EVALUATION STUDIES OF THE 1959 CENSUS OF AGRICULTURE

R. Hurley, T. Jabine and D. Larson
Bureau of the Census
Washington, D.C.

A. Content of the Evaluation Program of the 1959 Agriculture Census

After careful consideration of the resources available, the relative importance of needs for data on different types of measurement error, and past experience in the evaluation of agriculture census data, it was concluded that the Evaluation Program for the 1959 Census of Agriculture should consist of two projects:

1. A national Evaluation Survey designed primarily to obtain estimates of bias and its components for selected items.

2. An Enumerator Variability Study, designed to provide estimates of the enumerator contribution to nonsampling variability for selected items.

B. The Evaluation Survey[1]

1. Objectives

The objectives of the Evaluation Survey were as follows:

a. To provide national and regional measures of the accuracy of census results for a restricted number of important items. The items selected were

Number of farms
Total land in farms
Acres of cropland harvested
Acres of corn harvested
Acres of wheat harvested
Acres of cotton harvested
Number of farms classified by size, tenure, economic class and type

b. To provide data concerning factors associated with measurement errors, such as

(1) Cross tabulations of farms by "match status", i.e., missed in the Census, enumerated with errors, enumerated without error, and by farm characteristics likely to be associated with match status, such as size, tenure arrangements, residence of operator, number of separate tracts, etc.

(2) Classification of errors on farms enumerated in the Census by reason for error.

c. To conduct a test of procedures for the measurement of bias for selected livestock items. No livestock items had been included in the Evaluation Survey for the 1954 Census of Agriculture, mostly because satisfactory measurement techniques were not believed to be available for these items. For the 1959 program, it was decided to attempt to develop and test procedures for measuring the bias of:

Cattle and calves on hand
Hogs and pigs on hand
Litters of pigs farrowed, 12/1/58 to 11/30/59

2. Survey Design

The survey design of the Evaluation Survey may be summarized as follows:

a. An area sample was used to obtain information for farms missed in the Census. This sample had two parts -

(1) A rural area sample of 772 segments in 196 primary sampling units, comprising 348 counties. This was a self-weighting sample, with an overall sampling fraction of 1 in 1,500.

(2) An urban area sample, comprising seven-eighths of the segments in urban areas in the November 1959 sample for the Census Bureau's Current Population Survey.

b. A list sample was used to obtain information on places[2] which had been included in the Census. This sample also had two parts.

(1) A sample consisting of all enumerated places of less than 5,000 acres whose locations had been marked by Census enumerators within the boundaries of the 772 segments in the area sample. This sample included about 2,770 farms and nonfarm places.

(2) A sample of 562 Census farms of 5,000 acres and over, selected independently of the area sample.

c. The basic survey operations may be summarized as follows:

Step 1 - An independent enumeration designed to obtain detailed and accurate information for all farms in the area sample and all places operated by persons in the list sample.

Step 2 - Matching of the results obtained in Step 1 against the Census materials in order to tentatively identify farms missed in the Census and errors for places included in the Census.

Step 3 - Followup, as required, to clarify and check the results of Steps 1 and 2.

Step 4 - Final processing and tabulation of the results.

d. Estimates of net error were obtained by combining the results of the area and list samples. The area sample provided estimates of the component of net error arising from farms completely missed in the Census. The list sample provided estimates of the component of net error arising from errors in reporting and (to some extent) processing of data for farms included in the Census. Simple unbiased estimates were used in both cases.

The estimated totals (shown in column (2) of Table 1) were obtained by adding the estimated net error for each item to the corresponding published Census total.

3. Results

The principal results of the Evaluation Survey are shown in Tables 1 and 2. Except for the Census totals, the figures shown are subject to sampling error, and therefore should be interpreted with caution. Sampling errors for the estimates in Table 1 are shown in column (6) of that table. Some of the principal results which may be noted in Table 1 are as follows:

a. The Evaluation Survey estimate of the number of farms in the conterminous United States was 4,045,000, as compared with a Census total of 3,704,000. The difference, or net error, was 341,000, or 8.4 percent of the Survey estimate.

b. The estimates of net underenumeration for the principal acreage items were 6.0 percent of the estimated total for land in farms, 4.3 percent for acres of cropland harvested, 4.5 percent for acres of corn, 3.0 percent for acres of wheat and 9.1 percent for acres of cotton.

