

ECONOMIC CLASSIFICATION OF AGRICULTURAL REGIONS IN BENGAL

By A. GHOSH
Statistical Laboratory, Calcutta.

INTRODUCTION

The use of curves of concentration to obtain a representation of the distribution when the precise law of distribution is not known is already quite familiar. Designating by N_x that part of a population N possessing at least x of a characteristic and I_x the amount possessed out of total amount I

$$p_x = \frac{N - N_x}{N} \text{ and } q_x = \frac{I - I_x}{I}$$

we can obtain a relationship between p_x and q_x . Curves of this form were suggested for the representation of income simultaneously by Lorenz, Gini and others.

Since then, this type of curves of concentration have been used to measure distribution of industry, localisation of industry etc., in several cases. In each of these cases, however, one of the variables have been the density of population or, the ratio of the population in a region to the total and the other variable has been the proportion of industrial units or the proportion of wage earner etc., as the case may be, these being analogous to the original concept of Lorenz who related the percentage of population with a given income and the percentage of total income absorbed by that population.

In this paper* the curves of concentration have been used in a somewhat different way to measure the concentration of land owned in different ranges of ownership and the concentration of holding cultivated in these ranges as cumulative percentages of area owned under different acreage-intervals and the corresponding cumulative percentage of area cultivated in the same range respectively. This curve has been called the curve of relative concentration as in this case the concentration both of owned land and that of cultivated land has been considered. The two variables thus related are really the concentration of the volume of production and the concentration of the means of production respectively.

In this paper the curves of concentration obtained for the two variables as explained have been used to stratify different districts into groups with more homogeneous economic feature.

To analyse the economy of a region we always attempt to stratify the area into economically similar regions. Such process of stratification involves first the

*The data on which the concentration index is based, has been taken from a survey of Undivided Bengal carried out in 1945-46. The entire province was divided into cells each one mile square. 478 cells were picked up at random and surveyed fully. The survey covered 80,000 families of which the present data covers roughly 46,000, half of the families from each cell being selected at random. Sources of other data have been mentioned in the text. Sample survey unless otherwise mentioned relates to this survey only.

choice of control characteristics and second, some method of utilizing information on the control characteristics which will group the units into strata, each containing units relatively homogeneous with respect to the control characteristics. Such control characteristics usually chosen are non-quantitative classifications like geographic location, climate, soil character etc. or, if quantitative variables are chosen at all, they are by nature non-economic. It is very rarely that economic variables themselves are chosen for classifying regions of homogeneous economic features. This is because, usually, the economic variables we are interested in are large in number and selection of one among them in such a way as to get the best representation of them all is difficult. Extra economic variables are considered to be of a more fundamental nature and are therefore chosen as criteria of classification. It has been shown here, however, that even economic variables if they are of a fundamental importance as the index devised can be used as control characters successfully.

THE CONCENTRATION INDEX

The concentration of production and the means of production being of basic importance in agrarian economy, relationship of these two, it will be shown, may be used under certain conditions as a control characteristic for classifying regions homogeneous with respect to economic features which are usually considered to be of basic importance.

If $f_1(x)$ be the frequency of families owning at least x acres and if $f_2(x)$ be the average acreage cultivated by a family owning x acres

$$\text{Total area owned by families owning at least } x \text{ acres is } \int_0^x \frac{df_1(x)}{dx} x dx.$$

Similarly area cultivated by families owning at least x acres

$$= \int_0^x \frac{df_1(x)}{dx} f_2(x) dx.$$

Ratio of area owned to total

$$X = \frac{\int_0^x \frac{df_1(x)}{dx} x dx}{\int_0^M \frac{df_1(x)}{dx} x dx}$$

and ratio of area cultivated to total

$$Y = \frac{\int_0^x \frac{df_1(x)}{dx} f_2(x) dx}{\int_0^M \frac{df_1(x)}{dx} f_2(x) dx}$$

where $0 < x < M$.

ECONOMIC CLASSIFICATION OF AGRICULTURAL REGIONS

The nature of the functions $f_1(x)$ and $f_2(x)$ are of the form*

$$f_1(x) = a_1x^{b_1} \text{ and } f_2(x) = a_2x^{b_2}$$

It may be easily shown that X and Y can be related in the form

$$Y = a + b \log X$$

Curves of the Lorenz type on the two variables as defined above were fitted on 24 districts of Bengal.

