THE SAMPLE CENSUS OF THE AREA UNDER JUTE IN BENGAL IN 1940

By P. C. MAHALANOBIS
Indian Statistical Institute

EDITORIAL NOTE

At the instance of the Central Jute Committee, Professor P. C. Mahalanobis carried out a series of pilot studies from 1937 on the estimation of acreage and yield of Jute crop, which led to the organisation of a large scale sample survey in 1940 covering the whole province of Bengal. This was possibly the first large scale sample survey conducted anywhere in the world. Professor H. Hotelling who visited the Indian Statistical Institute in 1938 wrote in his report to the Central Jute Committee as follows :

> ···no technique of random samples has, so far as I can find, been developed in the United States or elsewhere, which can compare in accuracy or in economy with that described by Professor Mahalanobis.

The Jute survey was quickly followed by a series of large scale sample surveys for the estimation of acreage and yield of all important crops in Bengal and for the collection of socio-economic data through household enquiries. While conducting these surveys, Professor Mahalanobis was led to formulate the basic concepts and to lay down the logical foundations of the technique of large scale sample surveys. The first full length treatment of these foundations is given in the memoir on large scale sample surveys which appeared in 1944 in the *Philosophical Transactions of the Royal Society*, London (Series B, 231, 329-451).

The pioneering work of Professor Mahalanobis in the field of large scale sample surveys is widely recognised. He demonstrated the possibility of using large scale surveys to collect information quickly and economically with sufficient accuracy. He has made three notable contributions to the technique : (1) the concept of optimum design of the surveys (maximising precision at a given cost or minimising cost for assigned precision), (2) pilot surveys and (3) inter-penetrating net work of sub-samples (IPNS) and the control of non-sampling errors. The idea of pilot surveys aimed at the collection of information on cost and variance functions necessary for planning an efficient large scale survey was the fore-runner of the sequential method developed by the late Abraham Wald ten years later. In the introduction to his book *Sequential Analysis*, Late Abraham Wald wrote :

> The occasional practice of designing a large scale experiment in successive stages may be regarded as a forerunner of sequential analysis. ···
> ···A very interesting example of this type in the series of sample censuses of area of jute in Bengal carried out under the direction of P. C. Mahalanobis (1940).[1] Mahalanobis (1952)[2] has discussed the advantages of sequential sampling by considering what he terms, *historical* and *non-historical designs*.

[1] *Sankhyā*, **4**, 511-530

[2] *Sankhyā*, **12**, 1-7.

The IPNS, used in a proper way, is an ideal method for controlling and also for estimation of non-sampling errors which are inevitable in large scale sample surveys. The techniques developed in late thirties and early forties for survey of jute in Bengal and for the collection of socio-economic data through household enquiries were of great help in instituting the National Sample Survey, which is a nation-wide integrated, multi-subject survey being conducted in India since 1950.

Professor Mahalanobis prepared several reports for the Indian Central Jute Committee based on the results of his studies and surveys on the jute crop. Most of these reports are not easily accessible due to their having been marked confidential at the time of their publication and/or due to the passage of time. They contain a wealth of information on the theory and practice of sample surveys and it has been suggested to us that at least some of his earlier pioneering works should be republished and made available to research workers and sample survey technicians. As a first publication in this series, the "Report on the Sample Census of the Area under Jute in Bengal in 1940" prepared by Professor P. C. Mahalanobis in 1941 is being reprinted. It is proposed to follow this up with reprinting other earlier papers and reports on large scale sample surveys by Professor Mahalanobis and his associates at the Indian Statistical Institute. We hope that they are not only of historical but also of current interest, as the techniques developed in these papers are being increasingly used in modern surveys

The report on the Sample Census of the area under Jute in Bengal in 1940 is being reprinted as it was originally submitted with minor editorial changes and omitting some of the detailed tables which are not necessary for following the text of the report.

15 June, 1967.

C. R. Rao

CHAPTER 1 : GENERAL DESCRIPTION OF THE SAMPLE SURVEY IN 1940

History of the scheme

1. This report deals with the work done in the fourth (1940) season of the exploratory sample survey of the area under jute in Bengal. The scheme started in 1937 with a complete survey of 124 square miles, and a sample enumeration of 1,508 sample-units of size 5-acre and 14,159 random plots in two *thanas*[1] in 24 Parganas and Hooghly districts at a total cost of Rs. 6,830. The design of the survey was prepared in the Statistical Laboratory and the field work was in charge of the Director of Agriculture, Bengal; work on the statistical analysis was started in April 1938; and the Statistical Report was presented in August 1938.

2. In the meantime an *ad hoc* Jute Census Committee had been appointed and the scheme for an exploratory survey prepared by the Statistical Adviser was approved at the first meeting of the Committee held on the 6th July 1938. The field work was started on the 27th July and was finished on the 21st October; complete enumeration of crops on all plots was carried out in eight *thanas* in eight districts (Dacca, Mymensingh, Murshidabad, Pabna, Rajshahi, Rangpur, Nadia and Tipperah) covering 413.9 square miles; the random sample survey consisted of 7,888 sample units of three different sizes (1-acre, 4-acre, and 9-acre) and 2,540 random plots; and the total expenditure in the field branch was Rs. 18,876 exclusive of supervision but inclusive of Rs. 651 for the purchase of maps. Work in both field and statistical branches was carried out under the technical guidance of the Statistical Adviser and the administrative control of Mr. A. P. Cliff, Secretary, Indian Central Jute Committee. The Statistical Report, which was submitted on the 26th December 1938, was considered and the programme of work for 1939 was approved by the Jute Census Committee on the 6th February 1939.

3. In the third (1939) season complete enumeration was carried out in 683 square miles in seven *thanas* in three districts (Mymensingh, Rajshahi, and Pabna). The sample survey was carried out in 24 *thanas* covering 2,563 square miles, and consisted of 7,768 sample-units of size 1-acre, 2,260 grids of size 4-acre, 1,521 grids of 9-acre, 529 grids of 16-acre, and 314 grids of 36-acre or 12,311 grids of various sizes together with over 54,600 random plots. The training camp for field workers was opened on the 11th April; and the actual field survey was started on the 1st May and completed on the 30th August 1939. The total expenditure incurred for the field branch, which was under the administrative control of the Director of Land Records, Bengal, was Rs. 47,335 exclusive of cost of supervision and an expenditure of about Rs. 1,700 for maps. A block grant of Rs. 35,000 to the Indian Statistical Institute was sanctioned for the statistical portion of the work.

[1] *Thana* consists of a group of *unions* under the jurisdiction of a police station.

4. A preliminary report was prepared on the 1st October on the basis of which certain arrangements for preparatory work for 1940 were approved by the Jute Census Committee on the 10th October 1939. The Statistical Report, which was submitted on the 8th November 1939, was considered on the 13th December when the provisional programme was approved; fuller details were considered and approved by the Jute Census Committee on the 3rd April 1940. As in 1939 a block grant to the Indian Statistical Institute was sanctioned for the statistical work; it was also arranged that the work in the field branch would be carried out under the direct guidance of the Statistical Adviser but all payments for the field branch were to be made from the office of the Indian Central Jute Committee under the administrative control of the Secretary. This arrangement, which had worked quite smoothly in 1938, was convenient in every way.

AREA COVERED IN 1940

5. The sample survey in 1940 covered eight representative jute growing districts of Bengal, namely, Dacca (3.026) and Mymensingh (2.928) in Dacca Division, Tipperah (2.612) in Chittagong Division, Rangpur (2.628), Bogra (2.488) in Rajshahi Division, Jessore (1.518) and Nadia (1.096) in the Presidency Division. The statistical analysis of the material collected in 1939 had shown 4-acre to be on the whole the most economical size of the sample-unit. The best density (number of sample-units or grids per square mile) at the appropriate level of expenditure for 4-acre sample-units was calculated with the help of variance and cost functions for each of the eight districts. (These optimum densities are shown within brackets against each district, but it should be remembered that in the half-sample method the actual density in each half-sample was only half the density given within brackets).

6. The standard size of sample-units was 4-acre; but in order to study the cost and variance-functions, a comparatively small number of sample-units of size 1-acre, 2.25-acre, 6.25-acre and 9-acre were also used. The actual distribution of sample-units of different sizes in the different districts are given in Table 14.

7. In Rangpur cycle-units, that is, investigators provided with bicycles for expediting the work, were used for sample-units of size 16-acre with a thin density of one in eight square miles covering 405 square miles in four *thanas*. Another cycle-unit was used in three *thanas* in Rangpur district covering 488 square miles for enumerating 70 *mauzas*[1] picked at random with an effective density of 70 in 488 or about 1 in 7 square miles. Ordinary (non-cycle) units were employed to enumerate 100 *mauzas* in the form of two-half-samples picked at random over 1,246 square miles.

8. Information was collected for 41,345 sample-units of various sizes by the ordinary grid method covering 18,414 square miles. In Rangpur 2,139 square miles were covered by the special cycle and *mauza* units; the total area surveyed in 1940 was thus 20,553 square miles.

[1] *Mauza* is an area unit comprising of plots (parcels of land) and occupying an area of about 0.75 sq. mile on the average.

SAMPLE CENSUS OF THE AREA UNDER JUTE IN BENGAL IN 1940

PREPARATORY WORK IN THE LABORATORY

9. In each district the first thing necessary was to sort and arrange the *mauza* maps; to prepare sheet lists and compile a *mauzawar*[1] list of areas. The next stage was the partition of the area into two half-samples; and the distribution, by *thanas*, of the required number of sample-units into each of the half-samples. After this preliminary work was done, random numbers were copied or prepared and with their help sample-points were located at random on the *mauza* sheets; and the position of sample-units (grids) of various sizes were marked on the maps; and lists were then prepared of all the plots which fell within the sample-units. These lists were checked, and duplicate copies of *khasra*[2] sheets were prepared for the field branch.

10. The magnitude of the task may be gauged from the fact that it involved handling nearly 55,000 *mauza* sheets, locating about 42,000 random points, and listing 8,96,938 or nearly nine lakhs of individual plots. Naturally this preparatory work had to be organized on something like mass production lines, smooth working of which depended on the primary material (in the form of *mauza* and *thana* maps and *jurisdiction lists*[3]) being readily available. Unfortunately the supply of such primary material was not regular which appreciably hampered the work in the Laboratory and also caused unnecessary waste of time and expenditure.

11. However, the preparatory work, which was started on the 15th November 1939, was pushed on with all possible speed and the bulk of the material was made ready for the field branch by the end of March 1940; and the actual work was completed on the 5th May 1940. Credit is due to the computing section under the general charge of Jitendra Mohan Sengupta and Jitendra Nath Taluqdar for finishing this work in good time.

PREPARATORY FIELD WORK

12. Preparatory work was started by the field branch in the middle of October 1939 with a small number of workers under one chief inspector for collecting *mauza* maps from district record rooms and also for gathering local information regarding the approximate time for sowing and harvesting, *char areas*[4] etc. I understand that the total expenditure incurred for this purpose (exclusive of the price of maps) up to the end of March 1940 was about Rs. 7,800. I have not had any opportunity of examining the details of the preparatory field work; I am therefore unable to supply any analysis of costs. It was essential of course to send out men to district headquarters to fetch *mauza* maps; I am not sure, however, whether the expenditure for collecting local information and touring *char* areas was really necessary.

---

[1] *Mauzawar* : Mauzawise.

[2] *Khasra* is a record book consisting of detailed plots with their area and other particulars.

[3] *Jurisdiction list* is a list of *mauzas* within a *thana*. The identifying permanent serial numbers assigned to the *mauzas* in this list are known as *J. L. numbers*.

[4] *Char areas* are those obtained from river beds after the river has either changed its course or has dried up.

FIELD STAFF

13. As regards main operations on the field considerable difficulties were experienced owing to the delay in securing the services of a supervisor. In 1938 and in 1939 the field work was carried out with success under the effective charge of Mr. Nihar Chandra Chakravarti, Secretary, Board of Economic Enquiry, as Supervisor, whose services had been made available for this purpose by the Government of Bengal. The Committee had been informed by Government fairly early that Mr. Chakravarti's services could not be spared for the present scheme in 1940; but there was great delay in arranging the deputation of Mr. Sudhansu Kumar Banerji, Deputy Magistrate and Collector, who actually assumed charge as Supervisor on the 3rd April 1940. In the meantime it had been arranged that the field work in 1940 should be carried out, as in 1938, under the administrative control of the Secretary, Jute Census Committee, with the direct technical guidance of the Statistical Adviser; and at my suggestion Mr. Dhirendra Mohan Ganguli, Statistical Assistant in the office of the Indian Central Jute Committee, had been appointed Assistant Supervisor from the 27th March 1940. The first brunt of making detailed arrangement for the field survey (which included working out of the programme for different field units, printing various forms and schedules, and recruiting the field staff) fell on him; and the successful organization of the sample survey was possible mainly due to the efficiency and ability with which he completed this work. The whole of the accounts work was under the direct charge of the Supervisor, and the excellent arrangements he made for prompt payment of salaries and allowances to field workers helped materially in the smooth progress of the survey.

14. The whole area of about 21,000 square miles was divided into six blocks and twenty-one sub-blocks; two field units each consisting of from 6 to 8 investigators in charge of one inspector were posted in each sub-block; and about three sub-blocks made up a block under the charge of one chief inspector. Besides 15 agricutural overseers and 60 agricultural demonstrators of the Indian Central Jute Committee, whose services were made available for the present survey, an additional staff of 172 investigators, 19 inspectors and 6 chief inspectors were directly recruited; the total strength of the field staff thus consisted of 232 investigators, 32 inspectors and 6 chief inspectors under the Assistant Supervisor and Supervisor. A small staff consisting of an accountant, one head clerk and other clerks were employed in the field office for which rooms were rented at 249-D, Bowbazar Street. Among the staff originally appointed 41 investigators and 2 inspectors did not turn up at all and fresh recruitments had to be made in their place. During the period of about three months and a half of field work 32 investigators, 2 inspectors and 1 chief inspector resigned their posts; leave for more than one week on medical grounds had also to be granted to 20 investigators and 2 inspectors.

Field survey

15. Recruitment and training of field workers started in the middle of April and actual field survey commenced in the different districts in the first week of May. The regular grid survey was completed on the 10th July in Dacca and Nadia, 15th in Tipperah, 18th in Rajshahi, 25th in Rangpur, 28th in Jessore and on the 30th July in Bogra. The special (cycle and *mauza*) units started their work in the 2nd and 3rd week of June in the northern portion of Rangpur and completed the field enumeration on the 31st August 1940.

16. The investigators were provided with *mauza* maps and *khasra* lists of plots; and they had to locate the plots in the field, and record whether any jute was grown on them or not; in case jute was grown on only one portion of a plot, an *anna estimate of the area*[1] under jute was also entered in the *khasra*.

17. The inspectors were required to distribute the work among investigators and otherwise coordinate work in each sub-block. Besides this approximately ten per cent of the primary work was inspected and checked by the inspectors and the chief inspectors who constantly toured over the area under their respective charge; four additional inspectors were appointed for the exclusive duty of checking and inspecting the field enumeration at random; and the Supervisor and Assistant Supervisor between them inspected the work in all the districts. Mr. D. L. Mazumdar, Secretary, Indian Central Jute Committee, personally inspected the work in Tipperah and Jessore districts.

18. In spite of many difficulties connected with the recruitment, training and control of a large staff of workers, the field survey was carried out smoothly and finished in good time. Although the Supervisor had no previous experience of this type of survey, he quickly picked up the work and credit is due to him and the Assistant Supervisor for the successful completion of this year's survey. The Supervisor in his report dated 21st August 1940, specially mentioned the good work done by Messrs. Kanti Ranjan Banerjee, Saroj Chandra Chakravarti, Pranay Kumar Chatterjee, Bijay Kumar Das Gupta, Lokes Chandra Guha Roy and Jyotiprokas Sen.

Statistical analysis of the material

19. The *khasra* lists which were sent to the Laboratory directly from the field, began to come in from the 18th May. In the Laboratory these lists were sorted and arranged in the proper order and then compared with the original lists of sample-units. The next stage was "area extraction", that is, measurement of the area of plots under jute which was done with the help of transparent square-scales placed directly on *mauza*-maps; this was started on the 20th May. After the area of individual plots were measured the *anna* figures (for the portion of the plot under jute) given in the

[1] *Anna estimate of the area* is an eye-estimate of the proportion of area under jute in a plot (parcel of land) expressed in terms of *annas* taking the total area of the plot as one rupee consisting of 16 *annas*. Prior to the change over to decimal coinage, one-sixteenth of an Indian rupee was known as *anna*. For instance, if the plot is completely under jute, the *anna* estimate of crop area is taken as 16 *annas* and if only half the plot is under the crop, it is taken as 8 *annas*.

khasra lists were converted into acres, and actual tabulation started on the 21st May. From these primary records the next step was the calculation of the values of p (the proportion of land under jute) within each individual sample-unit; and of average values for *mauzas*, *unions*, *thanas*, etc., together with relevant standard errors. This completed what may be called the primary stage of the statistical analysis; some idea of its volume may be obtained from the fact that such calculations had to be done for about 42,000 sample-units involving 2,72,937 or a little less than three lakhs of individual plots which were under jute.

20. In the next stage, values for the two half-samples were compared which involved the calculation of appropriate standard errors, and values of certain statistical quantities which I have called (u), and probabilities of occcurrence $P(u)$[1] with the help of advanced statistical methods. The statistical results were gradually built up starting from individual sample-units and *mauzas* and carried through *unions*[2] and *thanas* to districts; its volume may be judged from the number of tables for unions included in the present report which however give only a condensed version of the whole material.

COST AND VARIANCE ANALYSIS

21. Besides the *khasra* lists, the time records and individual diaries kept by each field worker were classified and tabulated in detail for the study of the cost function. The next stage was the breaking down of the time records and cost into various components, such as enumeration, small journey, big journey, miscellaneous, indirect[3] etc.; calculations of various averages, and standard errors; and the preparation of two-way and summary tables some of which are reproduced in this report. After this primary analysis was over the material had to be studied analytically for which suitable methods of graduation had to be worked out by a laborious process of trial and success. From the statistical point of view the building up of the cost function was in many ways the most difficult problem and demanded strenuous work on the part of the senior staff of the Laboratory. Another part of the statistical analysis was concerned with the study of the variance function or the relation between the variability and the size of sample-units; this also presented many difficulties and required sustained analytical studies by senior statisticians.

22. After the variance and cost functions werer educed to a suitable mathematical form it was possible to take up the question of the best size and optimum density of sample-units in the different districts. The calculations were laborious and called for a good deal of mathematical skill and ingenuity.

23. The working time included in the cost analysis of statistical work was almost exactly 10,000 computer-days exclusive of all work done in connexion with the advanced mathematical analysis and the writing of the report. Even the preparation of the statistical tables required a good deal of careful work as it involved condensing a vast amount of primary material into a suitable form which would be

[1] These terms are explained in Chapter 2.

[2] *Union* is a group of about 10 to 12 *mauzas*.

[3] These terms are fully explained in Chapter 4.

easily intelligible to persons without technical knowledge of statistics. The mere typing, stencilling and printing of the report running to 275 pages[1] and including more than 200 pages of tabular matter involved a great deal of laborious work on the office staff of the Laboratory; the typing and stencilling took more than 6 men-months.

24. Great credit is due to Samarendranath Roy, Senior Statistician of the Laboratory, ably assisted by Birendranath Ghosh, Purnendu Kumar Bose, Anil Bhattacharyya, and others, for the difficult work connected with the advanced statistical analysis; to Jitendra Mohan Sen Gupta and Jitendra Nath Taluqdar for general supervision of the preparatory and statistical work in the Laboratory; to Sambunath Halder Nimai Charan Ghosh, S. Raja Rao, Rajendra Chandra Roy, Haribhajan Choudhury, Subodh Chandra Bhowmick, Krishna Chandra Banerji, Akhil Chandra Nandi, and others in the computing section; to Sudhir Kumar Banerjee for cost analysis and budget work; and to Naresh Chandra Mukherjee, Susanta Ranjan Guha, Jaladhar Sarma, Subodh Chandra Das Gupta and other for typing and general office work.

Crop-estimating survey

25. I had emphasized from the very beginning the need of starting systematic work on crop-cutting experiments, and had discussed this question in my first *Report on the Experimental Crop Census of* 1937 :

> "In the case of Jute, or any other crop, acreage figures by themselves will not enable any accurate forecast being made of the total yield. Crop-cutting experiments are essential for this purpose; and there is everything to gain in working on both the problems at the same time. The additional cost of supervision or inspection will be nominal, while large economies will be possible for the field work if the same set of enumerators are employed to "collect both sets of material." (*Proceedings, Jute Census Committee*, 1st September 1938, p. 69).

26. The Jute Census Committee at its meeting held on the 25th July decided that an exploratory crop-cutting experiment should be undertaken on a small scale this year. A suitable design was prepared in the Statistical Laboratory for this purpose; and a field survey was carried out in 13 *thanas* in 4 districts (Mymensingh, Dacca, Rangpur and Tipperah). The field staff consisted of 4 inspectors and 24 investigators who were appointed between the 27th and 30th July; actual work on the field started on the 3rd August and was completed on the 17th September 1940. Arrangements are being made for the statistical analysis of this material and a separate report will be submitted in due course.

Acknowledgements

27. I received the fullest cooperation of Mr. C. R. Nodder who officiated as Secretary of the Committee after the departure from India of Mr. A. P. Cliff. His successor, Mr. D. L. Mazumdar, I.C.S., has also helped me in every possible way; without his able cooperation it would have been impossible to carry out the sample survey so smoothly and with such efficiency and economy.

[1] Exclusive of summary.

CHAPTER 2: ACCURACY OF THE SAMPLE SURVEY

28. One of the chief objects of the Sample Survey in 1940 was to test the accuracy of the results attained in a practice by a field trial on a large scale. The information was therefore collected in the form of two independent half-samples (A) and (B), each of which consisted of roughly half the total number of sample-units. In each half-sample the sample-units were located at random over the whole area, and the field operations for each half-sample were carried out by different sets of investigators. The difference between the results obtained from the two half-samples will, therefore, furnish a good idea of the margin of error inherent in the present method.

COMPARISON OF HALF-SAMPLE BY UNIONS

29. The results by *unions* are shown in Tables 1.001–1.115 in which col. (1) gives the serial number, and col. (2) the name of the *union*. Relevant information for half-sample (A) is given in cols. (3)-(5) and for half-sample (B) in cols. (6)-(8); cols. (3) and (6) show the size of the sample-units (grid) in acres, and cols. (4) and (7) the number of such sample-units used in the two half-samples respectively; col. (5) gives the proportion (in percentage of land) under jute together with the standard error in the case of half-sample (A); and col. (8) the corresponding proportion (in percentage) under jute in half-sample (B). The difference between the two estimates is given in col. (9) together with the standard error of the difference; and this difference divided by its own standard error is given in col. (10). The corresponding probability of occurrence calculated from the Normal (Gauss-Laplacian) probability integral, is given in col. (11); in this column values of probability less than .01 but equal to or greater than .001 have been marked as [00*], while values less than .001 have been specifically marked as $< .001$.

30. I have explained in previous Reports and Notes (e.g., Section 4, paragraphs 21–44, and Section 5, paragraphs 54–57 of the *First Report on the Crop Census of* 1938) the statistical procedure for comparing the accuracy of two different estimates. Broadly speaking, if the difference between the two estimates is of the same order as the standard error of the difference then the two estimates must be considered to be in satisfactory agreement from the statistical point of view. For convenience of reference we shall call the difference between the two half-samples (A) and (B) divided by the standard error of the difference as u; these are the values shown in col. (10). The statistical agreement between the two half-samples is to be judged by the magnitude of u ; the smaller the value of u the better is the agreement between the two half-samples; and the larger the value of u the greater is the discrepancy between the two estimates.

31. The corresponding value of $P(u)$, given in col. (11), shows the probability of occurrence of u. If this probability is large, it means that the difference between the two estimates could easily have come about through chance errors of sampling;

TABLE 1. COMPARISON OF HALF-SAMPLES BY UNIONS

sl. no.	union: name	size in acre	half-sample (A): number	half-sample (A): percent under jute ± s.e.	size in acre	half-sample (B): number	half-sample (B): percent under jute ± s.e.	difference between half-samples: actual ± s.e.	difference between half-samples: divided by s.e. (=u)	approximate probability P(u)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
				District Mymensingh : Police Station—Khaliajuri						
1.	Mendipore	4.00	22	22.73± 5.10	4.00	36	3.00± 3.09	19.73± 5.99	3.29	0.00*
2.	Chakua	,,	46	4.48± 2.17	,,	37	0.14± 0.17	4.37± 2.18	1.99	.04
3.	Khaliajuri	,,	39	0.21± 0.22	,,	13	1.38± 1.36	−1.17± 1.37	−0.85	.40
4.	Panchat	,,	16	5.06± 2.53	,,	22	0.00± 0.00	5.06± 2.53	2.00	.04
5.	Baghatia	,,	16	0.81± 0.56	,,	23	0.00± 0.00	0.81± 0.55	1.47	.14
6.	Krishnapur	,,	13	2.62± 2.32	,,	21	0.38± 0.37	2.24± 2.35	0.95	.34
	total	,,	152	5.54± 1.19	,,	152	0.91± 0.74	4.63± 1.40	3.31	<.001
				Police Station—Madan						
1.	Kailatibaron	4.00	26	23.11± 6.70	4.00	12	18.36± 6.61	4.75± 9.41	−0.50	0.62
2.	Jahangirpur	,,	12	36.81± 7.12	,,	8	36.78±12.03	0.03±13.98	0.00	1.00
3.	Madan	,,	11	42.32±12.05	,,	17	7.85± 3.55	34.47±12.56	2.74	.00*
4.	Gobinda Sree	,,	19	16.42± 7.50	,,	23	6.25± 4.39	10.17± 8.69	1.17	.24
5.	Maghan	,,	3	11.00±11.00	,,	30	6.76± 4.52	4.24±11.89	0.36	.72
6.	Tiya Sree	,,	16	29.42± 7.81	,,	11	47.14±11.45	−17.72±13.86	−1.28	.20
7.	Naikpur	,,	19	39.87± 6.93	,,	10	36.23±11.12	3.64±13.10	0.28	.78
8.	Fatepur	,,	26	34.82± 8.14	,,	19	8.91± 6.12	25.91±10.18	2.54	.02
	total	,,	132	30.20± 3.07	,,	130	15.73± 2.54	14.47± 3.98	3.64	<.001
				Police Station—Kisoreganj						
1.	Latifabag	4.00	15	38.67± 5.39	4.00	15	39.93± 4.90	− 1.26± 7.29	−0.17	0.86
2.	Maijkhapan	,,	5	35.40±13.45	,,	22	33.95± 1.42	1.45±13.53	0.11	.91
3.	Josodal	,,	18	31.28± 4.90	,,	8	34.75± 8.00	− 3.47± 9.38	−0.37	.71
4.	Binnati	,,	12	35.83± 5.10	,,	12	29.17± 5.48	6.66± 7.49	0.89	.37
5.	Chauddasa	,,	15	26.53± 6.17	,,	7	26.86± 7.88	− 0.33±10.01	−0.03	.98
6.	Karshakariail	,,	20	39.45± 5.66	,,	4	47.00± 8.95	− 7.55±10.59	−0.71	.48
7.	Municipality				,,	20	33.50± 4.58			
	total	,,	85	34.56± 2.47	,,	88	34.32± 2.18	0.24± 3.30	0.07	.92

Note : Out of 115 tables only parts of Tables 1.058 and 1.059 are printed here as specimen.

on the other hand, if the probability of occurrence $P(u)$ is small, it indicates that the discrepancy cannot be reasonably ascribed to chance errors. If we glance through cols. (10) and (11) we find the values of u are comparatively small, and values of $P(u)$ are comparatively large showing on the whole good agreement between the two half-samples.

32. I may mention here that the standard error of the difference between the two estimates has been calculated on the basis of the classical formula, namely, $\sqrt{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)}$ where s_1^2 and s_2^2 are the two variances and n_1 and n_2 the size of the two half-samples respectively. This formula however can be used rigorously only when the two variances are the true (or population) variances of the two half-samples, that is, when P_1^2 and P_2^2 are used instead of s_1^2 and s_2^2 respectively. In actual practice we do not however know the population variances P_1^2 and P_2^2 and we are obliged to use the sample estimates s_1^2 and s_2^2 in their place. Unfortunately the exact distribution is not known when these sample estimates are substituted for the population variances; and this is why the probability of occurrence given in col. (11) is only an approximate estimate of the true probability. It is clear however that the probability of occurrence (calculated from the "'normal" probability integral) given in col. (11) is actually less than the corresponding true probability so that our approximate results are on the safe side in every case.

33. It will be noticed that our test really depends on the distribution of $P(u)$ which is shown in Table 2 for 1531 *unions* classified under different districts. In this table, col. (0.1) gives the range of $P(u)$; cols (1.1) and (1.2) the actual observed and expected number of $P(u)$ in the case of district Bogra, cols. (2.1) and (2.2) in the case of district Dacca and so on. Finally, cols. (9.1) and (9.2) show the observed and expected number of $P(u)$ for all eight districts taken together. If we compare the observed and expected values, either in the case of individual districts, or in the case of all eight districts taken together given in cols. (9.1) and (9.2), we find that smaller values of $P(u)$ less than 0.1 are comparatively more numerous than large values of $P(u)$ exceeding 0.9. This is however what is at least partly to be expected on account of the fact that the values of $P(u)$ used by us are only approximate and are actually smaller than the true values. In other words, from theoretical considerations we expect that the distribution of $P(u)$ will be shifted towards the lower end; that is, smaller values of $P(u)$ will be comparatively more numerous, which is fully confirmed in Table 2.

34. At the same time it will be seen that very low values of $P(u)$ are rare; out of 1531 values of $P(u)$ only 53 are less than 0.001 showing that highly improbable values are not too frequent. This is most reassuring; and on the whole we may conclude that the results are satisfactory although they may not be altogether ideal.

35. In order to make allowances for using the sample estimates of the variances (s_1^2 and s_2^2) in the place of population parameters (P_1^2 and P_2^2) we have also examined the application of Fisher and Behrens' test of significance in the case of the Tipperah

TABLE 2. FREQUENCY DISTRIBUTION OF VALUE OF $P(u)$ BY UNIONS

district→	Bogra		Dacca		Jessore		Mymensingh		Nadia		Rajshahi		Rangpur		Tipperah		8 districts	
range of $P(u)$	observed	expected	observed	expected	observed	expected	observed	expected	observed	expected	observed	expected	observed	expected	observed	expected	observed	expected
(0.1)	(1.1)	(1.2)	(2.1)	(2.2)	(3.1)	(3.2)	(4.1)	(4.2)	(5.1)	(5.2)	(6.1)	(6.2)	(7.1)	(7.2)	(8.1)	(8.2)	(9.1)	(9.2)
<.001	1	0.13	9	0.28	6	0.13	20	0.47	3	0.10	4	0.09	2	0.10	9	0.22	53	1.53
.001 — .01	6	1.18	14	2.54	8	1.20	26	4.25	5	0.89	7	0.84	7	0.86	9	1.99	82	13.72
.02 — .05	9	5.24	19	11.28	8	5.32	40	18.88	7	3.96	8	3.72	9	3.84	19	8.84	119	61.24
.06 — .10	14	6.55	23	14.10	12	6.65	35	23.60	13	4.95	9	4.65	9	4.80	16	11.05	131	76.55
.11 — .20	15	13.10	34	28.20	15	13.30	49	47.20	11	9.90	9	9.30	10	9.60	28	22.10	171	153.10
.21 — .30	11	13.10	32	28.20	17	13.30	40	47.20	10	9.90	6	9.30	10	9.60	21	22.10	147	153.10
.31 — .40	12	13.10	16	28.20	5	13.30	36	47.20	18	9.90	11	9.30	9	9.60	20	22.10	127	153.10
.41 — .50	11	13.10	25	28.20	10	13.30	45	47.20	6	9.90	9	9.30	14	9.60	20	22.10	140	153.10
.51 — .60	10	13.10	26	28.20	9	13.30	37	47.20	7	9.90	6	9.30	3	9.60	9	22.10	107	153.10
.61 — .70	10	13.10	18	28.20	14	13.30	37	47.20	7	9.90	5	9.30	6	9.60	22	22.10	123	153.10
.71 — .80	9	13.10	22	28.20	9	13.30	37	47.20	5	9.90	5	9.30	5	9.60	18	22.10	111	153.10
.81 — .90	9	13.10	20	28.20	12	13.30	35	47.20	3	9.90	7	9.30	5	9.60	12	22.10	103	153.10
.91 — .95	8	6.55	12	14.10	6	6.65	20	23.60	3	9.90	4	4.65	5	4.80	10	11.05	68	76.55
.96 — .99	6	5.24	11	11.28	2	5.32	12	18.88	—	—	3	3.72	2	3.84	8	—	44	61.24
1.00	—	1.31	1	2.82	—	1.33	3	4.72	1	—	—	0.93	—	0.96	—	—	5	15.37
	131	131.00	282	282.00	133	133.00	472	472.00	99	99.00	93	93.00	96	96.00	221	221.00	1531	1531.00

district. The relevant data are given in Table 3a in which col. (1) gives the name of the Police Station, col. (2) the number of *unions* for which the comparison is made, col. (3) the observed values of (u); col. (5) the values of $P(u)$ based on the normal probability integral; while col. (4) shows the five per cent level of (u) in accordance with Fisher and Behrens' formula, these figures were obtained by interpolation from P. V. Sukhatme's numerical table published in *Sankhyā* (Vol. 4, Part 1, December 1939).

