A SAMPLING SURVEY OF TAPPING ON SMALL HOLDINGS. (1939—40.)

BY

H. FAIRFIELD SMITH

I. History of the Survey

The project for a survey of small holdings was proposed in May 1938 at a meeting of the Small-Holders' Rubber Advisory Committee. Mr. A. Moore,* until May 1938 officer-in-charge of the Small-Holders' Advisory Service, in reporting on a visit to Kelantan and Trengganu "had expressed grave doubts as to the future of small holdings generally because of the improvident way in which the small-holder taps." It was pointed out that the survey done in 1931-32 (MEADS, 1933) "indicated that on the 'typical' small holding there were tappable reserves of bark enough to last nearly 7½ years, by the end of which time the natural renewal would have added further reserves." Nevertheless it was felt that, because the previous survey was based on holdings subjectively selected as typical of their areas and reasonably accessible by road, it might have failed to tell the whole story in that (1) it "gave no indication of the *variation* of bark reserves" among small holdings generally; and (2) it provided no objective criterion of the reliability of its estimates. Principally for the former reason, it being now considered important to obtain an estimate of the proportion of holdings approaching a state of being "tapped out" and only fit for replanting, it was decided to carry out a new and broader survey.

At a subsequent meeting bark reserves were defined for the purpose of this survey as area of bark thicker than 5 mm.

* Died in Thailand, December 1943.

and within 5 feet of the ground. Further it was requested that information be obtained on the thickness of bark being tapped by small-holders and the heights to which they were working.

The field work was done between March 1939 and March 1940 by the Rubber Instructors of the Rubber Research Institute under the direction of Mr. R. H. Meikle,† the Small-Holders' Advisory Officer, who reported progress in the *Annual Reports* (1938, p. 203; 1939, p. 256; 1940, p. 154). With the acquisition of punched-card machinery in immediate prospect it was uneconomic to attempt reduction of the data by hand. The war in Europe delayed receipt of the machinery for nearly a year, and with incidence of embodiment periods it was necessary to give precedence to the more regular work of the Institute, so that the detailed analysis of this work could not be taken up until the middle of 1941, although in the meantime a rough survey of the observations had indicated the general conclusions to be drawn with respect to re-planting advisory work. By November 1941 all the data had been put on cards, a preliminary study of the variation in the data had been carried out to determine the procedure for final analyses, reports on the frequency distributions of the variates had been written, and the principal data had been concentrated on four sets of summary cards. At that stage the Japanese interfered. During the ensuing four years reports, data schedules, the master code for card punchings and machinery were all destroyed. However most of the cards survived, except (unkind chance!) the principal set of summary cards, and it has been found possible to reconstruct most of the code. This report is based on hand-sorting of three secondary sets of summary cards. Although some questions cannot feasibly be answered without machinery for sorting the main bulk of cards, it has been possible to deal with most of the surviving data. As will be seen, some of the most interesting data have been either irretrievably lost or cannot now be correlated with other characters.

Much of the information is now out of date. The principal object of the present analysis is to pave the way for efficient design of any future survey, and Sections II, IV and V are directed to that purpose. Readers interested only in conclusions about the condition of small holdings need read only Section III, although supplementary information will be found in the other sections.

† Killed in action, January 1942.

The survey covered only small holdings of less than 25 acres.

II. Sampling, Field work and Records

Since the 1931-2 survey was not statistically controlled no information was available on the magnitude of variability likely to be encountered in the various strata to be sampled (States, districts, mukims, holdings, trees etc.). Consequently the numbers sampled at each stage represent only advance guesses for the distribution of available resources (*cf.* Sec. V), and were of course made before methods of sample surveys had been developed as they have been during the past decade.

For a general review of the principles of sampling, and the reason for describing procedure in detail, see YATES (1946). This survey used geographically stratified sub-sampling—stratified for States, sampled for districts and sub-sampled for holdings. The sampling fraction at each stage was variable, being determined by administrative convenience.

Twenty-five Rubber Instructors (R.I.) were available, and to each was assigned 20 holdings from the records of one land office (district) in his normal field of work. So far as can be recollected the intention of the original plan was to sample the whole of an Instructor's advisory district. Whether arguments of expediency were brought to bear, or whether there was a slip in carrying out the sampling plans at this stage, is not known; but it appears that the district assigned to each R.I. was the one in which he resided (whereas either a random sample of districts should have been taken—stratified sub-sampling, or each R.I's task should have been spread over the whole of his advisory district—stratifying districts as well as States). Except for this defect all other sampling is believed to have been done correctly. Mr. Meikle visited each land office, and by the usual procedure with a table of random numbers, selected 20 holdings of less than 25 acres from those in the books of the office. For each holding the lot number, mukim, kampong, owner's name and area of holding were entered on a prepared schedule. Seventeen of these have survived (including one for a district of Johore which was not subsequently surveyed); so that for 340 holdings we have the areas and names of owners. From the name can be deduced the race of the owner in the four main categories—Malay, Chinese, Indian, European. (The original classification showed eleven nationalities). Owing to

the loss of the code book, these data can be correlated with the other observations in only three districts (60 holdings).

Mr. Meikle accompanied the Rubber Instructor to one of the holdings to give personal instruction in sampling, full instructions for which the observer also had in writing. On another schedule, prepared for ease of tabulation, was recorded, partly from conversation with the owner, and where possible from observation:—

Nationality of owner
By whom tapped (owner, family or employee)
Nationality of tapper
Confirmation of area
Age of rubber
Number of trees per acre (estimated from average distance between trees)
Type of cultivation—cover crops
System(s) of tapping
Proportion of trees in tapping (as estimated by the observer)
Incidence of bad wounding
Incidence of mouldy rot
Incidence of root disease.

Except for one copy which was kept by the observer, these schedules have been destroyed, and the data from them (except systems of tapping) was punched only on the principal summary cards of which only 39 survive. Most of these data are therefore lost.

The next step was to select a sample of 21 trees. Using a pack of playing cards with face cards removed three random points in the holding were determined in the usual manner, detailed instructions being given as to how this was to be done. Measurements were then made on seven consecutive trees in a specified direction from each selected point, and if this led to the edge of a holding before seven trees were obtained further procedure was again detailed.

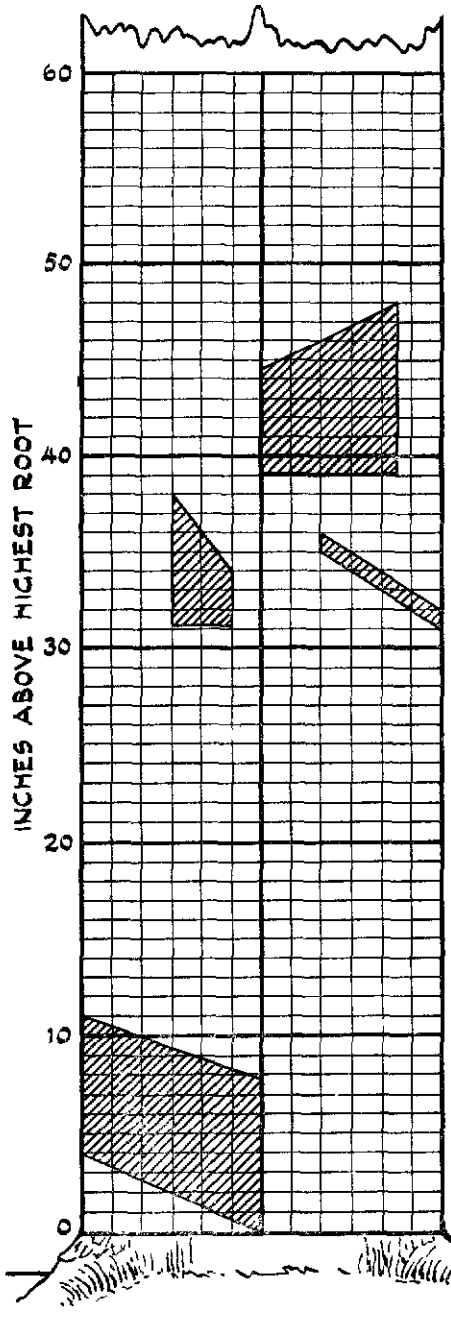
The measurements taken on each of the twenty-one trees were:—

(b) thickness of bark at three points along each tapping cut, in mm,

(h_m) maximum height at which tapping had ever taken place, in inches,

(a) the area of thin bark under 5 mm. in thickness, and below 60 inches,

(g) girth, at 60 inches above highest lateral root, to the nearest inch.



Date of Observations .....
A. R. I. .....
District .....
Mukim ....
Lot No. .....
Sample No. .....
Tree No. .....
Girth at 60 ins. (5 ft)
.....inches
Maximum height
ever tapped .....inches
Thickness of Bark
now in tapping
.....millimetres

The DIAGRAM represents the whole surface of a tree up to 60 inches (5 feet) above the highest root. This surface is divided vertically into 12 equal sections. Thus, for example, a one-third circumference cut extends over 4 of these vertical sections.

The horizontal lines represent intervals of one inch, so that each of the small rectangles on the diagram represents *one vertical inch* of bark over one-twelfth of the circumference of the tree.

The shaded areas represent untappable bark which is less than 5 millimetres thick.

(For use in Institute)
(a) No. of small rectangles **unshaded** .....
(b) *No. of vertical inches* of tappable bark on whole circumference, below 60 inches and thicker than 5 millimetres, equals
(a) ÷ 12 = inches

Fig. I—Tree diagram used to record observations.

The thickness of bark was taken with a Schlieper's gauge. Girths were measured with a cotton inch-tape; heights with a 5-feet pole marked in inches.

These observations were recorded on forms, one for each tree, as illustrated in Fig. 1, which explains the method and units for recording bark areas. The unit described on the figure as " Number of vertical inches on whole circumference below 60 inches " will henceforth be referred to simply as " circumference inches "—abbreviated ⊙". All subsequent references under the code letter *a* are to areas of *thin* bark in ⊙" per tree. (*cf.* remarks in Sec. III, 4).