Table 1 gives the observed and expected values of logarithms of percentage of area cultivated for one district (Bogra).

TABLE 1

Area owned in Acres	Cumulative	
	log y (observed)	log y (expected)
0	—	—
1	1.2514	1.2833
2	1.4051	1.4883
3	1.6512	1.6291
4	1.7327	1.7136
5	1.7915	1.7738
6	1.8310	1.8202
7	1.8580	1.8504
8	1.8720	1.8656
9	1.8894	1.8837
10	1.9035	1.8984
11	1.9129	1.9078
12	1.9253	1.9205
13	1.9334	1.9293
14	1.9419	1.9451
15	1.9460	1.9451
16	1.9524	1.9538
17	1.9568	1.9627
18	1.9640	1.9687
19	1.9675	1.9720
20	1.9686	1.9740
21	1.9709	1.9783
22	1.9747	1.9810
23	1.9770	1.9832
24	1.9794	1.9894
25	1.9794	1.9814
26	1.9820	1.9920
26 and above	2.0000	2.0174

It may be seen from a comparison of the expected and observed values that the nature of the fit obtained is quite satisfactory.

This points to the close connection between the two variables and it will be

*The nature of these functions as found empirically will be discussed by the author in a later communication.

seen later that the nature of relationship obtained between these two determine a large number of other characters of a region.

The integral of the curve, i.e., $\int_0^{100} y dx$ will measure the degree of concentration

by measuring the area between the abscissa and the curve. The bigger this area included under the curve the further away from production are the owners of the soil. The integration has been done from 10 to 100 as below 10 the graduation is not satisfactory.

The index of concentration thus gives the measurement of property production relation and therefore is the key to the basic structural nature of agricultural economy in any region.

The advantage of getting such a numerical device is that we can immediately get rid of vague general classification of different economic regions and classify straight by the most basic characteristic of an agrarian region which as it were combines the influence of other factors exogenous and endogenous in the most comprehensive form and this can be done in a mechanical, numerical form.

The constants of the equation as obtained for the 24 districts and the area under the curve is given in Table 2.

TABLE 2

Districts	a	b	Index (i.e. area under the curve)
Bakarganj	0.1367	0.9037	4564
Bankura	0.2459	0.8828	5351
Birbhum	1.1766	0.4007	8506
Bogra	0.3788	0.8194	5617
Burdwan	1.0016	0.4074	6412
Chittagong	1.1179	0.4425	6725
Dacca	1.0123	0.4908	6403
Dinajpur	1.2931	0.3438	6708
Fazidpur	0.2897	0.8583	5303
Hoogly	0.9042	0.8208	6015
Howrah	0.8688	0.8660	6230
Jalpaiguri	1.4165	0.2820	7001
Jessore	1.0650	0.4633	6487
Khulna	1.1612	0.4258	6702
Maida	1.0384	0.4924	6844
Midnapur	0.6886	0.6811	6540
Murshidabad	1.0655	0.4888	6249
Mymensingh	0.8996	0.7030	6843
Nadia	0.7809	0.8697	6826
Noakhali	0.2938	0.8553	6308
Pahna	1.1656	0.3967	6882
Rajshahi	0.9288	0.8373	6364
Rangpur	0.8030	0.8529	6274
Tipperah	0.2197	0.8021	6237
24-Parganas	0.7805	0.8182	6276

ECONOMIC CLASSIFICATION OF AGRICULTURAL REGIONS

The following table gives the correlation co-efficients of the values of this index for the 24 districts with the corresponding values of a large number of important variables in which an economist is generally interested.

TABLE 3

Variables	Values of the correlation co-efficients of index of concentration with the variable
Density (Census '41)	-.34*
Density (sample survey for only rural areas)	-.81**
Per cent of area cultivated by share-croppers	+.65**
Per cent of families share-cropping	+.72**
Per cent of cultivating owners (Census '31)	-.47*
Per cent of agricultural labourers (Census '31)	+.20*
Per cent of families cultivating (sample survey '45)	-.64*
Per cent of area utilized	-.43*
Per cent of cultivated area to cultivate	-.51**
Per cent of area double-cropped	-.32
Mean area per cultivator (sample survey '45)	+.61**
Mean bullock per acre (sample survey '45)	-.25
Value of crop per acre	-.42*
Value of crop per agriculturalist	-.51*
Mean loan per family	+.64**
Per cent of indebted families	-.45*
Proportion of rent to produce	+.43*
Rent per holding	+.50*

Significant at 5% level, **Significant at 1% level.