TABLE 3a. FISHER AND BEHREN'S TEST OF SIGNIFICANCE FOR UNIONS BY POLICE STATIONS IN DISTRICT TIPPERAH

name of the police station	number of unions	values of (u)		$P(u)$ (normal probability integral)
		observed	Fisher and Behren's five per cent	
(1)	(2)	(3)	(4)	(5)
Brahmanbaria	4	2.36	2.22	0.22
Brahmanbaria	16	2.12	2.22	.03
Laksham	7	2.35	2.17	.02
Faridganj	6	2.41	2.16	.02
Bancharampur	4	2.19	2.14	.03
Kashba	8	2.06	2.25	.04
Burichang	1	2.22	2.17	.03
Chandina	4	2.44	2.11	.02
Chandina	11	1.98	2.28	.04
Chandina	15	2.19	2.11	.03
Daudkandi	4	2.08	2.19	.04
Daudkandi	9	2.24	2.20	.03
Hajiganj	4	2.12	2.14	.03
Matlabbazar	6	2.29	2.16	.02

36. It will be noticed that the five per cent level according to Fisher and Behrens' formula is usually of the same order but slightly less than the observed values of (u); this shows that the probability of occurrence in our case is slightly below the five per cent level; and the values of $P(u)$ based on the normal probability integral given in col. (4) fully confirm this. We find then that the results given by using Fisher and Behrens' test are on the whole in agreement with the results obtained by calculating $P(u)$ from the normal probability integral. In view of this agreement, and also in view of the fact that Fisher and Behrens' test is itself restricted by a rather stringent condition (namely, that the ratio of the sample variances is constant) we have not considered it necessary to make a similar comparison for other districts. Our general conclusion is that the two half-samples (A) and (B) are on the whole in satisfactory agreement.

TABLE 3b. FISHER AND BEHREN'S TEST FOR DISTRICT

name of the district	number of police stations	values of (u) observed	values of (u) Fisher and Behren's five per cent	$P(u)$ (normal probability integral)
(1)	(2)	(3)	(4)	(5)
Jessore	21	1.96	1.982	0.05
Rajshahi	5	2.02	1.979	.04
Rangpur	29	2.02	1.972	.04
Mymensingh	2	2.05	1.982	.04
Mymensingh	8	1.98	1.982	.05
Mymensingh	30	1.99	1.978	.04

COMPARISON OF HALF-SAMPLES BY THANAS

37. The comparison of the two half-samples (A) and (B) for individual *thanas*[1] is shown in Table 4.01 in which col. (2) gives the name of the Police Stations, col. (3) the total area of Police Station in square miles; and col. (4)–(6) and (7)–(9) relevant information for the two half-samples (A) and (B) respectively. Cols. (4) and (7) give the size of the sample-units (grids) in acre, and cols. (5) and (8) their number; the percentage of land under jute in the case of half-sample (A) is given in col. (6), and the corresponding proportion of land under jute in half-sample (B) in col. (9). The difference between the two half-samples together with standard error is shown in col. (10); and the values of (u), that is, the difference between the two estimates divided by the standard error of the difference, in col. (11). Finally the approximate probability of occurrence $P(u)$, based on the normal probability integral, is given in col. (12). As in the case of *unions* it will be seen that the values of (u) given in col. (11) are usually small and the values of $P(u)$ are fairly large.

38. In Table 4.10 results are given for certain *thanas* in Rangpur district for which the information was collected in the form of two half-samples collected by cycle-units which were furnished with bicycles for expediting the field work. It will be seen that the values of (u) given in col. (11) are all small and less than 2, while the values of $P(u)$ given in col. (12) are all greater than 0.15; this shows satisfactory agreement between the two half-samples.

39. In district Rangpur arrangements were also made to carry out an enumeration of 57 *mauzas* selected at random, of which 25 were included in half-sample (A) and 32 in half-sample (B). The comparison given at the bottom of Table 4.01 shows that there was excellent agreement between the two half-samples. The propor-

[1] Preliminary results of the comparison of half-samples by thanas were given in Tables 1-8 of the Preliminary Report dated 26th August 1940. These results were subsequently checked and certain minor corrections were made which have been incorporated in the revised values given here in Table 4.01 —4.12.

TABLE 4.01. COMPARISON OF HALF-SAMPLES BY POLICE STATIONS

sl. no.	name of police station	area in sq. miles	half-sample (A)			half-sample (B)			difference between half-samples		approximate probability $P(u)$
			size in acre	number	percent under jute ± s.e.	size in acre	number	per cent under jute ± s.e.	actual ± s.e.	divided by s.e. ($=u$)	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
					district : Bogra						
1	Adamdighi	120	4.00	146	10.29±1.42	4.00	142	15.32±1.73	−5.03±2.24	−2.25	0.02
2	Bogra	148	4.00	187	19.06±1.73	4.00	188	15.72±1.73	3.34±2.43	1.37	.17
3	Dhunot	95	4.00	117	30.01±2.45	4.00	117	35.50±2.65	−5.49±3.56	−1.54	.12
4	Dupchanchia	63	4.00	77	3.85±0.72	4.00	77	7.61±1.42	−3.76±1.47	−2.56	.02
5	Gabtali	94	4.00	115	39.27±2.65	4.00	115	34.43±2.24	4.84±3.49	1.39	.16
6	Jaipurhat	92	4.00	111	31.29±2.65	4.00	111	32.69±2.65	−1.40±3.74	−0.37	.71
7	Kahaloo	93	4.00	114	22.53±0.65	4.00	114	4.50±0.96	−1.97±1.16	−1.70	.09
8	Khetlal	119	1.00	146	5.36±1.42	4.00	146	5.06±0.83	0.30±1.57	0.19	.85
9	Nandigram	103	4.00	122	0.39±0.14	4.00	122	0.83±0.22	−0.44±0.26	−1.69	.09
10	Panchbibi	108	4.00	133	18.26±1.73	4.00	133	15.05±1.73	3.21±2.61	1.23	.22
11	Sariakandi	200	4.00	244	24.11±2.00	9.00	245	26.87±1.73	−2.76±2.51	−1.10	.27
12	Sherpur	114	6.25	140	11.52±1.42	4.00	140	13.17±1.73	−1.65±2.20	−0.75	.45
13	Shibganj	122	4.00	149	20.27±1.73	2.25	149	18.36±1.73	1.91±2.45	0.78	.43
district Bogra (13 thanas)		1471	—	1801	17.14±0.51	—	1799	17.75±0.50	0.61±0.72	−0.85	0.40

Note : Out of 12 tables only Tables 4.01 and 4.10 are printed here as specimen.

TABLE 4.10. COMPARISON OF HALF-SAMPLES BY POLICE STATIONS

sl. no.	name of police station	area in sq. miles	half-sample (A)			half-sample (B)			difference between half-samples		approximate probability $P(u)$
			size in acre	number	percent under jute±s.e.	size in acre	number	percent under jute±s.e.	actual±s.e.	divided by s.e. (=u)	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
					district—Rangpur						
1	Fulchari	121	4.00	133	21.37±2.08	4.00	123	20.40±1.74	0.97±2.71	0.36	0.72
2	Gobindaganj	178	4.00	233	29.12±1.58	4.00	232	23.87±1.47	5.25±2.16	2.43	.02
3	Gaibandha	124	2.25	159	38.31±1.87	4.00	158	33.63±2.00	4.68±2.73	1.71	.09
4	Palashbari	74	4.00	97	41.01±2.00	4.00	97	32.02±2.08	8.99±2.88	3.12	.00*
5	Pirgachha	100	4.00	131	32.69±1.66	4.00	131	30.69±1.50	2.00±2.24	0.89	.37
6	Pirganj	159	4.00	209	18.18±1.50	2.25	209	18.43±1.46	−0.25±2.09	−0.12	.90
7	Sadullapur	90	4.00	118	40.97±1.97	4.00	118	38.26±1.94	2.71±2.76	0.98	.33
8	Saghata	88	4.00	100	39.48±2.22	9.00	94	33.60±1.69	5.88±2.79	2.11	.04
9	Sunderganj	161	4.00	200	36.54±1.61	4.00	204	40.78±1.35	−4.24±2.10	−2.02	.04
	nine thanas	1095	...	1380	31.74±0.60	...	1366	29.43±0.56	2.31±0.82	2.82	.00*
1	Domar	97		12	27.17±4.46		12	29.50±4.16	−2.33±6.10	−0.38	.71
2	Jaldhaka	126	16—	16	34.25±6.17	16—	15	32.53±3.00	1.72±6.86	0.25	.80
3	Nilphamari	135	acre	17	38.47±4.62	acre	17	46.82±3.53	−8.35±5.83	−1.43	.15
4	Saidpur	47	(cycle	6	20.50±9.22	(cycle	5	12.40±3.55	8.10±9.88	0.82	.41
	four thanas	405	unit)	51	32.37±5.47	unit)	49	34.23±3.52	−1.86±6.50	−0.29	.77
	ten thanas	1246	mouza	25	21.18±2.19	mouza	32	20.64±1.71	0.54±2.78	0.19	.85

tion under jute in half-sample (A) is 21.18 $\pm$ 2.19 per cent and 20.64 $\pm$1.71 in half-sample (B). The difference between the two estimates was thus only 0.54$\pm$2.78 per cent with a value of (u) equal to only 0.19. The corresponding value of $P(u)$ given in col. (12) is 0.85 which shows that the observed difference could easily have arisen from errors of sampling. It appears therefore that in this particular area of Rangpur the intensity of cultivation is sufficiently uniform to enable the *mauza* as a whole being used as the sampling unit.

40. The frequency distribution of $P(u)$ for Police Stations is shown in Table 5 in which col. (1) gives the range of $P(u)$ and cols. (2) and (3) the observed and expected number of $P(u)$ in each range. It will be noticed that smaller values of $P(u)$ less than 0.1 is more numerous than large values in excess of 0.9. This is just what we expect on account of the values of $P(u)$ being actually less than true values. It will be also noticed that very small values less than 0.001 are comparatively rare indicating that the agreement between the two half-samples is not unsatisfactory. In this case also we have used Fisher and Behrens' test in four districts and the results are given in Table 3b. The comparison of the five per cent level calculated according to Fisher and Berhens' formula with the observed values of $P(u)$ indicate that the agreement between the two half-samples is quite satisfactory.

TABLE 5. FREQUENCY DISTRIBUTION OF VALUES OF $P(u)$ BY POLICE STATIONS

range of $P(u)$	number of $P(u)$	
	observed	expected
(1)	(2)	(3)
less than — 0.001	14	0.16
0.001 — 0.01	10	1.43
0.02 — 0.05	19	6.36
0.06 — 0.10	16	7.95
0.11 — 0.20	19	15.90
0.21 — 0.30	15	15.90
0.31 — 0.40	10	15.90
0.41 — 0.50	11	15.90
0.51 — 0.60	13	15.90
0.61 — 0.70	5	15.90
0.71 — 0.80	12	15.90
0.81 — 0.90	6	15.90
0.91 — 0.95	5	7.95
0.96 — 0.99	3	6.36
1.00	1	1.59
total	159	159.00

SAMPLE CENSUS OF THE AREA UNDER JUTE IN BENGAL IN 1940

COMPARISON BY DISTRICTS

41. The consolidated results by districts are shown in Table 6 in which the district figures have been built up from the Police Station figures weighted by the respective areas of the Police Stations. Similarly the "eight district" figures have been built up from individual district figures weighted by the respective district areas. The arrangement of the columns is similar to Tables 1–4; col. (1) gives the name of the district; col. (2) the area in square miles covered in the present survey; col. (3) the total number of sample-units and col. (4) the per cent of land under jute in half-sample (A); col. (5) gives the number of sample-units, and col. (6) the corresponding per cent of land under jute in half-sample (B). The observed difference between the two estimates together with the standard error is shown in col. (7), and the value of (u) in col. (8); while the corresponding probability of occurrence $P(u)$, based on the normal probability integral, is given in col. (9). Finally the total number of sample-units in the two half-samples taken together is shown in col. (10), and the pooled estimates of the percentage of land under jute together with the standard error in col. (11); finally the proportional standard error (i.e., the standard error as a percentage of the proportion of land under jute) in col. (12).

42. In the case of Bogra, Dacca, Jessore, Mymensingh, Nadia and Rangpur (special units) the probability of occurrence is quite high and greater than 0.05 showing that the observed differences between the two half-samples could easily have occurred by chance. In Rajshahi and Rangpur (ordinary units) $P(u)$ lies between 0.01 and 0.001; while in Tipperah the two half-samples are seriously discrepant with $u = 6.37$, and $P(u)$ extremely small. In Tipperah personal bias or defective field enumeration in either or both of the two half-samples must be responsible for the big difference. For all eight districts taken together, the difference between the two half-samples is 0.69 ± 0.21 per cent which may be considered to be fairly satisfactory. The results for the whole survey with a difference between two half-samples of 0.63 ± 0.29 per cent and $u = 2.17$ and $P(u) = 0.03$ is quite satisfactory.

43. I may note here that, apart from errors of sampling, there may exist real differences in the intensity of cultivation in adjoining areas from which the pairs of half-samples were drawn. Even if the intensity of cultivation may be considered to be practically the same for pairs of half-samples we have to remember that the field enumeration was carried out by entirely different sets of investigators. Considering the extensive scale of operations covering more than 20,000 square miles, and the widely varying local conditions in different districts the agreement between half-samples actually attained in practice is not unreasonable.

44. I may mention here that even in the case of highly refined physical measurements of such fundamental constants as the velocity of light (c), the gravitational constant (G) and Planck's constant (h), the observations do not exactly conform to the theory of errors. W. A. Shewhart and W. E. Deming in their recent book on Statistical Methods have pointed out that these observations "are among the elite of

all physical measurements." But physicists pretty generally agree that for each of the three constants, the observed range of variation is so great as to be indicative of "constant" errors.[1] In fact they have stressed the difficulty of establishing statistically controlled condition to such an extent as to make the observed distribution of errors conform exactly to requirements of pure mathematical theory. If this is the state of affairs in physics we have no reason to feel dissatisfied with the agreement of half-samples in our estimates of the area under jute.

MARGIN OF ERROR

45. From the pooled estimates given in col. (11) of Table 6 we find that in the case of grid sampling the standard error varied from 0.20 per cent in Mymensingh to 0.38 per cent in Rajshahi with a value of 0.11 per cent for all eight districts taken together. The margin of error at the one per cent level may therefore be considered to be certainly less than two per cent; that is, the odds are more than hundred to one in favour of our being able to estimate the proportion of land under jute with a margin of error of 2 or 3 per cent.

46. The proportional error, however, naturally depends on the actual per cent of land under jute. For example, in the case of Nadia the standard error is 0.31 per cent while the estimated proportion of land under jute is only 5.64 per cent; this gives a proportional standard error of 0.31 in 5.64 or 5.5 per cent; while in the case of Mymensingh the standard error is 0.20 per cent against a value of 25.61 per cent of land under jute which gives a proportional standard error of 0.20 in 25.61 or only about 0.8 per cent. For all eight districts taken together the standard error is 0.11 per cent against 21.86 per cent of land under jute showing a proportional or relative margin of error of 0.50 per cent.

47. The actual difference between the two estimates is 0.69 per cent as shown in col. (7); that is, the actual observed difference between the two half-samples is 0.69 in 21.86 showing a relative divergence of about three per cent. Remembering that the margin of error is to some extent a matter of definition we may conclude that the effective uncertainty in the estimate of the proportion under jute in all eight districts taken together is not greater than, say, 3 or 4 per cent.

48. I may mention in this connexion that the Jute Census Committee at its sixth meeting was of opinion that a margin of error of five per cent would be considered adequate (Proceedings, 10th October 1939, page 5). Judged by this standard the results obtained this year by large scale operations extending over twenty thousand square miles must be considered to be quite satisfactory. In other words, the Sample Census may be expected to give results which will be sufficiently reliable for all practical purposes.

[1] Shewhart, W. A. and Deming, W. E. (1939): *Statistical Method from the View Point of Quality Control*, Graduate School of the Department of Agriculture, Washington, pp. 66-70.

TABLE 6. COMPARISON OF HALF-SAMPLES BY DISTRICTS

name of district	area in sq. miles	half-sample A		half-sample B		difference between half-samples		approximate probability $P(u)$	combined (A and B together)		
		no. of grids	percent under jute±s.e.	no. of grids	percent under jute±s.e.	actual ±s.e.	divided by s.e.=(u)		no. of grids	percent under jute±s.e.	proportional s.e.
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Bogra	1471	1801	17.14±0.51	1799	17.75±0.50	−0.61±0.72	−0.85	0.40	3600	17.45±0.36	2.06
Dacca	2652	3262	25.03±0.40	3208	23.90±0.40	1.13±0.57	1.98	0.05	6470	24.47±0.28	1.14
Jessore	1704	1263	10.47±0.44	1248	10.54±0.42	−0.07±0.60	−0.12	0.90	2511	10.51±0.30	2.85
Mymensingh	6306	7985	26.02±0.28	8295	25.24±0.27	0.78±0.39	2.00	0.05	16280	25.61±0.20	0.78
Nadia	1415	745	5.16±0.40	750	6.41±0.50	−1.25±0.64	−1.95	0.06	1495	5.64±0.31	5.50
Rajshahi	1235	1030	11.87±0.53	1020	13.94±0.53	−2.07±0.76	−2.72	>0.001	2050	12.91±0.38	2.94
Rangpur	1095	1380	31.74±0.60	1366	29.43±0.56	2.31±0.82	2.82	>0.001	2746	30.24±0.36	1.19
Tipperah	2536	3171	30.21±0.45	3022	26.20±0.44	4.01±0.63	6.37	<0.001	6193	28.62±0.35	1.22
8 districts	18414	20637	22.21±0.15	20708	21.52±0.15	0.69±0.21	3.29	>0.001	41345	21.86±0.11	0.50
Rangpur 16-acre (cycle)	405	51	32.37±5.47	49	34.23±3.52	1.86±6.50	−0.29	0.77	100	33.69±2.96	8.79
mouza (cycle)	1246	25	21.18±2.19	32	20.64±1.71	0.54±2.78	0.19	0.85	57	20.84±1.35	6.48
mouza	488	—	— —	—	— —	— —	—	—	70	26.74±1.86	6.96
Rangpur	2139	76	*23.92±2.13	81	*23.97±1.55	−0.05±2.63	−0.02	0.98	227	24.62±1.05	4.26
grand total	20553	20713	*22.35±0.22	20789	*21.72±0.19	0.63±0.29	−2.17	0.03	41572	21.51±0.15	0.70

*Exclusive of 488 sq. miles (*mouza* unit in Rangpur) for which half-samples were not available.

CHAPTER 3 : THE VARIANCE FUNCTION

49. We may now consider the variance function, that is, the relation between the variance and size of the sample units. Its form was practically settled by extensive model and field sampling work done in 1938 and 1939. We have again used a logarithmic equation which proved so successful in previous years :

$$\log(V_x)-\log(pq) = \log(\alpha)-g \cdot \log(x) \quad \ldots \quad (1.1)$$

or

$$\frac{V_x}{pq} = \frac{\alpha}{x^g} \quad \ldots \quad (1.2)$$

where x = size of sample-units;

p = proportion of land under jute

$q = 1-p$;

V_x = variance for sample-units of size × acre; and α^* and g are constants.

50. It will be remembered that we used 4-acre as the standard size of sample-units all over the area; we also used sample-units of other sizes (1-acre, 2.25-acre, 6.25-acre, and 9-acre) but we could not include a large number of these sizes for reasons of economy. Fortunately in the case of the two big jute districts, Mymensingh and Dacca, we have a sufficient number of grids of different sizes to enable the variance function being estimated with considerable accuracy. In other districts difficulties were caused by either all sizes of sample-units being not available, or the proportion of land under jute (p) varying too much from one region to another which made the graduation over the district as a whole rather uncertain; but the graduation for the eight districts taken together was not unsatisfactory.

51. The results are shown in Table 7 in which col. (1) shows the size in acre; col. (3) the observed values of log $(V_x)/pq$ while col. (4) shows the corresponding graduated values obtained from the variance function (1.1); col. (5) gives the difference between observed and graduated values; and this difference is expressed as percentage of the graduated values in col. (6). The values of (pq), (α), and (g) are also shown in each case. Comparing observed and graduated values in cols. (3) and (4) we find that on the whole the graduation is not unreasonable although there are appreciable discrepancies here and there.

52. I may mention in this connexion that the form of the variance function chosen by us is strictly valid in the case of homogeneous areas in which the intensity of cultivation (i.e., the value of (p) is sensibly constant. This means that it will be possible to use the variance function in a completely satisfactory manner only when the question of zoning is finally settled. It is intended to study this question in the 1940-41 scheme; in the meantime we are accepting a value of $g = 2.04$ in preparing our plans for the 1941 survey.

---

*In 1939 we had used a slightly different notation :

$V_x/pq=(1/bx)g$ where $(1/b)g = \alpha$ of our present notation.

TABLE 7. VARIANCE FUNCTION : OBSERVED AND GRADUATED VALUES

sample-units (grids)		values of log (v_x/pq)		difference (observed-graduated)	
size in acre	number	observed	graduated	actual	percentage
(1)	(2)	(3)	(4)	(5)	(6)
Mymensingh : $p=0.2561$; $\alpha=0.4947$; $g=0.2722$					
1.00	697	−0.31394	−0.30569	−0.00825	2.699
2.25	710	−0.42402	−0.40155	−0.02247	5.596
4.00	12988	−0.46700	−0.46957	−0.00257	−0.547
6.25	941	−0.50057	−0.52233	−0.02176	−4.166
9.00	944	−0.59983	−0.56543	−0.03440	−6.084
Dacca : $p=0.2447$ $\alpha=0.3632$ $g=0.2061$					
1.00	282	−0.40629	−0.43989	0.03360	−7.638
2.25	396	−0.51036	−0.51247	0.00211	−0.412
4.00	5160	−0.57085	−0.56397	−0.00688	1.220
6.25	257	−0.53243	−0.60392	0.07149	−11.838
9.00	375	−0.61877	−0.63656	0.01779	−2.795
eight districts : $p=0.2185$ $\alpha=0.4084$ $g=0.2404$					
1.00	1766	−0.35167	−0.38886	0.03719	−9.564
2.25	2447	−0.52796	−0.47352	−0.05364	11.328
4.00	32996	−0.53433	−0.53360	−0.00073	0.137
6.25	1944	−0.49771	−0.58019	0.08248	−14.216
9.00	2192	−0.64849	−0.61826	−0.03023	4.890

53. It is worth noting that, although the value of $g = 0.24$ has been obtained from sample-units lying between 1-acre and 9-acre grids, the variance function is giving reasonable results by extrapolation even in the case of random mauzas. The relevant material is shown in Table 8 in which col. (1) shows the average size of sample-units in acre; col. (2) the number of sample-units; and col. (3) the proportion of land under jute. The observed values of $100(v)/pq$, where v is the observed variance, is shown in col. (4); and the corresponding calculated value obtained by extrapolation from the variance function in col. (5). It will be noticed that even when the mauza as a whole is used as the sample-unit, and the average size of the grid is so large as 1075 acre, the calculated value of the variance is 7.2 per cent against an observed variance of 6.3 per cent; the result is clearly of the right order. Pending more detailed investigation after the question of zoning is at least approximately settled we may use the pooled estimate of $g = 0.24$ for our present calculations.

TABLE 8. CALCULATED VARIANCES FOR 16-ACRE GRIDS AND MAUZAS IN DISTRICT RANGPUR

average size of sample-units	number of sample-units	p=proportion under jute	values of (100 v/pq)	
			observed	calculated
(1)	(2)	(3)	(4)	(5)
16-acre (cycle units)	100	0.3351	16.2 p.c.	20.7 p.c.
509-acre (mauza cycle)	70	0.2674	11.2 p.c.	8.7 p.c.
1075-acre (mauza ordinary)	57	0.2087	6.3 p.c.	7.2 p.c.

COMPARISON WITH RESULTS OF THE SAMPLE SURVEY OF 1935

54. Last year in my report on the Sample Census of 1939 I had referred to the pioneer random sample survey carried out in 1935 on the initiative of Mr. H. P. V. Townend, I.C.S., then Rural Development Commissioner, Bengal. In this survey the size of sample-units was 40-acre, and altogether about 2,100 sample-units were scattered at random over 20 different districts in Bengal. I had some difficulty in tracing the original records, but with the help of Mr. Townend I finally succeeded in obtaining the primary material on the basis of which values of the variance for individual sample-units were calculated separately for each district.

55. These are shown in Table 9 in which col. (1) gives a serial number for convenience of reference; col. (2) the name of the district; col. (3) the number of sample-units of size 40-acre used in 1935; col. (4) observed values of (p) the proportion of land under jute; col. (5) values of (q) equal to ($1-p$); and col. (6) values of (pq). The observed variances are given in col. (9); and values of (v/pq) in col. (7), and of log (v/pq) in col. (8). Using the present variance function with $\alpha = 0.4084$, and $g = 0.2404$ we can calculate the variances for different districts and compare them with observed values. The graduated values are shown in col. (10), and may be compared with the observed values shown in col. (9). It will be noticed that, although the agreement is not entirely satisfactory, the variances calculated by our present formula are approximately of the same order as the observed variances.

56. As sample-units of only one size (40-acre) were used in 1935 it is not possible to calculate both g and α. We may, however, adopt a known value for one of these parameters and calculate the other from the observed variances given in Table 9. From physical considerations it appears desirable to use the value of $\alpha = 0.4084$ as known. We may first compare this with the value obtained in 1939. I have already mentioned that this year we have slightly changed the notation; our present α is simply $(1/b)g$ of the 1939 Report. From model sampling experiments we had found last year $b = 16.37$ with $g = 0.3328$; this corresponds to $\alpha = 0.3953$ which is very near the value of $\alpha = 0.4084$ found this year. For field surveys over a restricted area last year we had adopted $b = (2.6)^3$ with $g = 1/3$ which corresponds to $\alpha = 0.3846$ which differs from this year's value of 0.4084 by about 5 or 6 per cent which is not unsatisfactory. I may metnion here that, as stated in last year's report, the value of $(1/b)$ of $(\alpha)^1 g$ gives

a rough measure of the ultimate smallest unit of cultivation for jute; for $\alpha = 0.4084$ and $g = 0.2404$ this is .024 acre ; multiplying by 1 acre = 60.5 *cottah* we find that 1.45 or say one a half *cottah* represents the average lowest unit of cultivation for jute which is quite a plausible figure.

TABLE 9. VALUES OF VARIANCE PARAMETER (g) BASED ON SAMPLE SURVEY OF 1935

sl. no.	name of districts	no. of sample units	p=proportion under jute	$q=(1-p)$	pq	v/pq	log (v/pq)	variances observed	variances graduated	values of g districts	values of g accumulated
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
1	Mymensingh	415	.1660	.8340	.1384	.1575	– .8027	.0218	.0233	.2583	
2	Dacca	179	.1617	.8383	.1356	.2065	– .6851	.0280	.0228	.1849	
3	Tipperah	83	.1492	.8508	.1269	.2230	– .6517	.0283	.0214	.1640	
4	Faridpur	172	.1273	.8727	.1111	.1143	– .9420	.0127	.0187	.3432	
5	Rangpur	239	.1216	.8784	.1068	.1114	– .9531	.0119	.0180	.3522	
6	Pabna	119	.1162	.8838	.1027	.1714	– .7660	.0176	.0173	.2354	
7	Bogra	91	.0997	.9003	.0898	.1748	– .7575	.0157	.0151	.2301	
	group value									.2329	.2329
8	Dinajpur	130	.0789	.9211	.0727	.1087	– .9638	.0079	.0122	.3588	
9	Noakhali	64	.0738	.9262	.0684	.1038	– .9838	.0071	.0115	.3713	
10	Rajshahi	175	.0622	.9378	.0583	.1458	– .8362	.0085	.0098	.2792	
11	Bakerganj	87	.0591	.9409	.0556	.1709	– .7673	.0095	.0094	.2362	
12	Jessore	193	.0591	.9409	.0556	.1331	– .8758	.0074	.0094	.3039	
13	Nadia	190	.0542	.9458	.0513	.0702	–1.1537	.0036	.0086	.4774	
14	24 Parganas	121	.0486	.9514	.0462	.1580	– .8013	.0073	.0078	.2574	
	group value									.3263	.2896
15	Khulna	95	.0377	.9623	.0363	.2590	– .5867	.0094	.0061	.1235	
16	Murshidabad	137	.0339	.9661	.0328	.0884	–1.0536	.0029	.0055	.4149	
17	Hooghly	94	.0313	.9687	.0303	.1485	– .8283	.0045	.0051	.2743	
18	Maldah	35	.0303	.9697	.0294	.1497	– .8248	.0044	.0049	.2721	
19	Jalpaiguri	189	.0151	.9849	.0149	.0805	–1.0942	.0012	.0025	.4402	
20	Burdwan	36	.0037	.9963	.0037	.0540	–1.2676	.0002	.0006	.5480	
	group value									.3436	.3064

57. Adopting a value of $\alpha = 0.4084$ we can now calculate the values of g for each district which are shown in col. (10) of Table 9. It will be noticed that values of g calculated in this way are all appreciably less than unity thus definitely confirming a departure from binomial variation; but values for individual districts fluctuate considerably. It also appears that g varies inversely as p, that is, as the proportion under jute decreases the value of g increases.

58. In order to study this question in greater detail we arranged the districts in Table 9 in order of intensity of cultivation, that is, with decreasing values of p. Dividing the districts into three broad groups (1–7), (8–14) and (15–20) we have calculated values of g corresponding to each group. It will be noticed that in the first group (consisting of seven heavy districts with values of p of about ten per cent or more) the group value of g is equal to 0.2529 which is in satisfactory agreement with our adopted value of $g = 0.2404$. In the next group of seven districts (with the proportion under jute varying from about eight per cent in Dinajpur to a little less than five per cent in 24 Parganas) the group value of g is equal to 0.3263 which is appreciably higher than the value for the first group. In the third group of six districts (with value of p less than four per cent) we have $g = 0.3456$ which is still higher than that corresponding to the second group. The value of g for the first two groups taken together is 0.2896; and for all twenty districts pooled together $g = 0.3064$ for the material collected in 1935.

59. From the above analysis it appears that the value of g increases as the intensity of cultivation is decreased. This is quite plausible. The variance function adopted by us can be interpreted as arising from a correlation between values of p (proportion of land under jute) in neighbouring areas.[1] When the intensity of cultivation is high there is a much bigger chance of this correlation between the intensity of cultivation in neighbouring areas being large; on the other hand, when the intensity of cultivation is low the change of occurrence of jute is likely to be more independent. When this chance is completely independent, the value of g should of course be equal to unity in accordance with the binomial distribution.

60. I may mention here that for the eight districts in which the survey was carried out in 1940 the group value of g based on the sample survey of 1935, comes out as 0.2813 which is in fair agreement with the value of 0.2404 obtained in 1940. On the whole a comparatively low value of the order of 0.24 or 0.28 may be considered to be definitely confirmed by the results of the survey carried out five years ago. It is possible however that as we extend the work to districts with lower intensities of cultivation we shall get higher values of g. Broadly speaking, from the results of the 1935 survey, it would appear that the value of g for the proportion of area under jute in Bengal most likely lies between 0.25 and 0.35, or say roughly between one-fourth and one-third.

AUXILIARY INFORMATION : PROPORTION OF PLOTS UNDER JUTE

61. I shall now briefly consider certain auxiliary results obtained in the course of the statistical analysis which are useful in various ways. The total area and number of plots surveyed in 1940 in the different districts is shown in Table 10 in which col. (1) gives the name of the district; cols. (2)–(6) details of the number of sample units

[1] A fuller discussion is given in P. C. Mahalanobis (1940) : Sample survey of the acreage under jute in Bengal, *Sankhyā*, Vol. 4(4), pp. 511–530.

