The Pakistan Development Review
39 : 4 Part II (Winter 2000) pp. 793–806

Impact of Village-specific, Household-specific, and Technological Variables on Poverty in Punjab

TOSEEF AZID and SHAHNAWAZ MALIK

Rural poverty remains a serious problem in Pakistan, with more than 30 percent of rural population living in absolute poverty. In rural Pakistan there is a big gap between rich and the poor. While the stake of competition for position and status concerns the rich, the struggle for survival in the midst of increasing crises embarrasses the poor. The rural poor—the pauperised class—are week and powerless with inadequate command over resources relative to needs. In fact, the polarisation process which is making the rich richer and poor poorer is a consequence of poverty. Neither the poor nor the outside well wishers have the power to break the vicious deprivation trap. It is consensus that rural social structure is responsible for rural underdevelopment.

A number of attempts can be seen in the literature having discussed the different dimensions of the above phenomenon related to developing economies in general and to Pakistan in particular, e.g., [Ahmad (1993); Ali (1997); Allaudin (1975); Gazdar and Zaidi (1994); Jafri (1999); de Kruijk and Leeuwen (1985); Mahmood (1984); Mahmood (1999); Malik (1988); Malik (1996); Mujahid (1978); Naseem (1973, 1996); Shirazi (1995); Tahir and Ali (1999); Ali and Tahir (1999) and Bhatti *et al.* (1999)]. The present study is the continuity of the above efforts. However, this study discusses the enigma of poverty based on a village study.

Within above setting, a village survey has been conducted in the Southern Punjab. The study analyses the determinants of rural poverty in the above village and its relation with productive assets, tools and instruments, gender, rural employment, household size, earner size, education level, etc. The analysis explicitly takes into account such attributes as the village specific, household specific and technological variables.

Toseef Azid is Associate Professor, Department of Economics, Bahauddin Zakariya University, Multan. Shahnawaz Malik is Professor and Chairman, Department of Economics, Bahauddin Zakariya University, Multan.

Authors' Note: We would like to thank Dr Rehana Siddiqui Senior Research Economist, Pakistan Institute of Development Economics, Islamabad, for her valuable comments on this paper.

The study has three sections. Section I discusses the results of the logit model taking poverty as the dependent variable (poverty is one otherwise zero). In this section analysis presented is based on all income earners. In Section II an attempt is made to split up the whole series of households into different groups depicting the different level of per capita income which can differentiate the poor and the poorest. This is because in the dynamic state it is necessary to know that how many income groups exist in the village. For achieving this objective, an iterative methodology is developed. It may be helpful for the policy-makers to differentiate between poor and the poorest. Whereas Section III presents the concluding remarks.

SECTION I

As mentioned the study is based on a village survey conducted in Punjab containing 90 households. The village 'Wanda' (District Bhakkar, Punjab)[1] situated at a distance of 10 KM from the river Indus, which forms the boundary with the North-West Frontier Province, could be taken as a fair representative of the characteristics of the two provinces. The survey carried out in October/November 2000, makes no claim to being completely representative of rural Pakistan. We do feel, however, findings based on this sample, when broadly interpreted, can serve as useful generalisations [Malik (1996)]. The survey was a one shot exercise. Within the community, the objective was the total enumeration of household. The village had 90 households and 100 percent enumeration was obtained.

The risk of poverty for a household is taken to be a random event the occurrence of which is assumed to depend on village-specific (mainly infrastructural), technological and household specific variables. Empirical estimates of the risk of the poverty (or the probability of a household being poor conditional upon the set of explanatory variables) are obtained from a logit model.

The explanatory variables are classified into three categories: (1) village-specific, (2) technological, and (3) household-specific.

The list and description of the variables is given as:

Village-specific Variables

1. Tr If a bus or railway service exists between the nearest market and the village, the variable takes the value 1, and 0 otherwise.
2. Me If household visits the health centre, the variable takes the value 1, and 0 otherwise.
3. Cr If household availed the credit facility, the variable takes the value 1, and 0 otherwise.

Technological Variables

4. Tra If tractors are used by household, the variable takes the value 1, and 0 otherwise.

[1]See Appendix I.