The concentration index is thus shown to be correlated with a large number of characters which determine the economy of any region. Naturally, therefore, a classification by this index will assure a very large degree of homogeneity as regards these characters.

GEOGRAPHICAL FEATURES

In this section the geographical significance of the results obtained in the earlier section will be discussed in detail.

The districts have been arranged according to this index and roughly contiguous geographical regions marked out.

The four regions marked out are

Region I—Jalpaiguri, Malda, Dinajpur, Khulna, Chittagong.

Region II—Hooghly, Midnapur, Birbhum, Jessoro, Burdwan, Dacca,

Rajshahi, Raingpur, 24 Parganas, Murshidabad, Howrah

Region III—Pabna, Mymensingh, Nadia, Bogra.

Region IV.—Faridpur, Nonkhali, Tipperah, Bakarganj.

It will be seen that on the whole this index succeeds in giving a generally contiguous geographical area. Roughly four regions have been marked out.

Before discussing the specific features within these regions the main point to note is

(1) The areas developed earlier historically starting from the North to the West of Bengal are areas with a high index of relative concentration. They are also areas, which are generally deteriorating in agriculture.

(2) The more recently developed and still flourishing agricultural areas of East Bengal the more South we go have a relatively lower index of relative concentration.

Physical description of these two areas as given will be of use in finding out the importance and place of this index as also in understanding the factors behind these indices being in such an ordered form.

"Practically, the whole of Western and Central Bengal with the exception of the littoral tracts lies in a moribund delta, in which the activity of the rivers has almost ceased and decadent conditions of health and subsistence have checked the growth of population through successive decades. As contrasted with the moribund delta there is the active delta covering the greater portion of Eastern Bengal. Here the great rivers are building up and throwing up new alluvial formations rich in agricultural possibilities, while sometimes there is diluvion of land sufficiently old and supporting a teeming population. But here also the varying stages in the life and activity of the rivers are clearly discernible. The onward march of the delta building rivers to the east continues to this day. The rivers have become less active in the northern portion of this region than in the south eastern portion covering the littoral tracts, particularly in the vicinity of the Megna estuary. It is in these littoral tracts that the healthiest and most fertile regions of the Ganges Delta are situated at the present time and in these tracts the increase in the number of teeming population knows no bounds."

It is obvious that a higher index has generally been associated with the deteriorating regions and a lower one with the prospering regions.

DEMOGRAPHIC AND ECONOMIC FEATURES

It will now be shown that apart from geographical contiguity the index is useful in classifying the region by other factors demographic, economic and social.

Of these factors the most important one is the density of population of these regions. Sample density refers to the density of the population (only rural) in the area surveyed in our enquiry classified for the respective regions.

TABLE 4

Region	Mean index of concentration	Density per sq. mile (Census '41)	Density per sq. mile (Sample survey '46)
I	6444	563	364
II	6393	718	428
III	5791	825	460
IV	5182	1066	741

The cause usually associated with agricultural decadence in our country is said to be high density. This obviously underestimates other factors behind agricultural decadence like lack of good irrigation, lack of good technique, exhaustion of soil etc., as also the property relation existing on the soil. It will be seen that

ECONOMIC CLASSIFICATION OF AGRICULTURAL REGIONS

It is precisely the areas with low density now that have become backward and decadent in institutional factors with the owners of land more separated from actual cultivation while comparatively more densely populated areas are still maintaining self-sufficient agriculture. It is useful in this connexion to quote the generalisations made by other writers on this problem.

