

Socio-Economic Indicators in Rural Pakistan: Some Evidence

MAHMOOD HASAN KHAN and MAHMOOD IQBAL*

This study attempts to quantify the inter-provincial and interdistrict differences among villages with respect to the availability of and access to certain crucial inputs and services which affect the level of living and even quality of life in rural areas. Using the taxonomic method for rank ordering and clustering of regions, this comparative analysis should help in identifying the areas (provinces or districts) and activities (or services) in each area which may require immediate or special attention of policy makers and planners.

I. INTRODUCTION

Almost three-quarters of the population of Pakistan resides in about 38,000 villages or *mauzas*. An even higher proportion of the real poor of the country live in and around these communities. They are probably the most visible and legitimate "target groups" for rapid socio-economic development. Their individual efforts to break out of the "vicious circle" of poverty are frustrated not only by their own inability to acquire the needed information and resources from the market-place, but also by the woefully inadequate infrastructure and support services provided by the public sector. In fact, provision of inputs and public services can make all the difference in the life of many in rural areas.

As in many other underdeveloped countries, disparities among communities and individuals in Pakistan are often very striking. Casual observation provides overwhelming evidence of the generally inferior position of the rural folk *vis-a-vis* their urban counterparts. The "urban bias" is evident in almost all public services and activities, although rural constituency is much larger. But then villages, too, have equally visible and sometimes vast differences in each province and even within one *tehsil* or *taluka*. Finally, the disparities between individuals within a village are perhaps most disheartening. For these disparities, there are several historical and contemporary reasons.

*The authors are in the Department of Economics, Simon Fraser University, Canada. This study was completed when they were Visiting Professor and Visiting Economist at the Applied Economics Research Centre, University of Karachi, Pakistan. They are grateful to Dr. Tariq Siddiqui for his help in getting the relevant village statistics. Also, they thank the Editor and referees for their useful comments on an earlier draft.

The object here is limited to quantifying the inter-provincial and inter-district differences among villages with respect to the availability of and access to certain crucial inputs and services which most likely affect the income levels and even quality of life. This comparative analysis should help in identifying the areas (provinces or districts) and activities (or services) in each area which may require immediate or special attention of policy makers (politicians) and planners (bureaucrats or economists).

This study is, of course, not the first attempt to identify differences in the level of socio-economic indicators in various administrative units (districts/*tehsils* or *talukas*) of Pakistan. A large report by the Esesjay Consultants in the mid-Seventies [3] was among the few earlier attempts, although it was not published or circulated widely. However, it is important to stress here that our study is significantly different from these exercises and, in at least one respect, unique: the taxonomic method used here is demonstrably a stronger test than factor analysis or rank correlation used by others for rank ordering and clustering of regions or administrative units. Secondly, this study focuses on only those socio-economic indicators which have a bearing on the standard of living or quality of life in rural areas. Finally, we use the most recent (1980) data, which are also probably more complete than those used in earlier studies.

The paper is organized as follows. Section II contains a description of the socio-economic indicators and explanation of the numerical procedure we use for analysis. The results are discussed in Section III. Finally, the concluding comments, including the usefulness and limitations of this exercise, are given in Section IV.

II. METHOD AND DATA

We have selected 22 indicators for the inter-provincial and inter-district comparisons of rural life. We have divided the indicators into two groups.

Group A. Here the indicators show the availability of inputs and facilities which affect levels of production and employment in the village:

1. irrigation facility,
2. cottage industry,
3. sweet drinking-water,
4. electricity,
5. tractors, and
6. tubewells.

Group B. The indicators here show the distance at which the facilities or services are located from the village, which may determine the economic well-being and quality of life in the village:

1. metalled road,
2. railway station,

3. post office,
4. grain market,
5. fertilizer depot,
6. office of field assistant,
7. *tehsil/taluka* headquarters,
8. police station,
9. diesel pumps,
10. tractor workshop,
11. veterinary hospital,
12. civil hospital/dispensary,
13. bank,
14. primary school,
15. middle school, and
16. high school.

Here a word on choice of indicators is in order. We know that levels of production and income in a region depend on a host of factors. Some of these factors are used as inputs by individuals from the market-place; others are necessarily provided as infrastructural services or facilities by the public sector. Also, some factors affect the level of production directly, but others have a less direct impact. While each of these indicators has its specific (direct or indirect) role in determining the standard of living, the total impact is effected through a complex interactive process. In this study, we are assuming that the selected indicators are equally important. Of course, as we point out in the final section, the precise magnitude of the contribution of each indicator can only be tested by regression analysis on micro-level data collected at the village level.