TABLE 10. AREA AND NUMBER OF PLOTS SURVEYED IN 1940

name of district	number of sample-units						area in acres		number of plots		percent under jute	
	1-acre	2.25 acre	4-acre	6.25 acre	9-acre	total	surveyed	jute	total	jute	area	plots
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Bogra	146	149	2920	140	245	3600	941440	164281	71037	15401	17.45	21.68
Dacca	282	396	5160	257	375	6470	1697280	415324	104335	39294	24.47	37.66
Jessore	242	174	1831	—	264	2511	1090560	114618	50795	5885	10.51	11.59
Mymensingh	697	710	12988	941	944	16280	4034560	1033251	216178	85866	25.61	39.72
Nadia	—	76	1294	128	—	1495	905600	50080	26449	1627	5.53	6.15
Rajshahi	105	126	1628	147	44	2050	790400	102041	39096	7606	12.91	19.45
Rangpur	—	368	2284	—	94	2746	700800	213674	58568	20830	30.49	35.57
Tipperah	294	451	4891	331	226	6193	1623040	456886	106882	40980	28.15	38.34
8 districts	1766	2587	32996	1804	2192	41345	11784960	2575014	673340	217489	21.85	32.30
Rangpur	16-acre (cycle-unit)					100	259200	86858	7803	2813	33.51	36.05
Rangpur	random mauza (cycle-unit)					70	312320	83514	80006	19370	26.74	24.21
Rangpur	random mauza (ordinary)					50	797440	166426	135789	33265	20.87	24.50
	eight districts : total					41565	13153920	336798	896938	272937		

of different sizes and col. (7) the total number of sample units; col. (8) the total area and col. (9) the area under jute included within the sample units; col. (10) gives the total number of plots surveyed, and col. (11) the total number of plots wholly or partly under jute. The proportion of land under jute is shown in col. (12) while the proportion of plots wholly or partly under jute is given in col. (13).

62. It will be noticed that the proportion of plots under jute is appreciably larger than the actual proportion of land under jute. This is of course just what is to be expected in view of the fact that jute is grown on only a portion of a large number of plots.

AVERAGE AREA OF MAUZAS AND PLOTS

63. Information regarding the average size of *mauzas* and plots in acre is given in a summary form for the eight districts in Table 11 in which col. (1) gives a serial number for convenience of reference; col. (2) the name of the district; col. (3) the number of *mauzas* in each district; cols. (4) and (5) the total area of the district in square miles and acres respectively; col. (6) the number of plots per *mauza*; col. (7) the average size of plots in acre. From col. (7) we find that the average area of individual *mauzas* varies from about 320 acre in Dacca to 634 acre in Rangpur, while for all eight districts taken together the average size of a *mauza* is 458 acre. Similarly from col. (8) we notice that the average size of individual plots varies from 0.37 acre in Tipperah to 0.69 acre in Mymensingh while the pooled value is 0.47 or just less than half an acre. More detailed figures by thanas are given in Tables 12 in which the arrangement is exactly similar to Table 11.

TABLE 11. AVERAGE SIZE OF MAUZAS AND PLOTS BY DISTRICTS

sl. no.	name of district	number of mauzas	total area in		number of plots per mauza	average area (in acres)	
			square miles	acres		per mauza	per plot
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1.	Bogra	2500	1470.10	940863	979	376.35	0.3870
2.	Dacca	5313	2653.85	1698467	576	319.70	.5359
3.	Jessore	3726	2915.44	1865884	1297	500.80	.3828
4.	Mymensingh	7940	6335.92	4054925	683	510.70	.6920
5.	Nadia	2696	2838.94	1816917	1360	673.93	.4900
6.	Rajshahi	4738	2531.70	1620288	837	341.98	.3962
7.	Rangpur	3637	3602.36	2305510	1486	633.90	.4262
8.	Tipperah	4234	2535.74	1622874	1021	383.30	.3704
	eight districts	34784	24883.95	15925728	952	457.85	0.4671

TABLE 12. AVERAGE AREA OF MAUZA AND PLOTS BY DISTRICTS AND POLICE STATIONS*

(district—Bogra)

sl. no.	name of police stations	number of mauzas	total area in square miles	total area in acres	number of plots per mauza	average area (in acres) per mauza	average area (in acres) per plot
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1.	Adamdighi	228	119.53	76499	1118	335.52	0.3003
2.	Bogra	263	148.20	94848	1381	360.64	.2709
3.	Dhunat	91	95.48	61107	1740	671.51	.3862
4.	Dhupchanchia	115	62.72	40141	1131	349.05	.3087
5.	Gaptali	106	94.08	60211	1804	568.03	.3150
6.	Joypurhat	185	91.57	58605	789	316.78	.4074
7.	Kahalu	166	92.83	59411	1108	357.91	.3231
8.	Khetal	197	119.27	76333	894	387.48	.4335
9.	Nandigram	235	102.72	65741	489	279.75	.5713
10.	Penchbibi	222	107.52	68813	523	309.97	.5990
11.	Sariakandi	223	200.11	128070	1167	574.30	.4948
12.	Sherpur	224	114.41	73222	684	326.88	.4774
13.	Shibgonj	245	121.66	77862	800	317.80	.3979
	district Bogra	2500	1470.10	940863	979	376.35	0.3870

*Out of eight tables only one table is printed here as a specimen.

AVERAGE NUMBER OF PLOTS PER SAMPLE-UNIT

64. The average number of plots per sample-unit is shown in Table 13 in which col. (1) gives the name of the district; col. (2.1) the average number of 1-acre sample units, col. (2.2) the number of plots partly included within the sample-unit; col. (2.3) the average number of plots entirely included within the sample-unit; and col. (2.4) the total number of plots partly or entirely included within sample-units. Similar information for sample units of size 2.25-acre is given in cols. (3.1)–(3.4); for 4-acre sample-units in cols. (4.1)–(4.4); for 6.25-acre sample-units in cols. (5.1)–(5.4); and finally for 9-acre sample plots in cols. (6.1)–(6.4). The values for all eight districts pooled together are given at the bottom of the table.

65. We find that the total number of plots covered by sample-units for which the area has to be measured is on an average 6.56 in the case of 1-acre, 11.36 for 2.25-acre, 15.89 for 4-acre, 22.62 for 6.25-acre, and 28.85 for 9-acre sample-units. A simple graduation formula may be used to obtain these approximate results; if we take twice the area (in acre) plus the perimeter (in terms of root-acre) we get 6.0, 10.5, 16.0, 22.5 and 30.0 as the graduated values for 1-acre, 2.25-acre, 4-acre, 6.25-acre and 9-acre sample-units respectively which agree quite well with the observed values.

TABLE 13. AVERAGE NUMBER OF PLOTS PER SAMPLE-UNIT

name of district	1-acre				2.25-acre			
	N	part	full	total	*N*	part	full	total
(1)	(2.1)	(2.2)	(2.3)	(2.4)	(3.1)	(3.2)	(3.3)	(3.4)
Bogra	146	5.04	0.27	5.31	149	9.01	1.90	10.91
Dacca	282	6.22	0.26	6.48	396	8.71	1.17	9.88
Jessore	243	8.02	0.92	8.94	175	11.04	2.37	13.41
Mymensingh	700	5.17	0.17	5.34	713	9.15	1.10	10.25
Nadia	—	—	—	—	73	9.21	0.99	10.20
Rajshahi	106	6.71	0.40	7.11	126	9.52	1.38	10.90
Rangpur	—	—	—	—	368	11.47	3.73	15.20
Tipperah	307	7.11	0.87	7.98	454	9.32	1.62	10.94
eight districts	1784	6.14	0.42	6.56	2254	9.61	1.75	11.36

name of district	1-acre				6.25-acre				9-acre			
	N	part	full	total	*N*	part	full	total	*N*	part	full	total
(1)	(4.1)	(4.2)	(4.3)	(4.4)	(5.1)	(5.2)	(5.3)	(5.4)	(6.1)	(6.2)	(6.3)	(6.4)
Bogra	2926	13.37	5.69	19.06	140	16.11	12.99	29.10	245	21.85	14.36	36.21
Dacca	5202	12.23	3.39	15.62	258	16.76	7.01	23.83	376	16.49	8.51	25.00
Jessore	2007	14.97	5.81	20.78	—	—	—	—	106	20.74	13.42	34.16
Mymensingd	13014	10.13	2.16	12.29	825	15.00	6.52	21.52	909	16.21	6.75	22.96
Nadia	1301	13.23	4.49	17.72	182	13.78	5.38	19.16	—	—	—	—
Rajshahi	1646	13.52	4.68	18.20	149	16.46	8.70	25.16	44	22.27	15.75	38.02
Rangpur	2290	13.80	5.67	19.47	—	—	—	—	95	26.78	26.59	53.37
Tipperah	5027	12.66	5.19	16.85	345	15.63	8.08	23.71	230	20.90	15.36	36.26
eight districts	33413	11.95	3.94	15.89	1845	15.48	7.14	22.62	2007	18.36	10.49	28.85

CHAPTER 4 : THE COST FUNCTION

66. I discussed the general structure of the cost function in considerable detail in Section 4 of the *Report on the Sample Census of Jute in* 1939. With low densities, that is, with a small number of sample-units or grids per square mile, proportionally more time will be spent in journeys in moving from one sample-unit to another, on the other hand, with a high density or a large number of sample-units per square mile, less time will be spent on journeys so that more time will be available for actual field enumeration. For the same size of sample-units, the cost per square mile will however be less with widely scattered grids than with a large number of sample-units per square mile. The total cost as well as the sampling error per square mile will thus depend on both the size and density of grids.

67. In order to enable an adequate analysis being made of the cost of operations, daily time records were kept by each field investigator; these time records together with other material entered in the field diaries were tabulated and studied in considerable detail in the Statistical Laboratory.

68. In 1939 the total time spent in the field had been broken down into four broad groups. This year we have further split up the journey portion into two parts, namely "big" and "small" journeys. We thus have the following five components of the field cost :

(1) *Enumeration.* This consists of the time spent in identifying the plots included within the sample-units by reference to the *mauza* maps (on the scale of 16 inches to the mile); inspecting the plots; recording of the serial number of the plots wholly or partly under jute; and, on the case of plots partly under jute, entering the *anna*-estimate of the portion of land on which jute is grown.

(2) *Big journey.* This consists of the time spent in moving from one camp to another, camp for our purposes being defined as the place where the investigators spend the night.

(3) *Small journey.* This consists of the time spent in moving from camp to the sample-unit back from the sample-unit to camp; and other miscellaneous journeys undertaken in the neighbourhood of the camp in connexion with field work.

(4) *Miscellaneous.* This consists of the time spent in making preliminary arrangements, keeping time records and diaries, preparing copies of the field records and despatching the same to the inspectors or headquarters; time spent in receiving instructions etc.

The sub-total of (1) enumeration, (2) big journey, (3) small journey and (4) miscellaneous gives the actual net working time spent for direct productive purposes.

69. *Indirect.* The remaining portion of the day represents the indirect[1] time required for sleep, food etc., and also includes the time lost on account of sickness, cessation of work owing to draught, excessive rain or other unforeseen reasons; the

[1]In the 1939 Report we called this "overhead" which is however misleading as it properly refers to charges for supervision, office work etc.

indirect item thus represents the time spent for non-productive purpose. The total of the net working time plus the indirect time naturally adds up to 24 hours on each day; this was used as a check on the accuracy of the primary records.

ANALYSIS OF TIME REQUIRED FOR THE FIELD SURVEY

70. The primary material is shown in Table 14 in which col. (1) gives the name of the district with the density of grids (that is, the number of sample-units per square mile); the figure is repeated twice to indicate the density of each half-sample; col. (2.1) gives the number of samples per square mile; col. (2.2) the number of grids; and col. (2.3) the number of days spent by the investigators in collecting information for sample-units of size 1-acre. Similar information is given in cols. (3.1)–(3.3) for sample-units of size 2.25-acre and so on; and finally cols. (7.1)–(7.3) give the information for sample-units of all sizes taken together. The last line in this table gives the totals for all eight districts taken together; from cols. (7.1), (7.2) and (7.3) it will be noticed that altogether 15,774 days were spent by the primary investigators in collecting information for 41,345 sample-units of various sizes scattered over 18,416 square miles in eight districts.

71. The next Table 15 shows the total field time in net and gross man-hours per sample-unit classified by districts and size of grids. The arrangement of this table is similar to that of Table 14; col. (1) gives the name of the district and the density of sample-units; col. (2.1) the number of grids of size 1-acre; col. (2.2) gives the average number of man-hours (that is, net working hours spent by a single investigator) in collecting information for different sample-units of size 1-acre; while col. (2.3) gives the corresponding gross man-hours inclusive of the indirect time. In the same way information relating to the sample-units of other sizes are given in successive columns; and finally in cols. (7.1)–(7.3) average values are given for sample-units of all sizes taken together. In this table also the bottom line shows the average values for the eight districts taken together.

72. The time required for enumeration is shown in the same way in Table 16 ; the arrangement is so similar to that in the previous tables that a detailed description will be superfluous. It should be noticed that the man-hours given in this table are net working hours required for completing the enumeration for each sample-unit.

73. The time of small journey in net man-hours (that is, number of working hours spent by one single investigator) for each sample-unit is shown in the two-way Table 17; and similar information is given regarding the time of big journey in Table 18; and for miscellaneous work in Table 19. Finally the indirect time spent in each case has been shown in the form of percentages of the whole day (that is, of 42 hours) in Table 20. It will be noticed that the pooled average was 64.6 per cent which means that 15.5 hours per day were required on an average for indirect purposes, and 8.5 hours per day were available for productive work.

TABLE 14. JUTE CENSUS, 1940: FIELD WORK—NUMBER OF SQUARE MILES, SAMPLE-UNITS (GRIDS) AND MAN-DAYS BY DISTRICTS AND SIZE OF GRIDS

name of district with density of grids	1-acre			2.25-acre			4-acre			6.25-acres			9-acre			total		
	number of			number of			number of			number of			number of			number of		
	sq. miles	grids	days	sq. miles	grids	days	sq. miles	grids	days	sq. miles	grids	days	sq. miles	grids	days	sq. miles	grids	days
(1)	(2.1)	(2.2)	(2.3)	(3.1)	(3.2)	(3.3)	(4.1)	(4.2)	(4.3)	(5.1)	(5.2)	(5.3)	(6.1)	(6.2)	(6.3)	(7.1)	(7.2)	(7.3)
Nadia (0.548+0.548)	—	—	—	69	73	37	1224	1294	963	122	128	92	—	—	—	1415	1495	1092
Jessore (0.759+0.759)	162	242	136	118	174	113	1237	1831	1029	—	—	—	187	264	131	1704	2511	1409
Rajshahi (0.920+0.920)	63	105	43	76	126	49	981	1628	641	88	147	55	27	44	17	1236	2050	805
Bogra (1.224+1.224)	60	146	46	61	149	30	1193	2920	934	57	140	50	100	245	113	1471	3600	1173
Tipperah (1.306+1.306)	124	294	98	174	451	174	1998	4891	1829	150	331	110	90	226	102	2536	6193	2313
Rangpur (1.314+1.314)	—	—	—	142	368	97	909	2284	709	—	—	—	44	94	43	1095	2746	849
Mymensingh (1.464+1.464)	220	697	233	291	710	249	5098	12988	4465	346	941	321	351	944	404	6306	16280	5672
Dacca (1.513+1.513)	103	282	101	85	396	157	2254	5160	1955	71	257	100	139	375	148	2652	6470	2461
eight districts (1.123+1.123)	732	1766	657	1016	2447	906	14894	32996	12525	834	1944	728	938	2192	958	18414	41345	15774

TABLE 15. JUTE CENSUS 1940 : TOTAL FIELD TIME IN NET AND GROSS MAN-HOURS PER SAMPLE-UNIT (GRID) BY DISTRICTS AND SIZE OF GRIDS

name of district with density of grids	1-acre			2.25-acre			4-acre			6.25-acre			9-acre			total	average (for all sizes)	
	no. of grids	man-hours		no. of grids	man-hours		no. of grids	man-hours		no. of grids	man-hours		no. of grids	man-hours		no. of grids	man-hours	
		net	gross		net	gross		net	gross		net	gross		net	gross		net	gross
(1)	(2.1)	(2.2)	(2.3)	(3.1)	(3.2)	(3.3)	(4.1)	(4.2)	(4.3)	(5.1)	(5.2)	(5.3)	(6.1)	(6.2)	(6.3)	(7.1)	(7.2)	(7.3)
Nadia (0.548+0.548)	—	—	—	73	2.56	12.16	1294	4.62	17.82	128	4.02	17.52	—	—	—	1495	4.50	17.52
Jessore (0.759+0.759)	242	5.06	13.49	174	3.68	14.82	1831	3.46	13.47	—	—	—	264	4.18	11.86	2511	3.70	13.40
Rajshahi (0.920+0.920)	105	3.04	9.83	126	2.90	9.33	1628	2.60	9.44	147	3.35	8.82	44	3.42	9.27	2050	3.50	9.40
Bogra (1.224+1.224)	146	1.93	7.56	149	2.41	4.83	2920	2.87	7.68	140	2.52	8.57	245	5.00	1107	3600	2.94	7.82
Tipperah (1.306+1.306)	294	3.16	8.00	451	3.16	9.28	4891	3.26	8.96	331	2.89	7.98	226	3.48	10.78	6193	3.24	8.95
Rangpur (1.314+1.314)	—	—	—	368	2.50	6.33	2284	2.55	7.45	—	—	—	94	4.21	10.75	2746	2.60	7.42
Mymensingh (1.464+1.464)	697	2.73	8.03	710	2.98	8.42	12988	2.94	8.26	941	2.94	8.23	944	4.03	10.28	16280	3.00	8.37
Dacca (1.513+1.513)	282	2.48	8.60	396	3.60	9.52	5160	3.48	9.10	257	3.75	9.34	375	3.68	9.37	6470	3.47	9.13
eight districts (1.123+1.123)	1766	3.05	8.93	2447	3.04	8.86	32996	3.17	9.11	1944	3.11	9.01	2192	4.03	10.45	41345	3.20	9.15

TABLE 16. JUTE CENSUS 1940 : TIME FOR ENUMERATION IN NET MAN-HOURS PER SAMPLE-UNIT (GRID) BY DISTRICTS AND SIZE OF GIRDS

name of district with density of grids	1-acre		2.25-acre		4-acre		6.25-acre		9-acre		total no. of grids	average (for all sizes) man-hours
	no. of grids	man-hours	no. of grids	man-hours	no. of grids	man-hours	no. of grids	man-hours	no. of grids	man-hours		
(1)	(2.1)	(2.2)	(3.1)	(3.2)	(4.1)	(4.2)	(5.1)	(5.2)	(6.1)	(6.2)	(7.1)	(7.2)
Nadia (0.548+0.548)	—	—	73	0.68	1294	1.06	128	0.87	—	—	1495	1.03
Jessore (0.759+0.759)	242	1.15	174	0.99	1831	0.93	—	—	264	1.22	2511	0.98
Rajshahi (0.920+0.920)	105	0.77	126	1.06	1628	1.00	147	0.83	44	1.16	2050	0.98
Bogra (1.224+1.224)	146	0.58	149	1.07	2920	1.07	140	0.82	245	1.81	3600	1.09
Tipperah (1.306+1.306)	294	0.82	451	0.80	4891	0.96	331	1.02	226	1.28	6193	0.96
Rangpur (1.314+1.314)	—	—	368	0.69	2284	0.69	—	—	94	1.22	2746	0.71
Mymensingh (1.464+1.464)	697	0.84	710	1.06	12988	0.97	941	1.17	944	1.54	16280	1.01
Dacca (1.513+1.513)	282	0.72	396	0.86	5160	1.23	257	1.08	375	1.07	6470	1.17
eight districts (1.123+1.123)	1766	0.85	2447	0.91	32996	1.00	1944	1.06	2192	1.40	41345	1.01

TABLE 17. JUTE CENSUS 1940 : TIME OF SMALL JOURNEY IN MAN-HOURS PER SAMPLE-UNIT (GRID) BY DISTRICTS AND SIZE OF GRIDS

name of district with density of grids	1-acre			2.25-acre			4-acre			6.25-acre			9-acre			total	average (for all sizes)	
	no. of grids	man-hours	percent	no. of grids	man-hours	percent	no. of grids	man-hours	percent	no. of grids	man-hours	percent	no. of grids	man-hours	percent	no. of grids	man-hours	percent
(1)	(2.1)	(2.2)	(2.3)	(3.1)	(3.2)	(3.3)	(4.1)	(4.2)	(4.3)	(5.1)	(5.2)	(5.3)	(6.1)	(6.2)	(6.3)	(7.1)	(7.2)	(7.3)
Nadia (0.548+0.548)	—	—	—	73	0.87	7.2	1294	1.42	8.0	128	1.45	8.3	—	—	—	1495	1.42	8.1
Jessore (0.759+0.759)	242	1.62	12.1	174	1.32	8.9	1831	1.18	8.7	—	—	—	264	1.37	11.5	2511	1.25	9.3
Rajshahi (0.920+0.920)	105	0.87	8.8	126	0.53	5.7	1628	0.96	10.2	147	0.94	10.6	44	0.77	8.3	2050	0.92	9.8
Bogra (1.224+1.224)	146	0.78	10.3	149	0.60	12.5	2920	0.08	10.6	140	0.89	10.3	245	1.23	11.1	3600	0.84	10.7
Tipperah (1.306+1.306)	294	0.90	11.2	451	0.86	9.3	4891	0.84	9.4	331	0.56	7.0	226	1.14	10.6	6193	0.84	9.4
Rangpur (1.314+1.314)	—	—	—	368	0.66	10.4	2284	0.76	10.2	—	—	—	94	1.01	9.4	2746	0.76	10.2
Mymensingh (1.464+1.464)	697	0.69	8.6	710	0.81	9.6	12988	0.70	8.5	941	0.83	10.1	944	0.94	9.2	16280	0.73	8.7
Dacca (1.513+1.513)	282	0.72	8.4	396	1.03	10.8	5160	0.89	9.8	257	1.07	11.4	375	0.90	9.6	6470	0.90	9.9
eight districts (1.123+1.123)	1766	0.87	9.7	2447	0.84	9.5	32996	0.83	9.1	1944	0.87	9.7	2192	1.04	10.0	41345	0.85	9.3

TABLE 18. JUTE CENSUS 1940 : TIME OF BIG JOURNEY IN MAN-HOURS PER SAMPLE-UNIT (GRID) BY DISTRICTS AND SIZE OF GRIDS

name of district with density of grids	1-acre			2.25-acre			4-acre			6.25-acre			9-acre			total	average (for all sizes)	
	no. of grids	man-hours	percent	no. of grids	man-hours	percent	no. of grids	man-hours	percent	no. of grids	man-hours	percent	no. of grids	man-hours	percent	no. of grids	man-hours	percent
(1)	(2.1)	(2.2)	(2.3)	(3.1)	(3.2)	(3.3)	(4.1)	(4.2)	(4.3)	(5.1)	(5.2)	(5.3)	(6.1)	(6.2)	(6.3)	(7.1)	(7.2)	(7.3)
Nadia (0.548+0.548)	—	—	—	73	0.36	3.0	1294	1.16	6.4	128	0.81	4.6	—	—	—	1495	1.09	6.2
Jessore (0.759+0.759)	242	1.39	10.4	174	0.60	4.1	1831	0.69	5.2	—	—	—	264	0.75	6.3	2511	0.76	5.7
Rajshahi (0.920+0.920)	105	0.74	7.6	126	0.76	8.1	1628	0.92	9.8	147	0.62	7.1	44	0.79	8.6	2050	0.88	9.4
Bogra (1.224+1.224)	146	0.28	3.8	149	0.22	4.6	2920	0.50	6.5	140	0.44	5.2	245	0.68	6.2	6600	0.49	6.3
Tipperah (1.306+1.306)	294	1.00	12.5	451	0.64	6.9	4891	0.78	8.6	331	0.59	7.4	226	0.37	3.5	6193	0.75	8.4
Rangpur (1.314+1.314)	—	—	—	368	0.56	8.8	2284	0.54	7.4	—	—	—	94	0.59	5.5	2746	0.55	7.4
Mymensingh (1.464+1.464)	697	0.72	9.0	710	0.57	6.8	12988	0.71	8.6	941	0.49	6.0	944	0.75	7.3	16280	0.70	8.4
Dacca (1.513+1.513)	282	0.46	5.3	396	0.97	10.2	5160	0.75	8.2	257	0.69	7.4	375	0.94	10.0	6470	0.76	8.3
eight districts (1.123+1.123)	1766	0.78	8.7	2447	0.63	7.1	32996	0.72	7.9	1944	0.56	6.2	2192	0.73	7.0	41345	0.71	7.8

TABLE 19. JUTE CENSUS 1940 : MISCELLANEOUS WORK IN MAN-HOURS PER SAMPLE-UNIT (GRID) BY DISTRICTS AND SIZE OF GRIDS

name of district with density of grids	1-acre			2.25-acre			4-acre			6.25-acre			9-acre			total	average (for all sizes)	
	no. of grids	man-hours	percent	no. of grids	man-hours	percent	no. of grids	man-hours	percent	no. of grids	man-hours	percent	no. of grids	man-hours	percent	no. of grids	man-hours	percent
(1)	(2.1)	(2.2)	(2.3)	(3.1)	(3.2)	(3.3)	(4.1)	(4.2)	(4.3)	(5.1)	(5.2)	(5.3)	(6.1)	(6.2)	(6.3)	(7.1)	(7.2)	(7.3)
Nadia (0.548+0.548)	—	—	—	73	0.64	5.3	1294	0.99	5.5	128	0.89	5.1	—	—	—	1495	0.96	5.5
Jessore (0.759+0.759)	242	0.90	6.7	174	0.77	5.2	1831	0.66	4.9	—	—	—	264	0.84	6.8	2511	0.71	5.3
Rajshahi (0.920+0.920)	105	0.66	6.7	126	0.55	5.9	1623	0.71	7.5	147	0.96	10.9	44	0.70	7.5	2050	0.72	7.7
Bogra (1.224+1.224)	146	0.29	3.8	149	0.52	10.7	2920	0.49	6.4	140	0.36	4.2	245	1.28	11.5	3600	0.53	6.8
Tipperah (1.306+1.306)	294	0.44	5.6	451	0.86	9.2	4891	0.68	7.6	331	0.72	9.0	226	0.68	6.3	6193	0.69	7.7
Rangpur (1.314+1.314)	—	—	—	368	0.58	9.2	2284	0.55	7.4	—	—	—	94	1.39	13.0	2746	0.58	7.8
Mymensingh (1.464+1.464)	697	0.48	6.0	710	0.54	6.5	12988	0.55	6.6	941	0.45	5.5	944	0.80	7.8	16280	0.55	6.6
Dacca (1.513+1.513)	282	0.58	6.8	396	0.75	7.9	5160	0.62	6.8	257	0.91	9.8	375	0.77	8.2	6470	0.65	7.1
eight districts (1.123+1.123)	1766	0.54	6.0	2447	0.66	7.4	32996	0.60	6.6	1944	0.62	6.9	2192	0.86	8.2	41345	0.62	6.8

TABLE 20. JUTE CENSUS 1940 : FIELD WORK—INDIRECT TIME AS PERCENTAGE OF TOTAL TIME (24 HOURS) BY DISTRICTS AND SIZE OF GRIDS

name of district with density of grids	1-acre		2.25-acre		4-acre		6.25-acre		9-acre		total number of grids	average (for all sizes) percent
	number of grids	percent	number of grids	percent	number of grids	percent	number of grids	percent	number of grids	percent		
(1)	(2.1)	(2.2)	(3.1)	(3.2)	(4.1)	(4.2)	(5.1)	(5.2)	(6.1)	(6.2)	(7.1)	(7.2)
Nadia (0.548+0.548)	—	—	73	79.0	1294	74.1	128	77.1	—	—	1495	74.5
Jessore (0.759+0.759)	242	62.2	174	75.2	1831	74.3	—	—	264	65.1	2511	72.4
Rajshahi (0.920+0.920)	105	69.1	126	68.9	1628	61.9	147	62.0	44	63.1	2050	71.7
Bogra (1.224+1.224)	146	74.4	149	50.1	2920	62.6	140	70.6	245	54.8	3600	62.4
Tipperah (1.306+1.306)	294	60.5	451	66.0	4891	63.6	331	63.8	226	67.7	6193	63.8
Rangpur (1.314+1.314)	—	—	368	60.6	2284	65.7	—	—	94	60.9	2746	64.9
Mymensingh (1.464+1.464)	697	66.0	710	64.6	12988	64.4	941	64.2	944	60.8	16280	64.0
Dacca (1.513+1.513)	282	70.3	396	62.1	5160	61.7	257	59.8	375	60.7	6470	62.0
eight districts (1.123+1.123)	1766	65.8	2447	65.6	32996	64.6	1944	65.5	2192	61.5	41345	64.6

TIME OF FIELD WORK PER SQUARE MILE

74. For budget purposes it is more convenient to use the time required for field operations on a square mile basis; this can be easily obtained by multiplying the time in man-hours per sample-unit by the respective density (or number of sample-units per square mile). These values are shown in Tables 21–25; columnar arrangements in these tables are similar to those in Tables 15–19. Table 21 shows the total field time in net and gross man-hours per square mile by districts and size of grid; col. (1) gives the name of the district with the density of grids for each half-sample; col. (2.1) the number of grids; col. (2.2) the net man-hours per square mile; and col. (2.3) the gross man-hours per square mile for sample-units of size 1-acre. Similar information for sample-units of 2.25-acre, 4-acre, 6.25-acre, 9-acre, and finally for grids of all sizes taken together, are shown in cols. (3.1)–(3.3), (4.1)–(4.3), (5.1)–(5.3), (6.1)–(6.3) and cols. (7.1)–(7.3) respectively.

75. As only one single density (or number) of sample-units per square mile was used in any particular district, the information given by districts also automatically gives the information by densities. It will be noticed in this table that the time required for field work on a square mile basis increases very appreciably with the density as well as with the size of sample-units.

76. Table 22 shows the time required for enumeration in net man-hours per square mile by districts and size of grid; col. (1) as usual gives the name of the district with the corresponding density of sample-unit in each half-sample; col. (2.1) the number of grids; and col. (2.2) the net man-hours per square mile required for the enumeration portion of the work in sample units of size 1-acre; cols. (3.1)–(3.2), cols. (4.1)–(4.2), cols. (5.1)–(5.2), cols (6.1)–(6.2) and cols. (7.1)–(7.2) give similar information for grids of size 2.25-acre, 4-acre, 6.25-acre, 9-acre, and for sample-units of all sizes taken together respectively. The bottom line as usual gives the average values for each size of sample-unit for all eight districts or densities taken together. A glance at the marginal col. (7.2) and the bottom line shows that the time required increases directly with both density and size of sample-units.

77. Table 23 gives the time required for small journey in net man-hours per square mile by districts and size of grid; col. (1) gives the name of the district with the density of each half-sample; col. (2.1) the number of grids; col. (2.2) net man-hours per square mile for 1-acre grids. Similar information for 2.25 acre, 4-acre, 6.25-acre, and 9-acre, will be found in cols. (3.1)–(3.2). cols. (4.1)–(4.2), cols. (5.1)–(5.2) and cols. (6.2)–(6.2) respectively. The total number of grids of all sizes is given in col. (7.1); average net time for all sizes of grids taken together is given in actual man-hours per square mile in col. (7.2). The bottom line gives the average for all eight districts. In the case of small journey the time required per square mile is practically independent of the size of sample-units, but appears to decrease appreciably with the density of grids.

