

1944

A Design for Sampling Iowa Families

Bernice Brown
Iowa State College

Copyright © Copyright 1944 by the Iowa Academy of Science, Inc.
Follow this and additional works at: <https://scholarworks.uni.edu/pias>

Recommended Citation

Brown, Bernice (1944) "A Design for Sampling Iowa Families," *Proceedings of the Iowa Academy of Science*: Vol. 51: No. 1 , Article 34.
Available at: <https://scholarworks.uni.edu/pias/vol51/iss1/34>

This Research is brought to you for free and open access by UNI ScholarWorks. It has been accepted for inclusion in Proceedings of the Iowa Academy of Science by an authorized editor of UNI ScholarWorks. For more information, please contact scholarworks@uni.edu.

A DESIGN FOR SAMPLING IOWA FAMILIES

BERNICE BROWN

The sampling design described in this paper is one which was devised for a survey on Canning and Gardening in Iowa. This survey was initiated by the State Nutrition Council and M. L. Wilson, Director of Extension, United States Department of Agriculture. The sample was designed by the Statistical Laboratory and the investigation conducted by the Extension Service, Iowa State College.

The purpose of the survey was to make estimates of

1. Number of quarts of food canned, frozen, and stored by Iowa families
2. Number and size of 1943 gardens and plans for 1944
3. The amount of food spoiled and possible causes of spoilage
4. The amount of canning equipment needed for 1944.

The population of Iowa Families was divided into three strata: urban, rural non-farm, and rural farm. The number of occupied dwelling units in Iowa as shown in the 1940 census report is given in Table I.

Table I. Number of Occupied Dwelling Units in Iowa.

Stratum	No. Occupied Dwelling Units	No. of Places
Urban (2500 or more).....	312,393	89
Rural non-farm (less than 2500).....	161,077	1643
Rural farm	228,354	
	701,824	

The Census Bureau defines dwelling unit as the living quarters occupied by one household. For the purposes of this study the dwelling unit was used to represent the family. The sampling rate was 0.1 per cent of occupied dwelling units and was designed to get information from about 700 families. Since the variances in the state were unknown, the sampling was allocated according to only one condition, namely, that the number of sampling units in each stratum should be proportional to the total known number of dwelling units in that stratum.

SAMPLING FARM POPULATION

Since there are 228,000 farm families in the 99 counties in Iowa, it was planned to get about 228 interviews in the farm population. The farms were classified by county and the average size of the sampling unit in each county was 2.3 farms $\left(\frac{228}{99}\right)$. A count was made of the number of sampling units in each county and one sampling unit was

selected at random in each county. The sampling units varied from a single section to a quarter section depending upon the density of farms. County highway maps showing location of farmsteads were used. Every farm in a given county had an equal chance of being chosen in the sample from that county. However, the probability of a given farmstead appearing in the state sample varied inversely with the number of farms in the county. Since the County Extension Home Economists were asked to do the interviewing in their respective counties, the grouping of the farms into a single sampling unit seemed to be the only practicable plan. The efficiency of this grouping will be discussed later. Interviewers were instructed to visit every family residing in the area designated as the sampling unit. If a dwelling unit did not qualify as a farm according to census definition, it was marked by the enumerator as a rural non-farm dwelling, and was recorded as of the rural non-farm stratum. Attention may be directed to two necessary compromises in this design: (1) inequality of probability of selection of farms in the different counties and (2) only one sampling unit in each county.

SAMPLING RURAL NON-FARM POPULATION

More than 1600 towns in Iowa having a population under 2500 make up this stratum. It was not thought feasible to ask an interviewer to go into a town to take less than 3 interviews. The sampling rate was, therefore, designated as 3 out of 3000.

A listing was made of all towns in Iowa under 2500 population. The listing was by 9 geographic districts, the towns within districts arrayed according to population. The average size of family in this population group was reported by the 1940 census as 3.34. It follows that there would be approximately 3000 dwelling units in every 10,000 population. In order to get 3 interviews for every 3000 dwelling units, the list of towns was divided into groups containing approximately 10,000 persons each, and random numbers from 1 to 10,000 designated the towns in which interviews were to be taken. The probability of choosing a town was thus proportional to the number of inhabitants.