We adopt two procedures in this study. In the first, we establish a ranking order of provinces and districts. The second procedure involves grouping or clustering of districts with similarities. These methods are not new. They have been adapted from studies on inter-country differences [1;6].¹

(1) Rank Ordering of Provinces and Districts

An "ideal" area (province or district) is one which has the "best" values for each socio-economic indicator. The best values for the indicator are either the highest percentage of villages with facilities or services in Group A or highest percentage of villages with the least distance (up to 8 kilometers in each case) from

¹Comparative analysis of regions and countries has been done by methods like factor and discriminant analysis, canonical correlation and multiple regression. They all imply causal relationships when, in the complex interactive process of development, there is probably no one-way and unambiguous functional relationship between the so-called independent and dependent variables.

the facility or service in Group B. The difference between the "ideal" province or district and any observed province or district is called the Disparity of Development (*DD*) and is measured by:

$$(DD)_{i\lambda} = \left[\sum_{i=1}^N (x_{ij} - x_{\lambda j})^2 \right]^{1/2} \quad \dots \quad (1)$$

$$i = 1, 2, \dots, N$$

where x_{ij} = location of the *i*th province or district for the *j*th indicator; and

$x_{\lambda j}$ = location of the ideal province or district for the *j*th indicator.

The value of *DD* is normalized by a "critical" value of the distance from the available facility of the ideal province or district. This will give us the Level of Development (*LD*). The value of *LD* will range from zero (the most developed province or district) to one (the least developed province or district). The "critical" value (C_{λ}) of the distance is calculated by

$$C_{\lambda} = \frac{1}{N} \sum_{i=1}^N DD_{i\lambda} + 2 \left[\frac{1}{N} \sum_{i=1}^N \left\{ DD_{i\lambda} - \frac{1}{N} \sum_{i=1}^N DD_{i\lambda} \right\}^2 \right]^{1/2} \quad \dots \quad (2)$$

The calculated value of the "critical" distance from the "ideal" province or district is then used as the normalizer for determining the Level of Development (*LD*):

$$LD = \frac{DD_{i\lambda}}{C_{i\lambda}} \quad \dots \quad (3)$$

$$i = 1, 2, \dots, N$$

(2) Clustering of Districts

With this method, we can identify clusters of districts on the basis of similarities. Euclidean distances from one district to every other district can be estimated and clusters of "primary", "secondary" and "tertiary" groups can be identified. Districts belonging to different clusters are determined by the values of the Level of Development (*LD*). The values range from 0.0 to 0.49 for the "primary" cluster, from 0.50 to 0.74 for the "secondary" cluster, and from 0.75 to 1.0 for the "tertiary" cluster.

We will measure the rank orders and clusters separately for three groups of socio-economic indicators:

Group A: only on the basis of indicators in Group A;

Group B: only on the basis of indicators in Group B;

Group C: on the basis of a composite of indicators in Groups A and B.

The data on the 22 socio-economic indicators included in this study are from the village statistics published by the Agricultural Census Organisation as part of the 1980 agricultural census in Pakistan [8].

III. RESULTS AND DISCUSSION

We will here discuss the estimations of rank ordering of the four provinces and rankings and clustering of the fifty districts in Pakistan.

1. Rank Ordering of Provinces and Districts

We will first describe the numerical procedure we have used in establishing the ranks of provinces and districts on the basis of the three sets (groups) of indicators. For Group A indicators, the value of $DD_{i\lambda}$ is the under-root of the sum of squares of the difference between the percentage weighted averages of availability of facilities to villages in each district and in the ideal district. For Group B indicators, the value of $DD_{i\lambda}$ is likewise the under-root of the sum of squares of the differences between the percentage weighted averages of the location (within 8 kilometers) of the facilities from villages in each district and in the ideal district. An "ideal" district is one in which the highest proportion of villages have facilities of Group A and the least distance from facilities in Group B.

The weight for each economic or social indicator in both Groups A and B is calculated by finding the percentage share of villages of a province or district in total villages in the country. For the Group B indicators, each facility is again weighted on the basis of its location from the village at distances of up to one kilometer, over one to three kilometers, and from more than three to eight kilometers. An average is then taken of the weighted aggregate.

The ranking of provinces or districts on the basis of all indicators, Groups A and B together, is also based on the value of $DD_{i\lambda}$.

In each of the three rank orderings, based on indicators of Groups A, B and C, the lower the value of $DD_{i\lambda}$ of a district the higher its standing in the order. The value of *LD*, which is the adjusted value of *DD* by the "critical" distance (C_{λ}) ranges from zero to one. The closer the value of *LD* of a district or province to zero the higher its rank in the order.