TABLE 21. JUTE CENSUS 1940 : TOTAL TIME REQUIRED FOR FIELD OPERATIONS IN NET AND GROSS MAN-HOURS PER SQUARE MILE BY DISTRICTS AND SIZE OF GRIDS

name of district with density of grids	1-acre			2.25-acre			4-acre			6.25-acre			9-acre			total no. of grids	average (for all sizes)	
	no. of grids	man-hours		no. of grids	man-hours		no. of grids	man-hours		no. of grids	man-hours		no. of grids	man-hours			man-hours	
		net	gross		net	gross		net	gross		net	gross		net	gross		net	gross
(1)	(2.1)	(2.2)	(2.3)	(3.1)	(3.2)	(3.3)	(4.1)	(4.2)	(4.3)	(5.1)	(5.2)	(5.3)	(6.1)	(6.2)	(6.3)	(7.1)	(7.2)	(7.3)
Nadia (0.548+0.548)	—	—	—	773	2.81	1.34	1294	5.06	1.95	128	4.41	1.92	—	—	—	1495	4.90	1.92
Jessore (0.759+0.759)	242	7.68	2.05	1174	5.59	2.25	1831	5.25	2.05	—	—	—	264	6.28	1.81	2511	5.62	2.03
Rajshahi (0.920+0.920)	105	5.59	18.09	126	5.34	17.17	1628	6.62	17.37	147	6.16	16.23	44	6.29	17.06	2050	6.44	17.30
Bogra (1.224+1.224)	146	4.72	18.51	149	5.90	11.82	2920	7.03	18.70	140	6.17	20.98	245	12.24	2.72	3600	7.22	19.14
Tipperah (1.306+1.306)	294	8.25	20.90	451	8.25	24.24	4891	8.52	23.40	331	7.55	20.84	226	9.09	2.82	6193	8.46	23.38
Rangpur (1.314+1.314)	—	—	—	368	6.57	16.64	2284	6.70	19.58	—	—	—	94	11.06	2.84	2746	6.83	19.50
Mymensingh (1.464+1.464)	697	7.99	23.51	710	8.72	24.65	12988	8.57	24.19	941	8.61	24.10	944	11.77	3.02	16280	8.78	24.51
Dacca (1.513+1.513)	282	7.72	26.02	396	10.92	28.81	5160	10.53	27.54	257	11.35	28.26	375	11.14	28.35	6470	10.53	27.63
average districts and (8 densities) (1.123+1.123)	1766	6.85	20.06	2447	6.87	19.90	32996	7.12	20.46	1944	6.99	20.24	2192	9.05	2.36	41345	7.19	20.55

TABLE 22. JUTE CENSUS 1940: TIME REQUIRED FOR ENUMERATION IN NET MAN-HOURS PER SQUARE MILE BY DISTRICTS AND SIZE OF GRIDS

name of districts with density of grids	1-acre		2.25-acre		4-acre		6.25-acre		9-acre		total	average (for all sizes)
	no. of grids	man-hours	no. of grids	man-hours	no. of grids	man-hours	no. of grids	man-hours	no. of grids	man-hours	no. of grids	man-hours
(1)	(2.1)	(2.2)	(3.1)	(3.2)	(4.1)	(4.2)	(5.1)	(5.2)	(6.1)	(6.2)	(7.1)	(7.2)
Nadia (0.548+0.548)	—	—	73	0.75	1294	1.16	128	0.95	—	—	1495	1.12
Jessore (0.759+0.759)	242	1.75	174	1.50	1831	1.41	—	—	264	1.86	2511	1.49
Rajshahi (0.920+0.920)	105	1.42	126	1.96	1628	1.84	147	1.52	44	2.13	2050	1.81
Bogra (1.224+1.224)	146	1.42	149	2.62	2920	2.62	140	2.02	245	4.44	3600	2.67
Tipperah (1.306+1.306)	294	2.14	451	2.08	4891	2.51	331	2.67	226	3.35	6193	2.50
Rangpur (1.314+1.314)	—	—	368	1.82	2284	1.82	—	—	94	3.20	2746	1.87
Mymensingh (1.464+1.464)	697	2.47	710	3.11	12988	2.85	941	3.42	944	4.50	16280	2.97
Dacca (1.513+1.513)	282	2.19	396	2.59	5160	3.72	257	3.26	375	3.24	6470	3.54
average (eight districts and densities) (1.123+1.123)	1766	1.90	2447	2.04	32996	2.25	1944	2.38	2192	3.15	41345	2.28

TABLE 23. JUTE CENSUS 1940 : TIME REQUIRED FOR SMALL JOURNEY IN NET MAN-HOURS PER SQUARE MILE BY DISTRICTS AND SIZE OF GRIDS

name of districts with density of grids	1-acre		2.25-acre		4-acre		6.25-acre		9-acre		total	average (all sizes)
	no. of grids	man-hours	no. of grids	man-hours	no. of grids	man-hours	no. of grids	man-hours	no. of grids	man-hours	no. of grids	man-hours
(1)	(2.1)	(2.2)	(3.1)	(3.2)	(4.1)	(4.2)	(5.1)	(5.2)	(6.1)	(6.2)	(7.1)	(7.2)
Nadia (0.548+0.548)	—	—	73	0.95	1294	1.55	128	1.59	—	—	1495	1.56
Jessore (0.759+0.759)	242	2.45	174	2.00	1831	1.79	—	—	264	2.08	2511	1.90
Rajshahi (0.920+0.920)	105	1.59	126	0.98	1628	1.77	147	1.72	44	1.42	2050	1.70
Bogra (1.224+1.224)	146	1.91	149	1.48	2920	2.00	140	2.17	245	3.01	3600	2.05
Tipperah (1.306+1.306)	294	2.34	451	2.25	4891	2.20	331	1.46	226	2.99	6193	2.20
Rangpur (1.364+1.364)	—	—	368	1.73	2284	2.00	—	—	94	2.64	2746	1.99
Mymensingh (1.464+1.464)	697	2.02	710	2.36	12988	2.06	941	2.44	944	2.75	16280	2.14
Dacca (1.513+1.513)	282	2.19	396	3.11	5160	2.68	257	3.23	375	2.74	6470	2.71
average eight districts and densities (1.123+1.123)	1766	2.14	2447	2.06	32996	2.04	1944	2.12	2192	2.54	41345	2.08

78. Table 24 shows the time required for big journey in net man-hours per square mile by districts and size of grid; col. (1) gives the name of the district with the density of each half-sample; col. (2.1) the number of grids ; col (2.2) the net man-hours per square mile for sample-units of size 1-acre; cols. (3.1)–(3.2), cols. (4.1)–(4.2), cols. (5.1)–(5.2), cols. (6.1)–(6.2) and cols. (7.1)–(7.2) give the same information for grids of size 2.25-acre, 4-acre, 6.25-acre, 9-acre, and for grids of all sizes taken together ; the bottom line gives averages for each size of grid for all eight districts or densities. The connexion between the time required for big journey and size and density of grid is not quite obvious; this question has been considered later.

79. Table 25 gives the time required for miscellaneous work in net man-hours per square mile by districts and size of grids; col. (1) gives the name of the district with the density of each half-sample; col. (2.1) the number of grids; col. (2.2) the net man-hours per square mile for sample-unit of size 1-acre, cols. (3.1)–(3.2), cols. (4.1)–(4.2), cols. (5.1)–(5.2), cols. (6.1)–6.2) and cols. (7.1)–(7.2) give the same information for grids of size 2.25-acre, 4-acre, 6.25-acre, 9-acre and for grids of all size taken together; the bottom line gives the average for each size of grid for all eight districts or densities. The time for miscellaneous work is not affected appreciably by the size of grids, but decreases with the density which is quite plausible.

Graduation of time of field operations

80. The next important stage of the work was to fit suitable mathematical expressions to the different components of the cost. This required a great deal of laborious work on the part of senior statisticians and a large body of computers. The results were first plotted graphically to give rough indications of the nature of algebraic expressions which were likely to be useful. In each case a number of different mathematical formulae were tried, and numerical constants were calculated by least square methods; graduated values calculated from the fitted curves were then compared with the observed values. This comparison eliminated a large number of the trial curves; but in two or three cases alternative forms gave almost equally good fits. It will not serve any useful purpose to reproduce in this Report all the different mathematical equations which were given a trial for purposes of graduation. I am therefore quoting here only those equations which were finally selected for constructing the cost function; for convenience of reference I have, however, also given a few alternative forms which may prove useful when fresh field observations become available.

Graduated values of time for enumeration

81. The graduation of the time required for enumeration in net man-hours per grid was comparatively straightforward. In this case the volume of work would naturally increase with the size of the sample-units, but should be independent of the density of grids per square mile; and this was confirmed by the observed time-records. It was found that the following simple linear equation was quite adequate :

$$t_1 = e_0 + e_1(n) = 0.7589 + 0.0600(n). \quad \ldots \quad (1)$$

TABLE 24. JUTE CENSUS 1940 : TIME REQUIRED FOR BIG JOURNEY IN NET MAN-HOURS PER SQUARE MILE BY DISTRICTS AND SIZE OF GRIDS

name of districts with density of grids	1-acre		2.25-acre		4-acre		6.25-acre		9-acre		total	average (all sizes)
	no. of grids	man-hours	no. of grids	man-hours	no. of grids	man-hours	no. of grids	man-hours	no. of grids	man-hours	no. of grids	man-hours
(1)	(2.1)	2.2)	(3.1)	(3.2)	(4.1)	(4.2)	(5.1)	(5.2)	(6.1)	(6.2)	(7.1)	(7.2)
Nadia (0.548+0.548)	—	—	73	0.40	1294	1.27	128	0.89	—	—	1495	1.19
Jessore (0.759+0.759)	242	2.12	174	0.91	1813	1.05	—	—	264	1.13	2511	1.15
Rajshshi (0.920+0.920)	105	1.37	126	1.39	1628	1.69	147	1.15	44	1.46	2050	1.61
Bogra (1.224+1.224)	146	0.70	149	0.55	2920	1.21	140	1.08	245	1.67	3600	1.19
Tipperah (1.306+1.306)	294	2.61	451	1.68	4891	2.03	331	1.54	226	0.97	6193	1.96
Rangpur (1.314+1.314)	—	—	368	1.46	2284	1.44	—	—	94	1.54	2746	1.45
Mymensingh (1.464+1.464)	697	2.11	710	1.67	12988	2.09	941	1.44	944	2.20	16260	2.04
Dacca (1.513+1.513)	282	1.38	396	2.93	5160	2.27	257	2.09	375	2.84	6470	2.29
average (eight districts and densities) (1.123+1.123)	1766	1.76	2447	4.42	32996	1.63	1944	1.26	2192	1.64	41345	1.60

TABLE 25. JUTE CENSUS 1940 : MISCELLANEOUS WORK IN MAN-HOURS PER SQUARE MILE BY DISTRICTS AND SIZE OF GRIDS

name of districts with density of grids	1-acre		2.25-acre		4-acre		6.25-acre		9-acre		total no. of grids	average (all sizes) man-hours
	no. of grids	man-hours	no. of grids	man-hours	no. of grids	man-hours	no. of grids	man-hours	no. of grids	man-hours		
(1)	(2.1)	(2.2)	(3.1)	(3.2)	(4.1)	(4.2)	(5.1)	(5.2)	(6.1)	(6.2)	(7.1)	(7.2)
Nadia (0.548+0.548)	—	—	73	0.70	1294	1.08	128	0.98	—	—	1495	1.05
Jessore (0.759+0.759)	242	1.37	174	1.64	1831	1.00	—	—	264	1.28	2511	1.07
Rajshahi (0.920+0.920)	105	1.21	126	1.01	1628	1.31	147	1.77	44	1.28	2050	1.32
Bogra (1.224+1.224)	146	0.70	149	1.26	2920	1.20	140	0.88	245	3.12	3600	1.30
Tipperah (1.306+1.306)	294	1.17	451	2.23	4891	1.79	331	1.88	226	1.78	6193	1.79
Rangpur (1.314+1.314)	—	—	368	1.54	2284	1.45	—	—	94	3.66	2746	1.54
Mymensingh (1.464+1.464)	697	1.41	710	1.59	12988	1.60	941	1.32	944	2.34	16280	1.62
Dacca (1.513+1.513)	282	1.77	396	2.27	5160	1.87	257	2.76	375	2.32	6470	1.96
average (eigot districts and densities) (1.123+1.123)	1766	1.22	2447	1.48	32996	1.36	1944	1.39	2192	1.93	41345	1.39

The observed and graduated values are given in Table 26 in which col. (1) shows the size of grids in acre; col. (2) the number of grids surveyed; col. (3) the observed time for enumeration in net man-hours per grid; and col. (4) the corresponding graduated values calculated from the above equation. The next col. (5) shows the difference between observed and graduated values ; and col. (6) this difference expressed as a percentage of the graduated value. It will be noticed that the percentage differences are fairly small so that the fit may be considered reasonable.

TABLE 26. TIME FOR ENUMERATION IN NET MAN-HOURS PER GRID
(Seven Districts)

size of grids in acre($=x$)	number of grids	net man-hours per grid		difference (observed-graduated)	
		observed	graduated	actual	percentage
(1)	(2)	(3)	(4)	(5)	(6)
		$t_1=0.7589+0.0600(x)$			
1-acre	1620	0.87	0.82	+0.05	+6.10
2.25-acre	2298	0.90	0.89	+0.01	+1.13
4-acre	30076	1.00	1.00	0.00	0.00
6.25-acre	1804	1.08	1.13	−0.05	−4.43
9-acre	1947	1.35	1.30	+0.05	+3.85

GRADUATED VALUES OF TIME FOR SMALL JOURNEY

82. The time spent in small journey, that is, from camp to field, field to camp etc., should be independent of the size, and this was confirmed by time records. When the density of grids is high it may also be expected that the time of small journey per grid should be comparatively small; and this is what was actually observed. The following linear equation gave a fair fit :

$$t_2 = j_0 - j_1(y) = 1.5961 - 0.5673(y). \quad \ldots \quad (2)$$

83. The following exponential formula was also tried and gave a slightly better fit :

$$t_2 = a(y)^{-K} = 0.9752(y)^{-0.6080} \quad \ldots \quad (2.1)$$

Table 27 gives the observed and values graduated by both equations. In this table col. (1) gives the name of the district; col. (2) the values of y, that is, the density or number of grids per square mile; col (3) the total number of grids surveyed; col; (4) the observed time required for small journey, in net man-hours per grid; and col. (5) the corresponding graduated values; the difference between the observed and the graduated value is shown in actual hours in col. (6); and as a percentage of the graduated value in col. (7). It will be seen from col. (7) that the exponential equation gives slightly better fit; but the straight line also is not unsatisfactory. For our final calculations we have used the linear equation because of its simpler form.

TABLE 27. TIME FOR SMALL JOURNEY IN NET MAN-HOURS PER GRID

district	density of grids ($=y$)	number of grids	net man-hours per grid		difference (observed-graduated)	
			observed	graduated	actual	percentage
(1)	(2)	(3)	(4)	(5)	(6)	(7)
			$t_1 = 1.5961 - 0.5673(y)$			
Nadia	0.548	1495	1.42	1.29	+0.13	+10.1
Jessore	0.759	2511	1.25	1.17	+0.08	+ 6.8
Rajashahi	0.920	2050	0.92	1.07	−0.15	−14.0
Bogra	1.224	3600	0.84	0.90	−0.06	− 6.7
Tipperah	1.306	6193	0.84	0.86	−0.02	− 2.3
Rangpur	1.314	2746	0.76	0.85	−0.09	−10.6
Mymensingh	1.460	16280	0.73	0.77	−0.04	− 5.2
Dacca	1.513	6470	0.90	0.74	+0.16	+21.6
			$t_2 = 0.9752(y)^{-0.6080}$			
Nadia	0.548	1495	1.42	1.41	+0.01	+1.08
Jessore	0.759	2511	1.25	1.15	+0.10	+8.35
Rajshahi	0.920	2050	0.92	1.05	−0.13	−13.71
Tipperah	1.306	6193	0.84	0.83	+0.01	+1.59
Rangpur	1.314	2746	0.76	0.83	−0.07	−9.13
Mymensingh	1.464	16280	0.73	0.75	−0.02	−3.28
Dacca	1.513	6470	0.90	0.76	+0.14	+18.27

GRADUATED VALUES OF TIME FOR BIG JOURNEYS

84. It will be remembered that big journey represents the time taken for shifting from one camp to another. This item is of great importance in deciding the camp arrangements; and has therefore been analysed in considerable detail. Table 28 shows the time for big journey in net man-hours per grid as graduated by the following exponential formula :

$$t_3 = 0.7899(y)^{-0.312}. \qquad \ldots \quad (3.1)$$

Col. (1) shows the name of the district, col. (2) the density or the number of sample-units per square mile; col. (3) the number of grids surveyed ; col. (4) the observed time taken for big journey in net man-hours per grid; and col. (5) the corresponding graduated value. The difference between the observed and the graduated values is given in col. (6), and the percentage difference in col. (7). The time for going from one camp to another is naturally independent of the size of the grid, but depends on the density; the graduation is however not very satisfactory.

TABLE 28. TIME FOR BIG JOURNEY IN NET MAN-HOURS PER GRID

district	density of grids ($=y$)	number of grids	net man-hours per grid		difference (observed-graduated)	
			observed	graduated	actual	percentage
(1)	(2)	(3)	(4)	(5)	(6)	(7)
		$t_1 = 0.7899(y) - 0.312$				
Nadia	0.548	1495	1.09	0.95	+0.14	+14.26
Jessore	0.759	2511	0.76	0.86	−0.10	−11.64
Rajshahi	0.920	2050	0.88	0.81	+0.07	+8.24
Tipperah	1.306	6193	0.75	0.73	+0.02	+3.41
Rangpur	1.314	2746	0.55	0.73	−0.18	−24.12
Mymensingh	1.464	16280	0.70	0.70	—	−0.57
Dacca	1.513	6470	0.76	0.69	+0.07	+9.25

85. This is probably due to two important causes of variation, the mode of journey is heterogeneous, and includes such diverse ways as walking, bullock cart, country boat, motor bus, steamer or railways. The second cause is the intrinsic heterogeneity of the districts regarding means of communication. A joint analysis covering 20,000 square miles is however useful in giving a broad average picture of the situation.

86. In a provincial survey it will be necessary to cover many districts where the intensity of cultivation or proportion of land under jute is very low, of the order of only one or two per cent or less. In these districts it is clear that we can have only a small number of sample-units, that is, we must work with very low densities. In such cases the camps will probably have to be fixed at considerable distances apart, so that the time required for big journey that is, for moving from one camp to another, will be of great importance.

87. We have studied this question by analysing the time-records for big journey in terms of the distance travelled. The material is shown in Table 30 in which col. (1) gives the distance between camps, i.e., the length of the journey in miles; col. (2) shows the number of journeys undertaken in district Nadia, and similar information for other seven districts is given in cols. (3)–(9), and for all eight districts taken together in col. (10). As the total time required for these journeys are known it is easy to calculate the average time taken per mile for travelling different distances. These are shown in average hours per mile separately for the eight districts in cols. (11)–(18), and finally for all eight districts taken together in col. (19).

88. Further analysis of the same material for all eight districts combined is shown in Table 29. In this table col. (1) gives the range of the journey (that is, the distance between camps) in miles; col. (2) the total number of journeys; col. (3) the percentage frequencies of journeys; and col. (4) the accumulated percentage frequencies. It will be noticed from col. (4) that roughly half the journeys are less than 5 miles in distance, 76.6 per cent less than 10 miles, while 96.6 per cent are less than 25 miles.

TABLE 29. BIG JOURNEY : TIME IN MAN-HOURS PER LINEAR MILE

range of journey in miles	total number of journeys	per cent frequency of journeys	accumulated per cent frequency	time in hours per mile			percentage difference
				observed	graduated	difference	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			$t_1 = 1.3349(z) - 0.3869$				
0.1 — 1.0	300	4.5	4.5	1.85	1.68	+ 0.17	+10.12
1.1 — 2.0	914	13.8	18.3	1.10	1.13	— 0.03	— 2.65
2.1 — 3.0	834	12.5	30.8	0.89	0.93	— 0.04	— 4.30
3.1 — 4.0	682	10.3	41.1	0.76	0.82	— 0.06	— 7.32
4.1 — 5.0	549	8.3	49.4	0.73	0.74	— 0.01	— 1.35
5.1 — 6.0	460	6.9	56.3	0.65	0.69	— 0.04	— 5.80
6.1 — 7.0	430	6.5	62.8	0.69	0.64	+ 0.05	+ 7.81
7.1 — 8.0	358	5.4	68.2	0.66	0.61	+ 0.05	+ 8.20
8.1 — 9.0	283	4.3	72.5	0.63	0.58	+ 0.05	+ 8.62
9.1 — 10.0	276	4.2	76.7	0.63	0.56	+ 0.07	+12.50
10.1 — 15.0	821	12.4	89.1	0.55	0.50	+ 0.05	+10.00
15.1 — 20.0	353	5.3	94.4	0.41	0.44	— 0.03	— 6.82
20.1 — 25.0	149	2.2	96.6	0.37	0.40	— 0.03	— 7.50
25.1 — 30.0	73	1.1	97.7	0.36	0.37	— 0.01	— 2.70
30.1 — 40.0	87	1.3	99.0	0.28	0.34	— 0.06	—17.65
40.1 — 50.0	35	0.5	99.5	0.18	0.31	— 0.13	—41.94
50.1 & above	36	0.5	110.0	—	—	—	—
total	6640	100.0	—	—	—	—	—

89. The same Table 29 shows the results of the hours per mile required for journeys of various distances graduated by the following exponential formula

$$t = 1.334(z)^{-0.3869} \qquad \ldots \quad (3.2)$$

in which z is the length of the journey in miles, that is, the distance between camps. In this table col. (5) gives the observed time in hour per mile; col. (6) the corresponding graduated time; and col. (7) the difference between the observed and graduated values in hours; finally the difference is shown in the form of a percentage in col. (8). It will be noticed that quite naturally the time in hours per mile decreases as the length of the journeys increases. Remembering the heterogeneity of the material the graduation is not unsatisfactory.

90. Other results for big journey and camps are shown in Table 31 in which col. (1) gives the name of the districts, col. (2) the density or number of sample units per square mile; col. (3) the total number of square miles surveyed; col. (4) the total number of journeys; col (5) the total number of miles travelled; and col. (6) the total number of hours required for big journey.

91. The average length per journey in miles in each district is shown in col. (7); these are weighted averages based on detailed records for each district. It will be noticed that there is a certain amount of variation from district to district, from 6.35 miles in Dacca to 11.53 miles in Bogra. The weighted average time per journey in hours is shown in col. (8) of Table 31, here also the values fluctuate from 3.88 hours in Jessore to 6.66 hours in Bogra. The average time for shifting from one camp to another is 4.45 hours for eight districts taken together.

92. We can easily calculate the average haltage per camp by subtracting the total hours of journey from the total amount of time spent in each district, and dividing this by the number of journeys. The haltage per camp calculated in this way is shown in col. (9) in hours and converted into days in col. (10). It will be seen that the general average is 2.19 days indicating that most of the halts were of about two days with some of three or more days. The variation between districts is quite large; in Mymensingh the average haltage per camp is 1.87 days showing that many single night halts were made. Bogra, on the other hand, the average is 4.15 days showing that investigators spent several nights at each camp.

93. The average area (or circle) in square miles covered from each camp is shown in Table 31, col. (11). This varies from 2.27 square miles in Mymensingh and 2.42 square miles in Dacca to 5.55 square miles in Bogra; the general average is 2.77 square miles. As the average haltage is 2.19 days we find that the average area covered per day is about 1.22 square miles.

94. We can obtain some idea of the amount of travelling involved by calculating the number of linear miles travelled per square mile of the area surveyed. These are shown in col (12) of Table 31; the general average is 2.77 linear miles per square mile. The corresponding average hours for big journey per square mile are given in col. (13) with a general average 1.60 hours per square mile. This is the figure which we have finally accepted for building up the cost function.

95. The average number of hours required for travelling each linear mile is shown in col. (14). The variation in this column is partly due to the differences in the length of journey and partly to intrinsic differences in means of communication in different districts.

96. We find from the above analysis that in certain districts, e.g., in Dacca and Mymensingh, the average distance between camps was small and of the order of 6 or 7 miles; the average time taken for the journey was also small and of the order of 4 or 5 hours. In other districts like Bogra or Rangpur the distance between camps was much greater and of the order of 11 or 12 miles, with naturally a higher average time per journey of the order of 6 or 7 hours. The haltage per camp was quite small in Rajshahi, Mymensingh and Dacca where the investigators probably stayed only for a night or two at each camp; while in Rangpur or Bogra the average haltage was more than three or four days showing that the workers stopped at each camp for three or four nights or more. The average area covered while staying in the same camp was naturally quite small and a little over two square miles in Mymensingh and Dacca, and much higher and over 4 or 5 square miles in Rangpur or Bogra. We thus find that in some of the districts, camps were shifted frequently and were pitched at smaller distances apart; the haltage per camp and naturally the extent of ground covered was small. On the other hand, in Rangpur and Bogra the camps were fixed at much greater distances apart, haltage in each camp was longer, and the area commanded per camp was much greater. On the whole we find that the average distance between camps varied roughly from 6 to 12 miles, and the area commanded was about one and a quarter square mile per day.

TABLE 30. FREQUENCY TABLE OF BIG JOURNEY AND AVERAGE HOURS PER MILE BY DISTRICTS

range of journey in miles	number of journeys (frequency)									average hours per mile								
	Nadia 0.548	Jessore 0.759	Raj-shahi 0.920	Bogra 1.224	Tip-perah 1.306	Rang-pur 1.314	Mymen-singh 1.464	Dacca 1.513	eight districts 1.123	Nadia 0.548	Jessore 0.759	Raj-shahi 0.920	Bogra 1.224	Tip-perah 1.306	Rang-pur 1.314	Mymen-singh 1.464	Dacca 1.513	eight districts 1.123
(1)	(2)	(3)	(4)	(5)	(5)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)
0.1 — 1.0	10	20	7	1	58	2	154	48	300	.53	1.87	1.64	2.00	1.94	1.75	1.99	2.22	1.85
1.1 — 2.0	36	61	26	9	159	19	461	143	914	.96	.81	1.00	1.97	1.14	1.20	1.09	1.22	1.10
2.1 — 3.0	36	57	52	13	139	26	367	144	834	.73	.86	.77	.85	.88	.77	.87	1.05	.89
3.1 — 4.0	43	54	67	10	81	27	270	130	682	.66	.59	.69	.92	.67	1.08	.72	.97	.76
4.1 — 5.0	33	44	44	18	54	8	237	111	549	.62	.36	.37	.86	.73	1.48	.69	.86	.73
5.1 — 6.0	29	38	46	12	47	12	188	88	460	.44	.62	.58	.74	.88	.72	.61	.70	.65
6.1 — 7.0	25	26	40	36	46	27	176	54	430	.49	.54	.56	.92	.64	.83	.68	.80	.69
7.1 — 8.0	39	23	31	15	37	25	130	58	358	.56	.50	.57	1.23	.68	.80	.68	.90	.66
8.1 — 9.0	17	21	29	12	29	7	94	74	283	.43	.49	.45	.94	.98	.70	.64	.74	.63
9.1 — 10.0	16	19	11	11	39	21	109	50	276	.43	.35	.79	.64	.83	.72	.57	.73	.63
10.1 — 15.0	55	68	51	66	98	26	337	120	821	.43	.42	.63	.66	.61	.75	.53	.52	.55
15.1 — 20.0	29	26	22	22	64	14	122	54	352	.27	.34	.53	.60	.49	.78	.61	.52	.41
20.1 — 25.0	13	12	2	19	24	4	63	12	149	.35	.11	.13	.41	.35	.61	.39	.45	.37
25.1 — 30.0	7	6	—	13	9	9	25	4	73	.36	.21	—	.19	.28	.42	.52	.24	.36
30.1 — 40.0	26	3	—	8	13	7	24	6	87	.30	.12	—	.43	.23	.49	.20	.29	.28
40.1 — 50.0	3	1	6	—	5	5	15	—	35	.13	.23	.26	—	.18	.23	.13	—	.18
50.1 and above	1	13	5	—	10	7	—	—	36	.12	.19	.41	—	.59	.10	—	—	.30
total	418	492	439	265	912	246	2772	1096	6640	.41	.39	.56	.63	.61	.53	.62	.71	.58

TABLE 31. ANALYSIS OF TIME RECORDS RELATING TO BIG JOURNEY AND CAMPS

district	density of half-sample	total number of				average length per journey in miles	average time per journey in hours	haltage per camp		average area of circle in sq. miles per camp	average linear miles travelled per sq. mile	average hours for big journey per	
		square miles surveyed	journeys	miles travelled	hours of journey			in hours	in days			square mile	linear mile
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Nadia	0.548	1415	418	3951	1629	9.92	3.90	58.8	2.45	3.38	2.79	1.15	0.41
Jessore	0.759	1704	492	4442	1753	8.62	3.88	64.8	2.70	3.46	2.61	1.03	0.39
Rajshahi	0.920	1235	439	3392	1902	7.63	4.11	39.9	1.66	2.81	2.75	1.54	0.56
Bogra	1.224	1471	265	2928	1847	11.53	6.66	99.6	4.15	5.55	1.99	1.25	0.63
Tipperah	1.306	2536	912	7099	4327	7.71	5.09	55.8	2.32	2.78	2.80	1.71	0.61
Rangpur	1.314	1095	246	2987	1596	11.09	6.14	76.7	3.19	4.45	2.73	1.46	0.53
Mymensingh	1.464	6306	2772	19017	11348	6.73	4.11	45.0	1.87	2.27	3.02	1.80	0.62
Dacca	1.513	2652	1096	7113	5048	6.35	4.49	49.4	2.06	2.42	2.68	1.90	0.71
8 districts	—	18414	6640	50929	29450	7.56	4.45	52.6	2.19	2.77	2.77	1.60	0.58

TABLE 32. JUTE CENSUS 1940 : TIME OF BIG JOURNEY IN MAN-HOURS PER SQUARE MILE BY DISTRICTS AND SIZE OF GRIDS

district with density of half-sample	1-acre			2.25-acre			4-acre			6.25-acre			9-acre			average (grids of all sizes)		
	number of grids	number of square miles	man-hours	number of grids	number of square miles	man-hours	number of grids	number of square miles	man-hours	number of grids	number of square miles	man-hours	number of grids	number of square miles	man-hours	number of grids	number of square miles	man-hours
(1)	(2.1)	(2.2)	(2.3)	(3.1)	(3.2)	(3.3)	(4.1)	(4.2)	(4.3)	(5.1)	(5.2)	(5.3)	(6.1)	(6.2)	(6.3)	(7.1)	(7.2)	(7.3)
Nadia (0.548+0.548)	—	—	—	73	69	0.40	1294	1224	1.27	128	122	0.89	—	—	—	1495	1515	1.19
Jessore (0.759+0.759)	242	162	2.12	174	118	0.91	1831	1237	1.05	—	—	—	264	186	1.13	2511	1704	1.15
Rajshahi (0.920+0.920)	105	64	1.37	126	76	1.39	1628	981	1.69	147	88	1.15	44	27	1.46	2050	1235	1.61
Bogra (1.224+1.224)	146	70	0.70	149	62	0.55	2920	1194	1.21	140	57	1.08	245	100	1.67	3600	1472	1.19
Tipperah (1.306+1.306)	294	124	2.61	451	174	1.68	4891	1998	2.03	331	151	1.54	226	90	0.97	6193	2536	1.96
Rangpur (1.314+1.314)	—	—	—	368	142	1.46	2284	910	1.44	—	—	—	94	44	1.54	2746	1095	1.45
Mymensingh (1.464+1.464)	697	220	2.11	710	291	1.67	12988	5098	2.09	941	346	1.44	944	351	2.20	16280	6306	2.04
Dacca (1.513+1.513)	282	103	1.38	396	85	2.93	5160	2254	2.27	257	72	20.9	375	139	2.84	6470	2853	2.29
average of eight districts (1.223+1.223)	1766	743	1.76	2447	1017	1.42	32996	14896	1.63	1944	836	1.26	2192	937	1.64	41345	18516	1.60

97. After examining the question carefully we thought it best to accept the average number of linear miles travelled per square mile or rather the average number of hours required for big journey per square mile of area surveyed as a constant in our calculations for the cost function. We accepted the figure 1.6 hour per square mile given in col. (13) of Table 31 for this purpose. Details of the number of man-hours per square mile by districts and size of grids are shown in Table 32.

GRADUATION OF TIME FOR MISCELLANEOUS WORK

98. The time for miscellaneous work amounts to nearly a fifth of the total time required for field operations and is not by any means negligible. Various equations were tried in this case; the results for two linear fits are given in Tables 33 and 34 respectively. Table 33 shows the graduation of miscellaneous work in actual man-hours per grid by the following linear equation.

$$t_4 = 0.9976 - 0.2788(y) \qquad \ldots \quad (4)$$

where y as usual is the density or number of grids per square mile. In the Table 33 col. (1) gives the name of the district; col. (2) the density or number of grids per square mile; col. (3) the number of grids surveyed; col. (4) the observed values of net man-hours per grid required for miscellaneous work, and col. (5) the corresponding graduated values. Differences between observed and graduated values are given in actual man-hours in col. (6), and as percentages of the graduated values in col. (7).

TABLE 33. TIME FOR MISCELLANEOUS WORK IN MAN-HOURS PER GRID

district	density of grids ($=y$)	number of grids	net man-hours per grid		difference (observed-graduated)	
			observed	graduated	actual	percentage
(1)	(2)	(3)	(4)	(5)	(6)	(7)
			$t_4 = 0.9976 - 0.2788(y)$			
Nadia	0.548	1495	0.96	0.84	+0.12	+14.28
Jessore	0.759	2511	0.71	0.79	−0.08	−10.13
Rajshahi	0.920	2050	0.72	0.74	−0.02	− 2.70
Tipperah	1.306	6193	0.69	0.63	+0.06	+ 9.52
Rangpur	1.314	2746	0.59	0.63	−0.04	− 6.35
Mymensingh	1.464	16280	0.55	0.59	−0.04	− 6.78
Dacca	1.513	6470	0.65	0.58	+0.07	+12.07

99. A slightly better fit can be obtained by using the time required for miscellaneous work as a percentage of the total time for field work. The graduation of this percentage time by the following linear equation is shown in Table 34 :

$$t_4 = 22.6872 - 2.4058(y). \qquad \ldots \quad (4.1)$$

The arrangement of this table is exactly similar to that of Table 33. By comparing col. (7) of the Tables 33 and 34 it will be seen that the percentage time gives slightly closer fit. In view, however, of its simplicity we have used the linear fit in our calculations.