c. The relative net errors for acreage items, with the exception of cotton, are significantly smaller than the relative net error for farms.

d. Again with the exception of cotton, the estimates of relative net error for the 1959 and 1954 Censuses are quite similar for those items for which data are available from both years.

e. A considerable part of the net error in number of farms is accounted for by marginal farms. Approximately two-thirds of the net error in number of farms was accounted for by farms of less than 100 acres, and about three-fourths of the net error by farms with total sales of less than $2,500 in 1959.

f. The net errors for livestock and dairy farms were smaller than the net errors for other types of farms, especially those specializing in field crops. This is probably associated with the fact that farms in the latter group have a higher proportion of nonresident operators.

In Table 2, the percent distributions of Evaluation Survey (ES) farms by three match status classifications are shown for several different characteristics of farms and farm operators. The match status classifications used in this table are based on the extent to which these particular farms, as identified in the Evaluation Survey, were covered by the Census enumeration. The basic classification contains two groups - enumerated in Census and missed in Census - and the enumerated in Census group is sub-divided into "complete matches" and "partial matches", this classification depending on the extent to which Evaluation Survey and Census data for the farm differed. A study of these results will reveal several factors that are associated with failure to find and enumerate farms and, once a farm is located, failure to identify correctly the land included in the place.

Some of the kinds of farms that were most commonly missed were farms with nonresident operators (26.3 percent of the ES farms), small farms (24.3 percent of the ES farms under 10 acres and 17.1 percent of the ES farms with 10 to 49 acres), farms in enumeration districts which were not entirely rural in character (26.1 percent) and farms with operators under 25 years of age (22.6 per-

cent). The more separate tracts a farm had, the less likely it was to be missed. Partnership operations were less likely to be missed than individual operations; and farms operated by part owners and managers were less likely to be missed (4.7 percent) than those operated by full owners (12.1 percent), with tenant operated farms occupying an intermediate position (9.1 percent missed).

One of the questions asked of each farm operator interviewed during the part of the Evaluation Survey enumeration which took place after the Census was whether a Census questionnaire had been filled for his place. Of those who answered no to this question, an estimated 44.9 percent had actually been missed in the Census.

Some of the kinds of farms for which the Census enumerator was least likely to obtain correct information on acres in place were farms with land in more than one county (42.6 percent complete match, 48.2 percent partial match), farms with nonresident operators (30.3 percent complete versus 43.4 percent partial), farms of 1,000 acres and over (30.9 percent complete versus 64.2 percent partial), farms operated in partnership (45.2 percent complete versus 51.4 percent partial), farms whose most distant tract was at least 10 miles away from the operator's residence (34.2 percent complete versus 62.0 percent partial), farms with three or more landlords (33.6 percent complete versus 60.8 percent partial), and farms whose operators had three or more tenants (16.7 percent complete versus 75.0 percent partial).

The difficulty of determining total acres correctly clearly increased in proportion to the number of separate tracts in the place, in proportion to the numbers of landlords and tenants and in proportion to the size of the place. Apparently, it was considerably easier for the Census enumerator to determine the acreage of a place operated by a full owner, provided he found it in the first place, than it was to do this for a place operated by a part owner, manager or tenant. Changes in the acres in place which occurred during the Census year (1959) also increased the likelihood that the Census enumerator would get an incorrect figure for total acres.

C. The Enumerator Variability Study

1. Objectives

An experiment to measure the enumerator contribution to the non-sampling variability of census data was carried out in connection with the 1950 U.S. Census of Population. The results of this experiment had far-reaching effects on the design of procedures for the 1960 Census of Population. The 1959 Agriculture Census offered an opportunity to apply these measurement techniques to another kind of data and also to try an experimental design believed to be more efficient than the one used in the 1950 Population Census.

The purpose of the Enumerator Variability Study conducted in conjunction with the 1959 Agriculture Census enumeration was to obtain estimates, for selected items, of the nonsampling variability associated with census enumerators. Enumerator variability is the component of total enumerator error which tends to average out through compensating errors over the work of a large number of enumerators. This source of variability therefore has the greatest relative effect for census statistics which are based on the work of only a few enumerators, such as statistics for an enumeration area (one enumerator) or a county (usually a dozen or fewer enumerators), or, if the item occurs infrequently, for larger areas.