Results of the rankings are shown in Tables 1 and 2. Among provinces, the Punjab ranks first in all three orderings, with Sind taking the second, the N.-W.F.P. the third, and Baluchistan the last position. More importantly, villages in the Punjab are better placed than villages in other provinces in the case of each indicator, except

Ranking of Provinces of Pakistan

Province	Individual Indicators (A)						Individual Indicators (B)														Group				
	Irrigation facility	Cottage Industries	Sweet Drinking Water	Electrification	Tractor Owned	Tubewell Owned	D. to Mettalled Road	D. to Railway Station	D. to Post Office	D. to Grain Market	D. to Fertilizer Depot	D. to Office of F. A.	D. to Tehsil HQs.	D. to Police Station	D. to Diesel Pumps	D. to Tractor workshop	D. to Vet. Hospital	D. to Civil Hospital	D. to Bank	D. to Primary School	D. to Middle School	D. to High School	A	B	C
N.-W.F.P.	3	2	1	1	3	3	3	3	3	3	2	3	3	3	3	2	2	3	3	3	3	2	3	3	3
PUNJAB	2	1	2	3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
SIND	1	3	3	2	2	2	2	2	2	2	3	2	2	2	2	3	3	2	2	2	2	3	2	2	2
BALUCHISTAN	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4

Table 2

Measures and Rank of Districts of Pakistan on Socio-Economic Indices of Development

	$(DD\lambda)_A$	$(LD)_A$	RANK	$(DD\lambda)_B$	$(LD)_B$	RANK	$(DD\lambda)_C$	$(LD)_C$	RANK
Peshawar	83.7	0.46	9	625.6	0.65	11	631.1	0.69	11
Mardan	107.9	0.59	17	731.7	0.76	19	739.6	0.80	18
Kohat	117.5	0.65	21	269.6	0.80	23	778.5	0.85	23
D. I. Khan	130.3	0.72	25	767.7	0.80	23	778.6	0.85	23
Bannu	107.9	0.59	17	707.4	0.73	16	715.6	0.78	16
Abbottabad	155.5	0.85	32	704.0	0.73	16	720.9	0.79	17
Mansehra	151.5	0.83	31	729.2	0.76	19	744.7	0.81	19
Rawalpindi	155.6	0.85	32	552.5	0.57	6	573.9	0.62	7
Attock	111.2	0.61	19	722.0	0.75	18	730.5	0.79	17
Jhelum	134.0	0.74	26	617.1	0.64	10	631.5	0.69	11
Gujrat	95.3	0.52	14	430.7	0.45	3	441.1	0.48	3
Sargodha	86.0	0.47	15	584.8	0.61	8	591.0	0.64	8
Mianwali	107.6	0.59	17	693.7	0.72	15	701.9	0.76	15
Jhang	72.1	0.39	5	599.2	0.62	9	603.5	0.66	10
Faisalabad	75.7	0.41	6	445.6	0.46	4	451.9	0.49	4
Lahore	38.6	0.21	1	679.6	0.70	14	680.7	0.74	14
Kasur	62.9	0.34	4	651.7	0.68	12	654.7	0.71	12
Sheikhupura	52.1	0.29	3	562.8	0.58	7	565.2	0.61	6
Gujranwala	47.6	0.26	2	506.8	0.52	5	509.0	0.55	5

Table 2 - (Contd.)

	$(DD)_A$	$(LD)_A$	RANK	$(DD)_B$	$(LD)_B$	RANK	$(DD)_C$	$(LD)_C$	RANK
Sialkot	89.8	0.49	12	05.9	0.00	1	90.0	0.10	1
Multan	93.8	0.51	13	499.7	0.52	5	508.4	0.55	5
Vehari	81.8	0.45	8	667.0	0.69	13	672.0	0.73	13
Sahiwal	77.9	0.43	7	423.7	0.44	2	430.8	0.47	2
Muzaffargarh	111.7	0.61	19	599.0	0.62	9	609.3	0.66	10
D. G. Khan	125.8	0.69	23	719.3	0.75	18	730.2	0.79	17
Bahawalpur	96.6	0.53	15	661.7	0.69	13	668.7	0.73	13
Bahawalnagar	112.6	0.62	20	627.1	0.65	11	637.1	0.69	11
R. Y. Khan	109.1	0.60	18	584.3	0.61	8	594.4	0.65	9
Quetta	78.2	0.43	7	805.8	0.84	27	809.5	0.88	25
Pishin	146.7	0.81	29	803.7	0.83	25	816.9	0.89	26
Loralai	147.9	0.81	29	791.7	0.82	25	805.4	0.88	25
Chagai	165.1	0.91	34	816.5	0.85	28	833.0	0.91	28
Zhob	151.0	0.83	31	805.3	0.84	27	819.3	0.89	26
Sibi	155.0	0.85	32	796.4	0.83	26	811.3	0.88	25
Nasirabad	155.3	0.86	33	798.7	0.83	26	813.6	0.89	26
Kacchi	180.8	0.99	36	792.5	0.82	25	816.3	0.89	26
Lasbela	170.0	0.94	35	802.4	0.83	26	820.2	0.90	27
Hyderabad	85.0	0.46	9	757.3	0.79	22	762.0	0.83	21
Badin	137.2	0.75	27	773.6	0.80	23	785.6	0.85	23
Thatta	143.2	0.79	28	778.3	0.81	24	791.3	0.86	24

