Pakistan Journal of Applied Economics, (1988), Vol. VII No. 2, (101–108).

# SIMULTANEOUS DETERMINATION OF CHOICE TO ATTEND SCHOOL AND THE DEMAND FOR SCHOOL EDUCATION: A Case Study of Karachi, Pakistan

Salim CHISHTI and Akhtar LODHI*

The study, based on 20 thousand school age children, of which a little more than half attend school, identifies the factors which influence the decision-making with regard to sending the child to school. Using Heckman's procedure of selectivity bias correction, it also estimates the demand for education. An important finding of the study is that investment in female education is one of the most important factors in increasing the litracy rate.

Little is known with confidence about the demand for education in Pakistan. Public sector education at all levels is more than 90 per cent subsidized. The case for such a high subsidy rests primarily on affordability of education costs. However, in Karachi which has the highest literacy rate in the country, in spite of this high subsidy, 49 per cent of school age children do not attend school. On the other hand 30 per cent of those who attend school go to private school where the tuition fee ranges from Rs. 10 to Rs. 500, with an average of Rs. 72 per month. This study identifies factors related to the socio-economic background of households which influence the decision making with regard to the choice of sending the child to school. Pakistan is relatively behind in literacy rate among comparable third world countries. For instance, the literacy rates in India, China, and Sri Lanka are 43 per cent, 69 per cent and 89 per cent, whereas in Pakistan it is only 30 per cent. [The World Bank Atlas 1989]. In view of

*The authors gratefully acknowledge useful comments by Hafiz A. Pasha and the anonymous referees.

this, the issue is quite important for education policy. The study also examines the determinants of demand for school education. Again this issue of affordability of education services is important for appropriate pricing policies.

## Methodology and Data

The study is based on a sample of 6261 households covered in the socio-economic survey of Karachi carried out in 1987-88 by the Applied Economics Research Centre. It includes about twenty thousand school age children between the ages of 3 to 20 a little more than half of which attend school. The sample for this survey was drawn following a stratified random sampling technique. The first stage of stratification was the Analysis Zones (AZ). The Master Plan Department of Karachi Development Authority (KDA) has divided the metropolitan area of Karachi into 58 Analysis Zones (AZs). The total sample was distributed among these AZs in proportion to their population. These AZs have been further subdivided into 241 Master Plan Department Zones (MPD). In the second stage, within each AZ, the sample was assigned to these MPD zones in proportion to the number of residential plots as given in the 1985 Land-use Survey [KDA (1987)]. At the third stage of stratification, the sample assigned to each MPD zone was divided into strata with regard to housing typology which was again based on the 1985 land-use survey. Finally a systematic random sampling was followed.

Since the observable dependent variable is expenditure on education rather than quantity of education, the widely used approach in such cases is to estimate income elasticity of demand from the expenditure equation [Goodman (1988)]. More specifically the following equation has been estimated:

$$E = \beta_0 + \beta_1 (INCOM) + \sum_{i=2}^{K} \beta_i X_i + U \qquad (1)$$

E is the expenditure per school going child. Two versions of the above model have been estimated. One with only tuition fee per month as the dependent variable and the other with all education expenditures including tuition fee, books, transport cost, private coaching and other charges.

The variable 'INCOM' is per capita income of the household. Instead of taking income and household size as two separate variables we have used per capita income because the dependent variable represents education expenditure per child rather than that of the household. It may also be noted that in order to make the income of the renter and owner-occu-

pied households comparable, we have added imputed rents to the income of the owner-occupied households. The imputed rents have been obtained from the hedonic rent equations estimated by Lodhi (1989). These hedonic equations which relate the characteristics of rented properties to the market rental values have been estimated using the same data set as ours.

Vector X in the above equation consists of the variables, a brief description of which is in order:

1. Parental education significantly influences the demand for child's education [Kodde and Ritzen (1988)]. We have defined three dummies FD1, FD2 and FD3 corresponding to father's education level of primary school, secondary school and college, with no school education as the reference category. Three similar dummies MD1, MD2 and MD3 have been included for mother's education corresponding to the same education levels.
2. In view of the socio-economic status assigned to female within the family and in the society in Pakistan, girls' education is not considered as important as of boys'. About 32 per cent of the total school enrollment consists of girls. To take account of sex preference in demand for education a sex variable has been included with girl as a reference category.
3. There may be differences in expenditures on education with the level of education. Specially non-tuition expenditures may be higher corresponding to higher levels of education. To capture this differential effect we have included the number of years of schooling as a variable.
4. To account for the possible influence of neighbourhood effects on the demand for education we have included a dummy variable (*Pakki*) for *katchi-pakki abadi* (i.e., squatter-regular settlements) with *katchi abadi* being the reference category.
5. The education expenditure per child may also depend on the total number of children in a family attending school which has also been included as an explanatory variable.

