Discussion Paper 8008 November 1980

Distribution Flow of Education in Thailand

by

*Edita A. Tan and Wannasiri Naiyavitit**

with a Technical Note

by

Fe Lisondra

* The authors are respectively Professor of Economics, University of the Philippines, and Assistant Professor of Economics, Thammasat University.

Abstract

The paper investigates the degree of inequality of distribution of formal schooling in Thailand and how it intensifies as we move up the schooling ladder. It applies a constrained capital optimization model to explain the degree and movement of inequality. It shows how financial constraints on education choices result in an unequal distribution of choice sets facing the school-age children coming from different socio-economic and location classes. Location, particularly rural-urban categories, is found to exert the strongest influence on schooling attainment. This is shown in cross-tabulations and tests of the model.

The model is tested by logit method using cross-sections of individual observations from the National Statistical Office 1975 Survey of Youth from which we obtained values of probability of school attendance in lower elementary, upper elementary, high school and post high school. For "worst-off" children whose fathers had the lowest income, lowest education, the most number of brothers and sisters, living in the rural Northeastern region, the respective probabilities were .71, .53, .14 and .20. The corresponding figures for the "best-off" children whose fathers had B100,000 income and post-secondary education, who had only one brother/sister and who lived in Bangkok were .98, .99, .97 and .56. The paper also gives the inter-generational link in education capital from the coefficients of father's education and income. The distribution of education may be predicted by the model for any given distribution of children by socio-economic and location variables.

LIST OF TABLES

Income inequality has been a serious problem in a large number of developing countries including Asean members except Singapore. A high degree of inequality of between .45 and .50 Gini has persisted in the Philippines, Thailand and Malaysia during fairly high growth period of the past three decades. Recent attempts to solve the problem mainly through Malaysia's Bumiputra policy and the Philippine's land reform program have not made a perceptive impact so far. It might be argued that without applying drastic and consistent measures the inequality could be expected to remain serious. There are structural givens in underdeveloped economies which tend to perpetuate inequality and the poverty of those in the lower income brackets. The higher the degree of inequality in a low-income country, meaning the poorer are its poor, the more difficult it is to break the problem. The relative and absolute poverty of the lower income groups separates and locks them within their environment. Such environment tends to be socially (and economically) restrictive on individual choices so that opportunities for upward mobility are limited.

This is now a well-known phenomenon and it has attracted much interest in national and international bodies. This interest has been translated mostly in strong statements of concern and aims. Few LDCs have achieved success in poverty and inequality alleviation. There has been much research effort invested in this decade in producing and measuring inequality and to some extent, of poverty incidence. Fewer works have been done on social mobility or inter-generational transfer of physical and human capital. This paper is an attempt to describe the process of inter-generational transfer of education capital under specified market conditions affecting choice in two not atypical LDC countries - the Philippines and Thailand. A constrained optimization model in education choice is used as the framework of analysis. The model explains how inequality in income and in education interacts inter-generationally and traces the change (or lack of it) in inequality and poverty over time. The model is presented in Section 3 after a discussion of the state of inequality and poverty and their trend in Section 2.

Section 2. Descriptive Analysis of the Distribution of Education

Equity in the distribution of education means equality in education opportunity. This has been defined in varying degrees of strictness. It means each child of given ability and aptitude has an equal chance to pursue the education and training that suit him best. In strict terms the reference would be inherent ability and aptitude. There is an assumption that the distribution of inherent or inborn traits among the population is not affected by race, socio-economic, location and other environmental variables. In other words each child in the population has an equal chance to be born with a particular level of these inherent traits. Education of all forms - formal and non-formal - builds on these traits so that at later stages of life, it is not possible to distinguish inherent from acquired abilities. In fact performance in IQ tests is found not to be constant over time but to improve with acquired knowledge. In the looser sense equality in educational opportunity is measured in terms of measured ability which has already been affected by environment, socio-economic background and quality of formal schooling. In the absence of information on the distribution of inherent traits in the population equity in educational opportunities can be judged on the basis only of an a priori notion of equality in these traits. Our judgement, therefore, of degree of equality or inequality of educational opportunity in Thailand is based on this a priori distribution. The writers wish to believe that the distribution is equal, i.e., that each child of whatever background has the same probability distribution of possessing inherent traits as the probability distribution of the population to which he belongs. It means that he has a ten percent probability of being of superior scientific talent if the population has the same probability of producing this much scientific talent. Inequality of opportunity will be judged from the variation in the distribution of educational attainment between population groups. Under perfect equality each group should have the same distribution as the population distribution.