Continued -

Table 2 - (Contd.)

Dadu	129.8	0.71	24	764.3	0.79	22	775.2	0.84	22
Tharparkar	149.5	0.82	30	745.5	0.77	20	760.3	0.83	21
Sanghar	125.7	0.69	23	745.1	0.77	20	755.6	0.82	20
Sukkur	88.4	0.48	11	776.0	0.80	23	781.0	0.85	23
Khairpur	81.3	0.45	8	760.6	0.79	22	764.9	0.83	21
Nawabshah	86.3	0.47	10	718.1	0.74	17	723.2	0.79	17
Larkana	110.5	0.61	19	752.3	0.78	21	760.3	0.83	21
Shikarpur	100.0	0.55	16	778.9	0.81	24	785.3	0.86	24
Jacobabad	111.8	0.61	19	772.6	0.80	23	780.6	0.85	23
Karachi	124.3	0.68	22	808.4	0.84	27	817.9	0.89	26

Source: [8].

(A) - Availability of economic facilities in a district.

(B) - Closeness of district to economic facilities.

(C) - A & B both.

(a) irrigation facility in which Sind is first, (b) sweet drinking-water in which the N.-W.F.P. occupies the top place, and (c) electricity in which the N.-W.F.P. is first and Sind is second. All of these results are as expected. The Punjab has a higher proportion of villages in the *barani* areas which depend entirely on rainfall as against most villages in Sind which have access to perennial and non-perennial canals. That the N.-W.F.P. is ahead in sweet drinking-water and electricity in its villages is also well documented. However, somewhat surprising is the second position of Sind in the availability of electricity to its villages.

These aggregate (provincial) results are not detailed enough for rural planning in the four provinces. One useful approach for this is the ranking of all districts in the country. This will help not only in clarifying the aggregate results for provinces, but also in identifying the really backward and advanced districts at the national level. Similar exercise can be done in each province by ranking its districts and even *tehsils* or *talukas* separately: see [4].

Rank ordering, on the basis of indicators of Groups A, B and C, of all the districts of Pakistan is shown in Table 2. In all these orderings, the districts from the Punjab are among the top 20 percent. The bottom 20 percent in each case are the districts in Baluchistan. On the basis of Group A indicators, the top seven districts (Lahore, Gujranwala, Sheikhpura, Kasur, Jhang, Faisalabad, and Sahiwal) are all from the Punjab. The districts of Sind in the top 20 percent are Khairpur and Hyderabad. Peshawar from the N.-W.F.P. and Quetta from Baluchistan are also in this group. The bottom 20 percent are mainly from Baluchistan: Kacchi, Lasbela, Chagai, Nasirabad, Sibi, and Zhob. Rawalpindi is the only district from the Punjab in this group, and so is Tharparkar alone from Sind. Mansehra and Abbottabad from the N.-W.F.P. are also in the bottom 20 percent.

On the basis of the distance from facilities, the top 20 percent of the districts in Group B are all from the Punjab: Sialkot, Sahiwal, Gujrat, Faisalabad, Gujranwala, Multan, Rawalpindi, Sheikhpura, Sargodha, and Rahim Yar Khan. The bottom 20 percent are all from Baluchistan, except for Karachi from Sind.

Finally, on the basis of the composite of A and B, the top 20 percent districts are all in the Punjab and the bottom 20 percent are all in Baluchistan, except Karachi and Thatta from Sind. The ranking on the composite basis is not very different from that in Group B. The weight assigned to the distance from villages to economic or social services seems to have a large bearing on the overall ranking of a district.

