

Farm Business Survey - Statistical information

Sample representation and design

The sample structure of the FBS was re-designed starting from 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.

Farms are retained in the sample for several years and only 10% of the sample is replenished each year. Currently, this is achieved by releasing farms when they have been in the survey for 15 years. This, together with natural wastage, generates the required 10% vacancy per year. However, this procedure results in an age bias and is currently under review.

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.

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.

Sample Size

From 2010/11, the sample size for the England FBS is 1954 farms.

Farm Type (Standard Output Classification)	From 2010/11
Cereals	210
General cropping	168
Specialist fruit	36
Specialist glasshouse	85
Specialist hardy nursery stock	35
Other horticulture	35
Dairy (LFA)	51
Dairy (Lowland)	206
Grazing Livestock (SDA)	168
Grazing Livestock (DA)	56
Grazing Livestock (Lowland)	253
Mixed	162
Specialist Pigs	65
Specialist Poultry	81
EU regional Supplement	260
National Reserve	83
Total	1954

Government Office Region	From 2010/11
North East, Yorkshire & the Humber	178
North West	163
East Midlands	163
East of England	182
London & South East	130
West Midlands	122
South West	229
EU regional Supplement	260
National Reserve	83
Total	1510 ²

²The sample numbers for horticulture (specialist fruit, specialist glasshouse, specialist hardy nursery stock and other horticulture), Dairy (LFA), specialist pigs and specialist poultry are not collected on a regional basis.

Once recruited, farms may stay in the sample for several years (up to 15 years in some cases), so turnover is slow. About 90% of farms remain in the sample each year. So, about 60% of the current sample will have been in the sample for 5 years.

Sampling and Recruitment

Around 10% of the FBS sample is replaced each year due to the operation of the 15 year rule and 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 Single 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.

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. This represents the range of values that may apply to the figures. They mean that we are 95% confident that the true value lies within this range either side of the estimate. They are based on the Standard Errors (SE) multiplied by 1.96 to give the 95% confidence interval (95% CI). In addition confidence indicators may be provided alongside results which are based on the Relative Standard Errors (Standard Error/Estimate).

Non Sampling Errors

Non-sampling error may be subdivided as follows (see Statistics Canada Power from Data!):

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.

A project examining non-response in the FBS is going to be conducted throughout 2012 and early 2013 to examine the characteristics of the non-responders and the responders of the FBS. The outcome of this project will provide a greater understanding of the non-response, which will help inform the weighting model used for the FBS

Implementing new farm typology

The FBS has adopted a new classification of farm type based on Standard outputs for the 2010/11 accounting years. Please see the [UK farm classification document](#) for further information

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.

FBS methodological Review 2010

A statistical review of the FBS was undertaken by Dr James Brown at the University of Southampton in 2010. The review covered 9 topics: sub-sampling, smaller panel, non-response bias, 15 year rule, publication thresholds, year on year comparisons, stratification of survey, calibration weighting and outliers. A report detailing the outcomes and suggested areas of work was produced.