Fifty-two towns were chosen in this way as sampling units and three interviews were obtained in each town. A map of each town was secured or made, the blocks were numbered, and three blocks were selected at random. In most cases the interviewer was provided with a map indicating the blocks in which the interviews were to be taken. In towns over 1000 the average number of dwelling units per block was calculated from census data on number of dwelling units and the known number of blocks on the map. This would be in error to the extent that the number of blocks in a given town may have changed since the map was printed and the number of dwelling units may have increased or decreased since the 1940 census. However, the time and money which would be required in order to make a pre-listing of dwelling units and adjustments for out of date maps would have been prohibitive in this case. If the average number of *dwelling* units per block is designated as k , and a random number from 1 to

k is chosen as the first dwelling unit to be interviewed, the enumerator was instructed to take every k^{th} dwelling unit until the required number of interviews had been secured. In this way every block had equal probability of being selected in the sample and the probability of any dwelling unit being chosen for the interview was constant

$$\left(\frac{1}{k}\right).$$

In the towns under 1000 population the interviewer had to procure a map and in some cases it was necessary to make a rough map. In these cases the interviewer numbered the blocks, selected the required number of blocks at random, and designated the dwelling units to be interviewed by drawing a number at random. Every block in the town had an equal probability of being selected as one of the sub-sampling units but the probability of a given dwelling unit in the block being designated for an interview was inversely proportional to the number of dwelling units in the block.

SAMPLING THE URBAN POPULATION

Block statistics from the Census Bureau were used for the sampling in the four cities in Iowa which had a population of over 50,000 in 1930. Classification of dwelling units was geographic, by tracts or wards. The sampling rate of dwelling units was 1 in 1000. It was necessary only to pick random numbers from 1 to 1000 in order to designate the dwelling unit to be interviewed. Every dwelling unit had the same probability of being selected. For example, if the 6th dwelling unit in block 57 in ward 3 is the one designated by the random number, the interviewer was instructed to go to the northwest corner of block 57 and take the 6th dwelling unit east (clockwise) for the interview. If the random number drawn designated the 42nd dwelling unit in the block and the block statistics give the information that there are 44 dwelling units in that block, the interviewer was instructed to take the 3rd dwelling unit south from the northwest corner of the block (counter clockwise).

The selection of cities under 50,000 was made from a list of towns in the same way as the towns under 2500. There were two handicaps inherent in the sample design for this stratum: (1) there was no information as to the number of dwelling units in towns of less than 1000 population and (2) the sampling rate was 3 in 3000 instead of 1 in 1000.

RESULTS

The data in Table II give some indication of the representativeness of the sample.

Table II. Average Number of Persons Per Occupied Dwelling Unit.

Stratum	Ave. No. Persons per Occupied Dwelling Unit (1940 Census)	Ave. No. Persons per Occupied Dwelling Unit (Sample)
Urban	3.47	3.23
Rural non-farm	3.34	2.86
Rural farm	4.01	3.77
State	3.62	3.34

The Bureau of the Census reported that there had been a 9 per cent decrease in civilian population in Iowa during the period from April 1, 1940 to March 1, 1943. Assuming that there has been no appreciable change in the total number of dwelling units in Iowa, the average number of persons per dwelling unit in this sample is consistent with the census report in that it is about 8 per cent less than the average size as reported in 1940. This indicates that the sample was not biased on the number of persons per dwelling unit.

EXAMINATION OF DESIGN

There are two problems in regard to the sampling design that should be considered by an examination of the results of the survey. First, was the stratification and sub-stratification necessary and justified. Second, was the sub-sampling adequate.

The effect of stratification is shown by an examination of the means and variances in the several strata in Table III.

Table III. Means and Estimated Variances.

Stratum	Mean No. of Quarts Canned Per Family	Estimated Variance of Mean
Urban	165.1	23268
Rural non-farm	201.4	25525
Rural farm	297.8	30485

The estimated variances in the three strata do not differ significantly. Hence there was no apparent loss of efficiency from the allocation of the sampling according to the number of dwelling units alone. This is also shown by making a comparison of the optimum number of schedules designated under an $n\sigma$ allocation and the actual number of schedules obtained. This comparison is shown in Table IV.

Table IV. Comparison of Number of Schedules in Sample with Optimum Allocation.

Stratum	Optimum Allocation $n\sigma$	No. Schedules in Sample
Urban	295	300
Rural non-farm	159	155
Rural farm	247	237
	701	692

The effect of stratification is also shown by an examination of the mean squares for number of quarts canned as given in Table V.

Table V. Mean Squares and Degrees of Freedom for Number of Quarts Canned.

Variation	Degrees of Freedom	Mean Square
Between strata	2	1,202,000
Between dwelling units.....	689	26,163

The variation between strata is highly significant. The average number of quarts canned per family differs in the three strata. The computation of the estimated variance of an unstratified random sample shows that stratification increased the efficiency of the sampling 12.5 per cent.

SUMMARY

The results of the survey indicate that the sampling design was adequate for making estimates of the mean of the population. The stratification was effective and necessary. There is some evidence to show that it would have been as satisfactory to have taken more interviews in each of fewer counties and towns if such a procedure would have lowered the operational costs of conducting the survey.

IOWA STATE COLLEGE
AMES, IOWA