

GENERAL EXPLANATION Continued

\$100,000 or more and the lowest misclassification were those with sales of \$5,000 to \$39,999. Table 4 shows a distribution of mail size codes by value of sales. The misclassification is largely due to the difference in reporting unit and reference dates between the source records and the census.

Accuracy of the Estimates

Estimates of sampling variability expressed as standard errors are presented in tables 16 and 17. The chances are about two out of three that the difference between an estimate based on the coverage check sample and the result that would have been obtained by applying the coverage check procedures to all farms would be less than the sampling error shown. The chances are about 95 out of 100 that this difference would be less than 2 times the sampling error.

The standard error for the coverage check estimates of total farms expressed as percent of estimated total, was 1.2 percent at the United States level; ranges from 2.1 to 7.1 percent at the census division level; and 2.8 to 10.7 percent at the State level. Sampling errors were computed directly for the total, included, and missed components. The estimates of sampling error for the overcounted farms are based on a very small number of observations and are not considered reliable, therefore, they are not published. In addition, coverage estimates are presented only for census divisions or State groups where individual State estimates are not considered reliable and thus are not published separately. Estimates of sampling error for acres were not computed, but are considerably higher than for farms since the acre estimates are based upon a 1 in 10 subsample of segments.

There are several aspects of the coverage check procedures which make it probable that the estimates of net error are somewhat larger than the actual undercount. First, the difficulty of carrying out searching and matching procedures was great, and some of the farms corresponding to coverage check farms may not have been located. An intensive study completed in the 1969

Table 4. Percent Distribution of Mail Size Codes by Value of Sales

	Value of sales reported				
	\$100,000 and over	\$40,000 to \$99,999	\$5,000 to \$39,999	\$2,000 to \$4,999	Under \$2,000
Predicted value of sales.....	100.0	100.0	100.0	100.0	100.0
\$100,000 and over.....	143.7	6.9	1.4	0.5	0.4
\$40,000 to \$99,999.....	34.4	148.3	7.1	1.0	.7
\$5,000 to \$39,999.....	14.2	37.6	167.9	29.5	11.0
\$2,000 to \$4,999.....	4.4	3.9	14.1	138.7	22.2
Under \$2,000.....	3.3	3.3	9.5	30.3	165.7

¹Proportion within the predicted size class.

census evaluation program indicated about 5 percent of the farms classified as missing were actually included in the census.

Second, once a census farm corresponding to a coverage check farm was located, there was no systematic attempt to search the census files further for duplicate report forms, so that some cases of duplication in the census may have been overlooked. Duplicate cases which were found were normally adjacent in the file and in the same ZIP code area. To make a thorough search would have been costly in both time and money, as it would have been necessary in all matched cases, to make additional checks in adjacent and other ZIP code areas and counties.

The estimates of total farms from the coverage check sample are low in relation to data from other sources. For example, using the previous census farm definition, the coverage check estimate of 2.3 million total farms compares to the USDA estimate of 2.8 million farms. In addition, using the 1974 census farm definition the coverage estimate of 1.9 million farms included in the census compares with 2.4 million farms counted in the 1974 Census of Agriculture. The primary reasons for these lower coverage check estimates appear to be related to the sample design of the JES and the difference between enumeration dates for the census and the JES. The measurement base used for the coverage check was the JES area sample of farm operators living inside the segment boundaries. Although the JES contains some urban segments, there are indications that the part of the sample used for the coverage check may underrepresent farms operated by persons not living on their farms.

A special tabulation of farms in the coverage sample by residence was made for all States and compared with census counts. The census data indicated about 20 percent of the farm operators did not live on their farms. Census data for JES farm operators who were classified as included in the census in the coverage sample, indicated about 11 percent did not live on their farms. The proportion of coverage check nonresident operators missed in the census was about the same as for those included; therefore, this possible bias may not have a large effect upon the coverage proportions.

The enumeration for the JES took place in the latter part of May 1974, while most census data were collected during the first few months in 1975. The difference in reporting dates caused some matching difficulties when farms were sold, operators moved to different farms, or operators died. In general, farms dropped from the base sample as a result of these situations and there was no practical method available to add new farms into the base sample. The number of farms affected by the difference in dates has not been determined from the coverage sample; however, a previous study made in connection with the 1965 sample survey of agriculture indicated about 5 percent of the farms had changes of ownership or operators during a 1-year period.

An additional factor contributing to the low level estimates is the possible deficiency in the coverage estimates resulting from the fact that a little more than 4 percent of the cases were unclassified. Unclassified cases are those which were matched to the census mail list but the census report forms could not be located and also those not matched to