From these tree diagrams there was deduced in the laboratory, in addition to the observations recorded above, the number, height (h_t) length (l) and type (spiral or V) of cuts in tapping.

These forms have all been destroyed, but the data from them were recorded on punched cards, henceforth called tree cards, which survive, possibly complete. (An exact check of the full number surviving has not been possible because during the Japanese occupation they have been mixed with cards for other experiments).

Some peculiarities in the data presented below, for example the use of geometric means, are due to the loss of the principal summary cards and reconstruction from three other sets which were designed for exploratory studies in variation and methods of analysis. These carried most of their data as logarithms, with standard deviations, ranges and other measures of variability. It would be tedious to detail them and the steps taken to reconstruct the main data. Should anyone want further information it can be obtained from the Institute.

One R.I. completed only nine out of his assigned 20 holdings, and in another district three holdings were found with either seven trees or none—hence the number of holdings on which most statistics are based is 486.

Samples were taken in the following districts:

State	Districts sampled	Number of Districts in which rubber small holdings occur: Sampled	Total
Malacca ..	Central ..	1	3
Prov. Wellesley	Butterworth ..	1	3
Perak ..	Batang Padang, Kuala Kangsar, Parit, Lenggong, Larut ..	5	15
Selangor ..	Ulu Langat ..	1	8
Negri Sembilan	Seremban, Tampin, Rembau	3	6
Pahang ..	Raub, Temerloh, Pekan ..	3	6
Johore ..	Kota Tinggi, Johore Bharu, Pontian, Kluang, Batu Pahat ..	5	8
Kedah ..	Kuala Muda, Kulim ..	2	6
Kelantan ..	Kota Bharu, Ulu Kelantan	2	5
Trengganu ..		?	2

No clue has been discovered to the code for districts; therefore, while the cards can be grouped by districts it is not now known which actual geographical district within a given State is represented by any particular group of 20 cards.

III. General Survey of Information Obtained

Owing to the heterogeneity of observations, and to some of the characters (particularly area of thin bark and heights of cuts) being abnormally distributed, efficient statistical analysis presents many problems, some of which will be considered in Section IV, where will be indicated the way in which averages and standard errors have been calculated. In this section we attempt to present a general picture of the variation of each character over the country as a whole. Estimates of averages for each State and for the country, with standard errors to indicate the reliance which can be placed on each estimate, are shown in the margins of Tables I to XI.

While other technical details are deferred, one point should be noticed here. For nearly every character there are highly significant differences between States, and also significant, albeit smaller, differences between districts within States. In order therefore to obtain unbiased estimates for the averages of the whole country, the means for each state require to be weighted in proportion to the holdings in each. District means should also receive analogous weightings, but

as it is not known to which each refers this cannot be done. Both this and the failure to chose districts at random may introduce some bias, but it is not likely to be serious.

Two systems of weighting are possible according as we choose to regard the individual holding (or the owners of holdings) or the acreage (*i.e.* the power of small-holders to produce rubber) as the units of interest. Nearly all statistics of rubber growing are drawn up in terms of acreage, and such figures are easier to obtain. So in this we follow custom, and the figures presented (except races of owners and sizes of holdings) are weighted according to the area of sub-25-acre small holdings in each State, as published in the *Rubber Statistics Handbook*, 1939 (data for December 1938).

Within a district each holding is perforce given equal weight; but such evidence as is available indicates no correlation of any of the observations with size of holding. Characters may vary with race of ownership, but by the *method of sampling each race is automatically represented in* due proportion.

The estimates of ownership and of sizes of holdings, in so far as they may be taken to represent state averages, should be regarded as merely tentative; destruction of records of these characters having left only meagre samples for most states.

(1) ***Ownership***: The proportions of small holdings (by number, not acreage) owned by members of the three principal races were:

Malay	about	75	per cent.
Chinese	,,	20	,, ,,
Indian	,,	5	,, ,,

There were considerable differences in the proportions owned by Malays in different States; Trengganu and Kelantan being predominantly Malay. (Table I).

(2) *Size of Holdings*: The average size of Chinese holdings was 6.3 acres, of Malay holdings 3.2 acres. The distribution of sizes is shown in figure 2. All eleven holdings under one acre were in Kedah, Kelantan and Trengganu. Otherwise distributions for each state are roughly similar, albeit with differences in average size. (Tables II and III)

(3) *Girths*: Figure 3 shows the general distribution of mean girth per holding. Differences between averages for each district and State are highly significant (Sec. IV 3); Kelantan and Trengganu had a high proportion of holdings with girth less than 18 inches, whereas almost all holdings in Malacca, Province Wellesley, Selangor, Negri Sembilan and Kedah had girths greater than 18 inches. (Table IV)

FIG. 2.

HISTOGRAMS OF AREAS OF SMALL HOLDINGS.

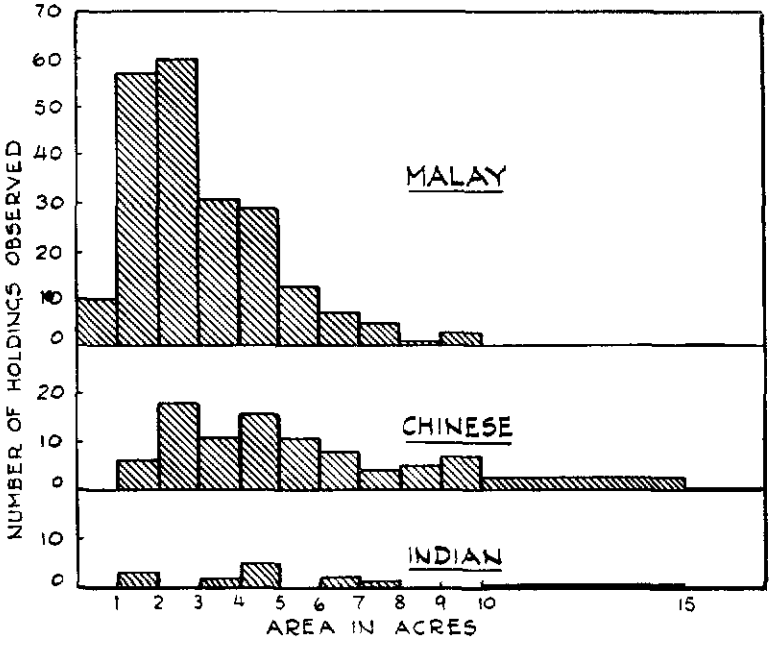


FIG. 3.

HISTOGRAM OF MEAN GIRTHS PER HOLDING.

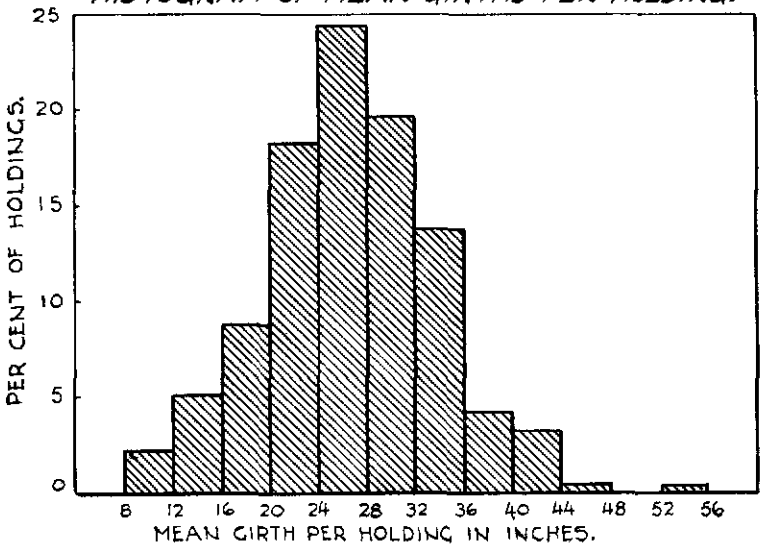


FIG. 4.

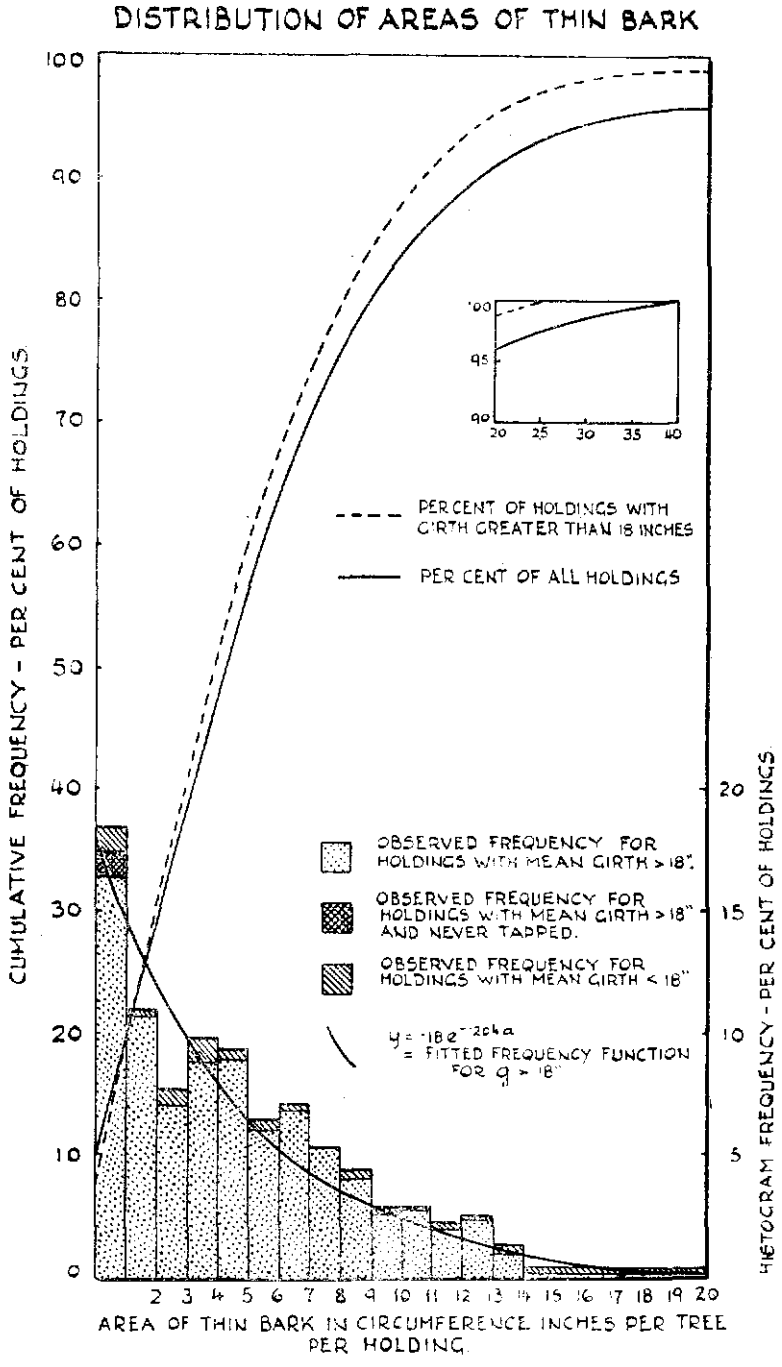


Fig. 4—Distribution of areas of thin bark. Histogram and Cumulative frequency curve.