5. El If household used electricity for farming, the variable takes the value 1, and 0 otherwise.
6. HYV Gross cropped area under HYV in acres.

Demographic Variables

7. HS Number of persons in a household.
8. $(HS)^2$
9. De Ratio of number of members (<14 and > 65 years) to HH size.
10. Age Age of head of the household.
11. $(Age)^2$
12. PR Ratio of number of workers to number of adults (i.e. household members > 14 years).
13. Fm Ratio of female workers to male workers in a household.
14. Ed The variable takes the value 1 if the highest educational level of any household member is higher than primary education, and 0 otherwise.

The dependent variable is 1 if household is under poverty line and 0 otherwise.

In the empirical estimations the effect of Tr variable is not measured because it is available to every respondent.

The logit model is used to study the above phenomenon. The explanation of the results is as below:

We estimated the effect of two village specific variables, i.e., credit (Cr) and medical (Me) facility used by the households. It is observed that the credit variable, Cr, possesses a significant coefficient, with a negative sign. As it is common in the rural areas that two sources of income namely, farm and non-farm are available. May be credit facility can enhance the efficiency of the inhabitants which in turn increase their farm as well as non-farm income. The coefficient of Medical (Me) also possessed a negative significant coefficient (–0.0723). It depicts that medical facility has negative effect on the poverty. A typical household in the rural areas may enhance his income through hiring out his wage labour. On the contrary, a prolonged illness will not allow him to increase his income level.

The next group of variables, tractor (Tra), electricity (El), and high-yield variety seeds (HYV), capture the effects of the adoption of the new technology on poverty among the rural households. All these have the significant coefficients with the expected negative sign. Their coefficients are –0.3461, –0.5341, and –0.1109 respectively.

The demographic variables, Hs, $(Hs)^2$, De, Age and $(Age)^2$ produce a few surprises. Note that first three variables have significant coefficients. The fact that the coefficient of Age and $(Age)^2$ are not significant suggests that across different age groups (of household heads) the risk of poverty did not vary. The structure of the rural society shows that instead of age the economic opportunities have the significant role in the growth of the income of the household. The coefficients of Hs

and $(Hs)^2$ are statistically significant, the signs are positive and negative and the values are 0.0327 and –0.0012 respectively. This implies that with an increase in household size the poverty will reduce. The coefficient of De (0.8650) has positive significant effect on the poverty. The coefficient of Pr (0.0034) is insignificant and gave the inconclusive results. Fm has a negative effect on the poverty, the value of coefficient is –0.1469. It implies that female earners among low income households supplement household income by working on nearby farms or in the relatively affluent homes as maids, subject to the constraints imposed by domestic cores, and religious and social considerations. Given the male participation rate, it is hypothesised that the higher the female participation, the higher the total household income and lower the risk of poverty.

Ed (–0.0143) has also negative effect on the poverty. It implies that the more educated has more potential to exploit the resources and technology.

From the above results it is concluded that variables belonging to each of three groups village specific, technological and household specific exercised a strong influence on the risk of being poor of understudy households. The risk of poverty is determined by a diverse set of factors belonging to these groups.

SECTION II

In this section an effort is made to estimate the different segments of the households according to their per capita income. It is observed that in a series of households' per capita income a jump is occurred after some observations. It implies that next observation is not of the present segment and belongs to the next segment of the household. One can observe that in a rural society a small difference in the income of the household makes no difference but a large difference essentially changes the social status. So on the basis of that it is assumed that economic and social status of the households belonging to each segment which are significantly different to each other is not similar.

Procedures for splitting up the data into different segments are available in the literature, but almost all the existing literature deals with time-series analysis. Because time is the most important variable which guides the analyst in splitting up the data into different segments, for example data before and after war, would naturally be split into three segments: peace, war and peace. Similarly data before and after on oil crises, would be split into two segments and so on. Similarly when economists deal with the techniques of production, they separate data into different sections according to time, and one segment is different to others due to some major events (such as war, innovation, any major political decision, a major social change, etc.) which are thought to have significant effects on the economic variables, e.g., [Brown and Popkin (1962)]. However, in the existing literature it is hard to find statistical techniques which can be used to split up cross-section data into different segments which are significantly different from each other. As discussed in the

introduction, an attempt is made to split up the series into different segments depicting the different level per capita income which can differentiate the poor and poorest. With this objective in mind an iterative methodology is developed in this section [Azid (1994)]. The main features of the methodology are as below:

(a) The data of the per capita income of each and every household is arranged according to the ascending order.
(b) The corresponding series of the income of the households is converted into the cumulative format then transformed into percentage.
(c) Take the difference of per capita income, from a larger value to the next lower, i.e., from the bottom to the top of the series (data is in ascending order).
(d) The difference is regressed on the relevant percentage of total income of the village.
(e) Delete the first observation and again regress on the percentage of the total income, and repeat the procedure.
(f) The 'jump mean of square' is obtained by taking the difference of the total sum of squares of the first regression from the next regression, divided by the degree of freedom (which is obviously equal to one).
(g) The 'error mean of squares' is the total sum of squares of the next regression divided by the total degrees of freedoms.
(h) The F-value is the ratio of 'jump mean of squares' to 'error mean of squares'. If the value is significant at these degrees of freedom, then continue the procedure, otherwise stop the iteration; there is no need to go further.
(i) The values between any two significant jumps are regarded as representing one segment.

By using the above methodology, different segments of the households based on their per capita income can be found, which are prevailing simultaneously, and a line can be drawn between various segments.

By applying the above technique fourteen segments are found in the given series of households' per capita income. Table 1 depicts the full description of the segments according to their first and last observation.

Table 1 depicts the fourteen segments which are statistically significant to each other based on the households per capita income (for the reconfirmation of the statistical significance of these segments we apply the Kruskal-Wallis H Test).[2] It is observed from the above table that than 28.8 percent households have only 13 percent share in the per capita income whereas 12 percent households have 46 percent (approximately). 86.6 percent households have less than 54.13 percent share in the per capita income. This gives us a skewed distribution of income with a Gini coefficient of 0.372. In the third segment (416.66-510.42) in which poverty line is

[2]See Appendix II.

Table 1

Description of the Segments Based on Households' Per Capita Income

No.	Size of Segment (Per Capita Income)	Number of HH	Share in the Total Income (%)
01	300.00	01	00.28
02	333.33	03	01.28
03	416.66-510.42	22	13.00
04	520.80-555.50	11	20.29
05	583.33-638.83	08	25.05
06	666.66-700.00	11	33.05
07	714.25-800.00	09	40.73
08	833.33-1071.42	13	54.13
09	1166.66-1833.33	06	68.81
10	4166.66	01	76.63
11	4583.33	01	81.79
12	5000.00	01	87.41
13	5500	01	93.67
14	6750	01	100.00

Source: Based on village survey conducted by the authors.

falling we found more cluster of the households, absolute number is 22. Most of the households are falling from 3rd to 9th segments. For the empirical analysis we will not consider the first two and the last five segments assuming them as abnormal observations.

Table 2 depicts the share of per capita income and number of households from 3rd to 9th segments.

Table 2

Households and Their Share in the Per Capita Income in the Selected Segments (%)

No.	Size of Segment	HH (%)	Share in the Total Income
03	416.66-510.42	24.44(6)	11.72(5)
04	520.80-555.50	12.22(4)	07.29(2)
05	583.33-638.83	08.88(2)	04.76(1)
06	666.66-700.00	12.22(4)	08.00(4)
07	714.25-800.00	10.00(3)	07.68(3)
08	833.33-1071.42	14.44(5)	13.40(6)
09	1166.66-1833.33	06.66(1)	14.68(7)

Source: Based on village survey conducted by the authors.
Note: Values in the parenthesis show their relative ranks.

An attempt is made to estimate the correlation between percentage of the household and their share in the per capita income. The estimate is positive (0.209) but not significant. It gives more strength to our hypothesis that every segment is independent to each other and has different social and economic characteristics. Owing to above, an attempt is made to see the effect of specific variables on these segments separately (segment nine is not examined because of the lower degree of freedom).