TABLE 34. TIME FOR MISCELLANEOUS WORK AS PERCENTAGE OF TOTAL TIME FOR FIELD WORK

district	density of grids ($=y$)	number of grids	percentage time		difference (obsrved-graduatd)	
			observed	graduated	actual	percentage
(1)	(2)	(3)	(4)	(5)	(6)	(7)
			$t'_4=22.6872-2.4058(y)$			
Nadia	0.548	1495	21.51	21.37	+0.14	+ 0.66
Jessore	0.759	2511	19.09	20.86	−1.77	− 8.48
Rajshahi	0.920	2050	20.44	20.47	−0.03	− 0.15
Tipperah	1.306	6193	21.21	19.55	+1.66	+ 8.49
Rangpur	1.314	2746	22.50	19.53	+2.97	+15.21
Mymensingh	1.464	16280	18.48	19.16	−0.68	− 3.55
Dacca	1.513	6470	18.57	19.05	−0.48	− 2.52

100. I may mention here that exponential curves were also tried; for miscellaneous work in man-hours per grid the following equation gives fairly good results :

$$t_4 = 0.6916(y)^{-0.3415}. \quad \ldots \quad (4.2)$$

For the percentage time the exponential equation was

$$t_4 = 20.26(y)^{0.0563}. \quad \ldots \quad (4.3)$$

In the actual range of density with which we are concerned the simple linear equation gives sufficiently reliable results.

ANALYSIS OF TIME-RECORDS FOR SPECIAL FIELD UNITS

101. In order to gain some idea regarding the most economical arrangement with widely scattered sample-units we had employed special cycle-units in Rangpur District for working with 16-acre grids with a density of 0.215 (i.e., 1 in 8 square miles) in each half-sample, and with *mauzas* picked at random with a density of 0.14 or a little less than one *mauza* in 7 square miles. We had also arranged for an ordinary field unit to work with *mauzas* selected at random with a density of 0.023 or 1 in 40 square miles in each half-sample.

102. Information regarding the time of operation of these special units is given in Tables 35 and 36. In the first table, col. (1) shows the nature of the special field-units; col. (2) the average size of sample-units in acre; col. (3) the total number of sample-units; col. (4) the total area covered; col. (5) the density or number of sample-units per square mile in half-samples. The area actually surveyed is shown as a percentage of the mauza in col. (6), and as a percentage of the total area in col. (7). It will be noticed that in the case of the 16-acre grids the proportion sampled was about 1 in 40 while in the case of the *mauza*-unit the proportion of the sample was about 1 in 8; and in the case of the non-cycle field-unit 1 in 16 on a *mauza* basis and 1 in 13 on area basis.

TABLE 35. FIELD WORK BY SPECIAL UNITS

nature of the special unit	average size of unit in acre	number of units	total area in sq. miles	density (number of units per sq. mile) in half-samples	area surveyed as	
					percentage of *mauza*	percentage of area
(1)	(2)	(3)	(4)	(5)	(6)	(7)
cycle : 16-acre	16	100	405	0.125+0.125	—	0.0024
cycle : *mauza*	509	70	488	0.14 (one sample)	0.12	0.1140
non-cycle : *mauza*	1075	57	1246	0.023+0.023	0.0601	0.0768

TABLE 36. ANALYSIS OF TIME FOR SPECIAL UNITS

name of components	cycle unit				non-cycle unit	
	16-acre grids		*mauzas* (509 acre)		*mauzas* (1075 acres)	
	no. of grids	man-hours	no. of *mauzas*	man-hours	no. of *mauzas*	man-hours
(1)	(2.1)	(2.2)	(3.1)	(3.2)	(4.1)	(4.2)
per sample-unit (16-acre or mauza)						
small journey	100	0.99	70	5.29	57	10.07
big journey	100	2.77	70	3.71	57	8.97
enumeration	100	2.67	70	20.14	57	26.04
miscellaneous	100	0.37	70	3.79	57	9.19
net total	100	6.80	70	32.93	57	54.49
indirect	100	12.64	70	67.50	57	110.50
gross total	100	19.44	70	100.43	57	164.99
per square mile						
small journey	100	0.25	70	0.74	57	0.46
big Journey	100	0.69	70	0.52	57	0.41
enumeration	100	0.67	70	2.82	57	1.19
miscellaneous	100	0.09	70	0.53	57	0.42
net total	100	1.70	70	4.61	57	2.49
indirect	100	3.16	70	9.45	57	5.05
gross total	100	4.86	70	14.06	57	7.54
miscellaneous (as percentage of total time)			70	11.36	57	17.06
big journey : time per linear mile			37*	0.33	51*	0.42
total man-hours		1944		7030		9404
total man-days		81		293		392

*37 and 51 are the total number of big journeys

10.3. Table 36 shows the analysis of time for special field-units; col. (1) gives the names of components; col. (2.1) the number of grids, col. (2.2) man-hours per grid for 16-acre cycle-unit; col. (3.1) the number of mauzas and col. (3.2) man-hours for the non-cycle mauza sample-units. Information is given both per sample-unit as well as per square mile. Time for miscellaneous work as a percentage of the total time is shown in the case of two mauza-units for facility of comparison; the time for linear mile for big journey has been also given. Finally the total man-hours and man-days are given at the bottom of the table for convenience of reference.

RELATIVE SHARE OF DIFFERENT COMPONENTS

104. We may now consider the total cost function for field operations as a whole excluding only the work done by the special field-units. The share of each component is shown in Table 37; col. (1) gives the name of the district; col. (2) the density of grids; col. (3) the time taken for enumeration as a percentage of the total time taken for field operations; col. (4) percentage time for small journey; col. (5) percentage time for big journey; and col. (6) percentage time for miscellaneous work.

TABLE 37. FIELD OPERATIONS : COMPONENTS AS PERCENTAGE OF TOTAL TIME

district	density of grids	enumeration	small journey	big journey	miscel-laneous	total	indirect
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nadia	0.548	22.9	31.6	24.2	21.3	100	74.3
Jessore	0.759	26.5	33.8	20.5	19.2	100	72.4
Rajshahi	0.920	28.0	26.3	25.1	20.6	100	62.7
Bogra	1.224	37.0	28.5	16.5	18.0	100	62.3
Tipperah	1.306	29.7	25.9	23.1	21.3	100	63.8
Rangpur	1.314	27.5	29.2	21.2	22.3	100	65.0
Mymensingh	1.464	33.8	24.4	23.4	18.4	100	64.2
Dacca	1.513	33.6	25.8	21.9	18.7	100	61.9
8 districts	1.120	31.6	26.6	22.2	19.4	100	65.1

Finally col. (8) gives time spent for indirect purposes (such as, sleep, taking food etc.) as percentage of the total time (or 24 hours). Omitting the indirect time, the share of each component is quite appreciable. Taking the average of all eight districts we find that enumeration is responsible for 31.6 per cent or a little less than one-third of the total time; small journey for 26.6 per cent or a little more than a quarter; big journey for 22.2 per cent or a little over one-fifth; and miscellaneous work for 19.4 per cent or a little less than one-fifth of the total time.

TIME OF FIELD OPERATIONS PER SQUARE MILE

105. We have discussed so far, with a few exceptions, the time per sample unit. It is now easy to convert these figures into net man-hours per square mile by simply multiplying by the respective densities or number of sample-units per

squar mile. The equations are given in Table 38 for convenience of reference. In these equations x is the size in acre of the sample-units, and y the density or number of sample units per square mile; and in equation (3.2), z is the length of the big journey (i.e., the distance between camps) in miles.

TABLE 38. EQUATIONS FOR COMPONENTS OF THE COST FUNCTION IN NET MAN-HOURS PER SQUARE MILE

x = size in acre of sample-units;
y = density (or number) of sample-units per square mile in each half-sample;
z = length of big journey (distance between camps) in miles in equation (3.2)

enumeration	$t_1 = c_0 \cdot y + e_1 \cdot xy = 0.7589y + 0.0600xy$	...	(1)*
small journey	$t_2 = j_0 \cdot y + j_1 \cdot xy = 1.5961y - 0.5673y^2$	...	(2)*
	$t_2 = a(y)^{1-k} = 0.9752(y)^{0.3920}$	...	(2.1)
big journey	$t_3 = j_2/2 = 0.8000$	...	(3)*
	$t_3 = a(y)^{1-k} = 0.7899(y)^{0.6880}$	...	(3.1)
	$t_3 = a(z)^k = 1.334(z)^{-0.3869}$	...	(3.2)
miscellaneous	$t_4 = m_0 \cdot y - m_1 \cdot y^2 = 0.9976y - 0.2788y^2$	...	(4)*
	$t_4 = a(y)^{1-k} = 0.6916(y)^{0.5865}$	...	(4.1)
percentage miscellaneous	$t_4' = e' - f' \cdot y = 22.6872 - 2.4068(y)$	...	(4.2)
	$t_4' = a(y)^{1-k} = 20.26(y)^{0.0583}$	...	(4.3)
statistical work	$t_5 = \frac{1}{2}' + l_0 \cdot y + l_0 \cdot xy = 0.1695 + 0.4414y + 0.669xy$	...	(5)*

*These equations have been used in the first form of the total cost function.

106. We have used linear equations (1) for enumeration and (2) for small journey, and we have taken the number of man-hours per square mile to be sensibly constant for big journey as shown in equation (3). Those equations which have been used for constructing the first form of the cost function have been marked by stars (*) in Table 38. Equations (2.1) and (4.2), on the other hand, have been used in the place of equations (2) and (4) for building up a second alternative form of the cost function. Equations (3.1), (3.2), (4.1) and (4.3) have been quoted here simply for convenience of reference for future investigations.

CONVERSION OF NET WORKING HOURS INTO RUPEES PER SQUARE MILE

107. It is now necessary to convert the cost in terms of net man-hours per square mile into rupees per square mile. For this we have to consider the actual expenditure. I am giving below in abstract form the expenditure incurred by the Field Branch in 1940 based on material obtained from the office of the Secretary, Indian Central Jute Committee.

A. Non-recurring			
(1) Price of maps	...	Rs. 11,660–0–0	
(2) Cost of mounting maps		4,715–4–9	Rs. 16,375–4–9
(3) Preliminary Field Work (1939–40)		...	7,800–0–0
Sub-total of Non-recurring expenditure :			Rs. 24,175–4–9
B. Overhead			
(4) Supervisor's salary	...	Rs. 2,086–3–0	
(5) Asst. Supervisor's salary		1,103–0–0	
(6) Supervisor's T. A.	...	1,023–3–0	
(7) Asst. Supervisor's T. A.	...	500–0–0	
(8) Supervisor's leave salary		295–9–0	
(9) Special clerk (I.C.J.C.)	...	191–2–0	
(10) Establishment (Field Office)	...	1,541–11–0	Rs. 6,743–12–0 (6 months)
C. Field Work			
(11) Salary of field staff	...	Rs. 35,432–11–3	
(12) T.A. for field staff	...	4,767–15–9	
(13) Sub-total of (11) and (12)	...	40,200–11–0	
(14) Contingency (Field Branch)	...	4,828– 7–9	
(15) Total Field Work	...	Rs. 45,029– 2–9	Rs. 45,029– 2–9
Total for sample Census	...	...	Rs. 75,948– 3–6
D. Crop-cutting (field work)	...	...	3,120– 5–6
Grand Total			Rs. 79,068– 9–0

108. It will be seen from the above statement that the cash expenditure for maps (item nos. (1) and (2)) was Rs. 16,375–4–9; and for (3) Preliminary Field Work (which will not be repeated this year) Rs. 7,800. This total non-recurring expenditure (A) of Rs. 24,175-4-9 need not be further considered in the present connexion.

109. The overhead (B) consists of items (4) Supervisor's salary; (5) Asst. Supervisor's salary; (6) Supervisor's leave salary; (9) salary of one special clerk employed in the office of the Indian Central Jute Committee for checking account; and (10) salary of staff in the office of the Field Branch. The total expenditure for six months from April to September 1940 was Rs. 6,743-12-0, or about Rs. 1,125 per month. Provision will have to be made on this basis in the budget for 1941; but this expenditure will be practically independent of the cost of actual field work.

110. The expenditure incurred for (C) Field Work is given in items (11)-(15). The salary and T.A. of the field staff came to Rs. 40,200-11-0; adding Rs. 4,828-7-9 for contingencies, the total expenditure was Rs. 45,059-2-9. This is the expenditure which depends on the actual volume of field work. (I may explain here that a sum of Rs. 4,000 provided for area extraction was transferred to the Statistical Section and has not been included in the above statement).

111. In order to convert net working hours into money values we are thus concerned with the total expenditure of Rs. 45,029–2–9 incurred directly for field work because our calculations of man-hours are based on the actual work done by the field staff. Table 39 gives the total number of pay-days and net working-days by districts and field units. In this table col. (1) gives the serial number of the block; col. (2) the name of the district, and col. (3) the serial number of the field unit. Actual dates of starting and completing the field work for each field unit are shown in cols. (4) and (5). The total number of pay-days, that is, the number of days for which the field staff was paid (the pay of one man for one day being counted as one pay-day) are shown in cols. (6), (7), (8) and (9), for Chief Inspectors, Check Inspectors, Inspectors and Investigators respectively. The net working days, that is, the actual number of days spent on the field in survey work as recorded in the field diaries are given for investigators in col. (10) with district totals in col. (11). The total number of investigators as shown in the diaries is given in col, (12). The total number of investigators who had actually worked in 1940 was 287 against a standard strength of 232; the excess is explained by the fact that some of the workers were transferred from one unit to another and were counted more than once, while a number of new workers had to be appointed to fill vacancies caused by resignation or leave.

112. The total number of pay-days given in col. (9) is invariably greater than the total number of net working days shown in col. (10) because a portion of the time was consumed in joining the post, receiving instructions, preliminary training etc. The total number of working days on which our cost analysis is based is 17,291; it is this total of 17,291 working days which has been broken down into component parts for which the actual expenditure incurred was Rs. 45,029-2-9. This gives us an effective rate of expenditure of Rs. 2,604 or Rs. 2-9-7 per net working day. Remembering that 65 per cent or about 15.6 hours per day was spent for indirect purposes such as sleep, taking food etc. and only about 8.4 hours per day were available for direct productive work we easily find that the actual cost of field work was almost exactly *annas* five per hour which is the rate at which we can now convert man-hours into cash values.

113. It will be noticed that the expenditure of Rs. 45,029-2-9 covered 19,128 pay-days although our rate is calculated on the basis of 17,291 net working days. This means that a balance of 1827 or a little more than 10 per cent is thus kept in reserve for joining time, receiving instructions and training etc. The time lost on account of sickness, leave, cessation of work owing to unforeseen reasons like drought, excessive rain etc., has of course been already automatically included in the deduction

TABLE 39. TOTAL NUMBER OF PAY-DAYS AND NET WORKING-DAYS BY DISTRICTS AND FIELD UNITS*

serial number of blocks	name of district	serial number of field units	date of		total number of pay-days				net working days		total number of investigators
			starting field work	completing field work	chief inspectors	check inspectors	inspectors	investigators	investigators	district total	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
I	Nadia	1	May 1	August 4	—	—	77	624	600		
	,,	2	May 1	July 11	—	—	83	628	592	1192	18
	Jessore	3	May 8	August 4	—	—	101	762	730		
	,,	4	May 1	August 4	113	—	101	771	708	1438	9
II	Rajshahi	5	May 2	July 26	—	—	87	508	414		
	,,	6	May 1	July 12	—	—	92	444	422	836	15
	Bogra	7	May 1	July 26	—	—	86	654	561		
	,,	8	May 3	July 27	—	—	85	632	617	1178	21
	Rangpur	9	May 2	July 26	—	—	81	474	426		
	,,	10	May 2	July 24	138	73	93	576	542	968	14
III	Mymensingh	11	May 1	July 14	—	—	86	631	594		
	,,	12	May 3	July 15	—	—	88	652	587		
	,,	13	May 1	July 12	—	—	82	560	527		
	,,	14	May 1	June 20	—	—	70	400	366		
	,,	15	May 1	June 28	—	—	68	512	446		
	,,	16	May 1	July 15	101	62	88	637	580		
IV	,,	17	May 8	July 20	—	—	90	686	588		
	,,	18	May 8	July 20	—	—	90	671	600		

* Out of two tables only one table is printed here as a specimen.

of 65 per cent for indirect purposes. The rate of conversion of five *annas* per net working hour thus makes full allowances for training, receiving instructions as well as for illness, unforeseen stoppage of work etc.; and may be safely used in making our calculations. I may also mention that the average rate of Rs. 2-9-7 per working day in 1940 compares quite favourably with the average rate of Rs. 2-14-9 in 1939 and Rs. 2-11-9 in 1938.[1]

Time required for statistical work

114. The statistical work may be divided into a number of broad groups. First of all we have a certain amount of preliminary work such as receiving and arranging *mauza* maps; preparing sheet lists; distributing sample-units in each half-sample in each *thana* etc. This time is idependent of the size or density of grids, but is determined only by the number of maps handled, that is, by the total area surveyed. From time records kept in the laboratory we found that on an average 0·1997 net man-hours per square mile was required for arranging and handling maps; and 0.1319 man-hours for distribution of samples and allocation of *mauzas* to half-samples giving a total requirement of 0.3388 man-hours per square mile.

115. Next there is another portion of work which depends purely on the total number of sample-units and is independent of the size of grids. A list of such items is shown in tabular form in Table 40 in which the figures are given in net man-hours per 100 sample-units. This non-variable work consists of (1) locating and marking sample-units (2.17 man-hours); (2) receiving the *khasra* list as they come back from the field and classifying and tabulating them (1.81 man-hours); (3) calculating averages of proportion under jute with standard errors (3.19 man-hours); (4) comparing results for half-samples and routine calculations for variance function (2.76 man-hours); (5) classification and primary tabulation relating to analysis of time records (5.83 man-hours); (6) tabulation of time records for statistical work and cost analysis (8.11 man-hours). The total requirement is thus 24.87 man-hours per sample-unit irrespective of the size of the sample-units.

116. We then have another portion of laboratory work which depends both on the size and number of sample-units; details are given in Table 40. This consists of (1) preparing lists of plots included within each sample-unit and checking these lists; (2) preparing *khasra* list for use by the Field Branch; (3) area extraction, or measurement of the area of individual plots under jute; and (4) conversion of *anna* estimates of the proportion of land under jute into acreages and preparing totals of grid figures. In all these operations the amount of work depends primarily on the total number of plots, and therefore increases appreciably as the size of sample-units is increased. It was found that the following simple linear equations gave a reasonable fit for this portion of the work.

$$t_5 = 44.14 + 6.69(x)$$

where t_5 is the time required for statistical work in man-hours per 100 sample-units of size x-acre.

[1] *Report on the Sample Census of Jute in* 1939, paragraphs 73 and 74 (Proceedings, Jute Census Committee, 13th December 1939, p. 32).

TABLE 40. TIME REQUIREMENTS FOR STATISTICAL WORK

(A) *preliminary work in man-hours per square mile*			
(1)	arranging maps ; preparing sheet lists etc.	0.1997	man-hours
(2)	distribution of grids by half-samples	0.1391	„
		0.3388	„
(B) *work independent of size of grids in man-hours per 100 grids*		-	
(1)	locating and marking sample-units	2.17	man-hours
(2)	classification and tabulation of *khasras*	1.81	„
(3)	calculating average and standard errors	3.19	„
(4)	comparing half-samples ; variance calculations	2.76	„
(5)	tabulation and analysis of time-records	5.83	„
(6)	cost analysis	8.11	„
		24.87	„

(C) *work depending on size of grids in man-hours per 100 grids*

	name of item	1-acre	2.25-acre	4-acre	6.25-acre	9-acre	16-acre
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1)	preparing list of plots with checking	10.56	11.52	19.35	26.44	28.92	42.03
(2)	preparing *khasra* for field branch	3.74	7.06	7.39	9.42	11.22	13.06
(3)	area extraction (jute plots only)	7.15	12.63	15.23	22.92	26.72	74.05
(4)	*anna* conversion and grid—totalling	2.21	3.83	4.14	6.87	8.17	11.29
(5)	total variable portion	23.46	35.04	46.11	65.65	75.03	140.43
(6)	plus non-variable portion	24.87	24.87	24.87	24.87	24.87	24.87
(7)	total man-hours per 100 grids	48.33	59.91	70.98	90.52	99.90	165.30

117. The observed and graduated values are shown in Table 41 in which col. (1) gives x, the size of grids in acre; col. (2) the number of grids surveyed; col. (3) the time required for the variable portion of statistical work in man-hours per 100 grids; and col (4) the corresponding graduated values. The difference between the observed and graduated value is given in col. (5); and the percentage difference in col. (6). It will be noticed that the percentage differences are usually low and less than 5 per cent which shows that the graduation is quite satisfactory and may be used in building up the cost function. The total cost function for laboratory portion of the work for each half-sample may therefore be written in the following form :

$$t = \text{net man-hours per sample-unit}$$
$$= l_0 + l_1(x) + 1/2y = 0.4414 + 0.0669(x) + 0.1695/2y$$

where x is the size of the sample-unit in acre, and y is the number of grids per square mile in each half-sample. Multiplying this by y the density of sample-units per square mile we immediately get the time required for statistical work in man-hours per square mile,

$$t_5 = 0.1695 + 0.4414y + 0.0669xy \qquad \ldots \quad (5)$$

for each half-sample.

TABLE 41. GRADUATION OF TIME REQUIRED FOR STATISTICAL WORK IN MAN-HOURS PER 100 GRIDS

size of grids (x)	number of grids	time for statistical work		difference	percentage
		observed	graduated	observed-graduated	difference
(1)	(2)	(3)	(4)	(5)	(6)
		$t_5 = 44.14 + 6.69(x)$			
1-acre	1766	48.33	50.83	−2.50	−4.9
2.25-acre	2587	59.91	59.28	+0.71	+1.2
4-acre	32996	70.98	70.91	+0.07	+0.1
6.25-acre	1804	90.52	85.96	+4.56	+5.3
9-acre	2192	99.90	104.37	−4.47	−4.3
16-acre	100	165.50	151.21	+14.09	+9.3

118. I may note here that the above equation does not include certain expenses which were incurred in 1940 but which may not be necessary in future. In 1939 the *mauza* maps were sent to the Statistical Laboratory without any properly prepared *chalan* and without the jurisdiction lists. This necessitated a great deal of work in arranging the maps in serial order, preparing lists of sheets and comparing them with the maps actually received, preparing list of transfers of *mauzas* between different police stations etc. As jurisdiction lists were not supplied to us relevant entries had to be copied from the office of the Director of Land Records which also caused a good deal of additional but unnecessary expenditure. From our records we find that the total time consumed in this work was nearly 2 man-hours per 100 square miles. I am omitting this portion of the expenditure in the hope that it will be avoided in future.

119. It should be noticed that the time given here for area extraction, that is, for the measurement of area of individual plots refer to only those plots which were actually reported to be under jute and which amounted to roughly one-third of the total number of plots. In the preliminary report submitted on the 26th August 1940 I had recommended that the area of all plots included within the sample-units should be measured in 1940-41 so as to expedite the preparation of final values of the area under jute; and this proposal was approved by the Jute Census Committee on the 5th September 1940. I have not included the additional amount of money which will be required for this purpose in the cost function but have given a separate estimate.

120. I may also mention that one important portion of the statistical work has not been included in the above equation. This the work required for the final analysis and the writing of the report. It is obvious that the time records have to be closed at some stage or other in order to start the more advanced analysis; and the time required for such advanced analysis cannot possibly be included in the analysis itself.

This year a number of statistical workers with roughly a staff of thirty computers have been working for more than one month on this final analysis. Some provision will have to be made for this purpose; but as this is essentially an overhead expense it need not be included in the cost equation.

121. Subject to these reservations we may use the above equation (5) for constructing the total cost function. But before doing so we have to convert the man-hours in the Laboratory into money values. As the work was done jointly by the entire staff of the Statistical Laboratory, an entirely independent estimate (as in the case of the Field Branch) is not possible. From extensive records in the Laboratory we know that on an average a pay-month consists of 144 net indexable working hours. This refers only to that portion of the work which is directly taken into consideration in the cost analysis; and is exclusive of all indirect items like receiving instructions, entering records, making arrangements for work, tiffin etc., and also for natural rest periods and pauses to overcome fatigue effects. We have used Rs. 3 per pay-day for Laboratory work as the standard rate of conversion against Rs. 2-9-7 for field work; this difference is due to the higher pay which has to be given in many cases to trained statisticians and computers and also partly due to expenditure for calculating machines and other appliances. I may mention that the average rate per pay-day calculated for all work done in the Laboratory is about Rs. 3-8-0 per pay-day; this however includes charges on account of purchase of books, heavy expenses for preparation of numerical tables and other expensive items sometime required for theoretical researches. I have therefore adopted Rs. 3 per pay-day as a reasonable estimate for the rate of conversion for work done in the Jute Census Scheme.

THE TOTAL COST FUNCTION

122. We are now in a position to write down the total cost function in rupees per square mile. The numerical values of the various constants have been given in Table 42; it will do however if I quote here only the final equations. The cost in rupees per square mile for field work for each half-sample is given by

$$\phi_f = .2500 + 1.0478 y_i + .0188 x_i y_i - .2645 y_i^2 \quad \ldots \quad (1)$$

where x_i is the size of sample-unit in acre, and y_i is the number of sample-units per square mile in each half-sample.

The cost in rupees per square mile for statistical work is given by

$$\phi_i = .1059 + .2759 y_i + .0418 x_i y_i. \quad \ldots \quad (2)$$

Adding the above two equations we get the total cost in rupees per square mile :

$$\phi_i = .3559 + 1.3237 y_i + .0606 x_i y_i - .2645 y_i^2. \quad \ldots \quad (3)$$

The total cost for the whole area is calculated by multiplying the area of each zone (A_i) by corresponding values of ϕ_i after substituting appropriate values of the size x_i and density y_i of sample-units.

123. The variance of the final estimate may be written in the following form:

$$V = \Sigma \frac{A_i p_i q_i \alpha}{y_i(x_i) g} \quad \ldots (4)$$

where p_i is the proportion of land under jute in the i-th zone; $q_i = 1-p_i$, and A_i, as usual, is the area of the i-th zone in sq. miles. Using analytic methods which have been explained elsewhere[1] it is possible to write the optimum values of the different variables when both the size and density of sample-units are selected in such a way as to give the lowest margin of error at any given value of the total expenditure. Relevant formulae are given below.

$$T(f) = \text{total cost of field work in rupees}$$
$$= \Sigma A_i(.2500+1.1778y_i-.3167y_i^2) \quad \ldots (5)$$
$$T(1) = \text{total cost of statistical work in rupees}$$
$$= \Sigma A_i(.1059+.5649y_i-.1152y_i^2) \quad \ldots (6)$$
$$T = \text{total cost in rupees} = T(f)+T(1)$$
$$= \Sigma A_i(.3559+1.7427y_i-.4319y_i^2). \quad \ldots (7)$$

This gives the cost per square mile as soon as the best density of sample-units is settled. We also have the following relation between the best size (x_i) and best density (y_i)

$$X_i = 6.912-2.762y_i. \quad \ldots (8)$$

124. In order to settle the particular values of (y_i) for different zones it is necessary to use an auxiliary parameter μ_i for which the equation is given below

$$\mu_i = y_i^2(x_i)^{g+1} = y_i^2\{6.912-2.762y_i\}^{1.2404} \quad \ldots (9.1)$$

This gives the relation between μ_i and the optimum values of x_i and y_i. The particular values of μ_i have to be settled from another equation given below

$$\mu_i = \frac{1.62p_iq_i}{\lambda}. \quad \ldots (9.2)$$

125. The procedure is now quite straightforward. We start with a suitable number of zones (which may be districts or portions of districts or groups of thanas etc.), and for each zone we have a definite value of p_i which is known, and $q_i = 1-(p_i)$ is also known for each zone. Suppose we have k zones; then if we select any trial value of the Lagrangian multiplier λ we shall immediately get k values of μ_i from equation (9.2). Corresponding to each value of μ_i we can then find the optimum values of x_i and y_i from equation (9.1). As soon as the optimum values of y_i is known, the cost per square mile is also known from equations (1), (2) and (3). Multiplying together A_i, the area in square miles belonging to each zone and the cost per square mile corresponding to the appropriate value y_i we can then get a value of T the total cost in rupees from equation (3).

126. In this way for each trial value of λ we shall get a corresponding value of T; and we can draw a graph showing the relation between T and λ. Once this curve is known it is possible to read off the value of λ corresponding to any assigned value of

[1] P. C. Mahalanobis (1940) : Sample survey of the acreage under jute in Bengal, *Sankhyā*, Vol. 4, Part 4, p. 520.

T, the total expenditure proposed to be incurred for the sample census. Using this particular value of λ it is then possible to find optimum values of both x_i and y_i that is, the best size and density of sample-units for each individual zone. Finally, it is also possible to divide the total cost into two portions for field and statistical work respectively with the help of equations (6) and (7).

127. In actual practice it is convenient to draw curves of the most economical cost per square mile (ϕ_i) the best size of sample-units (x_i) and the best density of sample-units (y_i) against different values of μ. Using a few trial values of λ it is then possible to reach the neighbourhood of values of T, that is, of the proposed total expenditure of the sample survey (exclusive of supervision and non-recurring expenses).

TABLE 42. THE FIRST FORM OF THE TOTAL COST FUNCTION

A_i = area in square miles;
p_i = proportion of land under jute;
x_i = size of sample-units in acre; $q_i = (1 - p_i)$.
y_i = density (or number) of sample-units per sq. mile

cost of field work in rupees per sq. mile $= \phi(f) = F_0+F_1y_i-F_2y_i^2+F_3x_iy_i$ ... (1)

cost of statistical work in rupees per sq. mile $= \phi(l) = L_0+L_1'y_i-L_2'y_i^2+L_3'x_iy_i$ ... (2)

total cost in rupees per sq. mile $= \phi = C_0+C_1y_i-C_2y^2+C_3x_iy_i$ ... (3)

$$\text{variance} = v_x = \psi(p_i\, q_i\, x_i) = \frac{\alpha p_i q_i}{(x_i)g},\ v = \Sigma \frac{A_i \alpha p_i q_i}{y_i (x_i) g} \quad \ldots (4)$$

$F_0 = j_2/2 = 0.2500$	$L_0 = l/2 = 0.1059$	$C_0 = F + L_0 = 0.3559$
$F_1 = e_0+j_0+m_0 = 1.0478$	$L_1 = l_0 = 0.2759$	$C_1 = F_1+L_1 = 1.3237$
$F_2 = j_1 + m_1 = 0.2645$	$L_2 = 0$	$C_2 = F_2 = 0.2645$
$F_3 = e_1 = 0.0188$	$L_3 = l_1\ 0.0418$	$C_3 = F_3+L_3 = 0.0606$
$\alpha = 0.4084$	$g = 0.2404$	$(1-g) = 0.7596$

for minimum V at given level of total expenditure we have

$T(f)$ = total cost of field work

$$= \Sigma A_i \left\{F_0+y_i\left(F_1+\frac{gF_2}{1-g}, \frac{C_1}{C_3}\right)-y_i^2\left(F_2+\frac{2g}{1-g}, \frac{F_2F_3}{F_3-L_3}\right)\right\}$$

$$= \Sigma A_i\,(0.2500+1.1778\ y_i-0.3167\ y_i^2) \quad \ldots (5)$$

$T(l)$ = total cost of statistical work

$$= \Sigma A\left\{L_0+y_i\left(L_1+\frac{gl_3}{1-g}.\frac{C_1}{C_3}\right)-y_i^2\left(\frac{g}{1-g}.\frac{2F_2L_3}{F_3+L_3}\right)\right\}$$

$$= \Sigma A_i\,(0.0559+0.5649\ y_i-0.1152y_i^2) \quad \ldots (6)$$

T= total cost in rupees

$$= T(f)+T(l) = \Sigma A_i\left\{C_0+\frac{C_1}{1-g}, y_i-C_2\left(\frac{1+g}{1-g}\right), y_i^2\right\}$$

$$= \Sigma A_i\,(0.3599+1.7426\ y_i-0.4352y_i^2) \quad \ldots (7)$$

x_i = best size of sample-units in i-th zone

$$= \frac{g}{1-g}\cdot\left(\frac{C_1-2C_2y_i}{C_3}\right)$$

$$= 6.912-2.762\ y_i \quad \ldots (8)$$

μ_i= auxiliary parameter$=y_i^2(x_i)^{1+g}$ ... (9.1)

$$= \frac{g\alpha p_i q_i}{C_3\lambda} = \frac{1.62 p_i q_i}{\lambda} \quad \ldots (9.2)$$

128. I may mention here that before using linear graduations for small journey and miscellaneous work a great deal of extremely laborious calculations were done with exponential curves. Equations relating to the second form of the cost func-

tion are given in Table 43. The first portion shows the graduation for the four components of the cost for field work. In the case of "enumeration" and "big journey" the equations are identical with those used in the first form of the cost function given in Table 42. In the case of "small journey" an exponential curve has been used in the second form; and in the case of "miscellaneous" work it has been changed into a linear curve with the miscellaneous time expressed as a proportion of the total time for field work. In Table 43 the name of the component is given in col. (1); the algebraic form of the equation in col. (2); numerical values in man-hours per square mile in col. (3); and finally numerical values in rupees per square mile in col. (4).