(4) *Bark Reserves*: The histogram of figure 4 shows the distribution of areas of thin bark (thinner than 5 mm.), and the upper curves show the proportion of holdings having less than any specified area. There are appreciable differences in the averages for each district and State. (Table VI)

Ten per cent of holdings had trees averaging less than 18 inches girth. Many of these had large areas of thin bark (Tables V and VI). Presumably this is mostly virgin bark suppressed by dense planting, but owing to loss of data on density of stands the presumption cannot be checked. Such holdings may improve with growth assisted by judicious thinning, but a proportion (perhaps, the three per cent with a greater than 15, or 4½ per cent with a greater than 10) may represent holdings which threaten to be failures. To these may be added another one per cent which also had *much virgin bark and girths between 18 and 21 inches.* Of the remaining 89 per cent of mature holdings (girth greater than 18 inches) it is indicated that

13 per cent had more than 10 circumference inches of thin bark,
4.7 „ „ „ 15 „ „ „
1.7 „ „ „ 20 „ „ „

The interpretation of these figures is however doubtful. The formal definition of bark reserves—*viz.* area below 60 inches height with bark thicker than 5 mm. (= 60 − a)—was intended to indicate reserves available on methods of tapping customary on small holdings. A height of 60 inches was chosen as representative of the height to which small-holders work, and this supposition proved to be correct (figure 5). But owing to the marked decrease in yield of seedling trees with increased tapping height a tapping system which necessitates going above 36 to 40 inches is not the most satisfactory. *It is now considered that more important* information would be what holdings could work below 40 inches.

The criterion of bark thickness was adopted on the assumption that small-holders were prepared to tap bark as thin as 5 mm., but the evidence indicates that in 1939 they were not tapping bark thinner than 6½ mm. (Sub-sec. 7, below. *cf.* Meads, 1933, Sec. V). Much bark thicker than 5 mm. is immature.* Such bark may nevertheless

---

* Meads, using a different type of gauge, observed that renewed bark 4 months after tapping was 3 to 4 mm. thick; at 16 months, 3.8 to 4.4 mm. (these figures being averages per state). A recent experiment on a small holding showed that it may reach 5 mm. (as measured by a Schlieper gauge) at 5 months after tapping.

be counted as reserves in the sense that it is likely to become tappable. But the criterion as used in the survey is defective in failing to distinguish between this immature but promising bark and bark which may never become better than 5-6 mm. It also fails to distinguish between thin virgin and renewed bark; and the value of the observations has been diminished by loss of data on wounding and disease which may seriously reduce effective bark reserves.

It is difficult to find a basis for comparison with the results of the previous survey (Meads, 1933), where bark reserves were defined as bark *immediately* tappable up to a height of 72 inches. Judgement of tappability depended both on maturity, assessed by hardness, and on variable thickness determined by observation of custom in each locality. For the 90 selected holdings the average bark reserve on this criterion was 36 circumference inches.

Both surveys overestimate available bark in that they include as reserves bark on poor yielding trees which holders may seldom or never tap.

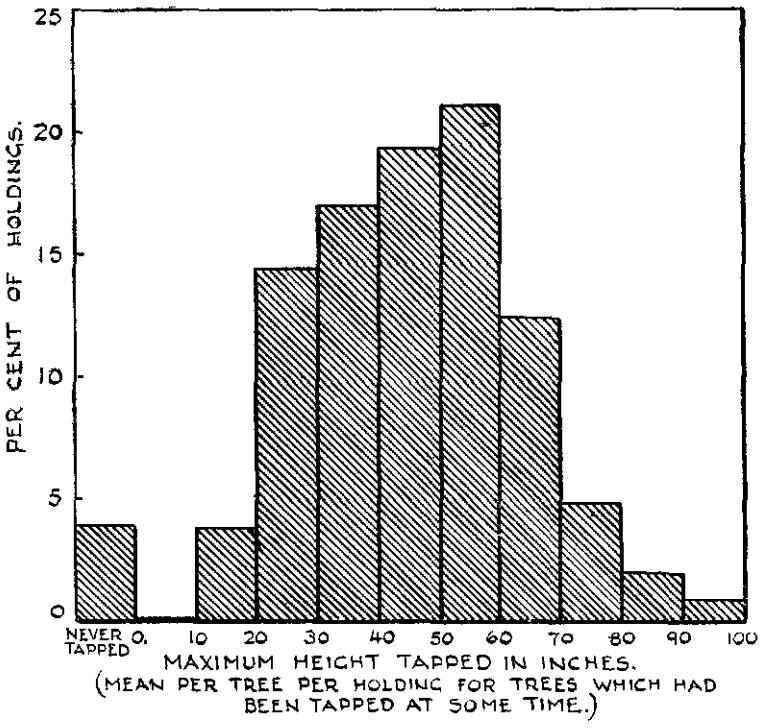
Perhaps the most that can be said is that it is indicated that the selected holdings studied more intensively by Meads do seem to have been fairly typical. Beyond that, if we assume that areas of thin bark above 40 inches might balance untappable bark thicker than 5 mm. below that level, it might be a reasonable guess to take (40-*a*) as representing roughly the amount of bark available for a complete tapping cycle at reasonable heights. On that basis 8 per cent of mature holdings had less than 28 circumference inches available. The 1.7 per cent of holdings with *a* greater than 20 circumference inches might be in danger of extinction from lack of tappable bark, and to them is to be added an unknown proportion of holdings (perhaps between 1 and 4 per cent) with permanently thin virgin bark.

For correlations of areas of thin bark with other characters see Sec. IV 4.

(5) *Maximum heights tapped*: Data at present available deal only with the average maximum heights tapped per holding, not with the maximum heights on individual trees. The distribution is shown in Figure 5 and Table VII. Averages per district differ considerably.

FIG. 5

HISTOGRAM OF MAXIMUM HEIGHTS TAPPED.



(6) *Incidence of Tapping*: Five per cent of all holdings had never been tapped—or 24 per cent of holdings with girth less than 18 inches, 0.7 per cent of those with girth greater than 18 inches. Four holdings (three in Perak and one in Kelantan) had practically no trees to tap.

Forty-three per cent of holdings had been tapped on practically all trees, the other 57 per cent showed a varying proportion of trees never opened (Tables VIII and IX). Holdings in Trengganu appear to begin tapping at least a proportion of the trees at smaller size than is the general practice in other states.

Only 59 per cent of holdings were in tapping when observed. On these the average number of trees being tapped was 71 per cent of the number which had been opened.

(7) *Thickness of bark in tapping*: The histogram in Figure 6 shows the distribution of mean thickness per holding of bark in tapping, and the curve shows the proportion of holdings tapping bark thicker than any given value. No holding was observed to be tapping bark thinner than 6.3

mm. average. Province Wellesley and Pahang had bark thicker than the general average, Trengganu had thinner bark (Table X).

FIG. 6

DISTRIBUTION OF BARK THICKNESS IN TAPPING.

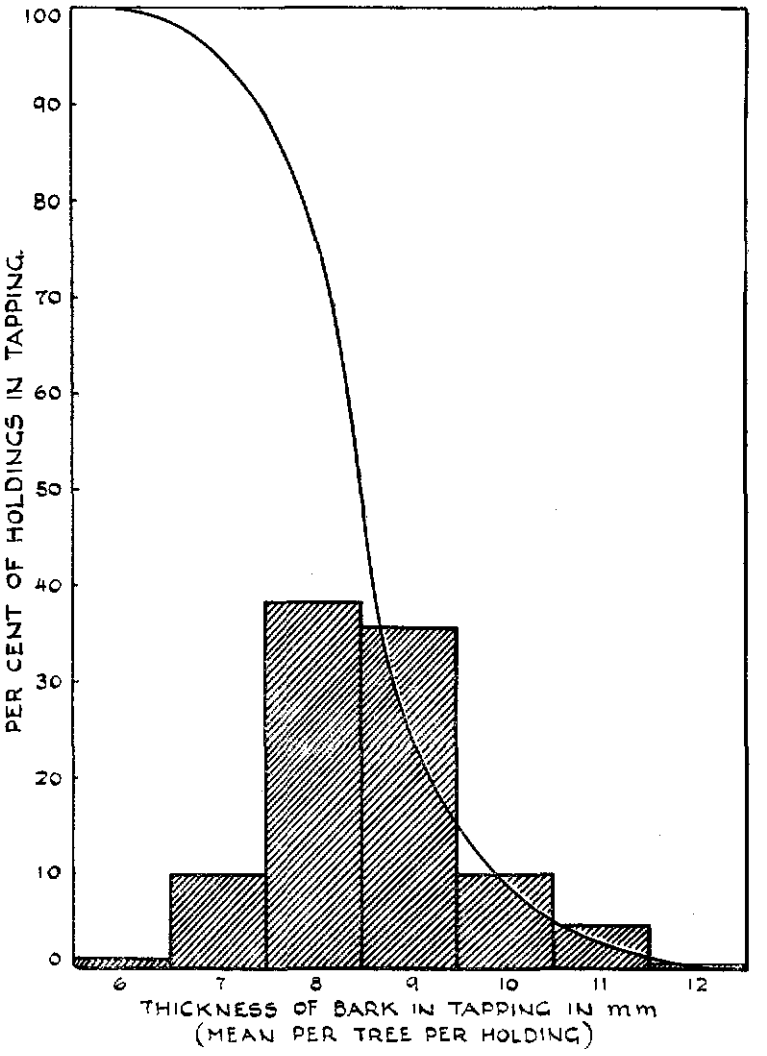


Fig. 6—Distribution of bark thicknesses in tapping. Histogram and Cumulative frequency curve.