Segment 3rd to 8th are classified as:

Segment Number	**Classification**
03	Poorest
04	Poorer
05	Poor
06	Lower Middle
07	Middle
08	Upper Middle

Table 3

Regression Results of the Classified Segments

Groups→ Variables↓	Poorest	Poorer	Poor	Lower Middle	Middle	Upper Middle
Me	0.6439	0.7342	0.5417	0.3218	0.4433	0.7735
	(2.07)*	(1.38)	(1.33)	(1.09)	(0.98)*	(2.65)*
Cr	0.3421	0.5620	0.9845	0.3322	0.5340	0.8541
	(1.07)	(2.38)*	(1.33)	(2.10)*	(2.76)*	(2.90)*
Tra	0.4208	0.6589	0.8703	0.0975	0.9418	0.0967
	(2.90)*	(2.90)*	(2.39)*	(2.65)*	(0.89)	(0.65)
El	0.7508	0.9357	0.7600	0.4090	0.0872	0.3008
	(2.09)*	(2.88)*	(2.76)*	(2.99)*	(2.66)*	(2.43)*
HYV	0.4458	0.6809	0.0089	0.0023	0.0234	0.0486
	(21.02)	(1.54)	(1.23)	(2.96)*	(2.67)*	(2.52)*
HS	–0.0965	–0.2987	–0.0721	0.6510	–0.9736	–0.8531
	(–2.07)	(–2.79)*	(2.19)*	(2.19)*	(–2.75)*	(2.35)*
De	–0.8934	–0.0954	–0.6408	–0.0842	–0.0780	–0.0630
	(–1.04)	(–2.54)*	(–2.23)*	(–2.69)*	(–1.53)	(–1.25)
PR	0.4398	0.4095	0.4093	0.0950	0.0933	0.0630
	(0.97)	(2.39)*	(0.93)	(1.09)	(2.86)*	(2.830*
Fm	0.2334	0.2392	0.2390	0.3003	0.2508	0.0390
	(2.04)*	(2.45)*	(2.73)*	(2.60)*	(1.09)	(1.35)

Table 3 presents results of an (OLS) analysis of the variations in per capita income of the households on the basis of distinct classified groups. This is an

alternative explanation of the relation between the set of variables and per capita income of the households. Most of the results are similar to those of the logit analysis and hence the findings of Section I are confirmed.

As expected the two village-specific variables, medical facility and credit, have positive effect on the per capita income of the households of all the classified groups. Medical facility seems to show strange positive effects in favour of lower income segments whereas the coefficients of credit seem to benefit more the higher income segments.

The technological variables, use of tractor, electricity and HYV have shown positive significant relationship with the per capita income irrespective of classified groups. The lower income groups seem to have benefitted from hiring the tractor services which is available to every household. On the other hand, HYV benefits more the higher income segments due perhaps to the reason that they have larger areas of land where HYV is used. The electricity benefits all uniformly.

Each of four household-specific variables namely, household size, dependency ratio, participation rate and female-male ratio produce coefficients with expected signs. The coefficient of household size and dependency ratio are inversely related whereas those of participation rate and female-male ratio are positively related to per capita income of households. However, a significant feature of the results is observed that each segment has its own magnitude of coefficient with different level of significance. This suggests that for a meaningful analysis different segments of a population of households may be analysed separately, e.g., Engel elasticity for each segment may be estimated which has its own significance in the economic literature [Mathur (1967)].

SECTION III

Summary and Concluding Remarks

Some general observations based on the major findings of the study are made here to put the discussion in perspective:

(a) Most of the variables belonging to each of the three groups of village-specific, technological and household-specific showed a strong influence on the risk of being poor for the village households.

(b) It has been shown that the probability of falling below the poverty line is lower for a village household with a larger area to cultivate for its own, a smaller number of dependents, greater participation in farm and non-farm work and a higher education level which increases the non-agricultural opportunities available to a village households. The other such variables are availability of credit and medical facilities to the households.

(c) As expected, the adoption of new technology in farming had a strong poverty reducing effect among the village households.

(d) On the contrary, the probability of falling below the poverty line is greater if the village population has fewer alternative opportunities for the labour households and hence fewer access to gainful employment.

(e) An attempt is made to split up the whole series of households into different income segments to differentiate poor and the poorest. This exercise enables us to know that the village income distribution is highly skewed with a Gini coefficient equal to 0.37 and a landholding Gini coefficient very close to 0.50. In such a setting the large income groups and land owners benefit at the expense of sections of small landowners tenants and agricultural labourers.