2. Design of the Study

Because the purpose of the study was the evaluation of enumerator variability in an agricultural census, the variability measured was not a "pure" enumerator variability but was the variability associated with actual census procedures. These procedures allowed some self-enumeration by respondents (questionnaires were distributed to farmers in advance of the Census enumeration with requests for completion prior to the enumerator's visit) and the usual editing of questionnaires during the census processing operation. Therefore the enumerator variability that was measured might be thought of as a residual variability still remaining after the possible diminishing effects of the above two factors.

The experiment was restricted to an area consisting of ten contiguous counties in the State of Indiana. Consequently, the results are not directly applicable to a nationwide census of agriculture. The area selected was one in which the enumeration of farms does not present any unusual or atypical problems.

To estimate enumerator variability, an interpenetrating sample design was used. Two weeks prior to the Census, listings were prepared of all the places with specified types of agricultural operations in the ten counties. There were about 17,800 such listings made in 104 enumeration areas (EA's). The listings

for each EA were divided into two random halves, which were then assigned, at random, to the two Census enumerators designated to work in the EA. In this way, the work of each of the two enumerators in an EA provided a basis for unbiased estimates of Census statistics for the EA.

From the resulting data, estimates of enumerator variability for 46 items were made for each EA.

The estimate used for a given EA was based on the difference of two variance components -- a "between" enumerator variance minus a "within" enumerator variance. Since each EA estimate was based on data from only two enumerators, the individual estimates were pooled over the 104 EA's in the ten counties to obtain more reliable overall measures of enumerator variability.

3. Results

Results from this study are presented in Tables 3 and 4, and Chart 1. The principal conclusions which have been drawn so far from the analysis of these results are as follows:

a. Levels of enumerator variability for most of the items studied are sufficiently high so that this factor requires careful consideration in the planning of future censuses and surveys.

b. The number of listings for which questionnaires were not obtained, and the number of nonresponses to particular items showed some of the highest levels of enumerator variability. This finding confirms results from the study of enumerator variability carried out during the 1950 Census of Population.

c. The levels of enumerator variability (expressed as a coefficient of variation) for attributes appear to be, on the average, about one-half of the sampling error for a 25 percent simple-random sample (see Chart 1). This statement is valid only for EA's of the approximate size used in this study (the mean number of listings per EA of places with specified types of agricultural operations was about 170). The experiment has not yielded conclusive information on how enumerator variability would be affected by changes in cluster size.

d. Even with the relatively large samples used in this study, it was not possible to obtain sufficiently reliable estimates of enumerator variability for variables, such as area, production and inventories. A satisfactory evaluation of enumerator variability for these items will require the development of improved sample design and estimation procedures.

FOOTNOTES

[1] Further information on the Evaluation Survey procedures and results may be found in the introduction to Volume II, General Report, for the 1959 Census of Agriculture. A detailed report on methodology, entitled "Checking the Accuracy of Area Statistics Obtained in the United States Censuses of Agriculture" may be had on request by writing to the Statistical Research Division, Bureau of the Census, Washington, 25, D.C.

[2] We use the term "place", rather than "farm", because the list sample represented all places for which questionnaires were filled, whether or not they qualified as farms.

Table 1.--ESTIMATES OF NET ERROR FOR SELECTED ITEMS FOR THE UNITED STATES: 1954 AND 1959 CENSUSES OF AGRICULTURE

(Numbers do not add exactly to totals in all cases due to rounding.)