(8) *Heights of cuts in tapping*: The distribution of average heights of cuts on 289 holdings tapping at the time of observation is given in Table XI. Apart from Trengganu, where all holdings had mean girth less than 24 inches, and had all mean tapping heights under 20 inches; differences between states are not significant.

The positions of cuts vary greatly within each holding, and the maximum heights would be better than the means to indicate the heights to which small-holders were working in 1939. These are not available on the summary cards, but combining with the means information on the ranges of heights observed (last line of Table XI), it is indicated that about 80 per cent of holdings were working below 45 inches. At least two holdings had cuts above 60 inches.

(9) *Tapping systems:* Observations on tapping systems are presented as they occurred in the observed sample without weighting for districts, states etc.

(i) *Lengths of cuts in tapping*:—The 289 holdings in tapping showed average lengths of cut per holding distributed as follows:—

Average length of cut in units of 1/12th circumference ..	3	4	5	6	7	8	10	12
Per cent of holdings ..	3	26	25	34	8	3	0.7	0.3

Kelantan and Trengganu showed an average length of cut of 0.53 circumference with no average per holding less than 5/12, the average for Pahang was 0.51 circumference, and for other states 0.41 circumference.

Lengths of cuts within single holdings tend to be very variable; the range between shortest and longest observed on 21 trees of each holding was distributed as follows:—

Range of cut lengths in units of 1/12th circumf.	0	1	2	3	4	5	6	7	8	9	10	11	12
Per cent of the 289 holdings in tapping ..	18	4	20	14	11	8	9	3	7	3	1	0.3	1

The range of lengths is correlated with average length of cut. Almost all the holdings (22 per cent) having a uniform or nearly uniform length of cut use ½ or ⅓ circumference cuts. Holdings employing long cuts use them only erratically on some trees.

(ii) *Type and number of cuts*:—The great majority of holdings use spiral cuts. V or multiple cuts when present usually occur only on some of the trees. The total distribution of numbers of holdings by proportion of trees in tapping with V or multiple cuts was as follows:—

Per cent of trees per holding	0	5	15	25	35	45	55	65	75	85	95	100	Total
V cuts ..	205	32	13	9	8	8	3	4	1	5	1		289
Multiple cuts ..	209	42	13	16	4	3	1	—	1	—	—		289

The averages per State were as follows:

	Av. percentage of trees in tapping with:— V cuts	Multiple cuts
Malacca ..	3	0
P. Wellesley ..	16	1
Perak ..	33	3
Selangor ..	0.4	4
N. Sembilan ..	0	4
I ahang ..	0.8	10
Johore ..	10	0.7
Kedah ..	6	0.8
Kelantan ..	2	21
Trengganu ..	3	6

(iii) *Frequency of tapping*: For information on frequency of tapping we have only the reports of observers' conversations with owners. The value of these reports is doubtful; such as they are, they indicate:

a.d.	daily	d.a.m.	
10	82	8	% of 216 holdings reported

Frequency of tapping shows no correlation with length of cut.

Miscellaneous: The following surviving scraps of data are tentatively presented, only because, except for Mead's selected holdings, no other information on these topics appears to be available.

(10) *Density of stand*:

Trees per acre ..	0	50	100	150	200	250	300	350	400	450	>500
No. of holdings	1	—	10	13	3	6	8	5	7	—	6

(11) *Age*: being not directly observable is unreliable. The data recorded for 56 holdings are distributed as follows:

Age in years	3-7	8-12	13-17	18-22	23-27
No. of holdings	6	16	13	20	1

(12) *Races of owners and tappers*: It has not been found possible to reconstruct the code for races of owners and of tappers; but it was indicated that owners and tappers tend to be of the same race, confirming the corresponding observation by Meads (1933).

IV. Analysis of Data

1. *Method of estimating means and their standard errors.* Each observation of a character x is regarded as made up of three independent parts:

$$x_{ijk} = \xi_i + \eta_{ij} + \varepsilon_{ijk}$$

where ξ_i represents the average for all holdings in the ith State,

η_{ij} represents the deviation from ξ_i of the mean of the jth district of the ith State,

ε_{ijk} represents the deviation from $\xi_i + \eta_{ij}$ of the mean of the observations in the kth holding of the jth district of the ith State.

Let V_s be an estimate of the variance of ξ_i, *i.e.* of the variance of State means,

V_d be an estimate of the variance of η_{ij}, *i.e.* of the variance of district means within States,

V_o be an estimate of the variance of ε_{ijk}, *i.e.* of the variance of means per holding within districts.

Part of the variance V_o will be due to sampling variance of trees within holdings, but because the sampling fraction of holdings within districts is very small, the proper contribution of tree variance to the error of State and country means will be automatically included in V_o which need not, for this purpose, be sub-divided further.

For s States, with d_i districts sampled in the ith state ($i = 1, 2 \ldots s$), and h_{ij} holdings observed in the ijth district ($j = 1, 2 \ldots d_i$), we obtain analysis of variance as follows:

	Degrees of Freedom	Sums of Squares
Between States ..	$(s-1)$	$\left\{ \sum_i \sum_j h - \frac{\sum_i (\sum_j h)^2}{\sum_i \sum_j h} \right\} V_s + \left\{ \sum_i \frac{\sum_j h^2}{\sum_j h} - \frac{\sum_i \sum_j h^2}{\sum_i \sum_j h} \right\} V_d + (s-1) V_o$
Between districts within States ..	$\sum_i (d_i - 1)$	$\left\{ \sum_i \sum_j - \sum_i \frac{\sum_j h^2}{\sum_j h} \right\} V_d + \sum_i (d-1) V_o$
Between holdings within districts ..	$\sum_i \sum_j (h_{ij} - 1)$	$\sum_i \sum_j (h-1) V_o$

from which estimates of V_s, V_d and V_o can be obtained.

In estimating State means (the code for districts being lost) districts have been weighted according to the numbers of holdings observed in each (h_{ij}),

$$\bar{x}_i = \frac{\sum_j h_{ij} \bar{x}_{ij}}{\sum_j h_{ij}} = \frac{\sum_j \sum_k \bar{x}_{ijk}}{\sum_j h_{ij}}$$

In these circumstances the error variance of $\bar{x}_i$ is estimated by (Smith, 1947)

$$V(\bar{x}_i) = \left\{ \frac{\sum_j h^2}{(\sum_j h)^2} + \frac{\sum_J A^2}{(\sum_J A)^2} - \frac{2\sum_j A}{d\sum_J A} \right\} V_d + \frac{1}{\sum_j h} V_o \quad \ldots\ldots(1)$$

where $\sum_j$ indicates summation over the d districts sampled

$\sum_J$ indicates summation over all districts in the state

A is the area of small holding rubber in each district

If h and A were both constant this formula would reduce to that given by Yates (1946) p. 19 with $f_2 = 0$. A further refinement would be to consider variation of V_0 in different districts, but in most of the work here reported it is sufficiently stable to be treated as constant, or has been made so by transformation.

To estimate country wide means, each State mean is weighted by the total area of small holdings in the State $A_i = \sum_J A)$ unless otherwise indicated.

$$\hat{\bar{x}} = \frac{\Sigma A_i \bar{x}_i}{\Sigma A_i}$$

and its estimated variance is given by

$$V(\hat{\bar{x}}) = \frac{\Sigma A_i^2 V(\bar{x}_i)}{(\Sigma A_i)^2}$$

V_s not being required in this formula since all states are sampled.

2. *Nationality of owner and sizes of holdings.*

Owing to destruction of records we have areas of holdings and nationalities of owners for 17 districts only, and even these can be correlated with other characters for only three districts. Furthermore the surviving records give a poor

sample of districts within States, Johore alone being well covered; consequently, proper estimates of the errors of means for each state are difficult to evaluate. The data are tentatively presented as (except for Johore and Mead's survey) no other data on ownership and size of holding have been published, and they provide a rough guide to the position.

Ownership: The direct records of nationality of owners are lost and the figures presented (Table I) have been deduced from the names of owners. The differences in proportion of Malays in the different States are highly significant. Smaller, but (as indicated by χ^2 tests) still significant differences occur between districts of the same State. For example in six districts in Johore the proportions of observed holdings owned by Malays were:

Kota Tinggi	..	40 *per cent*	± 11
Johore Bharu	..	15 ,, ,,	
Kukup	..	70 ,, ,,	
Kluang	..	45 ,, ,,	
Batu Pahat	..	60 ,, ,,	
Segamat	..	55 ,, ,,	

Race of owner can be correlated with tree characters only for one district, Malacca, with 14 Malay to 6 Chinese owners. No difference of mean girth or of areas of thin bark, can be demonstrated; but there is a suggestion that Malays may tap higher than Chinese. (Malay 60.8 inches, Chinese 46.9, difference 13.9 ± 7.0).

Areas of small holdings: Distribution of areas is markedly skew (Figure 2 and Table III) but the distribution of log area approaches normal. Most computations have therefore been carried out on logarithms, although the gain in efficiency is probably small.

The data being heterogeneous variation may be indicated better by variances for a few selected groups, than by the usual pooled analysis. The following are estimates of variance of $\log_e$ area, V_d is variance between means of districts as defined above, and V_0 is average variance within classes of the same racial ownership and district.