(f) An alternative explanation of the relation between village-specific, technological and household specific variables and per capita income of households has been provided using OLS analysis as the basis of distinct classified groups. Most of the results are similar to those of the logit analysis thus confirming those results.

APPENDICES

Appendix I

BACKGROUND TO VILLAGE SURVEY

The village (called 'Wanda' located in Punjab province) survey was conducted in 2000, for six continuous weeks. The survey was mainly based on a household questionnaire largely concerned with quantitative economic analysis. The format of the questionnaire was such that the information could easily be transformed on an individual basis. The modes of the data collection were the following:

(i) direct questioning of household head and other members;
(ii) extracting data from participant observation; and
(iii) interviewing of selected informants.

The survey was a 'one-shot' exercise, and repeated survey were not possible. The event of the recent past (agricultural data, etc.) had to be based on memory recall of respondents with cross checking from co-residents.

Within the community, the objective was the total enumeration of households. The village had 90 households and 100 percent enumeration was obtained. In general, households tended to have multiple attributes in terms of sectoral and organisational involvements. Data on production activities, income and employment were obtained.

The village consisting of 99 households is connected to the nearest town (called 'Darya Khan' at a distance of 8 miles) by a single metalled road. It was

electrified only two years ago and has educational facility upto the primary level. The primary health centre is located at distance of 3 miles.

The village agricultural land is plain and mostly cultivable. The land-tenure system consists of both owner-cropping as well as share-cropping. The main crops of the area are wheat, sugar cane, maize, sorghum and cotton.

Appendix II

KRUSKAL-WALLIS H TEST

Purpose: To determine whether the distributions in ranks for three or more independent samples differ significantly from those proposed for three or more populations.

Sampling Distribution: H statistics distributions are estimated by Chi-square.

Assumptions of the Test:

(i) independent and Random observations;
(ii) three or more independent samples; and
(iii) ordinal level of measurement (expressed as ranks) in dependent variable.

Typical Hypothesis:

$H_o : H = 0$

$H_1 : H \neq 0$

Tabular Statistic: H statistic as estimated by chi-square with d.f = K-1 from Table D_4 .

Test Statistic:

$$H = \frac{12}{N(N+1)}\left[\sum \frac{(\sum R_i)^2}{N_i}\right] - [3(N+1)]$$

where $\sum R_i$ = sum of ranks in each sample

N = number of scores in all samples combined

N_i = Number of scores in each sample.

REFERENCES

Ahmad, M. (1993) Choice of a Norm of Poverty Threshold and Extent of Poverty in Pakistan, Ministry of Finance, Islamabad. (Mimeographed.)

Ali, M. S. (1997) Poverty Alleviation: The Existing Situation. In Mohibul Haq Sahibzada (ed.) *Poverty Alleviation in Pakistan.* Institute of Policy Studies. 93–114.

Ali, S., and Sayyid Tahir (1999) Dynamics of Growth, Poverty and Inequality in Pakistan. *The Pakistan Development Review* 38:4.

Allauddin, T. (1975) Mass Poverty in Pakistan: A Further Study. *The Pakistan Development Review* 14.

Amjad, R., and A. R. Kemal (1997) Macroeconomic Policies and Their Impact on Poverty Alleviation in Pakistan. *The Pakistan Development Review* 36:1, 39-68.

Azid, T. (1994) Segments of the Directly Ordered Data: A Cross Section Analysis. Paper presented in the second Iranian Statistics Conference held at Ferdowsi University of Mashhad, Iran.

Besley, T. (1996) Political Economy of Alleviating Poverty: Theory and Institutions in Michael Bruno and Boris Pleskovic (eds.) *Annual World Bank Conference on Development Economics*, World Bank.

Bhatti, M. A. *et al*. (1999) A Sectoral Analysis of Poverty in Pakistan. *The Pakistan Development Review* 38:4.

Brown, M., and J. Popkin (1962) A Measure of Technological Change and Return to Scale. *Review of Economic and Statistics* 44, 402–414.

de Kruijk, H., and M. van Leeuwen (1985) Changes in Poverty and Income Inequalities in Pakistan during 1970s. *The Pakistan Development Review* 23,3-4. 407–19.

