiest 1% we have over seven times the number of observations we would have with a random sampling. In between, the rates are 0.94% for the 50th to 90th net worth percentile group, 1.67% for the 90th to 95th, and 2.55% for the 95th to 99th.

Some examples of the benefits of oversampling

We know from the EFF2002 that the wealthiest 10% of households hold 42% of net wealth in Spain, and invest 10% of their financial assets in unlisted shares. However, only 10% of these actually hold unlisted shares. Therefore, in a random sample of 5,143 (which is the size of the EFF2002 sample) we would expect to have 52 households owning unlisted shares, whereas the EFF2002 actually contains 292 such households. In the case of fixed income securities, the expected number would be even smaller, since only 4.5% of the households hold such assets. Hence we would expect around 23 households to represent that group, whereas there were 104 in the EFF2002.

Oversampling is also essential for precision and hypothesis testing in cross-country comparisons of certain routinely reported wealth statistics. For example, the percentage of wealth held by the top 1% is 13% in Spain. The standard error of this figure (with oversampling) is 1.6. If we calculate the bootstrap standard error that would have resulted from randomly sampling the Spanish population, the figure is 5.3.[2] In that case, the 95% confidence interval would be as large as the international variation of 20 percentage points found in the literature.

---

[2] Spanish population obtained from the EFF2002 sample and its population weights.