Johore only:

V_d for Chinese holdings	.040	(estimated from	5 d.f., P<.01)
V_d for Malay holdings	.023	(„ „	„ P>.05)
V_0 average all races	.260	(„ „	105 d.f.)

Kedah only:

V_0 for Chinese holdings	1.066	(„ „	9 „)
V_0 for Malay holdings	.368	(„ „	33 „)

All states:

V_0 average all races	.353	(„ „	297 „)

To interpret these figures in acres, approximately:

$$\left(\frac{\text{Standard deviation}}{\text{mean in acres}}\right)^2 = (e^V - 1) \quad = (V + V^2/_2 + \ldots\ldots)$$

Table II shows the estimated mean sizes for each race of owner in each State. Differences between State means are highly significant; the chief difference being that holdings in Trengganu and Kelantan are on average smaller than in other States. The average size of Chinese holdings is larger than of Malay ones. (Standard errors in Table II are only approximate—only those for Johore being properly evaluated —and may tend to be underestimated. *cf.* Sec. V 3).

So far as can be ascertained from the data of three districts for which the comparisons can be made, none of the tree characters are correlated with size of holding within races of ownership.

3 *Girths* (*g*). (Table IV). The distribution of girths in each district is sufficiently near to normal for ordinary statistical methods to be directly applied, and the variances between holdings within districts are homogeneous (Bartlett's test for homogeneity of variance gives—in the notation of Hartley and Pearson (1946)—M = 26.6, with k = 25, .3<P<.5) Analysis of variance gives

	d.f.	M.Sq.	P.	Estimates of Variances.
Between States ..	9	792.46	<.005	V_s = 13.96
Between districts within States ..	15	138.64	<.001	V_d = 5.481
Between holdings within districts ..	461	32.69		

4. *Bark reserves*—Areas of thin bark observed are not easily summarised. (i) Distribution is J shaped. (ii) Variances are heterogeneous, varying between districts as also between States. (iii) The relationship to girth, tapping

incidence, and other characters—such as age, density of trees, and soil conditions—introduces forms of heterogeneity difficult to summarise. When the relationship to these characters was appreciated, it became evident that summary methods were inadequate properly to describe the observations. To obtain a good picture more detailed analysis of the tree cards, correlating area of thin bark to girth of, and incidence of tapping on, individual trees would appear necessary. In the absence of machinery for handling the cards such procedure is not at present practicable, and the partitions of the data as presented below depend in part on subjective judgements based on frequency tables too extensive for publication.

For brevity we write a for area of thin bark (less than 5 mm.) in circumference inches per tree. Scrutiny of correlation tables of a: girth, for varying proportions of trees never tapped, revealed that the distribution of thin bark was distinctly different for holdings whose mean girth was less than or greater than 18 inches. Evidently many holdings with mean girth less than 18 inches must have had considerable areas of virgin bark thinner than 5 mm. This is borne out by the further observation that for all holdings with g less than 17.9 inches and a greater than 2.3 ⊙″, the range of a was greater than 9 ⊙″, and was in many cases 60 ⊙″ (*i.e.* the maximum possible range, indicating that some trees had all bark thicker than 5 mm. while some had all thinner than 5 mm.). In Tables V and VI the two girth groups are reported separately because the recorded observations evidently describe essentially different conditions. Four holdings with girth between 18 and 21 inches are classified with the smaller girth group because they also appear to have had much thin virgin bark, indicated by a greater than 23 ⊙″ and range 60 ⊙″, although one had never been tapped and three had 20 per cent of trees untapped.

Some holdings with small girth and thin bark may be young, and despite their apparent area of thin bark may represent reserves of new bark which will become satisfactory. Others may represent bad conditions of growth such that they will never mature satisfactory girth and bark. Owing to loss of records for age and density of stand it is not possible to discriminate between such groups. However probably all holdings with much thin bark are inferior, since well grown small trees need not be thin barked (in the sense of these observations)—*e.g.* four holdings with girth between 8 and 12 inches and partially tapped had a less than 3 ⊙″.

For the larger girth group, standard deviations between holdings within districts are related to their respective means approximately according to the regression

$$s = 0.8 + 0.5\ a,$$

but are otherwise homogeneous. Therefore to stabilise the variances a suitable transformation may be $\log_e (1.6 + a)$ (*cf.* Bartlett 1934, Cochran 1938). This transformation also renders the distributions within districts approximately normal except for the concentration of observations at $a = 0$. Approximate analysis of variance of the data thus transformed gives

	d.f.	M.Sq.	
Between States ..	9	6.45	*Pca* .066
Between districts within States ..	15	2.76	*P very low* *Vd* ? = .149
Between holdings within districts ..	394	.25	

showing marked differences between districts, and differences between States just approaching significance.

State means quoted in Table VI are direct arithmetic means. To obtain estimates of their standard errors from the above analysis of variance some approximations have been introduced—including in the general mean allowance for errors of the weights, which are estimated areas per State falling within the two girth groups. Although these estimated standard errors may not be strictly correct they are believed to be adequate to indicate the accuracy of the data.

Distribution of a might be best graduated by a " twisted-J " type of curve, but a simple exponential curve can describe it not too badly. Fig. 4 shows the curve

$$f = .180e^{-.204a}$$

where .204 = 1/â, and .180 = .204 × .886 because only 88.6 per cent of all holdings are included. An approximate test of goodness of fit indicates P almost .01 (approximate because the errors introduced by weighting of State frequencies have been ignored), the major deviations being a lower frequency of observations between 1 and 3 ⊙″, and excess at 6 to 9 ⊙″. Using this curve the estimated proportion of mature holdings having a within a specified range, say from a_1 to a_2, is

$$.886\ (e^{-.204a_1} - e^{-.204a_2}) = .886\ (10^{-.0887a_1} - 10^{-.0887a_2}),$$

an expression which may be easily evaluated from a table of logarithms. (Actually the deviations from the exponential

curve only appear after weighting; the crude frequency distribution of the 486 observations was fitted by the curve $f = .176\ e^{-.176a}$ with the test of goodness of fit showing P *ca.* .5. But this distribution does not hold within individual States where standard deviations are less than the means, instead of equal to them as they would be if it were applicable throughout).

Correlation of bark reserves with other characters—Correlations of a with other characters are very low; that is to say, for any given value of another character, a usually covers nearly its full range of variation: nevertheless a distinct trend can be detected in the average amounts of thin bark associated with some of the following characters.

(i) *Girth.*—Although a first survey indicated some negative correlation between area of thin bark and girth, this was shown to be due to the thin (virgin) bark on holdings with small mean girth. Within groups of holdings with mean girth less than or greater than 18 inches there is no correlation of a with g.

(ii) *Age.*—Only 16 surviving cards show estimates of age of trees; they suggest that about half of the old holdings above 18 years of age may have a greater proportion of thin bark than the younger holdings.

(iii) *Incidence of tapping.*—Although the correlation is very low, as the proportion of the trees which had been opened *decreases* from 100 to 75 per cent, the average area of thin bark *increases* steadily from 4.3 to about 7 ⊙″. For holdings with mean girth greater than 18 inches the amount of thin bark thereafter decreases to zero when no trees have been tapped; but for holdings with smaller trees it increases to an average of 16 ⊙″ per tree when no trees have been tapped. These observations suggest that omission of a tree from tapping is largely determined by thinness of bark. The point can be tested in more detail when facilities permit re-examination of the tree cards.

(iv) *Length of cut in tapping.*—Omitting 3 holdings using full spiral tapping (which have relatively low a), the correlation of area of thin bark with average length of tapping cut is $r = .12$ ($t = 2.00$, $P = .05$), and regression.

$$a = 3.12 + .585\ l$$

where l is in units of one twelfth circumference over the range 3 to 8.

Correlation of area of thin bark with range of tapping cut lengths is not significant ($r = .08$, $t = 1.32$). (This was computed because it was thought that variation in tapping cut lengths might perhaps be an indicator of careless tapping).

(v) *Type of Cut.*—There is no significant difference between averages of a for holdings using only spiral cuts and those using mixed spiral and V cuts.

(vi) *Multiple cuts.*—There is no detectable correlation between a and percentage of trees with multiple cuts.

(vii) *Frequency of tapping.*—*No significant differences* of a can be demonstrated between groups of holdings for which different frequencies of tapping were reported. Naturally holdings out of tapping had less thin bark than those in tapping.

Average a for 175 holdings not in tapping = 2.66 ± .274
" " 289 " in tapping = 6.68 ± .324
Difference 4.0 ± .424

Variability of bark reserves between trees within holdings (as measured by v, see sub-section 10) is slightly correlated with variability of girths (r = .18, P = .0001). It is not correlated with variability of maximum heights tapped.

5. *Maximum heights tapped* (h_m)

Obviously in considering a character such as maximum height ever tapped, the maxima on individual trees are of interest. This was however not indicated on the summary cards, and cannot meantime be recovered. This section therefore deals only with averages per holding of the maximum heights on trees which had been tapped at some time.

The analysis of variance gives

	d.f.	M.Sq.	
Between States ..	9	2,829	P = .15
Between districts within States ..	15	1,585	P very low V_d = 76.25
Between holdings within districts ..	440	186.88	

As with other characters there are marked differences between districts, but differences between States are not significant.

In computing the standard errors quoted in Table VII, small errors in the weights, due to proportions of holdings never tapped, were ignored. Their effect would be trivial.

Maximum height tapped would be expected to increase with age and girth of trees. Thirty eight surviving summary cards indicate that its correlation with girth is about r = .5, with regression

$$h_m = 6.7 + 1.46\ g$$

Only 15 of these cards show estimates of age of trees; they

		No. of holdings	Per cent of no. observed in each State	Estimated acreage '000 ac:
Perak ..	(in 3 out of 5 districts observed)	6*	6	13
N. Sembilan ..	(„ 2 „ 3 „ „)	3	5	3.4
Pahang ..	(„ 1 „ 3 „ „)	3	5	3
Johore ..	(„ 2 „ 5 „ „)	6†	6	16
Kelantan ..	(all in Ulu Kelantan)	7‡	17	9

* includes 3 holdings, all in one district, with either no rubber trees or only 7.

