II. Conceptual Framework

This section will begin with a brief review of literature on factors affecting scholastic achievement. Some major conceptual and methodological issues encountered by studies in this area will be discussed. The general conceptual framework and statistical method to be employed in our study will follow.

A. Brief Review of Previous Work

The first and perhaps the most influential and controversial study in this area is "Equality of Educational Opportunity" or generally known as the "Coleman Report".[1] The Report is a result of an attempt to determine the extent of racial and ethnic discrimination in schools in the United States. The best known findings of the Report are quantity and quality of school inputs (facilities, curriculum and personnels) have little or no effect on students' scholastic achievement; home environment and student peers are what count. The Report demonstrates that the unique variance in achievement attributed to school inputs is extremely small. The policy relevance of input-output studies has led to a rapid growth in number of analyses. For a comprehensive survey of educational input-output analysis in the United States, see Cohn and the 'Others[2]; for survey of international comparisons, see Simmons[3]; and for a review of research in developing countries, see Simmons and Alexander.[4]

---

1. James S. Coleman et. al., Equality of Educational Opportunity (Washington, D.C, U.S. Government Printing Office, 1966).
2. Elchanan Cohn, Stephen D. Millman and It-Keong Chew, Input-Output Analysis in Public Education, (Cambridge, Mass., : Ballinger, 1975), pp. 31-49.
3. John Simmons, How Effective is Schooling in Promoting Learning? A Review of the Research, International Bank for Reconstruction and Development, Staff Working Paper No. 200, (March 1975)
4. John Simmons and Leigh Alexander, The Determinants of School Achievement in Developing Countries : A Review of Research, The World Bank, (March 1976).

As economist entered this area, the relationships estimated became known as educational production functions. Samuel Bowles[1] pointed out that if education did have unique effect on labor productivity, it should be possible to relate the development of productive skills and allocation of scarce resources. The use of the concept of production function in educational production process helps establishing a more meaningful behavioral model and interpretation of the results. In setting school policy and in long-range educational planning, knowledge of educational production function is essential to efficient resource allocation. The theoretical concept of educational production function may be simple. There are, however, a serie of issues encountering researchers in attempting to estimate such relationships.

The conceptual problem begins with measurement of outputs. Outputs of schooling process could be multiple and, sometimes, interacting . Examples of schooling outputs are academic skill, political awareness, self-esteem, etc. A higher self-esteem may also improve scholastic achievement. If interaction exists, simultaneously bias will arise in the estimation of any single equation by the method of ordinary least squares.

Although modern study on education production tends to pay more attention to other measurements of outputs, the use of test scores as output measures still remains popular. Hanushek[2] has provided supporting arguments for the use of test scores as output measures. They are the most available and are also often been used to evaluate educational programs, and even to allocate funds. A more persuasive argument for the use of test scores relates to continuation in schooling. Test scores appear to have an increasing use in selecting individuals for further schooling.

---

1. Samuel S. Bowles,"Towards an Educational Production Function in Education, Income and Human Capital, ed. W. Lee Hansen, Studies in Income and Wealth No. 35 (New York : National Bureau of Economic Research, 1970), pp.11-61.
2. Eric A. Hanushek, "Conceptual and Empirical Issues in the Estimation of Educational Production Functions," The Journal of Human Resources 14 (Summer, 1979) pp. 351-388.

Inputs in the educational production process can be broadly classified into school and nonschool factors. The school factors include both human and physical resources. The human inputs typically emphasized in studies of this nature include teachers, administrators, secretarial, etc. Since a large fraction of schools' budgets is spent on teaching staffs, much attention has been directed to the efficacy of teachers. This implies a need to examine various potential productive attributes of teachers, including educational preparation, experience, talents, attitudes, and classroom practices. The physical inputs include building characteristics, quantity and quality of equipment, and other supporting physical facilities.

The nonschool inputs include general home characteristics of students and general characteristics of community environments that could affect educational outcomes. The formers are parental socioeconomic status, races, sexes, family sizes, etc., and the latters refer to degrees of urbanization, extent of poverty, racial compositions, etc.