Two sets of data are used to indicate the degree of inequality in the distribution of education in Thailand: a) attendance in school and the highest level of formal schooling attained by those who left the formal school system; b) performance in achievement tests administered to grade III and senior high school pupils. The population was grouped according to variables that are most likely to affect the distribution of opportunity: family income, father's education, location - by region, whether rural (non-municipal) or urban (municipal, metropolis).

First we look at some broad indicators of movement of distribution over time. Table 1 shows that the average level of school attainment has been rising from 1960 to 1975. This results from an increasing rate of attendance at each age group as shown in Table 1 below.

Table 1
Attendance Rate by Age Range
1960, 1975

Age Range	1960	Age Range at Any Program	1975
4-6		4-6	6.3
7-10 in grades 1-4	119.3	7-9	71.4
11-13 in grades 5-7	19.8	10-14	72.4
14-18 in high school and lower vocational schools	17.9	15-19	24.3
20-22 in college and technical schools	2.9	20-24	5.1

Source: National Statistical Office <u>1975 Survey of Youth</u>; <u>Report of the UNESCO Regional Advisory Team for Educational Planning in Asia Bangkok, 1965.</u>

Finer indicators of inequality are obtained from the data used in our paper, the 1975 survey of youth, the 1973 national achievement test of grade III pupils and the 1976 study of achievement of senior high school students who were admitted to college in 1977 by Nitongkorn and Vutisart.

School attendance was very much influenced by location and socio-economic variables as shown by the following cross-tabulations. In Table 2 attendance rate by location for each age range is given. As expected children in the Bangkok Metropolis had the highest rate at each age range or at its corresponding schooling stage - lower elementary, upper elementary, high school and post-high school. Attendance rate dropped between upper elementary to high school from 72 percent to 24 percent for all children, 90 percent to 60 percent for municipal children, and 70 percent to 18 percent for non-municipal children. For the last two stages, i.e., high school and post-high school, the corresponding drop was 24 percent to 5 percent, 59 percent and 20 percent, and 18 percent to 2 percent. Northeast, the poorest region, showed the lowest attendance rate at all levels and for both municipal and non-municipal area. Only 1.2 percent of its 20-24 year old youths attended school. Compare this to 22 percent for Bangkok.

Those who stopped schooling are not likely to go back to school so that their highest attainment as per survey time might be

Table 2

Rate of School Attendance of Children and Youth by Age, by Region, by Municipal/Village, 1975

	Total	4-6	7-9	10-14	15-19	20-24
Kingdom	38.9	6.3	11.4	72.4	24.3	5.1
1) Municipal	58.6	27.4	85.2	89.7	59.3	19.8
2) Non-Municipal	35.7	3.7	69.6	69.7	17.7	2.1
Bangkok	56.8	21.2	83.1	89.8	57.6	21.7
1) Municipal	59.2	25.6	85.0	91.0	61.0	24.0
2) Non-Municipal	46.2	6.1	76.0	84.2	40.5	9.7
Central	42.2	9.8	74.2	76.3	53.3	4.8
1) Municipal	57.1	29.6	84.2	88.2	53.3	11.9
2) Non-Municipal	40.6	8.1	73.2	75.1	23.3	3.9
Northern	39.5	6.9	77.5	71.4	19.2	3.0
1) Municipal	62.6	43.8	89.5	88.2	60.1	15.5
2) Non-Municipal	38.0	5.1	76.9	70.4	16.3	2.0
Northeast	29.9	1.1	63.2	61.6	12.3	1.2
1) Municipal	53.5	8.9	82.4	86.0	55.7	14.1
2) Non-Municipal	30.0	.5	62.5	60.7	10.3	.5
Southern	44.3	6.1	72.2	85.1	33.6	3.8
1) Municipal	58.8	27.6	86.4	89.0	59.4	11.7
2) Non-Municipal	42.5	3.9	70.6	84.6	30.1	2.8

Source: National Statistical Office (NSO), The Children and Youth Survey, 1975. Table C_1

Table 3
Percentage Distribution of Youth Not Attending School
by Age, Highest Level Attained, Municipal/Non-Municipal