† includes 1 holding with mean girth 27 inches and less than 5 per cent of trees tapped (*cf.* footnote Table IX).

‡ plus one holding with only 11 trees, arbitrarily omitted by observer from his records.

suggest only slight correlation with age, around a regression approximately

$$\hat{h} = 42 + .5\ y$$

for h in inches, and age (y) in years.

6. *Incidence of Tapping.*

(i) *Holdings never tapped.*—Of the 489 holdings observed, 25 had in effect never been tapped. These were distributed between States as shown on page 104.

Weighting the observations for each State as usual, they were distributed according to mean girths per holding as follows:—

Girth (inches) ..	(No rubber)	10	14	18	22	26	28	Total
Percent of total acreage of country in holdings never tapped ..	.7	1.6	1.0	.9	.4	.3		4.8

Combining this with Table IV we estimate that 24.1 per cent of holdings with mean girth less than 18 inches had not yet been tapped, 1.1 per cent of holdings with girth 18 to 28 inches (or 0.7 per cent of holdings with girth greater than 18 inches) had not been tapped.

(ii) *Trees never tapped.*—The proportions of trees per holding never tapped is difficult to summarise. The distribution is naturally J shaped, and is related to mean girth and bark thickness and probably also to variability of these. There are distinct differences between districts; but these cannot be correctly interpreted without correlation with girth or age, and simple analysis of covariance is precluded by the extremely skew distribution and curvature of regressions. Possibly the best method of analysis would be to transform percentages to probits, and compute regressions on girth and variability of girths in each district, but I have not considered it worth the time required to carry out a thorough analysis on such old data. Another approach would be to investigate the incidence of tapping relative to girth of individual trees, which could be done, if of interest, when card sorting machinery is available. Meantime Tables VIII and IX present a rough picture of the incidence of tapping.

The asterisks through Table IX indicate the average incidence relative to girth.

Inspection of correlation tables for each State of proportion of trees never tapped: girth indicates:—

In Malacca and Negri Sembilan nearly all trees had been tapped on all holdings of mean girth greater than 17 inches.

Province Wellesley had a lower proportion of trees tapped at all girths (all observed holdings had girth greater than 18 inches).

In other States the correlation was much as shown in Table IX.

Trengganu is distinctive in showing no holdings completely untapped despite a high proportion of holdings with girth under 18 inches, but the proportion of trees tapped in such holdings was not appreciably different from that in tapped holdings of similar girth in other States.

(iii) *Holdings in tapping.*—Only 59 per cent (289/486) of holdings were in tapping at the time of observation. Acreages under tapping fluctuated considerably during 1939 (quarterly reports in the Malayan Agricultural Journal) and the dates of observation in each district are lost; therefore detailed comparisons would be meaningless. The extremes however seem significant and may be of interest remembering that Trengganu had the smallest (and therefore presumably youngest) trees:

States	Per cent of holdings in tapping
Trengganu ..	90
Kelantan, Johore, Pahang ..	75
Perak, Selangor, N.S., Kedah	40 to 62
P.W. and Malacca ..	32

The number of trees being tapped in any one holding varied all the way from none to all. On holdings which were doing some tapping the average number was 71 per cent of the number which had at some time been opened.

7. *Thickness of bark in tapping.*

Computations have been done on logarithms of the observed data, and means and standard errors converted to the arithmetic scale by the formulae given by Finney (1941). This mode of calculation was adopted merely because of the accident that surviving data happened to be in this form; but actually it does give some gain in efficiency, because the distribution of the logarithms is more nearly normal than that of the observed data. Standard deviations are, however, so small relative to means that correction terms given by Finney for deviation from large sample theory are negligible (*e.g.* the largest correction to a State mean—those with only

6 or 7 holdings in tapping—is only .008), and for the same reason direct analysis of the original arithmetic data would have been quite satisfactory.

Bark in tapping was of course only observed on holdings tapping at the time of observation. Table X furthermore shows only 280, instead of 289, holdings because all the observations returned by one observer for a district in Perak were unreasonable (2.2 to 3.1. mm. for means per holding) and must be rejected. But for this one group of observations, having regard to the likelihood that different observers might press more or less heavily on the gauge, the data returned by different observers are pleasingly consistent.

Variances within States are homogeneous (M = 6.26, k = 10, P about .7). Analysis of variance of $\log_{10}b$ is given on page 108. It shows marked differences between States in that P.W. and Pahang had bark on average thicker than others, Trengganu had thinner bark, and differences between other States were no greater than differences between districts within States, which in turn are relatively smaller than for most other characters.

No holding observed was tapping bark on the average *thinner than 6.3 mm. It is not meantime practicable* to ascertain from the tree cards what was the thinnest observed in tapping on any single tree.

8. *Heights of cuts in tapping* were distributed as shown in Table XI. Analysis of variance gives

	d.f.	M.Sq.	P.
Trengganu v. mean of other States ..	1	1391	.001
Between other States ..	8	162	ca .07
Between districts within States ..	15	64.3	ca .25
Between holdings within districts ..	264	52.6	

V_d = .015

9. *Tapping systems.*—Observations on tapping systems were not considered worth detailed analysis. All the observations deemed worth presenting are given in Sec. III.

10. *Variability of observations between trees within holdings.*—As described in Section II the sample of trees from each holding was taken as three groups of contiguous trees. This was done primarily for the sake of simplicity of instructions to observers, but from memory of other work on estate fields (the reports of which have been destroyed) it is believed that this will give almost the same result as a

	d.f.	M.Sq.	P.	
Between P.W., Treng., Pahang and other states	3	.040043	.001	
Between other states .	6	.003987	—	
Between all states ..	9	.016007	—	
Between districts within states ..	14	.004040	ca .02	[V_d = .000171]
Between holdings within districts ..	256	.002046		

random sample of trees throughout a holding. The point requires examining but this cannot be done in the absence of machinery. Meantime we have on the summary cards estimates of the standard deviations of most characters over the 21 trees observed in each holding, and these form the subject matter of this section.

For all characters investigated these standard deviations were correlated with the magnitude of the character. It is unnecessary here to detail surviving data and their reconstruction. In effect we have for most characters the regression of the logarithms of the standard deviation ($\log_{10} s$) on the mean per holding for each character.

For the four characters, girth (g) in inches, area of thin bark (a) in circumference inches per tree, bark thickness in tapping (b) in mm., and maximum height ever tapped (h) in inches, these regressions were

$$\log s_g = .676 \log g - .126$$

$$\log s_a = .735 \log a + .274$$

$$\log s_b = 1.205 \log b - .977$$

$$\log s_h = 2.2 \log (h/10) - 1.2 (\log h/10)^2 + .13$$

Using deviations from these regressions as measures of variability (v) between trees within each holding, the following analyses of variance of v indicate the distribution of tree variability:

	d.f.	v(g) M.Sq.	d.f.	v(a) M.Sq.	d.f.	v(b) M.Sq.
Between States ..	9	.05767	9	.2542	9	.02868
Between districts within States ..	15	.03960S	15	.1761S	14	.02591
Between holdings within districts ..	461	.01375	429*	.03200	251†	.01847
Between holdings within districts adjusted for grouping ..		.01291		.02866		.01764
Theoretical sampling variance§ ..		.00495		.00495		.0117‡

S: P<.001.

* excluding 32 holdings for which no thin bark was recorded.

† excluding 9 incorrect observations as noted in Sec. IV, 7, and 5 holdings with only one tree in tapping.

§ deduced from figures given by Bartlett and Kendall (1946).

‡ only a rough approximation owing to variable number of trees observed in tapping. Estimated for the harmonic mean, n = 10.

V. Notes on sampling

1. *Number of trees to be sampled per holding.*

Section IV 10 shows that sampling variances within holdings are far from homogeneous. Ignoring this meantime, the Table on page 111 shows a comparison of variances within districts and the parts thereof ascribable to sampling within holdings.

Between 20 and 30 trees per holding would appear to be a satisfactory size of sample for a survey of such characters as those reviewed here. For observations on girth and bark thickness it might be worth while to consider varying the size of sample in proportion to the magnitude of the character in each holding. For area of thin bark the standard deviation of single tree observations may vary from zero to about 30 ⊙″; but holding variances, as well as tree variances, are closely correlated with the means, and when data are transformed as indicated the tree variance also becomes reasonably stable.

2. *Number of holdings to be sampled per district* in order to attain any assigned degree of accuracy, either of district means or, provided all districts be sampled (Sub-sec. 3), of State or country means can be deduced from the figures in the first row of the Table on page 111.

3. *Sampling districts.*

The analyses given in Sec. IV show that for all characters, except height of cuts in tapping, there were appreciable differences between districts. Since the number of districts is sufficiently small so that all could be sampled, it follows that in any future survey this should be done in order to eliminate this source of error from estimates of State means. The loss of information in this survey owing to incomplete sampling of districts has been serious.

As shown in Sec. IV 1 the error variance of State means is compounded of two parts, one due to variance between districts, and one due to variance between holdings within districts. For brevity equation (1) may be re-written

$$V(\bar{x}) = GV_d + HV_o$$

If all districts were sampled G would reduce to zero, and, provided the total sample was distributed between districts in proportion to their weights, $V(x)$ would reduce to HV_0. The ratio $(GV_d + HV_0)/HV_0$ is therefore a measure of the ratio by which the potential error variance has been inflated by failure to sample all districts. This ratio for each of the characters observed is shown in the following table on page 112.

	g (ins.)	$\log_e (1.6 + a)\ h_{,,}$ (ins.)		b (mm.)
Between holdings within districts ..	32.69	.25	187	.79
Average sampling variance for means of 21 trees within holdings§	2.57	.04	4.4	.19†
Range of sampling variance for means of 21 trees within holdings§ ..	0.15 to 35	.036 to .043*	0.1 to 48	.02 to 1.12†

* for a range of a from 2 to 15 ○″.