In attempting to evaluate the contribution of a given input to the output of a given student, it is often difficult to ascertain which portion of the input under discussion is relevant to the current investigation. In addition, there is a problem of accuracy of variable measurement. Individual and family characteristics are more in the form of stocks and, hence, are subject to less intertemporal variations than are school inputs. Thus, use of cross-section measurement of contemporaneous school factors clearly tends to underestimate the total effect of educational inputs on achievement. Another source of error of measurement is due to the aggregation of school inputs. Researchers frequently have information on individual students but information on schools are oftenly aggregate.[1]

Statistical estimations of empirical models frequently encountered the problem of multicollinearity. That is, there are high correlations among explanatory variables. The use of analysis of variance in the Coleman Report

---

1. Hanushek (1979) provides a lengthy discussion on the problem of error of measurement of variables in the educational production function studies.

was criticized by Bowles and Levin[1] as an inappropriate methodology due to interdependence among the right hand side variable. The analysis of variance judges the significance of the influence of explanatory variables upon the academic achievement by the additional increase in the variance explained, R^2. If two variables are highly correlated, the first entered will be assigned both its unique contribution to explained variance and its jointly explained variance with the other variable. Consequently the variable that enters later will cause the overall R^2 to increase by negligible amount. Changing the order in which the variables are entered will change the proportion of explained variance attributed to each.

A more appropriate techinque of estimation as suggested by many researchers, for example Bowles (1970) and Hanushek (1979), is multiple regression analysis. This method is designed to take into account correlations among exogenous variables. If these variables are uncorrelated, simple regression will suffice. However, with very high levels of correlation among explanatory variables, the coefficients estimated become imprecise, i.e., their standard error of estimates will increase.[2] Another advantage of the use of multiple regression is that the value of the coefficients will be more relevant to policy purpose than the adding contribution of the explained variance of each independent variable.

---

1. Samuel Bowles and Henry M. Levin, "The Determinants of Scholastic Achievement - An Appraisal of Some Recent Evidence", Journal of Human Resources 3 (Winter 1968), pp. 3-24.
2. Eric A. Hanushek and John E. Jackson, Statistical Method for Social Scientists (New York : Academic Press, 1977) pp. 133-137.

B. The General Conceptual Model

In our study of factors determining the academic achievement of students in the grade 3 and Mathayom Suksa 5, we will employ the following general model;

$$A_i = f(F_i, S_i, u_i)$$

That is the academic achievement of the ith student, A_i is determined by a vector of his or her family background characteristics, F_i, and a vector of variables describing the learning conditions in schools, S_i, and disturbance term, u_i.

The academic achievement of the grade 3 students under the study will be measured by their scores in the second test. For the MS 5 students, their academic achievement will be proxied by their grades in the final examination. A more detailed description on the dependent variables in these two studies will be presented in the section on data.

The vector of family characteristic variables includes students' own characteristics and family socioeconomic status. The former are varibles such as health condition, previous schooling, age, etc. The latter composes of father's occupations, parents' education, home locations, family sizes, numbers of siblings in school, exposure to media, absence records, textbook availability, etc. Various measurements of student's characteristics and home environment will be tried.

The vector of variables describing learning conditions at school or the school inputs composes of both school facilities and teacher characteristics. School variables are types, locations of school, sizes of school either measured by number of teachers or students, class sizes, student-teacher ratios, etc. Teachers' characteristics are measured by teachers' qualification, teachers' cognitive abilities, teachers' views of students' learning abilities, teachers' salaries, numbers of teaching years, teachers' attentiveness, etc.

Not all variables listed above are available in both sets of data. Since there is no a prior knowledge on either any specific list of determining variables or any functional form of the relationship. The "best" model in each study will have to be based on the results of statistical analyses.

C. Statistical Estimation Procedure

In order to evaluate the effect of educational inputs on the academic achievement, we employ "multiple regression analysis". The theory of educational production does not explicitly specify any functional form of the relationship. Most studies of this kind assume linear forms of relationship between outputs and inputs. However, the relationship between academic achievement and some variables may not linear. For example, there could be some economies of scale in class sizes or the effect of some inputs may depend on how much the other inputs are used. Therefore, various specifications of variables entering the model, such as polynomial forms or interaction among them, will be attempted.

In our conceptual model, we broadly classify inputs into two categories; namely family background and schooling characteristics. In estimating the effect of these variables, we shall run the equation twice. The block of variables in each category will take turn in entering the regression equation in the first step. Then, the remaining group of variable will be added in the second run. In so doing, we hypothsize that school characteristics do affect students' achievement as do family characteristics. If the school characteristic variables entered the regression equation prior to the family background variables, the estimated coefficient of the former should be significant. Due to high correlations among the two groups of variables, adding the latter in the second step will increase the standard error of the estimated coefficients of the school variables, and hence a decrease in the t-value pertaining to each variable. The adding contribution of the variance explained, or the increase in the numerical value of R^2, in the second step is expected to be low too. When the process of estimation is reversed, i,e. the family characteristics were introduced prior to the school variables, a result in the opposite direction is anticipated.