Item	Estimated total (000)	Census total (000)	Estimated net error		Sampling error of estimated percentage net error
			Amount (000)	Percent of estimated total	
(1)	(2)	(3)	(4)	(5)	(6)
Number of farms.....1959	4,045	3,704	341	8.4	1.2
1954	5,201	4,782	419	8.1	0.9
Farms by size:					
Under 10 acres.....1959	298	241	58	19.3	11.1
1954	588	484	104	17.7	4.1
10 to 49 acres.....1959	890	811	79	8.8	3.8
1954	1,364	1,213	151	11.1	2.2
50 to 99 acres.....1959	745	658	87	11.7	2.8
1954	925	864	61	6.6	1.9
100 to 219 acres.....1959	1,038	998	40	3.9	2.5
1954	1,271	1,210	61	4.8	0.9
220 acres and over.....1959	1,074	997	78	7.2	1.2
1954	1,053	1,011	42	4.0	0.9
Farms by economic class, 1959:					
Class I, II and III (Sales of $10,000 and over) .....number	817	794	23	2.8	1.0
Class IV and V (Sales of $2,500 to $9,999).....number	1,328	1,270	58	4.4	1.0
Class VI and other (Sales of $50 to $2,499) .....number	1,897	1,637	260	13.7	2.5
Farms by tenure, 1959:					
Full owner.....number	2,251	2,116	135	6.0	2.6
Part owners and managers.....number	931	830	101	10.8	2.7
Tenants.....number	863	758	105	12.2	2.5
Commercial farms by type, 1959:					
Cash grain.....number	451	398	53	11.7	2.4
Other field crop.....number	519	469	50	9.6	3.5
Vegetable, fruit and nut.....number	94	82	12	12.9	7.4
Dairy.....number	439	428	10	2.4	1.6
Other livestock, except dairy and poultry.....number	697	684	14	2.0	1.1
General.....number	224	212	12	5.4	2.2
Poultry and miscellaneous commercial.....number	149	140	9	6.3	2.0
Total commercial.....number	2,573	2,413	161	6.2	1.0
Land in farms.....acres 1959	1,191,706	1,120,158	71,548	6.0	0.9
1954	1,223,891	1,158,192	65,699	5.4	1.9
Cropland harvested.....acres 1959	325,110	311,285	13,824	4.3	0.9
1954	346,580	332,870	13,710	4.0	1.1
Corn harvested.....acres 1959	83,396	79,616	3,781	4.5	1.1
1954	80,886	78,123	2,763	3.4	1.2
Wheat harvested.....acres 1959	51,088	49,567	1,521	3.0	1.3
1954	54,263	51,362	2,901	5.3	4.1
Cotton harvested.....acres 1959	16,132	14,649	1,483	9.2	2.7
1954	19,026	18,854	172	0.9	1.5

Table 2: DISTRIBUTION OF EVALUATION SURVEY FARMS IN RURAL AREA SAMPLE BY MATCH STATUS, FOR SELECTED CHARACTERISTICS

Characteristic	Number of farms in sample	Percent of total	Percent distribution by match status			
			Total	Enumerated in Census		Missed in Census
				Complete match	Partial match	
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Total Evaluation Survey farms in sample	2,555	100.0	100.0	55.5	34.9	9.6
By age of operator						
Under 25	53	2.1	100.0	49.1	28.3	22.6
25 to 34	283	11.1	100.0	49.8	39.6	10.6
35 to 44	570	22.3	100.0	58.8	34.2	7.0
45 to 54	655	25.6	100.0	57.7	33.9	8.4
55 to 64	543	21.3	100.0	57.8	33.9	8.3
65 and over	394	15.4	100.0	53.0	34.8	12.2
Age unknown	57	2.2	100.0	24.6	47.3	28.1
By type of enumeration district						
A (entirely rural)	1,710	67.0	100.0	58.5	34.8	6.7
B (mostly rural)	665	26.0	100.0	51.0	36.2	12.8
C and D (urban)	180	7.0	100.0	42.8	31.1	26.1
By location of land						
One county only	2,414	94.5	100.0	56.2	34.1	9.7
More than one county	141	5.5	100.0	42.6	48.2	9.2
By residence of operator						
On place	2,258	88.4	100.0	58.8	33.8	7.4
Not on place	297	11.6	100.0	30.3	43.4	26.3
By number of tracts						
1	1,572	61.4	100.0	60.8	27.3	11.9
2	605	23.7	100.0	50.6	43.0	6.4
3	214	8.4	100.0	45.8	46.3	7.9
4	93	3.6	100.0	37.6	61.3	1.1
5 or more	71	2.9	100.0	31.0	66.2	2.8

Table 2: DISTRIBUTION OF EVALUATION SURVEY FARMS IN RURAL AREA SAMPLE BY MATCH STATUS, FOR SELECTED CHARACTERISTICS (cont.)