It should be noted that Sialkot, which has, overall, the first position among the districts of Pakistan, is in the top five districts, no matter which rank ordering is chosen. Sahiwal, Gujrat, Faisalabad and Gujranwala are the other four, and they are all in the Punjab. Chagai from Baluchistan ranks at the bottom in the country, and it is followed closely by other districts of Baluchistan in every ordering: Zhob, Kacchi, Lasbela, Sibi, and Nasirabad. The top and bottom rankings, no matter which

set of indicators one cares to choose, are not entirely surprising: the top districts are concentrated in the central and eastern regions of the Punjab and the bottom districts in the settled areas of Baluchistan.

2. Clustering of Districts

The position of a district is determined not only by its distance from the "ideal" district, as was done in ranking the districts, but also from its closeness (or Euclidean distance) to other districts. It is quite possible that similar districts may have different ranks. The "neighbours" of a district can be classed as primary, secondary and tertiary, based on the values of *LD*. The primary cluster has districts which have the lowest values of *LD*, or which are the "most developed", and are followed by the secondary cluster (the "moderately developed" districts), and the tertiary cluster (the "least developed" districts). In Table 3, these three types of clusters are given for each group of indicators, A, B and C.

The primary cluster comprises mainly the districts of the Punjab, particularly on the basis of indicators of Groups B and C. Only a few districts, like Peshawar (N.-W.F.P.), Quetta (Baluchistan), Hyderabad, Sukkur, Khairpur, and Nawabshah (Sind), find a place in the primary cluster on the basis of the indicators of Group A. The secondary and tertiary clusters in the three Groups (A, B and C) then have most other districts, particularly from Sind, the N.-W.F.P., and Baluchistan, in that order. The few districts of the Punjab in the primary clusters have their distant "neighbours" mostly in Sind and Baluchistan. Their close neighbours are also the districts from the Punjab.

If we determine the clusters of districts on the basis of individual indicators, as are shown in Table 4, the close neighbours may be quite dispersed geographically. For example, Jacobabad and Rawalpindi are the closest neighbours in irrigation facility, and Quetta and Lahore are the closest neighbours for electricity in villages. Similarly, the farthest neighbours in the clustering based on individual indicators could be in the geographic neighbourhood.

IV. LIMITATIONS AND POLICY IMPLICATIONS

The numerical procedures used in rank ordering and clustering of areas or regions are not without limitations. For one thing, the problem of collinearity is not entirely eliminated. If specific indicators have high correlation with each other, the individual effect of the collinear indicators cannot be assessed accurately. The second problem has to do with the implied assumption that all the selected indicators are equally important. This is a matter of judgement whether these indicators play equally important role in determining the production levels and quality of life in the countryside.

The policy implications of this exercise are almost self-evident. It is a powerful technique for identifying the relatively "backward" and "advanced" administrative or political units in the country. This should help policy makers and planners in allocating developmental resources to services and facilities in regions (or areas) where the deficiencies are evidently serious. That is not all. The taxonomic method can become an even more potent tool for planning purposes at the local level if the results on rank ordering and clustering of districts by selected socio-economic indicators are compared with evidence on levels of land and labour productivities, rural income per capita, and share of each district (or *tehsil/taluka*) in the level of income or production of the province (or district). This will, however, require a new set of data from recent sample surveys and aggregate figures from the 1980 agricultural census.²

REFERENCES

1. Anderberg, M. R. *Cluster Analysis for Applications*. New York: Academic Press. 1973.
2. Chaudhry, M. Ghaffar. "The Green Revolution and Income Inequality: Some Empirical Evidence from Rural Pakistan, 1960-1975". Ph. D. Thesis. University of Wisconsin. 1980.
3. Esesjay Consultants Ltd. *Identification of Under-developed Target Areas*. Study No. 1. Volume 1. Karachi. 1976.
4. Khan, M. H. *Agriculture in Sind: A Report to Sind Regional Plan Organisation*. July 1982. (Mimeographed)
5. Khan, M. H. *The Economics of the Green Revolution in Pakistan*. New York: Praeger Publishers. 1975.
6. Khan, M. H. and J. A. Zerby. "The Socio-Economic Position of Pakistan in the Third World". *Pakistan Development Review*. Vol. XX, No. 3. Autumn 1981.
7. Naseem, S. M. *Underdevelopment, Poverty and Inequality in Pakistan*. Lahore: Vanguard Publications. 1981.
8. Pakistan. Agricultural Census Organisation. *Pakistan Census of Agriculture 1980: Mouza Statistics*. Lahore. 1982.

²Most of the information on levels of productivity and income by provinces and districts is based on sample surveys in the late Sixties and early Seventies and on the agricultural census of 1972. See, for example, [2], [5] and [7].