"The important generalisation which can be made regarding the distribution of population in the Ganges Delta is that the density of population has a tendency to increase at a higher ratio in the fertile and healthy tracts of the active delta than in the decadent regions of the moribund delta, where both agricultural productivity and public health are at a low level. In Bengal 'during the last seven decades the relative distribution of population between north and south has changed comparatively little. The greatest changes of population distribution have been between east and west. The uninterrupted recession eastwards of the median point conveniently illustrates the proportionately larger growth of population in Eastern than in Western Bengal.' This is a striking statistical generalisation of very great import. It shows not only that at a particular point of time the density of population is higher in the active than in the moribund delta but also that the same dynamic natural factors, all of which are inseparably connected with the activity of the deltaic rivers are still operating to widen the differences in density through successive decades up to the present time. The moribund delta, *Rarh* as it was called in ancient times was the scene of activity of the deltaic rivers probably up to the sixteenth century. Here extensive areas had been brought under cultivation and land had always supported a high density of population. But the tide has definitely turned in comparatively recent times as a result of the shifting of the course of the Ganges. There has been widespread deterioration of soil owing to the slow death of the rivers: defective drainage has rendered vast tracts of the country insanitary and unfit for human habitation: population has been decimated by diseases like malaria and cholera and those who survive as victims of malaria have little vitality left in them to resist natural calamities like epidemic diseases and unfavourable natural conditions under which agriculture has to be practised. The result has been that agricultural productivity has declined, the death rate has increased, while at the same time there has been a steady migration of people to healthy and fertile regions.

"In the active delta on the other hand, there has been a steady growth of population through successive decades. Here there has been great extension of cultivation of new alluvial formation. Agricultural productivity has always been high owing to favourable natural condition. The death rate has been low because of the comparative immunity of the people to malaria and cholera, which are the scourge of the moribund delta. And it is a striking fact that although there is migration of people from the north to the littoral tracts in the south of the fertile land near the Meghna estuary in the south-east of it the increase of population in this region as a whole is due principally to a very great natural increase of population."

The prevalent idea is to associate a highly populated area with poverty and a sparsely populated area with comparative prosperity. But in reality as seen here

the opposite is the case. Property-production relation is generally more favourable in densely populated areas and is most unfavourable in sparsely populated decaying area.

The association of low density with decadent agriculture is a feature in backward or under developed regions where routine exploitation of soil over long periods retards the growth of population itself as the constant level of technique brings the Malthus law into operation.

Table 5 gives the concentration index with percentage of area share-cropped, leased and the percentage of share-cropping families. These factors will not give additional information but will help us to understand the physical significance of the index itself.

TABLE 5

Region	Mean Index of concentration.	Percentage of area cultivated by share croppers or lease holders.	Percentage of area share-cropped.	Percentage of area leased.	Percentage of families share-cropping.
I	6844	44.3	38.6	4.7	82
II	6395	34.8	29.0	6.8	66
III	6791	35.7	21.6	14.1	63
IV	6182	21.6	12.9	8.7	20

It may be concluded that tenure condition of areas with higher index is parasitical in its nature with owners away from cultivation, small owners cultivating more land in share-cropping while the reverse is the case for areas with lower index. Also that lease which implies better terms is more associated with a lower concentration index and that leasing is more taken in proportion to the area share-cropped than is the case under the share-cropping system.

Table 6 gives the general structure of ownership in these regions and the percent of area cultivated under different ranges of ownership.

TABLE 6

Region	Percentage of total families owning land in acres					Percentage of area owned by families with		Percentage of area cultivated by families owning	
	0	0-15	5-10	above 10	total	0-10 acres	above 10 acres	less than 10 acres	above 10 acres
I	41	41	11	7	100	52	48	76	24
II	37	49	10	4	100	67	33	84	14
III	26	60	9	5	100	70	30	83	17
IV	20	73	6	1	100	88	12	88	12

As is to be expected, of course, the higher the concentration index the more polarised in reverse directions are the distribution of area owned and area cultivated. While under capitalist condition concentration of ownership of the means of production and the concentration of production in a smaller number of hands go together here because of the backward condition of technique concentration of ownership does not go hand in hand with concentration of cultivation. Cultivation has more or less remained intact in small plots while the structure of ownership is highly polarised.

ECONOMIC CLASSIFICATION OF AGRICULTURAL REGIONS

The occupational structure of an area, it will be seen now, is very deeply affected by the specific property-production relation given by the concentration index.