Characteristic (1)	Number of farms in sample (2)	Percent of total (3)	Percent distribution by match status			
			Total (4)	Enumerated in Census		Missed in Census (7)
				Complete match (5)	Partial match (6)	
By size of farm						
Under 10 acres	148	5.8	100.0	47.3	28.4	24.3
10-49 acres	527	20.6	100.0	51.0	31.9	17.1
50-69 acres	183	7.2	100.0	54.6	31.7	13.7
70-99 acres	278	10.9	100.0	59.0	30.6	10.4
100-139 acres	277	10.8	100.0	60.3	33.6	6.1
140-179 acres	266	10.4	100.0	63.6	30.8	5.6
180-219 acres	171	6.7	100.0	59.7	33.9	6.4
220-259 acres	146	5.7	100.0	59.6	37.0	3.4
260-499 acres	335	13.1	100.0	57.9	39.4	2.7
500-999 acres	143	5.6	100.0	48.9	47.6	3.5
1000 acres and over	81	3.2	100.0	30.9	64.2	4.9
By type of operation						
Individual operator	2,409	94.3	100.0	56.1	33.9	10.0
Partnership	146	5.7	100.0	45.2	51.4	3.4
By tenure						
Full owner	1,374	53.8	100.0	63.3	24.6	12.1
Part owner or manager	621	24.3	100.0	43.1	52.2	4.7
Tenant	560	21.9	100.0	49.8	41.1	9.1
By distance from residence to most distant tract						
One tract only, non-resident, or information on distance not obtained	1,715	67.1	100.0	57.7	30.6	11.7
Less than 1.0 mile	327	12.8	100.0	54.7	38.2	7.0
1.0-4.9 miles	350	13.7	100.0	53.1	42.6	4.3
5.0-9.9 miles	84	3.3	100.0	41.7	53.6	4.8
10.0 miles and over	79	3.1	100.0	34.2	62.0	3.8
By number of landlords						
0	1,375	53.7	100.0	63.3	24.7	12.1
1	809	31.6	100.0	49.3	43.5	7.2
2	228	8.9	100.0	43.9	50.0	6.1
3 or more	143	5.8	100.0	33.6	60.8	5.6

Table 2: DISTRIBUTION OF EVALUATION SURVEY FARMS IN RURAL AREA SAMPLE BY MATCH STATUS, FOR SELECTED CHARACTERISTICS (cont.)

Characteristic	Number of farms in sample	Percent of total	Percent distribution by match status			
			Total	Enumerated in Census		Missed in Census
				Complete match	Partial match	
(1)	(2)	(3)	(4)	(5)	(6)	(7)
By number of tenants						
0	2,171	84.9	100.0	58.6	31.7	9.7
1	275	10.7	100.0	41.5	47.6	10.9
2	73	2.9	100.0	34.2	61.6	4.1
3 or more	36	1.5	100.0	16.7	75.0	8.3
By response to Evaluation Survey question on whether Census questionnaire was obtained for place						
Yes	1,732	67.8	100.0	61.8	35.2	3.0
Don't know	159	6.2	100.0	44.6	34.0	21.4
No	234	9.2	100.0	28.6	26.5	44.9
NA[1]	430	16.8	100.0	48.4	38.8	12.8
By changes in acres in place between Jan. 1 and Dec. 1, 1959						
Same on both dates						
No temporary changes[2]	2,128	83.3	100.0	57.8	32.0	10.2
Some temporary changes[2]	85	3.3	100.0	40.0	50.6	9.4
Larger on Dec. 1	250	9.7	100.0	43.6	49.2	7.2
Smaller on Dec. 1	92	3.7	100.0	46.7	50.0	3.3

[1] The large NA rate was due to the fact that this question could be asked only in the post-Census canvass. Operators in pre-Census segments were revisited only in those areas where livestock items were investigated.

[2] A temporary change occurred when land was rented in or rented out by the operator during 1959, but reverted to its owner prior to December 1, 1959.