† for means of 10.3 trees = the harmonic mean of numbers of trees in tapping per holding.

§ the sampling variances quoted are only approximate and may be underestimated as they are based on geometric means (and in the case of h no allowance has been made for trees never tapped).

Character	Ratio
Bark reserves ..	11
Maximum height tapped ..	9
Girth ..	4
Thickness of bark in tapping ..	2
Height of cuts in tapping ..	1
Area of holdings (Johore only) ..	2
Race of owner (Johore only) ..	1.1

The only observations for which an independent check is available are races of owners and sizes of holdings, and statistics for these have been published only for Johore. They may be used here to illustrate the importance of giving due attention to heterogeneity in a sample.

The Table on page 113 shows the frequency distribution of races of owners of the 20 holdings observed in each of six districts. If the 120 observations had been accepted as a homogeneous random sample of the whole State the proportion of Malays would have been reported as .475 ± .045, whereas more correct methods give .554 ± .057. The estimate for proportion of Malays appears to be low, but its 5 per cent confidence interval, .440 to .668, includes the expected value, .653.

The total error variance ($.057^2$) was built up as follows:

due to variance within districts (pq/n) ..	.00302
,, ,, between districts ..	.00022
,, error variance of the weights ..	.00004
	.00328

If all districts had been sampled in proportion to the number of holdings in each the first term would be .00205.

The areas of holdings differ appreciably according to the race of owner, and for most efficient sampling this feature also should be considered. If the Johore observations were treated as a homogeneous sample the mean area per holding would be estimated as 5.3 ± .30 acres. When heterogeneity between races and districts is considered the estimate is

District	Race of owner				Estimated* number of holdings in district (N)
	Malay	Chinese	Indian	Other	
Kota Tinggi ..	8	12	—	—	2,926
Johore Bharu ..	3	16	1	—	2,257
Kukup ..	14	6	—	—	7,593
Kluang ..	9	11	—	—	4,013
Batu Pahat ..	12	6	2	—	14,914
Segamat ..	11	9	—	—	3,606
Total unweighted ..	47 5	50.0	2.5	—	per cent
Total weighted by N ..	55.4	40.1	4.5	—	„ „
Expected†	65.3	30.4	3.2	1.1	„ „

* Estimated from the total acreage of small holdings per district (December 1938) divided by the average size of holding in each district as estimated from the survey data.

† According to numbers given in Rubber Statistics Handbook 1940, Table 48 (figures for December 1939).

5.13 ± .48*. Of the final error variance (.23) about 45 per cent is contributed by variance between holdings of the same race and district, 10 per cent by errors in estimating proportions of each race in each district, 45 per cent by variance between districts and ½ per cent by errors in the estimates of numbers of holdings per district (used as weights). If the number of holdings for each race in each district were known, and every group were sampled in proportion to its size, the error variance of mean area per holding estimated from 120 observations would be about .09 acre² instead of .23.

In view of the scrappiness of surviving data for States other than Johore the probable errors of country means (Table II) cannot be properly assessed.

4. *General considerations.*—To determine the most efficient form of sample for a survey, consideration must be given to the cost of each operation. Rubber instructors who carried out the field work of the 1939 survey, and still remain in the service, have been asked to give estimates of times taken. The replies, based on memory of events seven years ago, are very erratic. The following is a summary of the replies from 16 observers.

Time taken to find a specified holding		45 minutes to 3 days.
Average mileage per holding	..	7 to 50 miles.
Average excluding Johore and Kedah	..	17 miles.
Johore	..	36 „
Kedah	..	50 „

Time taken to obtain general information 10 minutes to 4 days. (*e.g. Some observers appear to have had much difficulty in locating* the owners).

---

* With an estimate of error compounded from several variances, one of which, estimated from only 5 degrees of freedom, contributes nearly half of the total variance, it is difficult to fix a confidence interval with certainty. This gives a further reason that an efficient survey should sample all districts to eliminate from the estimate of error, elements of variance which can be estimated only from few degrees of freedom. If we take t for 5 degrees of freedom as the most conservative value, the 95 per cent confidence interval indicated by the above figures is 3.90 to 6.36. Official statistics indicate an average size of 3.96 acres. For each race separately the survey estimates exceed official figures by approximately equal amounts (1.0 to 1.4 acres).

Time taken to observe 21 trees—½ to 10 hours.

Comments by observers on tapping systems (length, type and number of cuts) show reasonable, but far from precise correlation with corresponding data derived from the tree profiles. This merely confirms the importance, well known to workers on sample surveys, that all room for subjective judgement on the part of observers should, as far as possible, be eliminated from the observations to be recorded.

Acknowledgements

Obviously Mr. R. H. Meikle should have been a co-author of this paper. His work in building up the Small-Holders' Advisory Service will be recorded elsewhere. But I wish here to record my sincere regard for him as a colleague and deep regret that he did not survive to co-operate in writing up the results of this survey.

The share of the work done by the staffs of the S.H.A.S. and of the Statistics Section is obvious in the paper. In addition I am indebted to Mr. H. L. Barnett for supplying statistics of small holding acreages by districts.

Literature Cited

BARTLETT, M.S. (1936): The square root transformation in analysis of variance. *J. Roy. Stat. Soc. Suppl. 3*: 68-78.

BARTLETT, M.S. AND KENDALL, D.G. (1946): The statistical analysis of variance-heterogeneity and the logarithmic transformation. *J. Roy. Stat. Soc. Suppl. 8*: 128-183.

COCHRAN, W.G. (1938): Some difficulties in the statistical analysis of replicated experiments. *Emp. J. Exp. Agric. 6*: 157-175.

FINNEY, D.J. (1941): On the distribution of a variate whose logarithm is normally distributed. *J. Roy. Stat. Soc. Suppl. 7*: 155-161.

HARTLEY, H.O. AND PEARSON, E.S. (with tables by Thompson, C.M. and Merrington, M.) (1946): Tables for testing the homogeneity of a set of estimated variances. *Biometrika 33*: 296-304.

MEADS, H.D. (1933): Bark consumption and bark reserves on small rubber holdings in Malaya. *Dept. Agric. S.S. and F.M.S. Econ. Series Bul. 4.*

Rubber Statistics Handbook, 1939 and 1940. (Singapore).

SMITH, H. FAIRFIELD, (1947): Standard errors of means in sampling surveys with two-stage sampling. *J. Roy. Stat. Soc. 110*(2).

YATES, F. (1946): A review of recent statistical developments in sampling and sampling surveys. *J. Roy. Stat. Soc. 109*: 12-30.

Kuala Lumpur,
June 1947.

TABLE I

Distribution of nationalities of owners of small holdings in 17 districts.

State	Districts	Per cent owned by				No. of holdings observed
		Malay	Chinese	Indian	European	
Malacca ..	Central ..	70	30	–	–	20
Perak ..	Lenggong ..	90	5	5	–	20
N. Sembilan ..	Seremban, Rembau, Tampin ..	70	20	10	–	60
Pahang ..	Temerloh, Raub ..	62	33	5	–	40
Johore ..	6 districts ..	55	40	5	–	120
Kedah ..	Sungei Patani, Kulim ..	58	27	10	5	40
Kelantan ..	Kota Bharu ..	95	5	–	–	20
Trengganu ..	—	100	–	–	–	20
	Weighted total ..	75.2	20.1	4.4	.3	per cent

TABLE II

Average size of small holdings in acres; by race of owner and State.
(estimated from observations in districts indicated in Table I)

State	Race of owner				Weighted Mean	Standard errors (approx.)				
	Malay	Chinese	Indian	European		Malay	Chinese	Indian	European	Mean
Malacca ..	2.5	3.7	–	–	2.8	.42	.75	–	–	.37
Perak ..	3.7	8.9	3.5	–	4.0	.56	4.82	2.25	–	–
Negri Sembilan ..	2.5	3.8	6.4	–	3.1	.24	.69	2.06	–	.32
Pahang ..	4.3	6.5	8.0	–	5.2	.54	1.14	4.10	–	.80
Johore ..	4.2	6.7	5.3	–	5.1	.35	.56	2.30	–	.48
Kedah ..	3.2	8.5	3.8	1.7	4.6	.43	1.61	1.47	.65	.82
Kelantan ..	2.1	2.2	–	–	2.1	.30	1.21	–	–	.46
Trengganu ..	1.7	–	–	–	1.7	.25	–	–	–	.33
Weighted Mean ..	3.2	6.3	6.0	1.7	3.83					

TABLE III

Distribution of areas of 340 holdings

Size of holding (acres)	0	1	2	3	4	5	6	7	8	9	10....15....20....25			Total no. observed
Malay owners ..	10	57	60	31	29	13	7	5	1	3	—	1	1	218
Chinese „ ..	—	6	18	11	16	11	8	4	5	7	13	1	4	104
Indian „ ..	—	3	—	2	5	—	2	1	—	—	3	—	—	16
European „ ..	1	—	1	—	—	—	—	—	—	—	—	—	—	2
Total ..	11	66	79	44	50	24	17	10	6	10	16	2	5	340

TABLE IV

Distribution of mean girths per holding.

Mean girth at 5 ft. in inches	8	10	14	18	22	26	30	34	38	42	46	56	No. of holdings observed	Mean girth inches	St. error of mean	Total[3] acreage '00 acres
Malacca ..		—	—	1	7	5	3	4	—	—	—	—	20	24 16	2 63	485
Prov. Wellesley		—	—	—	6	2	8	4	—	—	—	—	20	25 80	2.72	167
Perak ..		—	3	2	10	22	32	20	3	3	2	—	97[1]	27.30	.94	2,188
Selangor ..		—	—	—	—	3	5	9	13	6	2	2	40	34.73	2.04	1,056
Negri Sembilan		—	2	—	2	20	17	11	5	1	1	1	60	27 85	1.37	687
Pahang ..		—	5	6	5	22	8	14	—	—	—	—	60	24 50	1.24	635
Johore ..		—	4	7	19	32	16	14	5	2	1	—	100	25.24	1.22	2,661
Kedah ..		—	—	—	4	7	5	10	2	1	—	—	29	28.26	1.64	654
Kelantan ..		1	7	14	7	3	4	1	2	1	—	—	40[2]	19.41	1.26	517
Trengganu ..		2	3	6	5	4	—	—	—	—	—	—	20	17.14	2.09	224
Per cent ..		.4	4.1	6.1	13 6	24.5	20.0	18 4	7.3	3.5	1.5	.7	100%	26.591	.534	9,274

(1) Plus 2 holdings with no rubber trees, and 1 with only 7 trees of average girth 5.5 inches.