TABLE 43. SECOND FORM OF THE COST FUNCTION

components of field cost	mean hours per sq. mile	rupees per sq. mile
(1)	(2)	(3)
enumeration $t_1 = y(e_0+e_1x)$	$y(0.7589+0.0600x)$	$y(0.2372+0.0188x)$
small journey $t_2 = yay^k = ay^{k+1}$	$0.9752y^{0.3920}$	$0.3048y^{0.3920}$
big journey $t_3 = j/2$	0.8000	0.2500
miscellaneous proportion $t_4' = t(e'+f'y)$	$t(0.2269-0.0241y)$	$t(0.2269-0.0241y)$

Total cost of field work per sq. mile $= \dfrac{ay^{k+1}+j/2+y(e_0+e_1x)}{1-e'-f'y}$ ... (1)

$$= \phi(f) = \frac{0.3048y^{0.3920}+0.2500+y(0.2372+0.0188x)}{0.7731+0.0241y} \quad \ldots (1.1)$$

Cost of statistical work per sq. mile $= L_0+L_1y+L_3xy$ ... (2)

$$= \phi(l)=0.1059+0.2759y+0.0418xy \quad \ldots (2.1)$$

Total cost in rupees per square mile for each half-sample $= \phi(f)+\phi(l)$

$$= \phi = \Sigma A_i \left\{ \frac{ay^{k+1}+j/2+y(e_0+e_1x)}{1-e'-f'y} + L_0+L_1y+L_3xy \right\} \quad \ldots (3)$$

$$= \Sigma A_i \left\{ \frac{0.3048y^{0.3920}+0.2500+y(0.2372+0.0188x)}{0.7731+0.0241} + 0.1059+0.2759y+0.0418xy \right\} \quad \ldots (3.1)$$

relation between optimum values of x and y:

$$x_i = \frac{2a(K+1)(1-\acute{e})y_i^k - 2af'y_i^{k+1} + 2e_0(1-\acute{e}) + yf' + 2L_1(1-\acute{e}-f'y^i)^2}{2L_3\left(\frac{1}{g}-1\right)(1-\acute{e}-f'y_2)^2 - 2\frac{e_1}{g}fy_i + 2(1-\acute{e})\left(\frac{1}{g}-1e_1\right)} \quad \ldots (7.1)$$

$$= \frac{0.1847y_i^{-.06080} - 0.0089y_i^{.03920} + 0.3547 + 0.4744(0.7731+0.0241y_i)^2}{0.2642(0.7731+0.0241y_i)^2 + 0.0038y_i + 0.0918} \quad \ldots (7.2)$$

$$\mu_i = \frac{p_iq_i\alpha g}{\lambda} = p_iq_i\,\frac{0.0982}{\lambda} \quad \ldots (8.1)$$

$$\mu_i = \left[\frac{e_1}{1-\acute{e}-f'y_i} + L_3\right] y_i^2\, x_i^{q+1}$$

$$= \left[\frac{0.0600}{0.7731+0.0241y_i} + 0.0669\right] y_i^2\, x_i^{21404} \quad \ldots (8.2)$$

129. The total cost of field work per square mile is given in equation (1) in algebraic form and with numerical values substituted in equation (1.1). The cost of statistical work per square mile is, of course, the same as before, but is quoted for

convenience of reference in equations (2) and (2.1). The total cost in rupees per square miles for each half-sample is given in algebraic form in equation (3) and in numerical form in equation (3.1). Finally the relation between x_i and y_i under optimum conditions, that is, when variance is minimum at any assigned level of expenditure, is given in algebraic form in equation (7.1) and in numerical form in equation (7.2). Finally the equation for the auxiliary parameter μ_i is shown in terms of zone constants in equation (8.1) and in terms of size and density of sample-units in equation (8.2).

130. We compared the two expressions for the cost function at five different values of $y = 0.1, 0.4, 0.5, 1$ and 1.5; and the two sets of results are given in Table 44 in which col. (1) gives the density; col. (2.1) values of t_1 the cost in rupees per square mile for enumeration; col. (2.2) values of t_2 the cost in rupees per square mile for small journey; col. (2.3) values of t_3 the cost in rupees per square mile for big journey; and col. (2.4) values of t_4 the cost in rupees per square mile for miscellaneous work; while the sub-total in col. (3) gives the cost in rupees per square mile for field operations; the cost in rupees per square mile for the statistical portion of the work is given in col. (4); and the total cost in rupees per square mile for both field and statistical work (but exclusive of overhead and supervision charges) in col. (5). The difference between the two estimates is shown in col. (6) in the form of a percentage.

131. It will be remembered that in 1940 we used densities ranging from 0.548 in Nadia to 1.513 in Dacca; the agreement between the two sets of values calculated by the two sets of equations is excellent for values of density lying between 0.5 and 1.5; the agreement is not unsatisfactory for a density of 0.4, but there is wide divergence at the very thin density of 0.1. So long as we do not go below 0.4 or even perhaps 0.3 it is therefore immaterial which particular set of equations is used in our calculations. In view of the great simplicity of the first set of equations based on linear graduations we decided to adopt it for our present calculations. But I am quoting the second set of equations based on exponential graduation for further investigation of the question relating to very low values of y, i.e., particularly thin densities.

TABULATED VALUES OF THE COST FUNCTION

132. Calculated values of $\phi(f)\ \phi(l)$, ϕ, $\mu_1 x$, y are shown in Table 45 in which the cost in rupees per square mile for field is given in col. (1), for statistical work in col. (2), and the total cost in rupees per square mile in col. (3). The next col.(4) gives the corresponding value of μ_i; col. (5) the best density (y_i); and col. (6) the best size of sample-units in acre x_i. These tabulated values are also shown graphically in the accompanying diagram. From these tabulated values (or from the graphs) it is easy to calculate the values of $\phi(f)$, $\phi(l)$, ϕ, x, y and for actual values μ.

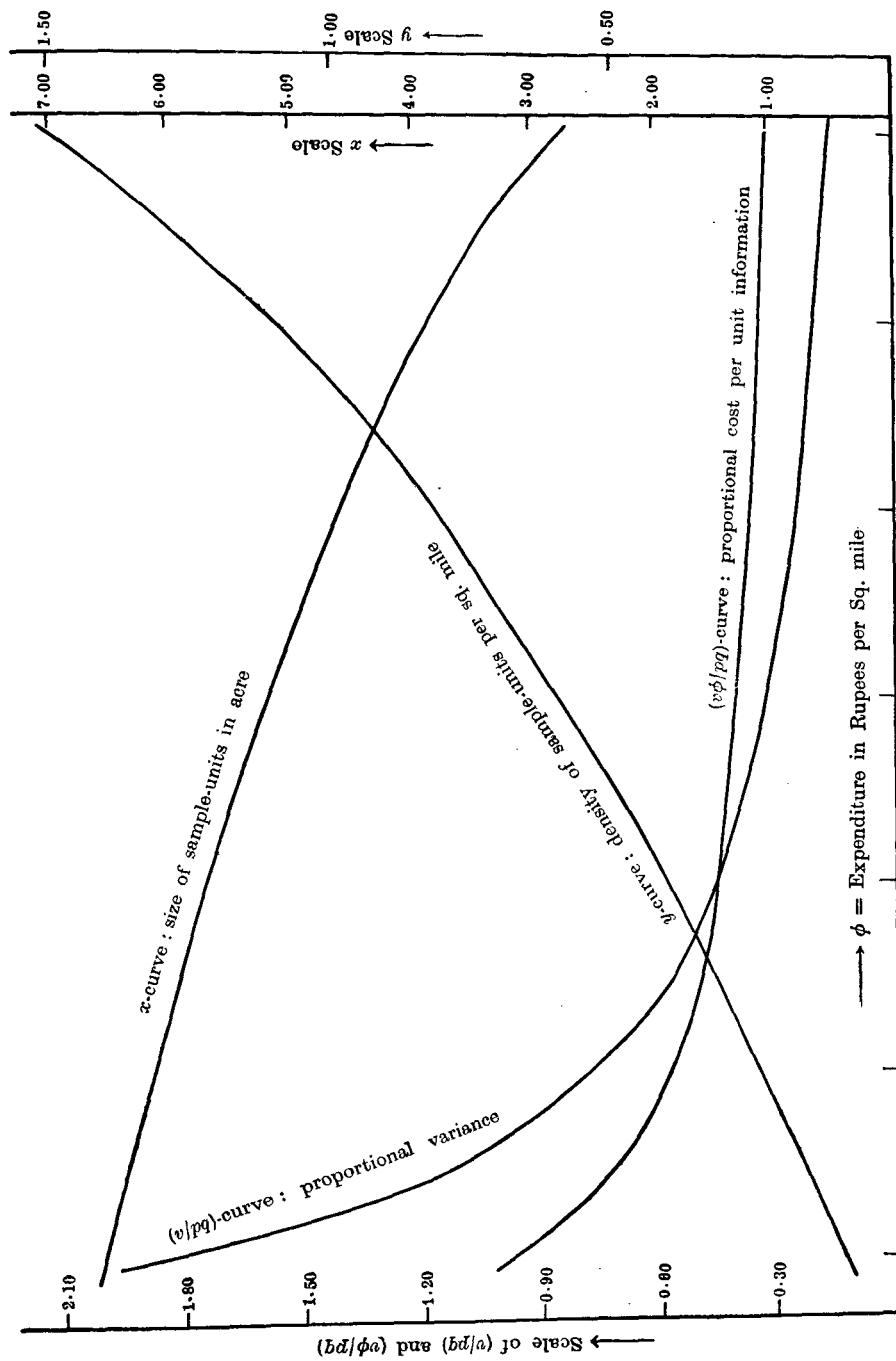
133. Finally proportional variances per square mile are given in col. (7); multiplying by appropriate values of $p_i\, q_i$ and A_i it is possible to calculate the value $V = \Sigma A_i\, p_i\, q_i(v_i)$ the variance of the final estimate. The last col. (8) gives the proportional cost in rupees per unit of information (i.e., the inverse of variance) as defined by R. A. Fisher. It will be noticed that at first the proportional rupee-rate per unit of information decreases rapidly, but soon becomes steady showing the law of diminishing returns.

TABLE 44. COMPARISON OF COST PER SQUARE MILE CALCULATED BY FIRST (ADOPTED) AND SECOND FORMS OF THE COST FUNCTION FOR SELECTED VALUES OF DENSITY OF SAMPLE-UNITS PER SQUARE MILE

number of sample-units per square mile in each half-sample	cost in rupees per square mile for two half-samples							
	field operations				sub-total for field work	statistical work	total cost	percentage difference between two sets
	enumeration	small journey	big journey	miscellaneous work				
y_i	(t_1)	(t_2)	(t_3)	(t_4)	($\phi(f)$)	$\phi(l)$	(ϕ)	
(1)	(2.1)	(2.2)	(2.3)	(2.4)	(3)	(4)	(5)	(6)
cost calculated by first (adopted) set of equations								
0.1	0.0625	0.0962	0.5000	0.0667	0.7194	0.2003	0.0197	
0.4	.2499	.3423	.5000	.2215	1.3137	.4657	1.7794	
0.5	.3124	.4101	.5000	.2682	1.4907	.5542	2.0449	
1.0	.6248	.6430	.5000	.4492	2.2170	.9966	3.2136	
1.5	.9372	.6986	.5000	.5430	2.6788	1.4390	4.1178	
cost calculated by second set of equations								
0.1	0.0625	0.472	0.5000	0.2344	1.0441	0.2003	1.2444	+26.2
0.4	.2499	.4256	.5000	.3263	1.5018	.4657	1.9875	+ 9.5
0.5	.3124	.4646	.5000	.3494	1.6264	.5542	2.1806	+ 1.6
1.0	.6248	.6096	.5000	.4412	2.1756	.9966	3.1722	− 1.3
1.5	.9372	.7146	.5000	.5070	2.6588	1.4390	4.0978	− 0.5

TABLE 45. BEST SIZE AND DENSITY OF SAMPLE-UNITS AT DIFFERENT LEVELS OF EXPENDITURE IN RUPEES PER SQUARE MILE

cost in rupees per sq. mile			auxiliary parameter	best no. of grids per sq. mile	best size of grids in acre	proportional variance per sq. mile	propotional cost per unit information
for field work $=\phi(f)$	for statistical work $=\phi(l)$	total cost $=\phi$	μ_i	$=y_i$	$=x_i$	v/pq	$v\phi/pq$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Rs. As.					
0.345	0.155	8	0.080	0.088	6.68	1.94	0.97
.385	.177	9	.165	.129	6.58	1.47	.82
.429	.197	10	.285	.168	6.48	1.14	.71
.473	.214	11	.400	.200	6.35	0.92	.64
.515	.235	12	.565	.240	6.23	.77	.58
.558	.254	13	.765	.278	6.13	.67	.54
.600	.274	14	1.000	.319	6.00	.58	.50
.642	.295	15	1.225	.362	5.90	.50	.47
.684	.316	1 0	1.475	.405	5.76	.46	.46
.725	.387	1 1	1.760	.450	5.04	.41	.43
.765	.359	1 2	2.100	.500	5.50	.38	.42
.811	.376	1 3	2.380	.538	5.38	.35	.41
.846	.403	1 4	2.800	.600	5.23	.32	.40
.887	.425	1 5	3.180	.652	5.10	.29	.38
.931	.444	1 6	3.625	.705	4.98	.27	.38
.969	.468	1 7	4.070	.760	4.82	.26	.37
1.008	.492	1 8	4.500	.820	4.65	.24	.36
1.047	.515	1 9	4.975	.885	4.48	.22	.35
1.086	.538	1 10	5.500	.950	4.29	.21	.34
1.126	.561	1 11	5.950	1.016	4.08	.20	.33
1.164	.586	1 12	6.450	1.094	3.89	.18	.32
1.199	.613	1 13	6.975	1.185	3.65	.17	.31
1.236	.639	1 14	7.400	1.275	3.38	.16	.31
1.271	.666	1 15	7.765	1.380	3.08	.16	.31
1.306	.694	2 0	7.950	1.500	2.77	.15	.31



Graph showing values of x (size), y (density of sample-units), v/pq (proportional variance), and $v\phi/pq$ (proportional cost of unit information) for different values of ϕ (expenditure in rupees per sq. mile).

CHAPTER 5 : THE PROBLEM OF ZONING

134. I should like to emphasize at this stage that our theory is based on homogeneous zones, that is, the cost and the margin of error are calculated on the assumption that the whole area has been divided into a number of zones in each of which the intensity of cultivation (i.e., the value of p) is practically constant. The work of demarcating the zones however can be started only after we have succeeded in collecting reliable information regarding the proportion of land under jute in each *thana* or possibly in units smaller than the *thana*. In this connexion we have also to remember that the proportion of land under jute in the same *thana* is by no means constant from one season to another; areas which are homogeneous in one season may cease to be so in another season. It is essential therefore to demarcate the zones on the basis of the most recent information available; it may also be necessary to make suitable changes in order to keep the zoning up to date.

Intensity of cultivation by police stations

135. Let us now consider the available material. We have first of all the values of proportion under jute for individual *thanas* obtained from the Sample Survey of 1935. So far as I can judge from internal evidence the material appears to be satisfactory. It is however old and refers to conditions existing five years ago. We have another set of figures based on the official Jute Registration of 1939. Finally for eight districts we have our own estimates based on the Sample Survey of 1940. I had written to the Agricultural Department for the *thana* estimates for 1940 but unfortunately these have not reached me at the time of writing this report.

136. I am giving all available material for individual *thanas* in Tables 46.01 -(46.18). In these tables col. (1) gives a serial number for convenience of reference; col. (2) the name of the Police Station; col. (3.1) the area of the Police Stations in square miles as given in the original records of the 1935 Sample Survey; col. (3.2) gives the total number of sample-units of size 40-acre used in 1935; and col. (3.3) the proportion of land under jute (in a percentage form, as calculated from the Sample Survey of 1935. Col. (4) gives the area of the Police Stations according to settlement records, and col. (5) the percentage proportion of land under jute record during the Jute Registration of 1939 and supplied to us from the office of the Chief Controller of Jute Registration, Bengal.

Correlation between different estimates

137. In order to compare the different sets of records we have calculated the coefficient of correlation between the different series; and the results are given in Tables 47 and 48. In Table 47[1] the results are given for the eight districts for which three series of values for the Sample Survey of 1935, Jute Registration of 1939, and the Sam-

[1]Not available.

SAMPLE CENSUS OF THE AREA UNDER JUTE IN BENGAL IN 1940

TABLE 46.01 PERCENTAGE OF AREA UNDER JUTE BY POLICE STATIONS: DIFFERENT ESTIMATES*

sl. no.	name of police station	1935 (40-acre) sample survey			settlement area in sq. miles	1939 registration per cent under jute
		area in sq. miles	number of grids	per cent under jute		
(1)	(2)	(3.1)	(3.2)	(3.3)	(4)	(5)
		district : Bakharganj				
1	Amtali	—	—	—	349	0.0
2	Babuganj	56	4	18.5	56	18.6
3	Bakarganj	—	9	—	141	0.0
4	Bamna	—	—	—	34	0.0
5	Banaripara	—	—	—	40	2.2
6	Bauphal	—	—	—	154	0.0
7	Baranadi	—	7	1.3	103	0.3
8	Barguna	—	—	—	122	0.0
9	Barisal	93	6	5.3	93	0.2
10	Betagi	—	—	—	59	0.0
11	Bhandaria	—	—	—	56	0.0
12	Bhola	131	9	0.7	131	2.6
13	Daulatkhan	178	12	2.4	178	2.3
14	Galachipa	—	—	—	309	0.0
15	Gournadi	119	8	20.8	119	27.1
16	Hizla	100	—	—	100	7.5
17	Jhalakathi	80	5	4.2	80	3.7
18	Kalapara	—	—	—	162	0.1
19	Kathalia	—	—	—	55	0.0
20	Kowkhali	—	—	—	40	0.0
21	Lalmohan	—	—	—	145	0.9
22	Mathbaria	—	—	—	118	0.0
23	Mehandiganj	123	8	9.3	123	10.0
24	Mirzaganj	—	—	—	57	0.0
25	Muladi	6824	5	13.0	68	12.5
26	Nalchiti	83	6	0.2	83	0.3
27	Nazirpur	—	—	—	95	2.2
28	Patharghata	—	—	—	90	0.0
29	Patuakhali	—	—	—	132	0.0
30	Pirojpur	—	—	—	75	0.0
31	Rajapur	58	4	0.1	58	0.0
32	Swarupkathi	—	—	—	72	0.0
33	Tazumuddin	—	—	—	112	0.0
34	Wazirpur	76	5	8.4	76	0.0

*Out of 18 tables only one table is printed here as a specimen.

ple Census of 1940, are available. Between these three series there are three different coefficients of correlation which are all given here. The name of the district is shown in col. (1); and after this the three series of results in cols. (2.1)–(2.4), cols. (3.1)–(3.4), and cols. (4.1)–(4.4) for correlation between the 1935 and 1939, between the 1935 and 1940, and between 1939 and 1940, respectively. The number of *thanas* included in the calculations is shown in cols. (2.1), (3.1) and (4.1) in each case; mean percentage of land under jute in cols. (2.2)–(2.3), (3.2)–(3.3), and (4.2)–(4.3); and the values of the coefficients of correlation in cols. (2.4), (3.4) and (4.4). It should be noticed that the number of Police Stations differs for the different series even in the same district; this is why the mean are also slightly different from one series to another.

138. Comparing the figures given in cols. (2.2) and (2.3) we find that the proportion of land recorded during registration in 1939 was higher than that found by the Sample Survey in 1935 in every district with the exception of Dacca for which, however, the records for 1939 are incomplete. The increase betwen 1935 and 1939 is not uniform but varies widely in different districts. In Nadia, Mymensingh, and possibly Dacca, the increase was comparatively small; while in the other districts the proportion of land under jute was nearly doubled.

139. The coefficients of correlation[1] are all positive and significant except in Dacca; in Bogra the value is high (+0.86), but otherwise lie between +0.55 in Rangpur and +0.69 in Nadia; the over-all value for all eight districts is 0.64. We find then that there was moderate association between the estimates obtained in 1935 and 1939, but the agreement was not very close.

140. If we compare the mean values of percentage of land under jute in 1935 and 1940 given in cols. (3.2) and (3.3) we find that the area under jute had practically doubled in Nadia, Jessore, Rajshahi, Bogra and Tipperah; in the heavier districts Rangpur, Mymensingh, and Dacca, the increase was very substantial although not quite double; for eight districts the average percentage increased from 13.3 per cent in 1935 to 23.1 per cent in 1940.

141. The coefficients of correlation given in col. (3.4) are appreciably higher than the corresponding coefficients given in col. (2.4). With the exception of Nadia they are all significant and range from +0.66 in Tipperah to +.903 in Rangpur. The pooled value of the correlation between 1935 and 1940 sample surveys is +0.81 against a correlation of +0.64 between 1935 sample survey and 1939 registration figures. The much closer agreement between the two sample series is re-assuring; and throws doubt on the reliability of the 1939 registration figures.

[1]Coefficients significant at the five percent level are marked with a single star (*), and at the one percent level with two stars (**)

SAMPLE CENSUS OF THE AREA UNDER JUTE IN BENGAL IN 1940

142. Finally the comparison between the 1939 registration and the 1940 sample census series shows an appreciable increase in the proportion of land under jute from 17 per cent in 1939 to 22 per cent in 1940. The coefficients of correlation are also high, suspiciously so in Bogra (+0.9633); with a pooled value of +0.82 which is about the same as the correlation between 1935 and 1940 series.

143. Table 48 shows the correlation between different estimates of the percentage of land under jute in 11 districts for which only two series of figures (for 1935 Sample Survey and 1939 Jute Registration) are available. The arrangement is similar to that in Table (47); col. (1) gives the name of the district; col. (2.1) the number of Police Stations for which records were available in both the series; col. (2.2) the mean percentage of land under jute in 1935, col. (2.3) the mean percentage recorded during registration in 1939; and col. (2.4) the co-efficient of correlation between the two estimates.

TABLE 48. CORRELATION BETWEEN DIFFERENT ESTIMATES OF PERCENTAGE OF LAND UNDER JUTE : ELEVEN DISTRICTS (1935 AND 1939)

name of district	number of police stations	mean percentage of land under jute		coefficient of correlation	standard deviation		coefficient of variation	
		1935	1939		1935	1939	1935	1939
(1)	(2.1)	(2.2)	(2.3)	(2.4)	(3)	(4)	(5)	(6)
Pabna	17	13.12	15.00	0.9057**	10.72	10.48	81.7	69.9
Khulna	15	4.06	5.26	.9476**	4.60	5.12	113.3	97.3
Jalpaiguri	17	2.06	3.70	.2406	1.74	3.46	84.5	93.5
Burdwan	4	1.00	3.00	—	—	2.30	—	76.7
Murshidabad	15	4.06	7.08	.1015	2.50	3.68	61.6	52.0
Bakherganj	13	6.84	7.30	.9714**	8.04	7.00	117.5	95.9
Noakhali	8	7.76	14.76	.7036*	5.44	10.50	70.1	71.1
Hooghly	18	9.88	8.66	.8796**	4.12	4.30	41.7	49.6
Dinajpur	16	8.00	12.38	.1863	4.44	10.80	55.5	47.2
24 Parganas	32	4.32	2.50	.8322**	5.96	2.60	133.3	104.0
Faridpur	25	13.16	21.16	.7096**	6.02	13.40	45.7	63.3
eleven districts	180	6.68	9.12	.7540**	6.90	9.94	103.3	109.0

144. The correlation is significant in six districts out of eleven at one per cent level and in one district at five per cent level. Among significant correlations, individual values lie between +0.70 in Noakhali and +0.97 in Backarganj, the pooled value for eleven districts is +0.75 showing a moderately high degree of association. The average proportion of land under jute in eleven districts increased from about 6.7 per cent in 1935 to 9.1 per cent in 1939; the increase in individual districts were, however, not regular.

145. In the same Table 48 the standard deviation in 1935 and 1939 are shown separately for each district in cols. (3) and (4) respectively. The corresponding coefficients of variation are given in cols. (5) and (6). These vary from about 42 per cent in Hooghly in 1935 to 133 per cent in 24-Parganas in the same year; the coefficient of variation for all eleven districts taken together was about 103 per cent in 1935 and 109 per cent in 1939.

146. The standard deviations and coefficients of variation for the intensity of cultivation in eight districts are given in Table 49. In this table col. (1) gives the name of the district; col. (2.1) the number of Police Stations included in the sample survey of 1935; col (2.2) the standard deviation of values of p in the different Police Stations; and col. (2.3) the coefficient of variation in each district in 1935. Similar information for the 1939 Registration figures is given in cols. (3.1), (3.2) and (3.3) respectively; and finally for Sample Census of 1940 in cols. (4.1), (4.2) and (4.3) respectively. It will be noticed that the values of standard deviation increased in recent years, the pooled value of eight districts being 8.78 in 1935, 10.72 in 1939 and 12.30 in 1940. Owing to the large increase in the proportion of land under jute, the coefficient of variation, however, cecreased from 75.2 per cent in 1935 to 68 per cent in 1939 and 53 per cent in 1940.

TABLE 49. STANDARD DEVIATIONS AND COEFFICIENTS OF VARIATION FOR THE INTENSITY OF CULTIVATION : EIGHT DISTRICTS (1935, 1939, 1940)

name of district	sample survey of 1935			registration 1939			sample census 1940		
	N.	s.d.	c.v.	N.	s.d.	c.v.	N.	s.d.	c.v.
(1)	(2.1)	(2.2)	(2.3)	(3.1)	(3.2)	(3.4)	(4.1)	(4.2)	(4.3)
Nadia	25	5.62	112.4	25	3.92	63.2	12	2.60	46.0
Jessore	23	3.82	70.2	23	6.42	61.5	12	5.18	47.1
Rajshahi	21	5.98	90.3	21	10.54	94.9	11	9.92	67.7
Bogra	12	8.14	85.7	13	11.62	71.5	13	13.06	74.1
Tipperah	16	7.42	49.9	20	10.40	40.8	20	15.90	54.6
Rangput	30	6.36	51.3	30	11.22	47.1	9	7.68	24.2
Mymensingh	50	8.08	46.7	50	8.10	42.1	50	11.30	40.3
Dacca	32	8.36	55.4	11	6.70	44.7	32	9.46	36.5
eight districts	209	8.78	75.2	193	10.72	68.0	159	12.30	53.0

PROVISIONAL DEMARCATION OF ZONES

147. The result of the correlational analysis shows that although there is moderate association between the different series the connexion is not very close. This indicates either that the proportion of land under jute has varied widely and irregularly in different Police Stations in different years, or that the different series of estimates are not all reliable. In this situation all we can do is to proceed with the work in a tentative manner. We are adopting the Registration figures of 1939 as the provisional basis for this purpose.

ZONING BY DISTRICTS

148. It will be remembered that so far we had treated each district as a zone by itself. A slight improvement is however possible by grouping together some of the districts in accordance with the intensity of cultivation. Table 50 gives the basic data for such zoning by districts. In this table col. (1) gives the serial number of the zone; col. (2) the name of the districts, some of which are grouped together within the same zone; col. (3) the zonal value p, the proportion of land under jute; and col. (4) the corresponding values of pq. The area in square miles of each zone is given in col. (5.1); and the area under jute in square miles in each zone in col. (5.2). The accumulated total area in square miles is shown in col. (6.1); and the accumulated totals of the area under jute in col. (6.2). Finally col (7) shows the accumulated jute area as a percentage of the total area under jute in the whole province.

1949. It will be noticed that nine zones with 36,669 square miles account for more than 90 per cent of the jute area; ten zones with 4.192 square miles cover 94.6 per cent; and eleven zones with 47,735 square miles cover 97.46 per cent of the total area under jute.

TABLE 50. BASIC DATA FOR ZONING BY DISTRICTS

no. of zone	name of district	p	pq	area (in square miles)		accumulated area (in square miles)		accumulated jute area as percentage
				total	jute area	total	jute area	
(1)	(2)	(3)	(4)	(5.1)	(5.2)	(6.1)	(6.2)	(7)
1.	Dacca	0.180	0.1476	2695	485	2695	485	12.06
2.	Mymensingh	.165	.1378	6360	1049	9055	1534	38.14
3.	Tipperah Faridpur	.147	.1254	4853	713	13908	2247	55.87
4.	Rangpur	.132	.1146	3612	477	17520	2724	67.73
5.	Bogra	.109	.0971	1479	161	18999	2885	71.73
6.	Pabna	.077	.0711	1818	140	20817	3025	75.21
7.	Rajshahi Noakhali	.055	.0520	4059	223	24876	3248	80.76
8.	Jessore	.045	.0430	2915	131	27791	3379	84.01
9.	Dinajpur Nadia Murshidabad	.035	.0338	8878	311	36669	3690	91.75
10.	Bakarganj	.022	.0215	3523	78	40192	3768	93.68
11.	24 Parganas Maldah Howrah	.015	.0148	7543	113	47735	3881	96.49
12.	Jalpaiguri	.022	.0215	2932	64	50667	3945	98.09
13.	Hooghly	.030	.0291	1188	36	51855	3981	98.98
14.	Khulna	.008	.0079	4689	38	56544	4019	99.93
15.	Burdwan	.001	.0010	2705	3	59249	4022	100.00

Note: Jute area according to five years' average of revised estimates.

ZONING BY POLICE STATIONS

150. We may now consider the zoning by groups of Police Stations. The basic data are given in Table 51 in which col. (1) shows the serial number of the zones; col (2) the range of p (the proportion of land under jute in each zone); col. (3.1) the mid-value or zonal value of the intensity of cultivation; and col. (3.2) the accumulated proportion under jute, that is, the value of p for all zones from No. (1) downwards up to any given stage pooled together. The number of Police Stations in each range of p is shown in col. (4); the area in square miles according to settlement records

TABLE 51. BASIC DATA FOR ZONING BY GROUPS OF POLICE STATIONS

zone no.	range of p proportion	proportion of jute		number of police stations	(area in sq. miles)		accumulated values of			percentage of total jute
		mid-value	accumulated		total (settlement)	under jute*	number police stations	total area in sq. miles	jute area in sq. miles	
(1)	(2)	(3.1)	(3.2)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	.52–.54		0.54	1	56	30	1	56	30	0.50
	.48–.50	0.49	.50	1	148	73	2	204	103	1.73
	.44–.46		.49	1	83	37	3	287	140	2.36
2	.42–.44		.46	2	225	97	5	512	237	4.00
	.40–.42	0.42	.44	3	272	112	8	784	349	5.89
3	.38–.40		.43	3	250	98	11	1034	447	7.55
	.36–.38	0.38	.42	2	255	94	13	1289	541	9.15
4	.34–.36	0.35	.40	5	518	181	18	1807	722	12.22
5	.32–.34	0.33	.38	7	641	212	25	2448	934	15.80
6	.30–.32	0.31	.36	8	973	302	33	3421	1236	20.90
7	.28–.30	0.29	.35	5	539	156	38	3960	1392	23.56
8	.26–.28	0.27	.33	13	119	322	51	5153	1714	29.02
9	.24–.26	0.25	.31	16	2006	502	67	7159	2216	37.52
10	.22–.24	0.23	.30	14	1511	348	81	8670	2564	43.41
11	.20–.22	0.21	.28	20	2455	516	101	11125	3080	52.15
12	.18–.20	0.19	.26	22	2470	469	123	13595	3549	60.10
13	.16–.18	0.17	.25	19	1968	335	142	15563	3884	65.77
14	.14–.16	0.15	.24	16	1942	291	158	17505	4175	70.71
15	.12–.14	0.13	.23	15	1637	213	173	19142	4388	74.31
16	.10–.12	0.11	.21	26	2902	319	199	22044	4707	79.72
17	.08–.10	0.09	.19	37	4132	372	236	26176	5079	86.03
18	.06–.08	0.07	.18	32	3493	245	268	29669	5324	90.17
19	.04–.06	0.05	.16	47	6294	265	315	34963	5589	94.66
20	.02–.04	0.03	.14	47	5571	167	362	40534	5756	97.49
21	.00–.02	0.01	.09	137	18715	148	499	59249	5904	100.00

*Jute area according to Jute Registration of 1939.

in col. (5); and the area under jute according to 1939 Registration figures in col. (6). The accumulated values of the number of Police Stations are shown in col. (7); the accumulated total area in col. (8); and the accumulated area under jute in square miles in col. (9). Finally the accumulated area under jute is shown in the form of a percentage of total jute area in col. (10).

