Farm Business Survey - Statistical Information

Sample representation and design

The sample structure of the FBS was re-designed for the 2010/11 accounting year. The coverage of the survey is now restricted to those farms which have at least 25,000 Euros of output.

The population of farms covered by the survey is stratified for sample selection purposes into 14 farm types and 7 regions. The FBS is based on a uniform sampling rate within each stratum. However, minimum publication thresholds are applied and some farm types are sampled at a higher rate to ensure adequate coverage for analysis. The FBS is a panel survey with around 93% of the sample retained annually.

For the 2004/05 to 2009/10 accounting years, rather than being defined by standard output, the minimum size for entry to the survey was defined by standard labour requirements. The FBS covered part time and full time farms - it did not cover spare-time farms (<0.5 SLR) or farms with negligible economic activity. For 2003/04 and earlier years the FBS covered all farms that had a minimum size threshold of 8 European size units (ESU)[1].

Sample Size

For 2015/16, the sample size for the English FBS is 1800 farms (see Table 1 for sample quotas by farm type). From 2016/17 onwards the sample size will be 1750 farms.

Sampling and Recruitment

Around 7% of the FBS sample is replaced each year due to natural turnover of the sample. Replacements are selected at random with uniform probability within each design stratum. Replacement addresses are obtained from the June Survey. Although the June Survey register is a register of holdings, these are amalgamated into businesses prior to selecting FBS replacements, using information from the Basic Payment Scheme and other sources. The business addresses are then sorted into random order within each design stratum before being despatched to FBS Research Centres. FBS Research Centres recruit replacement farms by systematically working down the address list in the order given.

[1] A European Size Unit (ESU) is a measure of the economic size of a farm business based on the gross margin imputed from standard coefficients for each commodity on the farm. The application of these standard coefficients results in the Standard Gross Margin (SGM) for a farm or group of farms. 1 ESU = 1200 SGM.

Table 1: Sample quotas for the 2015/16 FBS by farm type

Farm Type (Standard Output Classification)	**2014/15**	**2015/16**
Cereals	290	300
General Cropping	148	145
Specialist Fruit	50	45
Specialist Glass	60	40
Specialist HNS	35	35
Other Horticulture	50	35
All horticulture	195	165[(a)]
Dairy (LFA)	36	40
Dairy (Lowland)	169	125
Grazing Livestock (SDA)	120	120
Grazing Livestock (DA)	62	55
Grazing Livestock (Lowland)	250	251
Mixed	156	142
Specialist Pigs	60	60
Specialist Poultry	80	80
EU regional Supp.	253	240
National Reserve	81	77
Total	**1900**	**1800**

(a) There is a float of 10 horticulture farms for 2015/16 that have not been assigned to individual horticulture sub-categories.

Sampling Errors

Results from every sample survey will have a degree of sampling error because only part of the population is being used to estimate the value of a variable. The sampling error is the difference between the estimate derived from a sample survey and the 'true' value that would result if a census of the whole population were taken under the same conditions. Different samples will yield differing estimates for the same observation variable. Sampling error is quite separate from measurement and processing errors and biases due to non-random selection or response. It is a random error arising from random selection of part of the population.

It is not meaningful to consider the magnitude of the sampling error for an estimate in the case of a single individual sample, simply because the sample is the result of a random process and hence the sampling error is a random quantity. However, it is possible to know more about the distribution of sampling errors for the totality of possible samples of a given size (the sampling distribution), and hence how increasing or decreasing the sample size for a survey affects the sampling errors. What follows relates to this sampling distribution.

For most random sampling methods, the mean of the sampling errors for all possible samples of a given size is zero, which is why it is distinguished from the non-sampling errors referred to above which, in general, have non-zero mean. However, the range of the distribution of sampling errors can be estimated, and it tends to be narrower for larger samples than for smaller ones. In the limiting case, if the sample size is equal to that of the whole population, the sampling error will be zero.

The ‘average’ absolute magnitude of sampling errors (across all possible samples of a given size) is generally expressed not in terms of the mean absolute but as the ‘root mean square’ (RMS). The measurement expressed in this form is known as the ‘standard error’ of the observation variable. The general adoption of this measure is due to a curious statistical fact: for large populations, whatever the underlying distribution of the observation variable, the standard error is a simple function

$$\frac{S}{\sqrt{n}}$$

of just two things

the variance (S^2) of the observation variable in the population
the sample size (n)

This formula for the standard error can be estimated (approximately) from sample data. For the FBS the standard errors are calculated using a Taylor Series approximation.

Survey results are sometimes presented with estimates of standard errors, or alternatively in terms of confidence intervals. Farm Business Surveys results are presented with 95% confidence intervals. These show the range of values that may apply to the figures. They mean that we are 95% confident that this range contains the true value. They are calculated as the standard errors (se) multiplied by 1.96 to give the 95% confidence interval (95% CI). The standard errors only give an indication of the sampling error. For the Farm Business Survey, the confidence limits shown are appropriate for comparing groups within the same year only; they should not be used for comparing with previous years since they do not allow for the fact that many of the same farms will have contributed to the Farm Business Survey in both years.

Non Sampling Errors

Non-sampling error may be subdivided as follows:

- Coverage errors
- Non-response errors
- Response errors
- Processing errors
- Estimation errors
- Analysis errors

Any coverage errors in the FBS must be due mainly to imperfections in the sampling frame (June Survey). Coverage of particular sectors in the sampling frame is a problem, for example potato growers.

Minimising response (measurement) errors is the strongest area of quality management for the FBS. Processing errors are regarded as low-risk because of the self-checking nature of much of the farm management account and the high proportion of farms for which between-year checks can be applied.

Non Response

Although the FBS is designed to impose as little burden as possible on participating farmers, it is seeking sensitive data which some farmers might find intrusive. Before letting anyone trawl through the business accounts and other documents, a farmer needs to be convinced that he/she is getting sufficient in return. The refusal rate is high; around 90% of those approached who are in scope.

The potential population of non-respondents may have significantly different characteristics from the potential population of respondents, leading to bias in the estimates of the full population. Calibration weighting (see below) is used to reduce this bias, but is unlikely to completely remove it.

Weighting of the FBS results

The weighting of the FBS results is a two stage process with firstly an initial weight being produced and then this initial weight being adjusted via a calibration procedure. The weights are based on population data from the June business register (see section on Sampling and Recruitment) and are calculated for each design stratum. The Initial weights for the FBS are based on the inverse sampling fraction. Suppose for example there were 250 Cereal farms in the population and of these 50 were sampled then these 50 sampled farms would be given an initial sample weight of 5 (250/50). These weights are then adjusted (Calibration Weighting) so that they produce correct population totals for a series of calibration variables for which accurate population values are known from other sources. This ensures that the weights produce precise estimates of other variables, with little bias, despite the inevitable imperfections of the sampling strategy.