Table 3: 1959 CENSUS OF AGRICULTURE - ENUMERATOR VARIABILITY STUDY
ENUMERATOR COEFFICIENT OF VARIATION FOR AN EA FOR
"0-1" OR "0-1-NA" ITEMS

Item	Number of listings having item (1)	Proportion of listings having item P (2)	Enumerator coeff. of var. 1/ v (3)	Standard error of enumerator coefficient of variation 1/ $\hat{\sigma}_v$ (4)	Standard error of enumerator coefficient of variation 1/ $\frac{\hat{\sigma}_v}{v}$ as % (5)
Number of listings:					
1. A-1's taken	16,481	.926	.040	.011	27.5
2. A-1's meeting 1959 definition of a farm	14,725	.827	.033	.008	24.2
3. Refusals and not-at-homes	280	.016	1.021	.209	20.5
Tenure:					
4. Tenants	2,652	.149	.048	.166	345.8
5. NA on tenancy	139	.008	1.027	.204	19.9
Number of A-1's reporting acres of:					
6. Owned land	13,630	.766	.047	.020	41.1
7. Cropland	13,530	.760	.039	.018	44.7
8. Cropland in pasture	8,955	.503	.090	.015	17.0
9. Corn	12,255	.688	.037	.020	54.0
10. Soybeans	6,493	.365	2/	2/	2/
Number of A-1's reporting livestock and livestock production:					
11. Cattle	10,124	.569	.048	.007	14.2
12. Milkcows	6,476	.364	.070	.009	12.4
13. Hogs	8,646	.486	2/	2/	2/
14. Chickens	7,578	.426	.084	.004	5.1
15. Chicken eggs sold	5,613	.315	.110	.019	17.0
Off-farm work and other income:					
16. Having off-farm work	8,511	.478	.070	.007	10.0
17. Having 100 or more days of off-farm work	6,165	.346	.092	.005	5.4
18. Having off-farm income greater than farm sales	6,138	.345	.191	.007	3.7
19. NA for off-farm work	477	.027	1.000	.032	3.2
20. NA for off-farm income greater than farm sales	1,589	.089	.641	.041	6.4

1/ The form of these estimates is given in the Appendix.

2/ Estimate of variance negative.

Table 4: 1959 CENSUS OF AGRICULTURE - ENUMERATOR VARIABILITY STUDY
ENUMERATOR COEFFICIENT OF VARIATION FOR AN EA FOR
QUANTITATIVE ITEMS

Item	Total Value (1)	Average value per listing $\bar{x}$ (2)	Enumerator coeff. of var. 1/ v (3)	Standard Error of enumerator coefficient of variation 1/ $\hat{\sigma}_v$ (4)	Standard Error of enumerator coefficient of variation 1/ $\frac{\hat{\sigma}_v}{v}$ as % (5)
Number of acres:					
1. Acres of owned land.....	1,421,108	79.8	2/	2/	2/
2. Acres in the place.....	2,122,129	119.2	.035	.025	70.8
3. Acres of cropland.....	1,307,603	73.5	.058	.020	34.3
4. Acres of cropland pasture	205,208	11.5	.083	.055	65.8
5. Acres of corn.....	567,627	31.9	.096	.020	20.6
6. Acres of soybeans.....	209,799	11.8	.060	.071	118.5
Crop production:					
7. Bushels of corn.....	32,535,857	1828	.096	.040	41.7
8. Bushels of soybeans.....	4,625,036	260	2/	2/	2/
Livestock and livestock production					
9. Number of cattle.....	251,663	14	2/	2/	2/
10. Number of milk cows.....	74,046	4	.077	.034	43.8
11. Number of hogs.....	550,573	31	.045	.041	91.9
12. Number of chickens.....	2,566,351	144	2/	2/	2/
Total NA and answered-in-error entries: 3/					
13. Answered-in-error.....	1,220	.068	.647	.121	18.7
14. NA's excluding completely refused and not-at-home categories.....	9,698	.545	.640	.018	2.8