151. On this basis we find that eleven zones with 11,125 square miles account for 52.15 per cent of the total area under jute; fifteen zones with 29,669 square miles cover 90 per cent; and 19 zones with 34,963 square miles cover 94.66 per cent of the toal jute area; finally if we use twenty zones with 40,534 square miles more than 97 per cent of jute area will be covered.

152. It will be noticed that if we stop at zone No. (19) we shall have a balance of about twenty thousand square miles which contribute only a little over five per cent of the total area under jute; and if we stop at zone No. (20) we have only one zone No. (21) left with an intensity of cultivation of the order of one per cent covering 14,827 square miles which contributes only about two and a-half percent of the total area under jute. It may be necessary to devise special methods for sampling such zones with extremely low intensities of cultivation; I have discussed this question a little later.

153. It should be noted that the *thanas* included in a particular group of zones belong to more than one district. We may obtain a broad idea of the distribution from the two-way Tables 52 and 53 in which the number of Police Stations as well as the total area in square miles belonging to any particular zone is shown separately for each district. In this table, col. (0.1) gives serial number of the zone; sol. (0.2) the zonal value of p, the proportion of land under jute. The area in square miles together with the number of Police Stations is shown under each district in successive columns; for example, col. (1.1) gives the area, and col. (1.2) the number of Police Stations for each zone in district Dacca; col. (0.4) gives the sub-total; and col. (0.5) the grand total for 21 districts.

154. In many cases Police Stations included within a particular zone belong to the same district, some of these Police Stations may occur in compact blocks in which case they can be all grouped into a single zone. On the other hand, it will also be seen that the same zone in most cases contain Police Stations belonging to different districts; in these cases each thana (or each group of *thanas* occurring in a compact block) will have to be treated as a separate zone. Actual demarcation into zones will necessarily have to be made on a geographical basis with the help of detailed examination of maps. Provision has been made in the budget for 1940-41 for this purpose.

TABLE 52. DISTRIBUTION OF *THANA*-ZONES BY DISTRICTS : AREA IN SQUARE MILES WITH NUMBER OF POLICE STATIONS

zone no.	zonal	Dacca		Mymensingh		Faridpur		Tipperah		Rangpur		Bogra		Pabna		Rajshahi		Noakhali		Jessore		Murshidabad		sub-total (1—11)	
	p	area	*N*	area	*N*	area	*N*	area	*N*	area	*N*	area	*N*	area	*N*	area	*N*	area	*N*	area	*N*	area	*N*	area	*N*
(0.1)	(0.2)	(1.1)	(1.2)	(2.1)	(2.2)	(3.1)	(3.2)	(4.1)	(4.2)	(5.1)	(5.2)	(6.1)	(6.2)	(7.1)	(7.2)	(8.1)	(8.2)	(9.1)	(9.2)	(10.1)	(10.2)	(11.1)	(11.2)	(0.31)	(0.32)
1	0.49	—	—	—	—	139	2	148	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	287	3
2	.42	192	2	—	—	67	1	—	—	—	—	—	—	—	—	158	1	—	—	—	—	—	—	417	4
3	.38	—	—	—	—	87	1	141	1	74	1	—	—	35	1	—	—	—	—	—	—	—	—	337	4
4	.35	—	—	202	1	28	1	63	1	90	1	135	1	—	—	—	—	—	—	—	—	—	—	518	5
5	33	—	—	41	1	—	—	94	1	356	3	95	1	—	—	—	—	55	1	—	—	—	—	641	7
6	.31	—	—	306	3	—	—	239	2	287	2	—	—	—	—	141	1	—	—	—	—	—	—	973	8
7	.29	..	—	129	1	62	1	—	—	88	1	200	1	60	1	—	—	—	—	—	—	—	—	539	5
8	.27	163	2	114	2	127	1	328	3	147	2	92	1	—	—	103	1	—	—	—	—	—	—	1074	12
9	.25	135	2	470	3	392	3	678	5	229	2	—	—	102	1	—	—	—	—	—	—	—	—	2006	16
10	.23	152	2	717	6	42	1	88	1	98	1	—	—	414	3	—	—	—	—	—	—	—	—	1511	14
11	.21	46	1	919	9	64	1	118	1	588	3	—	—	170	1	—	—	278	2	272	2	—	—	2455	20
12	.19	345	4	734	6	356	3	173	1	263	2	122	1	—	—	—	—	—	—	338	3	—	—	2331	20
13	.17	289	4	722	4	—	—	—	—	658	7	—	—	—	—	—	—	—	—	180	2	—	—	1849	17
14	.15	278	4	258	2	—	—	207	1	401	3	256	2	—	—	—	—	130	1	—	—	95	1	1625	14
15	.13	364	1	132	1	111	2	—	—	167	1	114	1	—	—	178	2	60	1	267	2	—	—	1093	11
16	.11	137	1	532	4	108	1	—	—	166	1	120	1	130	1	524	5	—	—	—	—	301	3	2018	17
17	.09	497	5	—	—	345	3	—	—	—	—	—	—	405	3	168	2	—	—	348	3	229	2	1992	18
18	.07	121	1	426	3	340	3	—	—	—	—	—	—	245	3	54	1	166	1	425	4	640	6	2417	22
19	.05	—	—	526	3	—	—	153	1	—	—	182	2	143	2	661	5	155	1	506	3	142	2	2468	19
20	.03	209	2	132	1	47	1	108	1	—	—	93	1	114	1	—	—	62	1	579	4	108	1	1452	13
21	0.1	67	1	—	—	—	—	—	—	—	—	70	1	—	—	544	4	622	6	—	—	575	5	1878	17
	total	2695	32	6360	50	2315	25	2538	20	3612	30	1479	13	1818	17	2531	22	1528	14	2915	23	2090	20	29881	266

TABLE 53. DISTRIBUTION OF THANA-ZONES BY DISTRICTS : AREA IN SQUARE MILES WITH NUMBER OF POLICE STATIONS

zone no.	zonal p	Nadia		Hooghly		Dinajpur		Malda		Jalpaiguri		Bakerganj		24-Parganas		Khulna		Burdwan		Howrah		sub-total (12-(21)		grand total	
		area	N	area	N	area	N	area	N	area	N	area	N	area	N	area	N	area	N	area	N	area	N	area	N
(0.1)	(0.2)	(12.1)	(12.2)	(13.1)	(13.2)	(14.1)		(14.2)	(15.1)	(15.2)	(16.1)	(16.2)	(17.1)	(17.2)	(18.1)	(18.2)	(19.1)	(19.2)	(20.1)	(20.2)	(21.1)	(21.2)	(0.41)	(0.42)	(0.51)	(0.62)
1	0.49	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	287	3
2	.42	—	—	—	—	80	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	80	1	497	5
3	.38	—	—	—	—	168	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	168	1	505	5
4	.35	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	518	5
5	.33	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	641	7
6	.31	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	973	8
7	.29	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	539	5
8	.27	—	—	—	—	—	—	—	—	—	—	119	1	—	—	—	—	—	—	—	—	119	1	1193	13
9	.25	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2006	16
10	.23	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1511	14
11	.21	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2455	20
12	.19	—	—	—	—	—	—	—	—	—	—	56	1	—	—	83	1	—	—	—	—	139	2	2470	22
13	.17	38	1	—	—	81	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	119	2	1968	19
14	.15	—	—	—	—	207	1	—	—	—	—	—	—	—	—	110	1	—	—	—	—	317	2	1942	16
15	.13	181	1	45	1	250	1	—	—	—	—	68	1	—	—	—	—	—	—	—	—	544	4	1637	15
16	.11	134	1	227	4	141	1	—	—	259	2	123	1	—	—	—	—	—	—	—	—	884	9	2902	26
17	.09	596	5	—	—	969	8	—	—	96	1	—	—	260	3	219	2	—	—	—	—	2140	19	4132	37
18	.07	445	3	78	1	159	2	118	1	—	—	193	2	83	1	—	—	—	—	—	—	1076	10	3493	32
19	.05	542	6	287	4	763	6	—	—	280	2	—	—	181	2	341	4	267	2	165	2	2826	28	5294	47
20	.03	527	5	223	2	407	3	275	2	1230	6	524	5	446	5	304	2	102	1	81	3	4119	34	5571	47
21	.01	377	3	328	6	723	5	1371	12	1067	6	2440	22	4297	28	3632	12	2336	20	266	6	16837	120	18715	137
total	.09	2840	25	1188	18	3948	30	1764	15	2932	17	3523	33	5267	39	4689	22	2705	23	512	11	29368	233	59249	499

CHAPTER 6 : THE BEST SIZE AND DENSITY OF SAMPLE-UNITS

155. We shall now discuss the most important question of the proper choice of the size and density of grids; that is, we must settle what particular size and density of grids should be used in each zone in order that the error of the final estimate may be reduced to a minimum for any assigned level of total expenditure.

156. The best size and density of grids at different levels of expenditure for district-zones are shown in Tables 54–57. In these tables the serial number of the zone is shown in col. (1) and refers to the same series of zone numbers in Table 50; the proportion of land under jute is shown in col. (2); the accumulated total area in squatre of miles in col. (3.1); and the accumulated area under jute as a percentage the total area under jute in the province in col (3.2). The best size of sample-units is given in col. (4); and the corresponding best density of sample-units per square mile in each half-sample in col. (5). The total number of grids in each zone is shown in col. (6.1), and the accumulated totals in col. (6.2). The cost in rupees for the field portion of the work is given in col. (7.1); for the statistical work in col. (7.2); and the total cost is shown in col. (7.3) for each zone separately. The accumulated cost up to any particular zone inclusive is shown for field work in col. (8.1), for statistical work in col. (8.2); the total accumulated cost is given in col. (8.3). The values of calculated standard error are shown in col. (9.1), and accumulated values upto any zone inclusive in (9.2), finally the proportional standard errors are shown separately in col. (10.1), and accumulated values up to any particular zone inclusive in col. (10.2).

157. In Table 54, for a total expenditure of about one lakh and twenty thousand rupees, the best size of sample-units varies from 4.56 acre in district-zone No. (1) to 6.71) acre in district-zone No. (15) As the expenditure is increased to about one lakh and thirty thousand in Table 55 the best size is reduced to 4.08 acre in zone No. (1), but remains practically unchanged in the thinner zones. Increasing the total expenditure to one lakh and forty thousand rupees in Table 56 the best size is further reduced to about 3.5 in zone No. (1), but remains unchanged as before in the thinner regions. Finally with a total expenditure of about one lakh and forty-four thousand rupees in Table 57 the best size of grids becomes 3.12 acre in zone No. (1).

158. It will be noticed that the best density in the heavier zones increases appreciably with increase of expenditure; for example, in zone No. (1) the best density of sample-units per square mile in each half sample increased from 0.858 in Table 54 to 1.374 in Table 57. In the thinner zones the actual change is small,

TABLE 54. BEST SIZE AND DENSITY OF GRIDS (IN HALF-SAMPLES) AT DIFFERENCE LEVELS OF EXPENDITURE FOR DISTRICT-ZONES

zone no.	proportion under jute	accumulated		size of grid x_i	half-sample deisity y_i	total number of grids		cost in rupees			accumulated cost in rupees			$\sqrt{v_i}$		$100\times\sqrt{v_i}/A_i\rho_i$	
		total area in square miles	jute area percentage			zone	accumulated	field	statistical	total	field	statistical	total	zone	accumulated	zone	accumulated
(1)	(2)	(3.1)	(3.2)	(4)	(5)	(6.1)	(6.2)	(7.1)	(7.2)	(7.3)	(8.1)	(8.2)	(8.3)	(9.1)	(9.2)	(10.1)	(10.2)
1.	0.180	2695	12.06	4.56	0.858	4624	4624	5525	2722	8247	5525	2722	8247	8.12	8.12	1.67	1.67
2.	0.165	9055	38.14	4.66	.815	10366	14990	12784	6232	19016	18309	8594	27263	12.31	14.75	1.17	.96
3.	0.147	13908	55.87	4.70	.765	7426	22416	9366	4562	13928	27675	13516	41191	10.58	18.15	1.48	.81
4.	0.132	17520	67.73	4.89	.716	5172	27588	6718	3287	10005	34393	16803	51196	8.98	20.25	1.88	.74
5.	0.109	18999	71.73	5.12	.645	1908	29496	2647	1243	3890	37040	18046	55086	5.55	20.99	3.44	.73
6.	0.077	20817	75.21	5.43	.529	1924	31420	2891	1345	4236	39931	19391	59322	5.76	21.77	4.11	.72
7.	0.055	24876	80.76	5.70	.442	5512	36932	5804	2679	8483	45735	22070	67805	8.03	23.21	3.60	.71
8.	0.045	27791	84.01	5.81	.394	2298	39230	3877	1807	5684	49612	23877	73489	6.51	24.10	4.96	.71
9.	0.035	36669	91.75	5.95	.346	6144	45374	10831	5149	15980	60443	29026	89469	10.75	26.39	3.46	.71
10.	0.022	40192	93.68	6.16	.267	1882	47256	3805	1761	5566	64248	30787	95035	6.15	27.09	7.93	.72
11.	0.015	47735	96.49	6.28	.223	3364	50620	7467	3395	10862	71715	34182	105897	8.11	28.28	7.17	.73
12.	0.022	50667	98.09	6.11	.267	1566	52186	3166	1467	4633	74881	35649	110530	5.58	28.83	8.65	.73
13.	0.030	51855	98.98	6.23	.321	762	52948	1378	653	2031	76259	36302	112561	3.76	29.07	10.55	.73
14.	0.008	56544	99.93	6.51	.154	1444	54392	3704	1735	5439	79963	38037	118000	5.59	29.60	14.90	.74
15.	0.001	59249	100.00	6.71	.043	232	54624	1596	703	2299	81559	38740	120299	2.71	29.73	100.37	.74

TABLE 55. BEST SIZE AND DENSITY OF GRIDS (IN HALF-SAMPLES) AT DIFFERENT LEVELS OF EXPENDITURE FOR DIFFERENT ZONES

zone no.	proportion under jute	accumulated: total area in square miles	accumulated: jute area percentage	size of grid x_i	half-sample density y_i	total number of grids: zone	total number of grids: accumulated	cost in rupees: field	cost in rupees: statistical	cost in rupees: total	accumulated cost in rupees: field	accumulated cost in rupees: statistical	accumulated cost in rupees: total	$\sqrt{v_i}$: zone	$\sqrt{v_i}$: accumulated	$100\times\sqrt{v_i}/A_i y_i$: zone	$100\times\sqrt{v_i}/A_i y_i$: accumulated
(1)	(2)	(3.1)	(3.2)	(4)	(5)	(6.1)	(6.2)	(7.1)	(7.2)	(7.3)	(8.1)	(8.2)	(8.3)	(9.1)	(9.2)	(10.1)	(10.2)
1.	0.180	2695	12.06	4.08	1.021	5504	5504	6091	3018	9109	6091	3018	9109	7.53	7.53	1.55	1.55
2.	0.165	9055	38.14	4.25	0.962	12236	17740	13801	6933	20734	19892	9951	29843	11.46	13.71	1.09	0.89
3.	0.147	13908	55.87	4.43	0.895	8686	26426	10191	5047	15238	30083	14998	45081	9.84	16.88	1.38	.75
4.	0.132	17520	67.73	4.59	0.840	6068	32494	7332	3612	10944	37415	18610	56025	8.35	18.83	1.75	.69
5.	0.109	18999	71.73	4.82	0.746	2206	34700	2810	1376	4186	40225	19986	60211	5.19	19.54	3.22	.68
6.	0.077	20817	75.21	5.22	0.617	2244	36944	3127	1509	4636	43352	21495	64847	5.35	20.26	3.82	.67
7.	0.055	24876	80.76	5.51	0.504	4092	41036	6291	2923	9214	49643	24418	74061	7.53	21.61	3.37	.66
8.	0.045	27791	84.01	5.66	0.449	2618	43654	4198	1953	6151	43841	26371	80212	6.12	22.46	4.66	.66
9.	0.035	36669	91.75	5.82	0.391	6942	50596	11630	5505	17135	65471	31876	97347	10.12	24.63	3.26	.67
10.	0.022	40192	93.68	6.07	0.305	2134	52730	4051	1868	5919	69522	33744	103266	5.76	25.30	7.43	.67
11.	0.015	47735	96.49	6.21	0.248	3742	56472	7845	3620	11465	77367	37364	114731	7.69	26.44	6.79	.68
12.	0.022	50667	98.09	6.07	0.303	1776	58248	3372	1554	4926	80739	38918	119657	5.25	26.96	8.14	.68
13.	0.030	51855	98.98	5.91	0.359	852	59100	1497	701	2198	82236	39619	121855	3.58	27.19	15.66	.68
14.	0.008	56544	99.93	6.43	0.176	1650	60750	4267	1922	6189	86503	41541	128044	5.25	27.69	13.94	.69
15.	0.001	59249	100.00	6.70	0.052	282	61032	1677	730	2407	88180	42271	130451	2.46	27.80	91.11	.69

TABLE 56. BEST SIZE AND DENSITY OF GRIDS (IN HALF-SAMPLES) AT DIFFERENT LEVELS OF EXPENDITURE FOR DISTRICT ZONES

zone no.	proportion under jute	accumulated: total area in square miles	accumulated: jute area percentage	size of grids x_i	half-sample density y_i	totl number of grids: zone	totl number of grids: accumulated	cost in rupees: field	cost in rupees: statistical	cost in rupees: total	accumulated cost in rupees: field	accumulated cost in rupees: statistical	accumulated cost in rupees: total	$\sqrt{v_i}$: zone	$\sqrt{v_i}$: accumulated	$100\times\sqrt{v_i}/A_ip_i$: zonea	$100\times\sqrt{v_i}/A_ip_i$: ccumulated
(1)	(2)	(3.1)	(3.2)	(4)	(5)	(6.1)	(6.2)	(7.1)	(7.2)	(7.3)	(8.1)	(8.2)	(8.3)	(9.1)	(9.2)	(10.1)	(10.2)
1.	0.180	2695	12.06	3.51	1.234	6652	6652	6549	3369	9918	6549	3369	9918	6.99	6.99	1.44	1.44
2.	0.165	9055	38.14	3.76	1.137	14462	21114	14946	7696	22642	21495	11065	32560	10.73	12.80	1.02	0.83
3.	0.147	13908	55.87	3.91	1.034	10066	31180	11210	5484	16694	32705	16549	49254	9.32	15.84	1.31	.70
5.	0.132	17520	67.73	4.23	0.980	7080	38260	7802	3973	11775	40507	20522	51029	7.81	17.66	1.64	.65
5.	0.109	18999	71.73	4.57	0.845	2500	40760	3047	1479	4526	43554	22001	65555	4.93	18.33	3.06	.63
6.	0.077	20817	75.21	5.02	0.690	2508	43268	3309	1600	4909	46863	23601	70464	5.09	19.03	3.64	.63
7.	0.055	24876	80.76	5.35	0.557	4522	47790	6576	3125	9701	53439	26726	80165	7.23	20.36	3.24	.63
8.	0.045	27791	84.01	5.51	0.490	2856	50646	4518	2070	6588	59757	28796	86753	5.87	21.18	4.47	.63
9.	0.035	36669	91.75	5.71	0.427	7582	58228	12518	5771	18289	70475	34567	105042	9.74	23.32	3.13	.63
10.	0.022	40192	93.68	5.99	0.335	2360	60588	4263	1973	6236	74738	36540	111278	5.48	23.95	7.07	.64
11.	0.015	47735	96.49	6.15	0.272	4104	64692	8222	3771	11993	82960	40311	123271	7.36	25.06	6.50	.65
12.	0.022	50667	98.09	5.99	0.335	1964	66656	3548	1642	5190	86508	41953	128461	4.99	25.55	7.73	.65
13.	0.030	51855	98.98	5.80	0.397	944	67600	1568	1749	2317	88076	43702	130778	3.41	25.78	9.57	.65
14.	0.008	56544	99.93	6.35	0.201	1884	69484	4361	2016	6377	92437	45718	137155	4.91	26.24	13.09	.65
15.	0.001	59249	100.00	6.70	0.052	282	69766	1677	730	2407	94114	46448	139562	2.49	26.36	92.22	.66

TABLE 57. BEST SIZE AND DENSITY OF GRIDS (IN HALF-SAMPLES) AT DIFFERENT LEVELS OF EXPENDITURE FOR DIFFERENT ZONES

zone no.	proportion under jute	accumulated: total area in square miles	accumulated: jute area percentage	size of grids x_i	half-sample density y_i	total number of grids: zone	total number of grids: accumulated	cost in rupees: field	cost in rupees: statistical	cost in rupees: total	accumulated cost in rupees: field	accumulated cost in rupees: statistical	accumulated cost in rupees: total	$\sqrt{v_i}$: zone	$\sqrt{v_i}$: accumulated	$100\times\sqrt{v_i}/A_ip_i$: zone	$100\times\sqrt{v_i}/A_ip_i$: accumulated
(1)	(2)	(3.1)	(3.2)	(4)	(5)	(6.1)	(6.2)	(7.1)	(7.2)	(7.3)	(8.1)	(8.2)	(8.3)	(9.1)	(9.2)	(10.1)	(10.2)
1.	0.180	2695	12.06	3.12	1.374	7406	7406	6791	3585	10376	6791	3585	10376	6.70	6.70	1.38	1.38
2.	0.165	9055	38.14	3.54	1.221	15532	22938	15328	7950	23278	22119	11535	33654	10.39	12.36	0.99	0.81
3.	0.147	13908	55.87	3.87	1.101	10686	33624	11307	5727	17034	33426	17262	50688	9.06	15.32	1.27	.68
4.	0.132	17520	67.73	4.10	1.015	7332	40956	8163	4009	12172	41589	21271	62860	7.70	17.16	1.61	.63
5.	0.109	18999	71.73	4.45	0.889	2630	43586	3106	1538	4644	44695	22809	67504	4.81	17.82	2.98	.62
6.	0.077	20817	75.21	4.90	0.715	2600	46186	3381	1637	5018	48076	24446	72522	5.01	18.51	3.57	.61
7.	0.055	24876	80.76	5.28	0.584	4740	50926	6819	3207	10026	54895	27653	82548	7.03	19.80	3.15	.61
8.	0.045	27791	84.01	5.46	0.520	3032	53958	4547	2157	6704	59442	29810	89252	5.71	20.61	4.35	.61
9.	0.035	36669	91.75	5.65	0.451	8008	61966	12962	5948	18910	72404	35758	108162	9.46	22.67	3.04	.61
10.	0.022	40192	93.68	5.95	0.348	2452	64418	4298	2043	6341	76702	37801	114503	5.39	23.31	6.95	.62
11.	0.015	47735	96.49	6.12	0.283	4270	68688	8373	3847	12220	85075	41648	126723	7.22	24.40	6.38	.63
12.	0.022	50667	98.09	5.95	0.348	2040	70728	3577	1701	5278	88652	43349	132001	4.90	24.89	7.60	.63
13.	0.030	51855	98.98	5.76	0.413	982	71710	1651	749	2400	90303	44098	134401	3.35	25.11	9.40	.63
14.	0.008	56544	99.93	6.32	0.212	1988	73698	4548	2063	6611	94851	46161	141012	4.78	25.56	12.74	.64
15.	0.001	59249	100.00	6.69	0.060	324	74022	1758	758	2516	96609	46919	143528	22.29	25.66	84.81	.64

Optimum values for thana zones

159. We may now consider the detailed calculations for the best size and density of grids at different levels of expenditure for thana-zones. Relevant figures are given in Tables 58–61 in which col. (1) gives the serial number of the zones corresponding to the serial numbers given in col. (1) of Table 51. The proportion of land under jute is given in col. (2); the accumulated total area in square miles in col. (3.1); and the accumulated area under jute is shown as a percentage of the total area under jute in col. (3.2). The next col. (4) gives values of x_i, the best size of sample-units for each zone; and col. (5) the corresponding values of y_i the best density (or number) of sample-units per square mile for each half-sample in each zone. The total number of grids in each zone is shown in col. (6.1); and the accumulated totals in col. (6.2). The cost in rupees for each zone is shown separately for field work and statistical work in col (7.1), and (7.2), and the total cost in col. (7.3). Accumulated cost in rupees is given for field work in col. (8.3). The value of the standard error of the final work in col. (8.1), for statistical work in col. (8.2), and the total cost in col. (8.3). The value of the standard error of the final estimate is shown for each zone separately in col. (9.1); and the accumulated values for zones up to any stage inclusive in col. (9.2). Finally the value of proportional error is given for individual zones in col. (10.1), and in an acumulated form for all zones inclusive of any particular stage in col. (10.2).

160. Table 57 gives the relevant figures corresponding to a total expenditure of about one lakh and twenty thousand rupees for all 21 zones and about Rs. 88,000 for 19 zones. Similar information is given for a total expenditure of about one lakh and twenty five thousand rupees for 21 zones and Rs. 93,000 for 19 zones in Table 59; for a total expenditure of one lakh and thirty thousand rupees for 21 zones and about Rs. 97,000 for 19 zones in Table 60; and finally for a total expenditure of one lakh and thirty one thousand for 21 zones and about Rs. 98,000 for 19 zones as shown in Table 61.

161. The best size and density of grids depend on the total expenditure. In Table 58 the best size varies from 4.05 acre in the first zone to 5.91 acre in zone No. (19), and 6.50 acre in zone No. (21). With a higher total expenditure in Table 59 the best size of grids is reduced to 3.69 in zone No. (1), but is still about 6 acre in zone Nos. (19) and (20), and 6.48 acre in zone No (21). With a higher expenditure in Table 60 the best size is still further reduced to 3.08 acre in zone No. (1), but remains practically unchanged in this zones. Finally with highest expenditure shown in Table 61 the best size is reduced to 2.7 acre in zone No. (1) without any appreciable change in the thinnest regions. We find then that on the basis of *thana*-zones the effective range of sample-units is likely to lie between say about 3 and 6 acre. This fully confirms the tentative results obtained in 1939.

162. As regards density it will be seen that the best value in zone No. (1) increases from 1.03 in Table 58 to 1.51 in Table 61; in the thinner zones the change is comparatively small, for example, in zone No. (19) the half-sample density increases from 0.36 in Table 58 to 0.41 in Table 61.

TABLE 58. BEST SIZE AND DENSITY OF GRIDS (IN HALF-SAMPLES) AT DIFFERENT LEVELS OF EXPENDITURE FOR *THANA*-ZONES

zone no.	proportion under jute	accumulated: total area in square miles	accumulated: jute area percentage	size of grid x_i	half-density y_i	total number of grids: zone	total number of grids: accumulated	cost in rupees: field	cost in rupees: statistical	cost in rupees: total	accumulated cost in rupees: field	accumulated cost in rupees: statistical	accumulated cost in rupees: total	$\sqrt{v_i}$: zone	$\sqrt{v_i}$: accumulated	$100\times\sqrt{v_i}/A_ip_i$: zone	$100\times\sqrt{v_i}/A_ip_i$: accumulated
(1)	(2)	(3.1)	(3.2)	(4)	(5)	(6.1)	(6.2)	(7.1)	(7.2)	(7.3)	(8.1)	(8.2)	(8.3)	(9.1)	(9.2)	(10.1)	(10.2)
1.	0.49	287	2.36	4.05	1.03	592	592	654	325	979	634	325	979	3.19	3.19	2.27	2.27
2.	.42	784	5.89	4.10	1.01	1004	1596	1118	557	1675	1772	882	2654	4.18	5.25	2.00	1.50
3.	.38	1289	9.15	4.19	0.98	990	2586	1126	551	1677	2898	1433	4331	4.19	6.72	2.18	1.24
4.	.35	1807	12.22	4.28	.95	984	3570	1129	560	1689	4027	1993	6020	4.23	7.94	2.33	1.10
5.	.33	2448	15.80	4.33	.93	1192	4762	1385	679	2064	5412	2672	8084	4.68	9.21	2.21	0.99
6.	.31	3421	20.90	4.38	.91	1770	6532	2072	1022	3094	7484	3694	11178	5.72	10.85	1.90	.88
7.	.29	3960	23.56	4.46	.88	948	7480	1132	555	1687	8616	4249	12865	4.24	11.65	2.71	.84
8.	.27	5153	29.02	4.52	.86	2052	9532	2470	1216	3686	11086	5465	16551	6.23	13.21	1.93	.77
9.	.25	7159	37.52	4.64	.82	3290	12822	4052	1986	6038	15138	7451	22589	8.05	15.47	1.61	.70
10.	.23	8670	43.41	4.76	.78	2358	15180	3007	1435	4442	18145	8886	27031	6.94	16.95	2.00	.66
11.	.21	11125	52.15	4.82	.76	3732	18912	4714	2307	7021	22859	11193	34052	8.66	19.04	1.68	.62
12.	.19	13595	60.10	4.93	.72	3556	22468	4619	2248	6867	27478	13441	40919	8.57	20.88	1.83	.59
13.	.17	15563	65.77	5.03	.68	2676	25144	3601	1713	5314	31079	15154	46233	7.52	22.19	2.25	.57
14.	.15	17505	70.71	5.15	.63	2464	27590	3476	1612	5088	34555	16766	51321	7.36	23.38	2.53	.56
15.	.13	19142	74.31	5.24	.59	1932	29522	2767	1309	4076	37322	18075	55397	6.56	24.28	3.08	.55
16.	.11	22044	79.72	5.41	.54	3134	32656	4643	2177	6820	41965	30252	62217	8.46	25.71	2.65	.55
17.	.09	26176	86.03	5.55	.49	4050	36706	6322	2934	9256	48287	23186	71473	9.66	27.47	2.60	.54
18.	.07	29669	90.17	5.73	.42	2934	39640	4890	2271	7161	53177	25457	78634	8.52	28.76	3.48	.54
19.	.05	34963	94.66	5.91	.36	3812	43452	6565	3123	9688	59742	28580	88322	9.64	30.34	3.64	.54
20.	.03	40534	97.49	6.15	.27	3008	46460	6072	2786	8858	65814	31366	97180	8.90	31.61	5.33	.55
21.	.01	59249	100.00	6.50	.15	5614	52074	15346	7112	22458	81160	38478	119638	12.68	34.06	6.77	.57

TABLE 59. BEST SIZE AND DENSITY OF GRIDS (IN HALF-SAMPLES) AT DIFFERENT LEVELS OF EXPENDITURE FOR *THANA*-ZONES

zone no.	proportion under jute	accumulated: total area in square miles	accumulated: jute area percentage	size of grids x_i	half-sample density y_i	total number of grids: zone	total number of grids: accumulated	cost in rupees: field	cost in rupees: statistical	cost in rupees: total	accumulated cost in rupees: field	accumulated cost in rupees: statistical	accumulated cost in rupees: total	$\sqrt{v_i}$: zone	$\sqrt{v_i}$: accumulated	$100 \times \sqrt{v_i}/A_i p_i$: zone	$100 \times \sqrt{v_i}/A_i p_i$: accumulated
(1)	(2)	(3.1)	(3.2)	(4)	(5)	(6.1)	(6.2)	(7.1)	(7.2)	(7.3)	(8.1)	(8.2)	(8.3)	(9.1)	(9.2)	(10.1)	(10.2)
1.	0.49	287	2.36	3.69	1.17	672	672	683	350	1033	683	350	1033	3.03	3.02	2.15	2.15
2.	.42	784	5.89	3.77	1.14	1134	1806	1173	596	1769	1856	946	2802	3.97	4.99	1.90	1.43
3.	.38	1289	9.15	3.87	1.10	1112	2918	1177	596	1773	3033	1542	4575	3.99	6.40	2.08	1.18
4.	.35	1807	12.22	3.97	1.06	1098	4016	1191	596	1787	4224	2138	6362	4.03	7.56	2.22	1.05
5.	.33	2448	15.80	4.03	1.04	1334	5350	1461	725	2186	5685	2863	8548	4.47	8.78	2.11	0.94
6.	.31	3421	20.90	4.12	1.01	1966	7316	2180	1089	3269	7865	3952	11817	5.48	10.35	1.82	.84
7.	.29	3960	23.56	4.21	0.98	1056	8372	1191	588	1779	9056	4540	13596	4.06	11.12	2.60	.80
8.	.27	5153	29.02	4.31	.94	2242	10614	2577	1276	3853	11633	5816	17449	5.99	12.63	1.86	.74
9.	.25	7159	37.52	4.40	.91	3650	14264	4253	2106	6359	15886	7922	23808	7.71	14.79	1.54	.67
10.	.23	8670	43.41	4.50	.87	2630	16894	3128	1556	4684	19014	9478	28492	6.62	16.21	1.90	.63
11.	.21	11125	52.15	4.61	.83	4076	20970	4959	2455	7414	23973	11933	35906	8.34	18.23	1.62	.59
12.	.19	13595	60.10	4.73	.79	3902	24872	4841	2371	7212	28814	14304	43118	8.24	20.00	1.76	.56
13.	.17	15563	65.77	4.86	.74	2912	27784	3739	1811	5500	32553	16115	48668	7.23	21.27	2.16	.55
14.	.15	17505	70.71	5.01	.69	2680	30464	3554	1709	5263	36107	17824	53931	7.05	22.41	2.42	.54
15.	.13	19142	74.31	5.14	.64	2096	32560	2914	1375	4289	39021	19199	58220	6.32	23.28	2.97	.53
16.	.11	22044	79.72	5.28	.58	3366	35926	4875	2293	7168	43896	21492	65388	8.16	24.67	2.56	.52
17.	.09	26176	86.03	5.45	.52	4298	40224	6529	3057	9586	50425	24549	74974	9.37	26.39	2.52	.52
18.	.07	29669	90.17	5.64	.46	3214	43438	5100	2375	7475	55525	26924	82449	8.19	27.63	3.35	.52
19.	.05	34963	94.66	5.85	.38	4024	47462	6882	3230	10112	62407	30154	92561	9.39	29.18	3.55	.52
20.	.03	40534	97.49	6.09	.29	3232	50694	6184	2897	9081	68591	33051	101642	8.61	30.42	5.15	.53
21.	.01	59249	100.00	6.48	.16	5988	56682	15720	7299	23019	84311	40350	124661	12.29	32.81	6.57	.55