Gaiha, R. (1985) Poverty, Technology and Infrastructure in Rural India. *Cambridge Journal of Economics* 9, 221–245.

Gazdar, H., S. Howes, and S. Zaidi (1994) Pakistan: Recent Trends in Poverty. Mimeo. Background Paper for Pakistan Poverty Assessment, World Bank Islamabad.

Jafri, S. M. Younas (1999) Assessing Poverty in Pakistan. In *A Profile of Poverty in Pakistan*. Islamabad: Mahbubul Haq Center for Human Development and UNDP.

Mahmood, M. (1999) Macro Explanations of Poverty. In *A Profile of Poverty in Pakistan*. Islamabad: Mahbubul Haq Center for Human Development and UNDP.

Mahmood, Z. (1984) Inequality in Pakistan: An Analysis of Existing Evidence. *The Pakistan Development Review* 23:2-3, 365–76.

Malik, M. H. (1988) Some New Evidence on the Incidence of Poverty in Pakistan. *The Pakistan Development Review* 27:4, 509–515.

Malik, S. (1996) Determinants of Rural Poverty in Pakistan: A Micro Study. *The Pakistan Development Review* 35, 171–187.

Mathur, P. N. (1967) An Appropriate System of Deflation of Sectoral Income in a Developing Economy: Illustrated by the Industrial Sector of the Indian Economy (1950-51 to 1962-63). *The Review of Income and Wealth, Journal of the International Association for Research in Income and Wealth*. Income and wealth Series 13:1

Mujahid, G. B. S. (1978) A Note on Measurement of Poverty and Income Inequalities in Pakistan: Some Observations and Methodology. *The Pakistan Development Review* 17:4.

Naseem, S. Mohammad (1973) Mass Poverty in Pakistan: Some Preliminary Findings. *The Pakistan Development Review* 13:4.

Naseem, S. Mohammd (1996) The Informal Sector and Urban Poverty Alleviation in Pakistan. In *ESCAP*: *Role of Informal Sector in Poverty Alleviation.* New York: United Nations.

Shirazi, N. S. (1995) Determinants of Poverty in Pakistan. *Pakistan Economic and Social Review* 33:1&2, 91–101.

Tahir, Syed, and Salman Syed Ali (1999) Growth with Equity: Policy Lessons From Experience of Pakistan. In *Growth with Equity.* Bangkok: ESCAP.

Comments

The study by Dr Azid and Dr Malik is based on information of 90 households of a village. It is divided in two parts. In the first part, the authors apply a Logit Model to examine the role of technological, village and household-specific variables on probability of being poor. The authors conclude that most of the variables have expected signs and the results are statistically significant. In the second part, the authors divide the data by income segments and conclude that household-specific and village-specific characteristics affect poverty.

The paper is very interesting in terms of methodology adopted and the conclusions drawn. However, I would like to give few suggestions to clarify the results.

(1) It will be useful to add some discussion on the characteristics of the village, like the population size, education and health status of the population and the major occupations, land holdings and others.
(2) It will be interesting if the authors add a brief discussion for the rationale of applying the segmentation approach in the second part. It will be useful to know how this approach can improve the understanding about the poverty.
(3) It is not clear whether the unit of observation used in the regression analysis is household or individual. Furthermore, how the segmentation of highest group was done as there is only one household in that group?
(4) How the poverty line is determined? Is it national poverty line or the village-specific poverty line?
(5) The results presented in Table 3 are quite confusing. Sometimes it contradicts the results presented in the earlier tables. For example, the results for the household size variable are different in the two parts but no explanation is provided for this change.
(6) Occupational choice is an important variable for determining individual's earnings. The authors may include this variable in the model.
(7) Based on Table 3, the authors conclude that village and individual characteristics are important determinants of income in each segment. But it does not tell us whether the differences in coefficients across segments are statistically significant or not.
(8) The authors should add the number of observations for each regression equation in Table 3.

Incorporating these details in the paper will improve the exposition of the paper and the researchers and policy-makers will be able to benefit from the findings.

Rehana Siddiqui

Pakistan Institute of Development Economics,
Islamabad.