Inclusion of only school going children in explaining the demand for education may have a sample selection bias because a large number of school age children in our sample do not attend school. In order to correct for this bias, two-stage Heckman procedure has been used [Heckman (1976); Maddala (1983)]. This procedure consists of estimating the probability of a non-zero observation using a logit (or probit) estimation and then, conditional on the observation being greater than zero, estimates the main equation by OLS correcting for truncation bias. More specifically in the first stage we have estimated the following logit model:

$$SCHGO = \beta_0 + \beta_1(INCOM) + \beta_2 Sex + \beta_3 MTONG + \sum_{i=1}^{3} \tau_i MD_i + \sum_{i=1}^{3} \delta_i FD_i \quad (2)$$

Where

SCHGO = one if the school age child attends school, zero other wise.
SEX = one if the child is a boy, zero otherwise.
MTONG = one if the mother tongue is Urdu or Punjabi, zero otherwise.
All other variables have been defined above.

Equation (2) is used to estimate Mill's ratio which is then used in stage 2 to estimate equation (1) by OLS correcting for the selection bias.

A linear specification of equation (1) has been estimated. This is because it usually fits better than other functional forms and is derivable from Stone-Geary type utility function.

The income elasticity of education expenditure may be obtained from equation (1) However, because the probability of sending the child to school in equation (2) is also a function of income, estimation of income elasticity of expenditure from equation (1) will yield biased results [Goodman (1988)]. It has to be adjusted for the probability that increased income will motivate some households to start sending their children to school who did not do so earlier. Therefore, in our system of equations the 'total' income elasticity after this adjustment is given by:

$$\tilde{n} = \tilde{n}_1 + \tilde{n}_2$$

Where $\tilde{n}_1 = \delta E/\delta Y \bullet Y/E$; $\tilde{n}_2 = \delta f/\delta Y \bullet Y/f$.

Y being income and $\delta f/\delta Y$ the partial derivative of the logit probability function obtained from equation (2).

## Results

The logit probabilities are presented in Table 1. All the coefficients have the expected signs and have been estimated with high precision (as indicated by high t-values). The results indicate that parents education (both mother and father) significantly increases the probability of a child to attend school and the higher the level of parent's education the higher is the probability. Income also plays a significant role in the decision. As hypothesized, being a boy entails a much higher probability of going to school. Ethnic background, proxied by mother tongue is also important.

Children with Urdu and Punjabi mother tongue have higher odds than children with other ethnic backgrounds to go to school.

The results corresponding to the two versions of the demand equation, one with tuition fee and the other with total expenditure on education as the dependent variable are presented in columns 1 and 2 of Table 2. With one exception all the variables in both the regressions have the expected signs and all of them are estimated with high precision as indicated by t-values. The coefficients corresponding to parental education background indicate that the higher the level of parents' education the more they tend to spend on their children's education. It is particularly interesting that MD2 and MD3 are higher than FD2 and FD3 in both the equations. This means that the influence of mother's education is stronger than father's education on demand. Our results also indicate that expenditure on boys education is higher than on girls. Neigh-bourhood effects proxied by *katchi-pakki abadi* dummy variable indicate that *katchi abadi* residents tend to spend less on their children's education. The variable 'school' which represents the number of children in a household which are going to school, takes a negative sign. It indicates that increasing the number of school going children in a household adversely affects the per capita expenditure on education.

Income coefficient in both the models are significant although quite different from each other. The reason of their being different is that tui-

TABLE 1

Logit Results of a School age Child attending School

| Variable | Coefficients | t-ratios |
|---|---|---|
| Intercept | −1.11 | −30.08 |
| INCOM | 0.0004 | 7.76 |
| MD1 | 0.32 | 7.97 |
| MD2 | 0.59 | 11.83 |
| MD3 | 0.28 | 3.40 |
| FD1 | 0.44 | 9.75 |
| FD2 | 0.63 | 15.84 |
| FD3 | 1.07 | 18.61 |
| SEX | 0.31 | 10.29 |
| MTONG | 0.38 | 11.97 |
| No. of observations | | 19644 |
| Log likelihood | | −12556 |
| Chisquared (9) | | 2114.4 |