TABLE 60. BEST SIZE AND DENSITY OF GRIDS (IN HALF-SAMPLES) AT DIFFERENT LEVELS OF EXPENDITURE FOR *THANA*-ZONES

zone no.	proportion under jute	accumulated: total area in square miles	accumulated: jute area percentage	size of grid x_i	half-sample density y_i	totl number of grids: zone	totl number of grids: accumulated	cost in rupees: field	cost in rupees: statistical	cost in rupees: total	accumulated cost in rupees: field	accumulated cost in rupees: statistical	accumulated cost in rupees: total	$\sqrt{v_i}$: zone	$\sqrt{v_i}$: accumulated	$100\times\sqrt{v_i}/A_ip_i$: zone	$100\times\sqrt{v_i}/A_ip_i$: accumulated
(1)	(2)	(3.1)	(3.2)	(4)	(5)	(6.1)	(6.2)	(7.1)	(7.2)	(7.3)	(8.1)	(8.2)	(8.3)	(9.1)	(9.2)	(10.1)	(10.2)
1.	0.49	287	2.36	3.08	1.40	804	804	712	384	1096	712	384	1096	2.83	2.82	2.01	2.01
2.	.42	784	5.89	3.32	1.30	1292	2096	1238	641	1879	1950	1025	2975	3.78	4.71	1.81	1.35
3.	.38	1289	9.15	3.47	1.24	1252	3348	1237	631	1868	3187	1656	4843	3.81	6.06	1.99	1.12
4.	.35	1807	12.22	3.63	1.19	1232	4580	1243	637	1880	4430	2293	6723	3.85	7.18	2.12	0.99
5.	.33	2448	15.80	3.72	1.16	1488	6068	1513	782	2295	5943	3075		4.28	8.35	2.02	.89
6.	.31	3421	20.90	3.82	1.12	2180	8248	2267	1168	3435	8210	4243	12453	5.25	9.87	1.74	.80
7.	.29	3960	23.56	3.95	1.07	1154	9402	1245	625	1870	9455	4868	14323	3.90	10.61	2.50	.76
8.	.27	5153	29.02	4.04	1.05	2482	11884	2708	1348	4056	12163	6216	18379	5.76	12.07	1.79	.71
9.	.25	7159	37.52	4.16	1.00	4012	15896	4473	2227	6700	16636	8443	25079	7.40	14.16	1.48	.64
10.	.23	8670	43.41	4.29	0.94	2840	18736	3279	1617	4896	19915	10060	29975	6.39	15.53	1.84	.61
11.	.21	11125	52.15	4.42	.90	4420	23156	5180	2553	7733	25095	12613	37708	8.05	17.49	1.56	.57
12.	.19	13595	60.10	4.57	.85	4150	27306	5064	2470	7534	30159	15083	45242	7.99	19.23	1.70	.54
13.	.17	15563	65.77	4.71	.79	3110	30416	3838	1889	5727	33997	16972	50969	7.02	20.47	2.10	.53
14.	.15	17505	70.71	4.79	.74	2874	33290	3709	1787	5496	37706	18759	56465	6.84	21.58	2.35	.52
15.	.13	19142	74.31	5.01	.69	2260	35550	2996	1440	4436	40702	20199	60901	6.10	22.43	2.87	.51
16.	.11	22044	79.72	5.17	.63	3656	39206	5137	2408	7545	45839	22607	68446	7.91	23.78	2.48	.50
17.	.09	26176	86.03	5.36	.56	4628	43834	6694	3181	9875	52533	25788	78321	9.10	25.46	2.45	.50
18.	.07	29669	90.17	5.55	.49	3424	47258	5344	2480	7824	57877	28268	86145	7.93	26.67	3.24	.50
19.	.05	34963	94.66	5.80	.41	4342	51600	7147	3388	10535	65024	31656	96680	9.08	28.17	3.43	.50
20.	.03	40534	97.49	6.06	.31	3454	55054	6462	2452	8914	71486	34108	105594	8.32	29.38	4.98	.51
21.	.01	59247	100.00	6.45	.17	6364	61418	16095	7486	23581	87581	41594	129175	11.92	31.71	6.37	.53

TABLE 61. BEST SIZE AND DENSITY OF GRIDS (IN HALF-SAMPLES) AT DIFFERENT LEVELS OF EXPENDITURE FOR *THANA*-ZONES

zone no.	proportion under jute	accumulated: total area in square miles	accumulated: jute area percentage	size of grid x_i	half-density y_i	total number of grids: zone	total number of grids: accumulated	cost in rupees: field	cost in rupees: statistical	cost in rupees: total	accumulated cost in rupees: field	accumulated cost in rupees: statistical	accumulated cost in rupees: total	$\sqrt{v_i}$: zone	$\sqrt{v_i}$: accumulated	$100 \times \sqrt{v_i}/A_i p_i$: zone	$100 \times \sqrt{v_i}/A_i p_i$: accumulated
(1)	(2)	(3.1)	(3.2)	(4)	(5)	(6.1)	(6.2)	(7.1)	(7.2)	(7.3)	(8.1)	(8.2)	(8.3)	(9.1)	(9.2)	(10.1)	(10.2)
1.	0.49	287	2.36	2.70	1.51	866	866	738	399	1137	738	399	1137	2.75	2.75	1.96	1.96
2.	.42	784	5.89	3.06	1.40	1392	2258	1257	666	1923	1995	1065	3060	3.68	4.59	1.76	1.31
3.	.38	1289	9.15	3.34	1.29	1302	3560	1257	647	1904	3252	1712	4964	3.76	5.93	1.96	1.10
4.	.35	1807	12.22	3.50	1.23	1274	4834	1264	647	1911	4516	2359	6875	3.81	7.05	2.10	0.98
5.	.33	2448	15.80	3.63	1.19	1526	6360	1538	789	2327	6054	3148	9202	4.23	8.22	2.00	0.88
6.	.31	3421	20.90	3.72	1.15	2238	8598	2316	1177	3493	8370	4325	12695	5.19	9.72	1.72	0.79
7.	.29	3960	23.56	3.87	1.10	1186	9784	1256	636	1892	9626	4961	14587	3.86	10.46	2.47	0.75
8.	.27	5153	29.02	3.95	1.06	2530	12314	2756	1372	4128	12382	6333	18715	5.71	11.91	1.77	0.70
9.	.25	7159	37.52	4.05	1.01	4052	16366	4554	2226	6780	16936	8559	25495	7.37	14.01	1.47	0.63
10.	.23	8670	43.41	4.23	0.97	2932	19298	3324	1511	4835	20260	10070	30330	6.31	15.36	1.82	.60
11.	.21	11125	52.15	4.37	0.91	4468	23766	5273	2578	7856	25538	12648	38186	8.01	17.33	1.55	.56
12.	.19	13595	60.01	4.42	0.87	4298	28064	5138	2519	7657	30676	15167	45843	7.90	19.04	1.68	.54
13.	.17	15563	65.77	4.64	.82	3228	31292	3956	1948	5904	34632	17115	51747	6.91	20.26	2.07	.52
14.	.15	17505	70.71	4.81	.77	2990	34282	3729	1845	5574	38361	18960	57321	6.71	21.34	2.30	.51
15.	.13	19142	74.31	4.98	.70	2292	36574	3028	1457	4485	41389	20417	61806	6.06	22.18	2.85	.51
16.	.11	22044	79.72	5.14	.63	3656	40230	5166	2437	7603	46555	22854	69409	7.88	23.54	2.47	.50
17.	.09	26176	86.03	5.32	.57	4710	44940	6487	3223	9710	53042	26077	79119	9.01	25.21	2.42	.50
18.	.07	29669	90.17	5.52	.50	3494	48434	5379	2515	7894	58421	28592	87013	7.85	26.40	3.21	.50
19.	.05	34963	94.66	5.76	.41	3442	52776	7253	3388	10641	65674	31980	97654	9.07	27.91	3.43	.50
20.	.03	40534	97.49	6.05	.31	3454	56230	6574	3008	9582	72248	34988	107236	8.32	29.13	4.98	.51
21.	.01	59249	100.00	6.45	.17	6364	62594	16469	7486	23955	88717	42474	131191	11.92	31.47	6.37	.53

163. If we compare the proportional errors given in col. (10.2) in the two sets of tables for the district-zones and the *thana*-zones, it is clear that, with any assigned expenditure, sampling on the basis of *thana*-zones is far more efficient in the sense of giving an appreciably lower proportional error. This of course is what is to be expected from theoretical considerations. The more homogenous the zones the lower will be the margin of error. Obviously, it is possible to demarcate the zones more accurately in working with *thanas* than with districts, this improvement in zoning is immediately reflected in the appreciably lower coefficients of error. It is clear therefore that the zoning by *thanas* must be the basis for our work in 1941.

ZONES WITH LOW INTENSITIES OF CULTIVATION

164. The accumulated proportional standard errors given in col. (10.2) show one curious feature which may be noted here; the value increases in each case in zone Nos. (20) and (21). This suggests that the limit of economic sampling is probably reached in zone No. (19); and that exceptionally thin zones like Nos. (20) and (21) may require special treatment.

165. We may conveniently consider at this stage the general question of zones with very low intensities of cultivation of the order of two or one per cent or less. It is clear that we cannot afford to give a large number of sample-units to such areas. In other words we must work with low values of y, the density of grids per square mile. It will be remembered that we intentionally used certain special field units to collect information regarding the cost of operations for working with sample-units of very low densities. A detailed comparison of time records and standard errors show that units of size 16-acre were definitely more efficient than *mauzas* as sample-units.

166. This definitely excluded the *mauza* or very large units from the sampling technique. If we confine our attention to grids of the size with which have been working we find a comparatively restricted range of choice. It will be remembered that the equation connecting the best size and the best density of sample-units is given by

$$X = 6.912 - 2.762y$$

in which x is the size of the sample-units in acre, and y the number of sample units per square mile. It will be remembered however that the equation involves a factor $g/(1-g)$. We have reasons to believe that the values of g lie between one-fourth and one-third. In our present calculations we have adopted the value of $g = 0.24$ or a value of $g/(1-g)$ of about 0.3; even if g becomes as high as 0.33 values of $g/(1-g)$ will only increase to 0.5. This means that the values of x may possibly increase by say 30 per cent but not more.

167. From the above equation it is clear that when y is very small or nearly zero the largest value of x or the best size of grids will be, say, about 7 acre; adding one-third as the effect of a possible increase in the value of g, we have an upper limit of about 9-acre for the best size of sample-units. On the other hand, with a high density of y, say of the order of 2 in each half-sample, the value of x will become 1.4; this shows that the best size of grids is not likely to be smaller than 1-acre. The results of extensive field operations in 1940 therefore fully confirm our provisional range of from 1 to 9-acre as the most economical size of sample-units. Confining our attention to the range of densities from 0.3 to 1.5 the effective range of the size of sample-units lies between 3 and 6.acres. We find then that there is not much hope of being able to cover the thin areas by simply increasing the size of the grid.

TWO-STAGE SAMPLING

168. A word of caution is also necessary at this stage. The cost function; used by us is based on field operations with densities lying between roughly 0.5 to 1.5; we do not know whether extrapolation beyond these limits will give reliable estimate. In fact, from broad physical considerations it would appear that with low values of the density, that is, with sample-units scattered at great distances apart, a radical change in the type of sampling may prove more efficient.

169. We have seen in a previous section that camps on an average were fixed roughly at distances apart of from 6 to 12 miles. In the present method of sampling the whole of the ground lying in the neighbourhood and betwen different camps were surveyed, we may call this a single-stage sampling for convenience of reference.

170. A second alternative is open to us. We may for example locate a number of camps or centres at distances apart of approximately say 7 miles, or 12.5 miles or 18 miles; in this case we shall get on an average one camp per 50 square miles, 160 square miles or 324 square miles respectively. Remembering that the average size of a *thana* is about 160 square miles, this means that we shall get say three camps per *thana* in the first case, one camp per *thana* in the second case, and about one camp in two *thanas* in the third case. The investigators may be asked to halt at each camp for a number of days, 4 or 5 days or more, and cover a certain area which may be say 5 or 10 or 20 square miles or more in the neighbourhood of each camp with a suitable second-stage size and density of sample-units. In this plan we have got the first-stage density of individual camps which will be very low and of the order of say 1 in 50 square miles, 1 in 160 square miles or 1 in 320 square miles. In the neighbourhood of each camp we have also a second-stage density of grids of usual type of the order of say 0.5, or 1 etc.

171. In this method we do not make any attempt to cover the whole area within each zone with one uniform density of sample-units. We rather try to break up the region into a suitable number of circles surrounding each camp; and carry out a more intensive survey within the restricted area of these circles. This is what I have called the two-stage method of sample survey. From physical considerations it appears to me that in the case of large areas of ten or twenty thousands of square miles with extremely low intensities of cultivation it may be a better plan to adopt this two-stage technique. This is what I am recommending in general terms for the programme of work in 1941. That is, our plan will be to cover about 35,000 or 40,000 square miles by the present method of one-stage sampling, and the remaining 15,000 or 20,000 square miles by the two-stage method.

172. From col. (10) of Table 51 it will be seen that forty thousand square miles upto zone No. (20) proposed to be surveyed by the one-stage sampling technique cover about 97.5 per cent of the total area under jute. The remaining twenty thousand square miles contribute only two and a half per cent to the total area under jute. Even if the percentage error in this portion is comparatively high, its contribution to the total error will be quite small; for example, if we can estimate the acreage under jute in this twenty thousand square miles with such a high percentage of proportional error as 20 per cent, the contribution to the error of the total estimate will not be more than half of one per cent from this zone.

CHAPTER 7 : PROGRAMME OF WORK FOR 1941

Question of policy : future plan

173. We have now to consider the programme of work for 1941. This, however, cannot be settled purely on statistical or mathematical considerations for questions of policy must inevitably play a large part in deciding the general nature of the plan. It will be remembered, for example, that during the last three years provision was made in the budget for preparatory work for the next season. Obviously if it is intended to continue the sample census it is necessary to make a similar provision for 1942. Actual requirements will be comparatively small, because the heavy expenditure for maps will not be required any longer, but arrangements will have to be made for the preparatory work in the statistical section to start by December 1941. On the other hand, if the scheme is to be definitely terminated after the sample survey of 1941, there is no need of making any provision for preparatory work for 1942; and the whole of the available amount may be utilized for the work in 1941.

Date of preparation of estimate

174. The second important question is the date by which the forecast should be made ready in 1941. The calculation of p for individual grids requires the previous measurement of the area of individual plots falling within the grid. Up till now, for reasons of economy, we have measured the area of only those plots which were reported to be under jute. The cost shown in Tables 54—61 were prepared on this basis. It was however decided by the Jute Census Committee in September 1940 that the area of all plots included within the sample-units should be measured beforehand in order to expedite the preparation of the actual forecast in 1941.

175. We may now consider the additional expenditure which will be required for this purpose. The average number of plots per sample-unit has been already given in Table 13; these were 6.56 in the case of 1-care, 11.36 for 2.25-acre, 15.89 for 4-acre, 22.62 for 6.25 acre, and 28.85 for 9-acre grids. We also know from laboratory records that the area of about 33 plots can be measured per hour on an average. The time required for area extraction of all plots irrespective of the fact whether jute is grown on them or not is thus 0.219 man-hour for 1-acre, 0.379 man-hour for 2.25-acre, 0.530 man-hour for 4-acre, 0.754 man-hour for 6.25 acre, and 0.962 man-hour for 9-acre sample-units.

176. We have, however, already included in our calculations the time required for measuring the area of the plots on which jute was grown. Deducting the portion already included, and smoothing the material, we get the following linear equation for the additional time required for full extraction ;

$$t = 0.07 + 0.06(x) \text{ man-hours per grid}$$

where x is the size in acre of the sample-unit. With sample-units of size 3, 4, 5, or 6-acre, and densities ranging from 0.5 to 1.0, the additional cost will come to something of the order of 0.3 or 0.4 rupee per square mile which is by no means negligible.

177. In this situation a compromise is possible. We may arrange full extraction in the heavy zones in which the jute proportion is large; but measure only the plots under jute in thin zones which will enable substantial savings being made.

178. Everything depends on the date by which the forecast is required. It is of the utmost importance therefore for the Jute Census Committee to come to some decision on this point. Once this is known it will be possible to calculate the additional cost for area extraction. This additional expenditure will have to be included in the grant for statistical work as it will be more convenient and more economical to start measuring the area of individual plots along with the preparatory work.

Crop-cutting experiments

179. Thirdly, there is the question of crop-cutting experiments. I have always emphasized the need of improving the estimates of the yield of jute per acre *pari passu* with the improvement in the forecast of area under jute. I did not know however whether it was intended to include the scheme for crop-cutting experiments in 1941 within the total allotment of Rs. 1,80,000 provided in the budget for the present scheme. I had therefore prepared a tentative programme without taking into consideration the question of crop-cutting experiments. However on the 5th November 1940 I received Mr. D.L. Mazumdar's D. O. letter No. 2859/40 dated the 4th November 1940 in which I was asked to include crop-cutting experiments in the programme for 1941. This necessitates radical changes in the budget. Personally I am strongly in favour of Mr. Mazumdar's proposal; but here also various considerations of policy are involved which it is not possible for me, as a technician, to decide.

180. The most important question is whether the work on crop-cutting proposed to be carried out in 1941 is to be considered as an integral part of a scheme extended over say, three years, or it is to be considered as complete by itself. From the statistical point of view an extended programme is certainly desirable, as this will enable a sound technique being evolved on the basis of gradually expanding exploratory work. This is what is being done in the case of sugarcane and wheat in an U.P. Scheme financed by the Imperial Council of Agricultural Research, and in the case of cotton in an all-India Scheme financed by the Indian Central Cotton Committee.

181. If the idea of an extended scheme is approved I should suggest providing something of the order of twelve thousand rupees for crop-cutting experiments in 1941. (I may mention here that this is the order of the expenditure proposed to be incurred in the first year in the sugarcane scheme in U.P.). On the other hand if it is intended to have only a single-year scheme it will be obviously necessary to spend a good deal more than this amount to collect adequate information for a total area of the order of sixty thousand square miles.

Zones with low intensities of cultivation

182. Then there is the question of zones with low intensities of cultivation. I should suggest that about nine or ten thousand rupees should be spent in two-stage sampling to cover about fifteen or twenty thousand square miles. I have already

explained the basic idea underlying this method; details will naturally have to be worked out after the demarcation of zones. I may mention here that in preparing the design for the crop-cutting experiments on Jute this year I had used a two-stage method; the time-records for crop-cutting experiments may therefore be expected to furnish a good deal of useful basic information required for planning a similar two-stage sample-survey for jute areas. A portion of the crop-cutting experiments on sugarcane in the United Provinces has been designed in accordance with the two-stage plan; primary records of these surveys will be shortly available and will naturally supplement the information obtained from the work on jute.

183. I propose therefore to prepare the detailed design for the two-stage sample survey of the area under jute after I have had the opportunity of studying the time-records of similar two-stage work on jute, and possibly sugarcane. I hope, however to get the programme ready in time for its being considered by the Jute Census Committee in March or April 1941.

THE MAIN AREA-CENSUS

184. We now come to the main item, namely the grid-sampling of the area under jute in the whole province of Bengal. It is necessary in this connexion to remember the possibility, or rather the certainty, of the intensity of cultivation decreasing appreciably either through administrative action or through the operation of ordinary economic forces.

185. The best size and density of grids as well as the total cost and margin of error depend on the actual values of the proportion of land under jute in different zones. If there is any change in the value of p (the proportion of land under jute) and $q\ (=1-p)$ in all or many of the zones it is inevitable that the margin of error and other calculations will have to be modified. Broadly speaking it may be stated that the margin of error will increase if the intensity of cultivation decreases.

186. I understand that there is a possibility of the cultivation of jute being compulsorily restricted in 1941 by administrative orders. In case this is done, or in case the intensity of cultivation decreases through other causes, the margin of error will inevitably increase. In the absence of any knowledge regarding the magnitude of the fall in the values of p it is not possible to make an accurate forecast of the change in the margin of error. It will be obviously necessary to keep a little reserve in hand. This is why I consider it desirable to use slightly higher densities and a wider range of sizes of sample-units in order to be on the safe side. I have prepared the plan on the basis of standard size of grids of 4-acre and 5-acre together with a sprinkling of grids of 3-acre and 6-acre in the outlying zones. As regards densities, I have used for the bulk of the work the values 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 with a sprinkling of 0.3 and 1.0 for each half-sample. Taking everything into consideration it appears to be desirable to spend on the main area-census something like a lakh and a quarter of rupees inclusive of the expenditure for both field and statistical work but exclusive of the cost of maps, supervision and other overhead. A tentative scheme has been prepared on this basis and is shown in Table 62. This question is however intimately linked with the plan for the future. If provision has to be made for preparatory work for 1942 then we may have to do with a lower figure; this question has been discussed under budget provision.

187. About Rs. 5,000 should be held in reserve for unforeseen expenses and for preparing a consolidated report on the whole scheme. We have already spent about Rs. 2,70,000 in the first four years of the present scheme. If another Rs. 1,80,000 is spent next year the total expenditure will come to about Rs. 4,50,000. In case it is decided to terminate the scheme it will be obviously desirable to have a full, connected and critical account of the work preserved in a permanent form. This will be helpful not only in dealing with the question of sample surveys in general but will enable a sample census of jute being undertaken in future without any difficulty. The writing of this Final Report will take at least three months, possibly four or five; and necessary staff will have to be provided during this period. The amount spent for this purpose will be fully worthwhile. On the other hand, a small saving here is likely to prevent the results achieved in the present scheme being profitably utilized in future.

THE BUDGET POSITION

188. It will be remembered that the total allotment for the fifth season of the present scheme, that is, for the provincial survey in 1941, is one lakh and eighty thousand rupees. To this may be added the savings in the budget for 1939-40 amounting to about Rs. 7,000, I believe. This gives us a total figure of Rs. 1,87,000. Out of this amount Rs. 7,500 was ear-marked for zoning and model sample experiments and Rs. 7,500 for experiments with optical arrangements for using micro-film photographs of mauza maps at the meeting of the Jute Census Committee held on the 5th September 1940. An expenditure of about Rs. 13,000 has also been incurred for the purchase and collection of maps. I have also arranged with the Chief Controller of Registration, Bengal, to supply us with the original records of the jute registration in 1939. He has informed me that freight and forwarding charges would come to about Rs. 1,000. This gives a total of Rs. 29,000 already spent or ear-marked.

189. We may next consider the overhead expenses. For the statistical work it is usual to set aside Rs. 8,000; and we may adopt the same figure. For the Field Branch we have seen that the overhead expenses amounted to Rs. 6,743-12-0 in six months or Rs. 1,125 approximately per month. It has been suggested that the Field Branch should start work from the beginning of February of 1941 and continue till the end of September, that is, for 8 months. On this basis we have to provide for overhead expenses for eight months at the rate of Rs.1,125 per month which comes to exactly Rs. 9,000. I have proposed that Rs. 12,000 should be provided for crop-cutting experiments in 1941; and Rs. 5,000 should be held in reserve for unforeseen expenses and for writing the report.

190. The total of all the above amounts comes to Rs. 63,000. Deducting this we have an available balance of Rs. 1,24,000. In case it is desired to make any provision for preparatory work for 1942 it will be necessary to allot, say, ten or fifteen thousand rupees for this purpose depending on the scale of operations. (This should be sufficient as no heavy expenditure on account of maps or preliminary field work will be required next year). This will leave one lakh and ten or fifteen thousand rupees available for the sample survey in 1941. If no provision has to be made for 1942 then the whole of Rs. 1,24,000 may be utilized for this purpose. The abstract budget has been prepared on this basis and is given in Table 63. Necessary adjustments will be made after the questions of administrative policy discussed in this chapter have been decided by the Jute Census Committee.

TABLE 62. PROPOSED DISTRIBUTION OF SIZE AND DENSITY (IN HALF-SAMPLES) FOR WORK IN 1941

serial number of zones	proportion under jute	accumulated: total area in sq. miles	accumulated: jute area percentage	size of grid (x)	density of grid (y)	total no. of grids: zone	total no. of grids: accumulated	cost in rupees: field	cost in rupees: laboratory	cost in rupees: total	accumulated cost in rupees: field	accumulated cost in rupees: laboratory	accumulated cost in rupees: total	remarks
(1)	(2)	(3.1)	(3.2)	(4)	(5)	(6.1)	(6.2)	(7.1)	(7.2)	(7.3)	(7.1)	(8.2)	(8.3)	(6)
										Rs.	Rs.	Rs.	Rs.	
1	0.49	287	2.36	4.0	1.0	574	574	721	410	1,131	721	410	1,131	It is proposed to include 20 zones in one stage sampling scheme at a total cost at about Rs. 1,15,000; and to spend about Rs. 10,000 for two-stage sampling in zone no. 21. The total expenditure is to be about Rs. 1,25,000.
2	.42	784	5.89	4.0	1.0	594	1168	1247	587	1,834	1,968	997	2,965	
3	.38	1289	9.15	4.0	1.0	1010	2178	1268	722	1,990	3,236	1,719	4,955	
4	.35	1807	12.22	4.5	0.9	932	3110	1249	705	1,954	4,485	2,424	6,909	
5	.33	2448	15.80	4.5	0.9	1154	4264	1545	873	2,418	6,030	3,297	9,327	
6	.31	3421	20.90	4.5	0.9	1751	6015	2345	1325	3,670	8,375	4,622	12,997	
7	.29	3960	23.56	4.5	0.9	970	6985	1299	734	2,033	9,674	5,356	15,030	
8	.27	5153	29.02	4.5	0.9	2147	9132	2875	1625	4,500	12,549	6,981	19,530	
9	.25	7159	37.52	5.0	0.8	3210	12342	4514	2455	6.969	17,066	9,436	26,499	
10	.23	8670	43.41	5.0	0.8	2418	14760	3445	1953	5,398	20,508	11,389	31,897	
11	.21	11125	52.15	5.0	0.8	3928	18688	5598	3171	8,769	26,106	14,560	40,666	
12	.19	13595	60.10	5.0	0.7	3458	22146	5212	2810	8,022	31,318	17,370	48,688	
13	.17	15563	65.77	5.0	0.7	2755	24901	4153	2239	6,392	35,471	19,609	55,080	
14	.15	17505	70.71	5.0	0.6	2330	27231	3729	1930	5,659	39,200	21,539	60,739	
15	.13	19142	74.31	5.0	0.6	1964	29195	3143	1627	4,770	42,343	23,166	65,509	
16	.11	22044	79.72	5.5	0.5	2902	32097	5078	2583	7,671	47,421	25,749	73,170	
17	.09	26176	86.03	5.5	0.5	4132	26229	7231	3678	10,909	54,652	29,427	84,079	
18	.07	29669	90.17	5.5	0.4	2794	39023	5379	2550	7,929	60,031	31,977	92,008	
19	.05	34963	94.66	6.0	0.4	4235	43258	8259	4044	12,303	68,290	36,021	1,04,311	
20	.03	40534	97.49	6.0	0.3	3343	46601	7465	3331	10,796	75,755	39,352	1,15,107	
21	.01	59249	100.00	6.5	0.15	5614	52215	18528	6886	25,414	94,283	46,238	1,40,521	

TABLE 63. JUTE CENSUS SCHEME : DRAFT BUDGET FOR 1940-41

(A)	Field Branch				
	(1) Price of Maps	Rs. 11.000			
	(2) Cost of collection	2,000			
				Rs.	
	Preliminary expenses			13,000	
	(3) Overhead (Supervisor, Asst. Supervisor, Field Office, contingency)			9,000	
	(4) Staff for field work (including T. A.)			82,000 (A)	
					Rs.
	Total Field Branch				1,04,000
(B)	Statistical Branch (Block grant)				
	(5) Zoning and Model Sampling	7,500			
	(6) Photographic experiments	7,500			
				Rs.	
	Special work (non-recurring)			15,000	Rs.
	(7) Statistical work for Area Census			42,000(B)	57,000
(C)	General				
	(8) Freight charges for Jute records			1,000	
	(9) Overhead (Statistical Adviser etc.).			8,000	
	(10) Provision for Crop-cutting work			12,000	
	(11) Reserve			5,000	26,000
	Grand total				1,87,000

N.B.—The actual budget for the area-census (exclusive of non-recurring and overhead charges) is given by the sum of item (4) Rs. 82,000 for field work and item (7) Rs. 42,000 for statistical work or Rs. 1,24,000 altogether.

SUMMARY OF PROGRAMME FOR 1941

191. We may now summarize the chief features of the programme of work for 1941 as described above :

(1) In the main area-census the sampling technique will be arranged, as in previous years, in the form of two half-samples, information for which will be collected by entirely different sets of investigators. One of our important objects will be to find out, by comparing the results for the two independent half-samples, with what effective accuracy of a sample census carried out on a full provincial scale throughout Bengal. We believe that the margin of error for the full provincial survey in 1941 will not exceed the limit of five per cent as settled by the Jute Census Committee. The calculation of the error may however be rendered uncertain in case the intensity of cultivation decreases appreciably. We have kept a certain amount of reserve in hand as a provision against this contingency; and it is hoped that results of the two half-samples will be in sufficiently close agreement to enable the field estimate being used for practical forecasting purposes.

(2) Secondly, we intend to gather information regarding variations in the cost of field operations in different zones and districts so that it may be possible to settle on an objective basis the most economical size and density of grids in different zones.

(3) Another important object is to study, by actual experimentation, the efficiency of a two-stage sampling technique in areas in which the proportion of land under jute is very low and is of the order of two or one per cent or less.

(4) It is also intended to study as far as possible the most economical organization and distribution of field units, and related questions of inspection and supervision in different zones.

(5) It is proposed to carry out crop-cutting experiments on a moderately large seale, as a part of an extended scheme, with a two-fold object : (a) to determine the yield of jute per acre under actual conditions of cultivation; and (b) to collect basic information required for developing an efficient technique for this purpose.

It is intended in this connexion to study how the crop-cutting work can be best co-ordinated and integrated with the work of the area census.

(6) The question of zoning will be studied in detail with the help of maps and available records of the intensity of cultivation. I have already made arrangements with the Chief Cntroller of Registration, Bengal, for securing the original primary records of the registration in 1939. Freight charges for transferring the records from district headquarters to Calcutta and cost of storage is estimated at about Rs. 1,000 which is being provided in the budget.

(7) Provision has been made for model sampling experiments for further study of the variance function. We have seen that the value of the variance parameter g is likely to lie between 0.24 and 0.33; there is reason to believe, however, that g increases as the intensity of cultivation is decreased. In order to improve efficiency of planning it is necessary to study the question in greater detail; and, if possible, to find out the functional relation between g (the variance parameter), and p (the portion of land under jute). Model sampling experiments are proposed to be undertaken for this purpose.

(8) Experiments have been started to find out whether an optical method can be used for the preparatory work in the Statistical Laboratory. Microphotographs of mauza maps have been already taken on cinema films. Attempts are now being made to project these microphotographs on a screen, and use the screen image directly for locating and marking the position of grids and also for measuring the area of plots included within the grids. It is hoped to settle this question one way or the other in the course of next year's work.