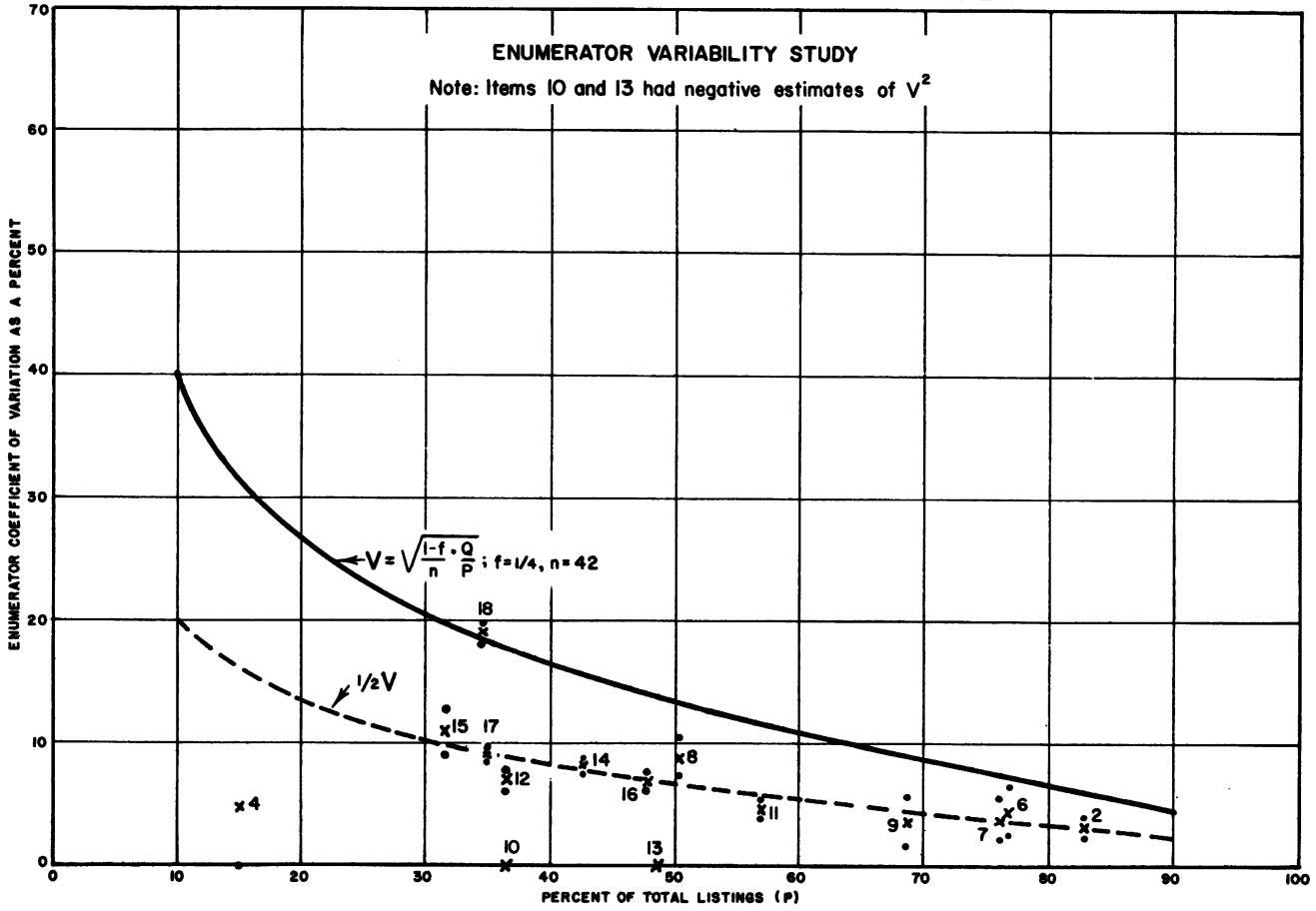
1/ The form of these estimates is given in the Appendix.

2/ Estimate of variance negative.

3/ The original A1 entry was not acceptable according to Census editing rules. These included cases of internal inconsistencies, duplications, extreme values, etc.

CHART I

ENUMERATOR COEFFICIENT OF VARIATION FOR AN EA



APPENDIX:
The Estimate of Enumerator Variability

The Estimate of v^2

A listing of places was first made by means of the normal type of canvass. (It will be noted that the normal type of canvass introduces some geographical stratification into the listing). Within each EA, the listings were divided into groups of ten consecutive listings and each group was randomly divided into two enumerator assignments of size five each.

