Proceedings of Statistics Canada Symposium 2001
Achieving Data Quality in a Statistical Agency: A Methodological Perspective

EVALUATING FRAME COVERAGE AND ACTIVITY OF MICRO-BUSINESSES IN SAWMILLING IN THE UK

Paul Smith[1]

ABSTRACT

The Sawmill Survey is a voluntary census of sawmills in Great Britain. It is limited to fixed mills which use domestically-grown timber. Three approaches to assess the coverage of this survey are described:
(1) a sample survey of the sawmilling industry from the UK's business register, excluding businesses already sampled in the Sawmill Survey, to assess the undercoverage in the list of known sawmills
(2) a non-response follow-up using local knowledge of regional officers of the Forestry Commission to estimate for sawmills which do not respond (mostly the smaller mills)
(3) a survey of small-scale sawmills and mobile sawmills. Many of these businesses are micro-enterprises.
The three approaches are synthesised to give an estimate of the coverage of the original survey compared to the total activity identified, and to estimate the importance of micro-enterprises to sawmilling activity in Great Britain.

KEY WORDS: Coverage; multiple frames.

1. INTRODUCTION

1.1 Sawmilling in Great Britain

Forestry and primary wood processing is a small industrial sector in the UK, but one in which there is considerable interest because of government policy to increase the area of woodland, and for the UK to supply a larger proportion of its own timber requirements. In England, Scotland and Wales (which together make up Great Britain), responsibility for undertaking surveys to measure these sectors lies with the Forestry Commission (FC), a government department with responsibility for matters related to forestry. [Forestry is, however, an area where responsibility is devolved, so that decisions about forestry in Scotland and Wales are made by the ministers in the Scottish Parliament and the Welsh Assembly respectively.]

Information on sawmilling activity in Great Britain (GB) is collected annually by the FC in two surveys which are intended to cover all fixed sawmills in GB, and to provide information mainly on the processing of British-grown roundwood – these are called the Sawmill Survey and the Round Fencing Survey. This restriction to domestic production reflects the main focus of the survey to give evidence to assist policy making to support wood processing sustained by domestic resources. It is different from the main business surveys run by the Office for National Statistics (ONS) in several respects:

- it is voluntary;
- it looks to include all sawmills regardless of the industrial classification of their parent company (that is, it is an activity-based survey), but interest is restricted to the processing of British-grown timber;
- it uses a list which has been built up over many years by experts in the industry (particularly representatives of various Industry Associations), which is not linked to any administrative system.

[1] Paul Smith, Economics and Statistics Unit, Forestry Commission, 231 Corstorphine Road, Edinburgh, EH12 7AT, UK. Current address: Office for National Statistics, Cardiff Road, Newport, NP10 8XG, UK.

These factors mean that the frame from which the survey is undertaken changes rather slowly as businesses cease to trade, with occasional exercises to add new sawmills to the list. The ONS has a business register, the Inter-Departmental Business Register (IDBR) (Perry 1995) which is updated regularly from two administrative sources (ref), but which classifies businesses according to their main activity, which means that sawmills are not always identifiable from the parent business's classification.

Because the survey is voluntary, the response rate is around 60%, but many sawmills have not responded for several years. This makes estimation of their activity difficult – in some cases the problem is associated with incomplete address information, and it is hard to tell whether the sawmill is still extant.

The coverage of the survey is intentionally limited to the larger sawmills, to avoid placing an undue burden on the smallest businesses, known as micro-businesses. Mills producing less than $1000m^3$ sawn timber annually and all mobile mills have been excluded.

1.2 Evaluating sawmilling activity

Three complementary activities have been undertaken in 2001 to try to better assess the overall sawmilling activity within GB. Firstly, an extract of the businesses classified to sawmilling on the IDBR was obtained, and matched against the list of businesses on the FC's sawmill survey frame. A sample of the unmatched businesses was sent a questionnaire to determine whether they were indeed sawmills, and how much British-grown roundwood they processed if so.

Secondly, a frame was constructed of small wood processors, including mobile sawmills and mills sawing only small quantities of timber in a year. In GB it is thought that most of the small processors saw hardwood, whereas the larger processors mainly saw softwood, so the survey was named the Small Hardwood Processors Survey.

Thirdly, the network of Forest District offices which is part of the FC's organisation was contacted, and each Forest District Manager (FDMs) was asked about the sawmills in their area which had not responded to two or more years of the sawmill survey. Because these districts are quite small, the FDMs have a lot of local knowledge (and appropriate local contacts) and provide a useful source of information, particularly on whether a mill is still trading, but also anecdotal evidence on how much processing is being done.

All these approaches have been used to make an estimate of the total activity in sawmilling in GB, and this has been compared with the estimates which have been obtained using the standard approach. Section 2 describes the approaches used for the three studies in more detail, section 3 gives the results of the surveys, and section 4 assesses the quality of the coverage of sawmilling activity in Great Britain.