tion fee is only 0.97 per cent of income whereas education expenditures are about 3 per cent of income. The income elasticity of tuition fee and of total education expenditure ($\tilde{n}_1$) estimated from the demand equation at the mean level comes to 0.44 and 0.41 respectively. The income elasticity estimated at the mean level of the estimated logit function ($\tilde{n}_2$) is 0.08. Therefore, the total income elasticity of tuition fee and of education expenditure are 0.52 and 0.49 respectively. These elasticities are quite low. The major explanation is that, for Karachi, lower level of education is a necessity. In our sample 80 per cent of school age children either have attended or are attending school. Another reason may be a strong preference for neighbourhood schools. In our sample, 65 per cent of both the primary and secondary school children walk to their school. From among those who use transport 77 per cent take less than 30 minutes to reach. Dearth of expensive schools in the neighbourhood may have been reflected in the low income elasticity. Part of the explanation may also be found in the unfulfilled demand for good schools. In Karachi, affordability is not enough to get admission in an expensive school. It may be men-

**TABLE 2**

Results of the Regressions Corresponding to the Demand for Education

| Variable | Tuition Fee only | | All Expenditures | |
|---|---|---|---|---|
| | Coefficients | t-ratios | Coefficients | t-ratios |
| Intercept | −18.67 | −3.42 | −28.47 | −1.84 |
| Income | 0.028 | 28.57 | 0.08 | 39.25 |
| MD1 | −1.28 | 0.96 | 13.98 | 4.58 |
| MD2 | 21.95 | 11.81 | 33.31 | 8.15 |
| MD3 | 42.49 | 20.32 | 57.26 | 13.21 |
| FD1 | 3.40 | 1.92 | 6.68 | 1.85 |
| FD2 | 9.28 | 4.66 | 17.62 | 4.34 |
| FD3 | 27.58 | 9.87 | 46.45 | 7.95 |
| Years | −1.06 | −7.12 | 2.75 | 9.38 |
| PAKKI | 13.42 | 13.53 | 25.04 | 12.72 |
| SEX | 9.15 | 8.40 | 15.05 | 6.80 |
| SCHOOL | −1.43 | −5.83 | −3.79 | −7.78 |
| Mill's Ratio Inverse | 19.51 | 5.26 | 43.50 | 3.35 |
| No. of observations | | 10,005 | | 10,005 |
| Adjusted $R^2$ | | 0.31 | | 0.35 |

tioned here that this result is quite robust as several other versions of the model including separate estimation for public and private schools, and separate estimation for primary and secondary schools were also tried. Dropping variables like parents education did not change the income elasticity either.

The only coefficient which switches from a negative to a positive sign in the two versions corresponds to the variable relating to the number of years spent in school. It is small in the first version of the model because there is not much variation in the tuition fee with regard to class level. It is negative because in our sample relatively a larger proportion of secondary level children go to public school. This may be because of the parent's preference to send younger children to the neighbourhood schools and also because the resource requirements to make significant quality differentials between private and public schools at the secondary level are relatively high and therefore there are fewer secondary level private schools. It is positive in the second version of the model because the expenditure on books, private tuition, transport etc., increases as the class size increases.

## Conclusion

The purpose of the study is to estimate the demand for school education. To eliminate sample selection bias Heckman two stage procedure has been adopted. The results indicate that income, parents' education, child's sex, neighbourhood of residence and ethnic background significantly influence the decisions regarding the choice of sending the child to school as well as determining the expenditure on education. Income elasticity of demand for education turns out less than one. Investment in female education seems to be crucially important in increasing the literacy rate. The demand for primary and secondary education in Karachi is inelastic with regard to income.

*Applied Economics Research Centre*
*University of Karachi*

## References

Goodman, A.C., 1988, An econometric model of housing price, permanent income, tenure choice and housing demand, Journal of Urban Economics, 23(3).

Heckman, J., 1976, The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator of such models, Annals of Economic and Social Measurement, 5(a).

Kodde, D.A., and J.M.M. Ritzen, 1988, Effects of parental education on the demand for higher education, Journal of Human Resources, 23(3).

Karachi Development Authority, Master Plan and Environment, 1987, Land-use Survey Report (1985), Karachi.

Lodhi, Akhtar, 1989, Housing demand in developing countries: A case study of Karachi, Unpublished MAS Thesis, Applied Economics Research Centre, University of Karachi.

Maddala, G., 1983, Limited dependent and qualitative variables in econometrics, New York: Cambridge University Press.