TABLE 7

Region	Mean Index of concentration	Percentage of earners to total (Census '31)			Percentage of families depending on Agr. Lab. (Floud's Commission)	Percentage of families cultivating	
		Cultivating owners	Tenant cultivators	Agr. labourers		Sample Survey	Census '41
I	8541	27.6	9.3	22.8	22.4	44	63
II	8295	31.9	6.2	22.7	25.3	54	67
III	8791	40.0	6.8	18.8	24.1	62	70
IV	8182	52.6	4.0	15.4	15.5	70	77

It may be seen that the census of 1931 and the survey of 1946 as also the Floud's Commission are unanimous about (1) the relatively higher density of cultivators in the regions with lower index of concentration, (2) the relatively higher density of labourers with the higher index of concentration. The correlation coefficient actually calculated for the index of concentration with census percentage of cultivators is -0.47 and with the percentage of agricultural labourers it is -0.54 . The significance of these factors is to be clearly noted. A higher concentration index implies that the agrarian economy of the region has lost its self-sufficient peasant character and capitalist methods of cultivation has made its appearance in these places. Further that, people have been forced out of their dependence on the soil. Also that landless pure agricultural labour has become far more frequent along with the growing pauperisation of the peasantry.

Areas with low concentration index on the other hand still have self-sufficient peasant base not differentiated nor divorced from the soil. Thus areas where a high concentration index has been obtained are precisely the areas where decadence in agriculture and destruction of self-sufficient peasant economy is moving together.

The question of self-sufficient peasant economy and its destruction can best be studied in the background of the nature of development of the area. Table 8 gives the different regions and the percentage of area developed, area double-cropped, area left fallow etc., which are the best indicators of the intensity of utilisation and cultivation.

TABLE 8

Region	Utilised	Culturable waste	Unculturable waste	Fallow	Orchards	Double cropped	Cultivable to cultivated
I	41.6	9.9	17.2	3.5	7.7	14.2	51.8
II	49.2	7.8	16.5	2.2	4.9	16.0	62.0
III	65.3	8.0	15.9	1.6	3.6	30.0	64.5
IV	75.0	4.1	21.1	1.4	4.4	29.3	92.0

Combining the results of this table and the previous one it may be concluded that the areas with a lower index of concentration are precisely the areas which are being fully utilised and cultivated because of favourable conditions, where waste and uncultivable fallow are lower and double cropping is high. But orchards and other cash crops like betels etc., are cultivated more generally in areas with a high index, while subsistence crops are more generally the rule in areas with lower index. This is another feature of the increasingly capitalistic development associated with a high index.

The whole information is summarised adequately through a productivity measure and other features of prosperity of these regions as given in Table 9.

TABLE 9

Region	Mean area per tiller in acres	Mean bullock per acre	Value of crops per acre in Rs.		Mean loan per family in Rs.	Percentage of indebted families
			Food's Commission	Season & Crop Reports		
I	7.3	0.40	45	44	46.0	34.7
II	5.1	0.40	46	51	66.2	51.0
III	5.0	0.29	51	49	65.5	56.4
IV	2.5	0.22	56	54	95.2	58.8

A lower concentration index thus means a higher productivity and more credit but considerably smaller size of holdings. Thus while on an average these areas are more prosperous there is less polarisation and prosperity is comparatively more evenly distributed. Peasantry is more creditworthy but their units of cultivation are not large. Larger holdings associated with a more capitalist organisation is present in the areas with high index where the bulk of the peasantry is actually worse off. The average bullock unit also follows the same trend.

In Table 10, the rates of exploitation in different areas as revealed by the *rent per holding* (i.e., the amount paid by a tenant to a landlord) proportion of rent to produce, cess per acre etc., is given.

TABLE 10

Region	Proportion of rent to produce	Rent per holding in Rs.	Debt per family in Rs.	Cess per acre in Rs.
I	30	51.5	33.7	0.35
II	33	48.4	51.0	0.45
III	18	38.4	56.4	0.30
IV	24	29.0	58.8	0.28

It will be seen that rack-renting and more intense exploitation and taxation goes on in the areas with a higher index and is in fact inversely related to population. Thus what makes for intense exploitation is not population pressure so much as the extent of disintegration of the peasantry. Whatever force helps the disintegration of the peasantry changes the concentration index. Degree of exploitation is highest precisely where due to natural factors agriculture is decaying and the poor peasantry is finding it hard to maintain an existence.

The concentration index thus based on land owned and land cultivated very adequately summarises a whole set of relevant economic factors and yet unlike physical and other characters are precisely and objectively measurable.

REFERENCES

- GANDOLY, B. N. (1938): *Trends of agriculture and population in the Ganges valley*. Methuen & Co. Ltd. London.
- LORENTZ, M. O. (1903): Methods of measuring the concentration of wealth. *Quarterly Publications of the American Statistical Association*, 9 (New Series No. 70), 209-219.

Paper received: 1 March, 1949.