2. METHODS

2.1 Supplementary coverage from the IDBR

The IDBR is updated from administrative records for valued added tax (VAT) and income tax (PAYE). The main survey unit is the enterprise, which may cover a number of local units (sites). Enterprises are classified according to the activity in which they have the largest value added, usually estimated by the largest sales. This means that in trying to find businesses with sawmills, it is difficult to locate any where sawmilling is only a small part of their activity, because they will be classified to a different industry.

For this exercise, we restricted our attention to the main heading containing sawmills within the Standard Industrial Classification 1992 (SIC92), which is *Sawmilling and planing of wood, impregnation of wood*. All the extant businesses classified to this heading on the IDBR were extracted, and matched against the FC lists for the sawmill and round fencing surveys. Unmatched units were divided into four strata – those with "sawmill" in their name or address, and then large, medium and small businesses for the remainder. All businesses with "sawmill" in their name or address, and all the largest businesses were sampled, and smaller proportions of the medium and small businesses. Units with even a tenuous match were excluded from the sampling, to avoid the possibility of any mill receiving two questionnaires.

The IDBR records auxiliary information about businesses from the administrative sources, and as well as using the employment size in the stratification, it was further used in a ratio estimator (Cochran 1977) for the number of businesses and volume of wood sawn. This is particularly important for the stratum of businesses with “sawmill” in their name where size was potentially variable.

2.2 Small Hardwood Processors Survey

The Small Hardwood Processors Survey needed a different source for population information, since many of the businesses of interest are very small, and so have no requirement to register with either of the administrative systems used to update the IDBR. Different sources were used in different parts of GB. In Scotland information on small sawmills was provided by the National Office of the Forestry Commission. In England and Wales information came from a number of organisations and initiatives, mostly with a regional focus, who support the forestry and primary wood processing sector. These provided information on potential small processors known to them.

These lists were variable in detail, but were used as if they were a complete frame. Again the businesses were matched against the Sawmills Survey and Round Fencing Survey lists, and also against the IDBR list so that sending out several questionnaires to the same business was avoided. A simple random sample of the unmatched sawmills was used, since there was no auxiliary information for stratification.

2.3 Local knowledge of non-respondents

The various surveys of sawmills are all voluntary, and the response rate, although reasonable, gives the potential for bias in the estimates. Of particular concern is a small proportion of sawmills which has not responded for several years; in these cases it is often unclear whether the business still exists. In these cases non-respondent values have been estimated by carrying forward their previous response where there is one, or by weighting otherwise.

For all businesses which had not responded in two or more years, information was sought from the Forest District Managers (FDMs) (Forest Districts are local areas of organisation of the Forestry Commission) on the existence and activity of sawmills. In general, Forest Districts are sufficiently small for the FDMs to maintain a reasonable working knowledge of the forestry and wood processing businesses in their area. FDMs were provided with lists of the non-responding sawmills, and asked to say whether they were still operating, and if so, to give an estimate of their consumption and production.

3. SURVEY RESULTS

3.1 Matching, population sizes and sampling

The matching of the IDBR extract and the Sawmill/Round fencing surveys produced 154 matched units, and 319 unmatched units in GB. Of these 319, 33 had “sawmill” in their name or address and the remaining 286 were divided into 9 large business (employment >100), 76 small and medium-sized enterprises (employment 10-99) and 210 micro-enterprises (employment <10). 1:2 small and medium-sized enterprises and 1:10 micro enterprises were sampled.

Only 26% of the Sawmill Survey population and 27% of the Round Fencing Survey population were classified to *Sawmilling and planing of wood; impregnation of wood* on the IDBR, which indicates that it would not be straightforward to run such a sectoral survey from the IDBR. The rather small overlap of the sawmill survey and IDBR populations is surprising. No attempt has been made to find out whether the remaining three-quarters of the units is on the IDBR but classified differently, or whether they have no need to register with the administrative systems and hence are not on the IDBR.

A small number of Woodland Initiatives approached to obtain lists of small processors did not participate, resulting in some small geographical undercoverage. From the composite list for the Small Hardwood Processors Survey in England and Wales, 126 matches or possible matches were detected. The remaining 316 potential sawmills were sampled 1:2, and all of the 16 small processors identified by the FC National Office for Scotland were included.

3.2 Response

The response for the sample from the IDBR was 61%, but for the Small Hardwood Processors Survey response was only 31%. In both cases, but particularly for the small hardwood processors, it is difficult to know whether the non-respondents have chosen not to respond, or have failed to respond because they felt the questionnaire was not applicable. Both questionnaires began with questions designed to discover whether any sawing was taking place, to promote response from all sampled units.

Non-response was a particular problem for the largest businesses from the IDBR, where only 2 responses were received, and neither respondent was engaged in processing British-grown timber.

In one case in the IDBR survey, a respondent said that they sawed roundwood, but did not estimate the volume. A ratio imputation based on other positive responses was used to compensate for this item non-responder.