(2) Plus 1 holding with only 11 trees (records omitted by observer).

(3) Rubber Statistics Handbook (1939) figures for December 1938; used as weights to evaluate country means in this and subsequent tables.

TABLE V

Distribution of area of thin bark (a) relative to tapping incidence and mean girth (g)
Numbers of holdings observed with a ⊙

a: circ. inches per tree	0	4	8	12	16	20	24	28:...40	Total	Mean $(\bar{a})$
Girth > 18″										
Never tapped	7	—	—	—	—	—	—	—	7	1.0
50-95% never tapped ..	4	2	1	—	—	—	—	—	7	3.6
5-50% „ „ ..	102	62	30	10	2	4	—	—	210	5.2
All trees tapped	106	64	17	6	2	—	—	—	195	4.3
Girth < 18″										
Never tapped	2	1	2	2	1	2*	1	3	14	17.1
50-95% never tapped ..	7	1	2	2	3	1	1	1	18	11.0
5-50% „ „ ..	7	9	3	4	1	2	3***	2	31	11.8
All trees tapped	3	—	1	—	—	—	—	—	4	3.5

* Four holdings with $18 < g < 21$, and range of $a = 60$ ⊙″.

Distribution of (a) by States.

Numbers of holdings with a ⊙"—(Note: different class intervals from Table V)

(a)		0	2	4	6	8	10.....15.....25.....40			Total	Mean (a)	St. error of mean
$g > 18''$: Malacca	..	9	7	2	1	—	—	—	—	19	2.4	.97
P. Wellesley	..	5	9	4	1	—	1	—	—	20	3.4	1.45
Perak	..	55	16	10	4	3	4	—	—	92	2.6	.34
Selangor	..	7	11	8	9	2	2	1	—	40	5.0	1.61
N. Sembilan	..	31	11	9	5	1	1	—	—	58	2.8	.57
Pahang	..	13	7	13	8	5	2	1	—	49	4.9	.92
Johore	..	12	8	16	14	13	20	5	—	88	7.6	1.42
Kedah	..	3	3	6	7	3	5	1	—	28	6.9	1.56
Kelantan	..	3	6	3	4	—	—	—	—	16	3.9	.85
Trengganu	..	2	1	1	3	1	1	—	—	9	5.4	1.77
										419		
$g < 18''$: Malacca	..	1	—	—	—	—	—	—	—	1	.6	—
Perak	..	2	2	1	—	—	—	—	—	5	2.5	—
N. Sembilan	..	—	—	—	—	—	—	1	1	2	28.1	—
Pahang	..	1	1	2	—	3	2	1	1	11	10.3	—
Johore	..	—	2	—	—	1	3	3*	3	12	16.5	—
Kedah	..	—	—	—	—	—	—	—	1*	1	25.6	—
Kelantan	..	2	2	3	1	2	3	8*	3*	24	14.5	—
Trengganu	..	3	3	2	2	—	—	—	1	11	6.3	—
										67		
Weighted total % $g > 18''$	..	28.1	15.8	14.8	12.1	6.6	9.1	2.1	—	88.6	4.90	.494
Weighted total % $g < 18''$	..	1.5	1.8	1.1	.4	1.0	1.5	2.2	1.9	11.4	12.71	—

* Four holdings with girth between 18 and 21 inches.

TABLE VII

Numbers of holdings showing maximum heights tapped as indicated.

State	Never tapped	*Maximum height tapped in inches—Mean per holding of trees which have been tapped at some time*										Total number tapped	Mean height	St. error of mean
		0	10	20	30	40	50	60	70	80	90	100			
Malacca ..	—	—	—	1	2	4	5	4	2	2	—		20	56.4	9.09
P. Wellesley ..	—	—	—	2	9	6	3	—	—	—	—		20	40.0	9.47
Perak ..	3	—	3	12	18	23	23	10	2	2	1		94	46.1	3.11
Selangor ..	—	—	—	1	4	3	11	15	5	—	—		40	58.7	7.16
N. Sembilan ..	3	—	1	4	9	6	18	13	6	—	—		57	52.3	4.68
Pahang ..	3	1	4	19	13	6	13	—	1	—	—		57	36.0	4.16
Johore ..	5	—	6	18	17	21	17	7	5	3	1		95	44.1	4.26
Kedah ..	—	—	—	1	4	12	8	4	—	—	—		29	48.4	5.33
Kelantan ..	7	—	4	13	8	3	2	1	1	1	—		33	34.3	4.24
Trengganu ..	—	—	2	7	5	5	1	—	—	—	—		20	33.0	6.89
Weighted total % ..	3.9	0.1	3.8	14.4	17.0	19.4	21.2	12.5	4.9	2.0	0.8		465	46.4	1.84

TABLE VIII

Incidence of tapping by States (numbers of holdings).

Per cent of trees which had been tapped	100	95	80	60	40	20	5	0	Total	Total per cent (1)	Total per cent (2)
Malacca ..	17	3	—	—	—	—	—		20	0.7	0.7
P. Wellesley ..	2	4	6	6	2	—	—		20	31.2	31.2
Perak ..	39	36	12	5	—	2	3		97	11.9	14.6
Selangor ..	25	13	2	—	—	—	—		40	3.9	3.9
N. Sembilan ..	47	9	—	—	1	—	3		60	2.2	7.1
Pahang ..	9	27	14	5	2	—	3		60	18 8	22.9
Johore ..	32	42	15	3	1	2*	5		100	12.5*	17.7
Kedah ..	18	10	1	—	—	—	—		29	4.1	4.1
Kelantan ..	9	12	5	4	2	1	7		40	19.6	34.5
Trengganu ..	1	8	6	1	4	—	—		20	27.6	27.6
Weighted per cent of total holdings	43.0	34.9	11.6	3.9	1.6	1.2	3 9		486		14.5

(1) Estimated per cent of trees never tapped on holdings in which tapping has occurred.

* Omitting from the mean one holding which is abnormal, *cf* Table IX.

(2) Estimated per cent of trees never tapped over all holdings.

TABLE IX

Incidence of tapping relative to mean girth (numbers of holdings).

Per cent of trees which had been tapped	No rubber	*Mean girth per holding—inches.* 8	10	14	18	22	26	30	34	38	42	44. . .58	Total
100 ..	—	—	1	3	13	48	42	53	*22	*9	5	3	199
93 ..	—	—	1	6	15	24	32*	27	7	4	1	—	117
83 ..	—	—	1	4	10	*20	11	—	1	—	—	—	47
74 ..	—	—	2	5	*8	13	10	4	—	1	—	—	43
64 ..	—	—	1	*3	5	7	1	1	—	—	—	—	18
55 ..	—	—	1*	4	5	3	1	2	—	—	—	—	16
45 ..	—	—	2	3	2	1	—	—	—	—	—	—	8
36 ..	—	3*	4	1	—	1	—	—	—	—	—	—	9
26 ..	—	—	1	1	1	—	—	—	—	—	—	—	3
17 ..	—	—	1	1	1	—	—	—	—	—	—	—	3
7 ..	—	—	1	—	—	—	1†	—	—	—	—	—	2
0 ..	3	—	8	5	5	3	—	—	—	—	—	—	24
Total ..	3	3	24	36	65	120	98	87	30	14	6	3	489

†*a* = 8 ⊙″ despite only one tree observed which has been tapped. Evidently thin virgin bark.

* indicate regression of per cent of trees tapped on girth.

TABLE X

Distribution of bark thicknesses in tapping (mm.).

Bark thickness mm. (Av. per holding)	6	7	8	9	10	11	12	Total no. of holdings	Mean mm.	St. error of mean
Malacca ..	—	1	3	1	1	—	—	6	8.20	.526
P Wellesley ..	—	—	—	1	3	2	1	7	10.15	.459
Perak ..	—	6	18	12	1	1	—	38*	8.26	.181
Selangor ..	—	—	9	13	1	2	—	25	8.71	.202
N. Sembilan ..	—	—	9	8	7	—	—	24	8.87	.216
Pahang ..	—	2	10	19	10	5	2	48	9.17	.157
Johore ..	—	8	28	27	6	3	—	72	8.48	.154
Kedah ..	1	1	3	6	2	—	—	13	8.48	.281
Kelantan ..	1	5	12	9	1	1	—	29	8.22	.191
Trengganu ..	2	3	10	3	—	—	—	18	7.75	.267
Weighted total per cent ..	1.0	10.0	38.5	35.8	10.0	4.2	0.5	280	8.51	.078

* Excluding one district (9 holdings tapping)

TABLE XI

Numbers of holdings with mean height of cuts in tapping as indicated.

Average height of cuts in tapping (inches)	0	10	20	30	40	50	Mean	St. error of mean
Malacca .	—	4	1	1	—		20.0	2.94
Province Wellesley ..	—	3	4	—	—		20.7	2.74
Perak ..	3	24	19	1	—		18.8	1.06
Selangor ..	1	7	16	1	—		21.8	1.45
Negri Sembilan ..	3	11	8	2	—		18.8	1.48
Pahang ..	14	26	5	3	—		14.4	1.05
Johore ..	9	33	29	—	1		18.2	0.86
Kedah ..	3	4	6	—	—		17.3	2.12
Kelantan ..	8	13	8	—	—		15.0	1.35
Trengganu ..	11	7	—	—	—		8.9	1.71
Weighted total, per cent.	13.1	45.4	38.3	2.9	0.4		17.2	0.47
Average range of heights within holdings (inches)	16	27	38	—	—			