For the i-th EA (i = 1, 2, ..., 104) within the experiment the average value per listing for a given item would be

$$\bar{x}_i = \frac{\sum_{h=1}^{k} \sum_{c=1}^{b_i} \sum_{j=1}^{\bar{\bar{n}}} x_{ihcj}}{k b_i \bar{\bar{n}}}$$

where $\bar{\bar{n}} = 5$ = the number of listings in each group of ten assigned to one enumerator

b_i = the number of groups of ten in the i-th EA

$k = 2$ = number of enumerators assigned to the i-th EA.

The estimate of enumerator variance for $\bar{x}_i$ is $C_i - D_i$

where
$$C_i = \frac{\sum_{h}^{k} (\bar{x}_{ih} - \bar{x}_i)^2}{k - 1}$$

$$D_i = \frac{\sum_{h}^{k} \sum_{c}^{b_i} \sum_{j}^{\bar{\bar{n}}} (x_{ihcj} - \bar{x}_{ihc})^2}{k b_i^2 \bar{\bar{n}}(\bar{\bar{n}} - 1)}$$

$$\bar{x}_{ih} = \frac{\sum_{c}^{b_i} \sum_{j}^{\bar{\bar{n}}} x_{ihcj}}{b_i \bar{\bar{n}}}$$

$$\bar{x}_{ihc} = \frac{\sum_{j}^{\bar{\bar{n}}} x_{ihcj}}{\bar{\bar{n}}}$$

The estimate of enumerator variance for a total

$$x_i = \sum_{h}^{k} \sum_{c}^{b_i} \sum_{j}^{\bar{\bar{n}}} x_{ihcj}$$

is

$$N_i^2 (C_i - D_i)$$

where $N_i = k b_i \bar{\bar{n}}$ = the number of listings in the i-th EA.

Using information from all EA's within the experiment, an estimate for the relative enumerator variance for an EA total is

$$v^2 = \frac{\frac{1}{L} \sum_{i=1}^{L} N_i^2 (C_i - D_i)}{(\frac{1}{L} \sum_{i=1}^{L} x_i)^2}$$

where $L = 104$ = the number of EA's in the experiment.

The values of v for selected items are recorded in column (3) of Tables 3 and 4.

The Variance of v

To estimate the variance of v, the 104 EA's were divided randomly into three groups of 35, 35, and 34 EA's. For the g-th random group a v_g^2 was computed. Then an estimate for the variance of v would be

$$\hat{\sigma}_v^2 = [\frac{1}{G} \frac{1}{G-1} \sum_{g=1}^{G} (v_g - \bar{v})^2]$$

where $G = 3$.

However for many of the items some of the estimates of v_g^2 were negative so the above estimate, $\hat{\sigma}_v^2$, could not be used. In its place an approximation[1] of the following type was used.

$$V_v^2 \doteq \frac{1}{4} V_{v^2}^2$$

$$\frac{\sigma_v}{v} \doteq \frac{1}{2} \frac{\sigma_{v^2}}{v^2}$$

$$\hat{\vartheta}_v \doteq \frac{1}{2} \frac{\hat{\sigma}_{v^2}}{v}$$

where $\hat{\sigma}_{v^2} = \{\frac{1}{G} [\frac{1}{G-1} \sum_{g=1}^{G} (v_g^2 - \bar{v}^2)^2]\}^{\frac{1}{2}}$

The $\hat{\vartheta}_v$ for each item appears in column(4) of Tables 3 and 4.

Comparison of the Enumerator Variability with Sampling Variability

Chart 1 provides a basis for comparing the enumerator coefficient of variation for the (0, 1) variables with the corresponding coefficient of variation for a 25 percent simple random sample without replacement. For each of the 13 items normally covered[2] in a Census a v is plotted along with the points $v \pm \hat{\vartheta}_v$. The curve which represents one-half of the $\hat{\vartheta}_v$ coefficient of variation for a 25 percent simple random sample from an EA with 168 listings provides a fairly good fit for these items.

FOOTNOTES TO APPENDIX

[1] See Morris H. Hansen, William N. Hurwitz and William G. Madow, "Sample Survey Methods and Theory, Vol. II; Theory," Chap. 10, sec. 1, John Wiley and Sons, Inc., New York, N. Y., 1953.

[2] (0, 1) variables excluded were: (1) Al's taken, (2) Refusals and not-at-homes, (3) NA on tenancy, (4) NA for off-farm work, (5) NA for off-farm income greater than farm sales.