3.3 Results

The survey of sawmills from the IDBR indicated that an estimated total of more than 50 sawmills did in fact undertake sawing of roundwood of British origin. This is based on information from valid responses and an assumption that non-response was ignorable. However, there must be a suspicion of non-response bias – there can be *post hoc* explanations for non-respondents preferentially *not* to saw roundwood (not returning the questionnaire if it is not relevant), or indeed for the smaller businesses not to return a questionnaire if they *do* saw roundwood as a perceived tax-avoidance measure (in fact there is no exchange of statistical survey data from the FC to tax authorities, and the confidentiality of the data was emphasised in the questionnaire). If non-response is presumed to indicate an answer of "no sawmilling activity", then the estimated number of businesses is reduced to around a third of its original, but the volume is around two thirds of the originally estimated volume (that is, non-respondents tend to be the smaller businesses).

Table 1: Results from the IDBR supplementary coverage survey; estimated totals are totals of businesses appearing on the IDBR but not the Sawmill Survey frame.

Characteristic	Estimated total	
	(assuming ignorable non-response)	(assuming that non-response indicates no sawmilling activity)
Number of businesses sawing roundwood	53.1	18.5
No of businesses sawing British roundwood logs	53.1	18.5
No of businesses sawing British small roundwood	32.8	4.3
No of businesses sawing imported roundwood	9.2	6.3
Number of fixed mills	59.0	22.6
Number of mobile mills	1.6	1.1
Volume of timber (000m^3 underbark)	118.6	72.2

The Small Hardwood Processors' Survey has an even more pronounced non-response problem. The reasons for non-response here are even less clear – in many cases there may be no sawmilling activity, or it may be a very small component of a business (for example a furniture maker). Equally, there is the possibility that small businesses don't want to give information on their activity because of a (false) perception of the survey as providing information to the tax authorities.

Table 2: Results from the small hardwood processors' survey.

Characteristic	Estimated total	
	(assuming ignorable non-response)	(assuming that non-response indicates no sawmilling activity)
Number of fixed sawmills	130	39
Number of mobile sawmills	141	42
Purchases of sawlogs (m^3)	45,714	13,749
Purchases of small roundwood (m^3)	6,714	2,289
Inputs not purchased (m^3)	156,886	47,750

4. EVIDENCE ON NON-RESPONDERS

4.1 Results

23 of 27 FDMs responded to the request for information on sawmills which had not responded to the Sawmill Survey. Most of the changes were to sawmills in the smaller size categories (categories are based on total sawnwood production from the survey), which is in line with the targetted follow-up for this survey which aims to get responses from all the largest sawmills. Table 2 shows these results, with sawmills allocated to size categories based on the survey (or imputed) information. Softwood production is shown as an example variable – it is economically the most important area

Table 3: Unweighted sample totals from the 2000 Sawmill Survey before and after updating imputations for non-responding sawmills with information based on local knowledge of Forest District Managers, and the estimated bias induced by carrying forward previous years' values.

Size category (m^3 sawnwood production)	Number of sawmills			Softwood production (m^3)		
	Before	After	Estimated bias (%)	Before	After	Estimated bias (%)
0-<1000	161	131	+23%	47,079	30,765	+53%
1000-<5000	94	100	–6%	193,867	210,776	–8%
5000-<10000	30	35	–14%	157,189	183,201	–14%
10000-<25000	18	18	0%	265,169	264,025	0%
25000-<50000	12	12	0%	405,151	405,151	0%
50000+	13	13	0%	1,065,368	1,065,368	0%
Total	328	309	6%	2,133,824	2,159,287	–1%

Some of the differences between smaller and medium sawmills is because of switching between size categories – a tendency for medium sized businesses to be larger and smaller businesses to be smaller than their imputations is noticeable. Overall, carrying forward previous year's information on sawmills seems to lead to a slight overestimate of the number of sawmills in the UK and almost no difference in the level of sawn softwood production.

5. DISCUSSION

The non-response follow-up indicates that there is some over-estimation of the number of extant sawmills in Great Britain, but that this has only a small effect on the estimates of sawnwood production from the Sawmill Survey. The Coverage Survey based on information from the IDBR suggests that an additional 18% of sawmills not known to the Forestry Commission are operating, but overall these additional sawmills consume only about 3% more timber inputs. The small hardwood processors likewise consume only a small additional amount of the total inputs to sawmilling (+4%), but they are a very significant proportion of *hardwood* processing (+80%), which suggests that for a complete coverage of sawmilling there should be some effort to measure their output. The two surveys taken together suggest that there is 6-7% undercoverage of the sawmilling sector based on inputs.

Measuring the output of microbusinesses on a regular basis is, however, likely to be difficult because they do not have the resources to respond to questionnaires. The low response in this survey is already indicative of a reluctance to provide information, and the most that could be collected with a reasonable hope of usable response is a total activity figure on a multi-annual periodicity.

This study, despite reservations about the quality of the data because of the low response rates, shows that the Sawmill Survey provides a reasonable estimate of *overall* activity in GB, but that certain components (particularly hardwood processing) are affected by undercoverage. When using these as evidence for adjusting policies, the results should be treated appropriately.

REFERENCES

Cochran, W. (1977) *An introduction to sampling theory*. Wiley.

Perry, J. (1995) *The Inter-Departmental Business Register*. Economic Trends vol. **505**, November 1995.