	Zero Level	Lower Education	Upper Elementary	Lower Secondary	Upper Secondary	Teacher Training	Academic	Other Post Secondary
Age Range								
Tota Kingdom								
7- 9	99.2	.8	-	-	-	-	-	-
10-14	.5	98.4	1.1	-	-	-	-	-
15-19	.2	90.7	5.8	2.6	.1	.3	-	-
20-24	.3	83.8	5.3	5.0	.8	2.4	.3	1.8
Municipal								
7- 9	100.0	-	-	-	-	-	-	-
10-14	2.1	93.6	4.3	-	-	-	-	-
15-19	.2	66.5	19.5	11.3	1.1	.6	-	-
20-24	.9	48.7	15.2	16.3	3.2	3.5	2.1	10.0
Non-municipal								
7- 9	99.2	.8	-	-	-	-	-	-
10-14	.2	98.6	1.0	-	-	-	-	-
15-19	.2	93.0	4.6	1.9	-	.3	-	-
20-24	.3	89.7	3.7	3.1	.4	2.3	-	..5
Total Youth	16,656,570							
7- 9	3,645,810							
% na	28.6							
10-14	5,408,880							
% na	24.9							
15-19	4,412,360							
% na	73.1							
20-24	3,189,520							
% na	91.1							
% in municipal	14.5							

na is not attending

Source: NSO 1975 Survey of Youth, Table 4, p. 5.

taken to be their lifetime attainment. Some of these have completed their programs such as academic, teacher training and vocational. The highest attainment of the current cohort over their life can therefore be estimated from their attendance rate. We need not make this estimate since the survey reports this information directly. (Please see Table 3.)

It is seen that majority of school leavers even in the older ages of 20 to 24 completed lower elementary levels only with many very young out-of-school youth aged 7-9 never having attended school. This number amounted to 22 percent of the total population of this age range in 1975.

Family background is another important determinant of school attendance. We obtained from the raw data borrowed from the National Statistics Office (NSO) the following tables showing attendance rate by age, by family income. The effect of income is strongest in the three lowest income brackets to which 64 percent of Thai youths belonged. The difference in attendance rate increases as age or schooling level rises. The lowest income group, for instance, to which 25 percent of youth population belonged, had attendance rates of 67 percent, 62 percent, 24 percent and 10 percent for ages 6-10, 11-14, 15-19, and 20-24. At the highest income bracket the corresponding rates were 91 percent, 93 percent, 77 percent, and 38 percent. Please note that the attendance rate reported by the NSO is lower than in our cross-tabulation since

Table 4
Attendance Rate by Age, by Family Income
1975

Family Income (in Baht)	Total	6-10	11-14	15-19	20-24
0- 9,999	43.8	67.1	62.0	24.1	10.1
10,000-14,999	57.2	72.6	75.7	41.8	14.2
15,000-19,999	62.2	77.8	70.2	57.1	20.5
20,000-29,999	65.0	82.4	88.1	59.7	22.6
30,000-39,999	65.4	84.1	90.6	62.2	22.9
40,000-49,999	66.7	88.4	90.8	67.9	24.2
50,000-59,999	64.5	85.8	89.2	63.5	26.6
60,000-69,999	64.8	89.5	88.6	64.2	29.1
70,000 and above	69.2	91.2	92.8	76.7	37.5
Weighted Average	57.9	75.4	75.6	49.3	20.6

Source: Our cross-tabulations of raw data of the NSO 1975 Survey of Youth.

Table 4a
Percentage Distribution of Population by Age

Family Income	Total	6-10	11-14	15-19	20-24
0- 9,999	24.77	30.52	24.95	22.18	18.41
10,000-14,999	22.63	25.30	22.89	22.18	18.15
15,000-19,999	17.12	16.69	20.39	16.42	14.30
20,000-29,999	12.85	10.60	12.22	14.56	14.93
30,000-39,999	6.64	5.38	5.95	7.37	8.69
40,000-49,999	4.31	3.24	3.76	4.70	6.38
50,000-59,999	3.24	2.30	2.82	3.57	5.11
60,000-69,999	2.43	1.70	1.98	2.66	4.04
70,000 and above	6.01	4.26	5.02	6.36	9.99
Weighted Average	100.00	100.00	100.00	100.00	100.00

Source: Our cross-tabulations of raw data of the NSO 1975 Survey of Youth.

our data consisted of unmarried children of the head of households only. The difference in attendance rate between the two reports reflects the effect of relationship and civil status on school attendance. This effect is more precisely shown in the empirical test of the model. Cross-tabulations with other socio-economic variables are not given here since they should have the same pattern as that of income.

b. Distribution of Learning at Grade III and High School V.

The quality of education offered each child differs so that amount learned also differed. It depends on his total environment. The educational content of environment varies substantially according to location and socio-economic background. Schools differ in quality of instruction and available instructional materials. Family income exerts further effect on learning by determining the physical and mental health of the child. It is not uncommon in poor situations in LDCs that children have not enough to eat. Hunger must have a direct or immediate impact on the rate of learning in school. Persistent malnutrition would have longer and more profound effects on the whole process of education, including its effect on the development of brain cells during infancy. This variation in learning environment is reflected in an unequal distribution of measured scholastic achievement among population groups. The chapter by Nitongkorn and Vutisart evidences this very clearly. For this introductory chapter we present some descriptive statistics used in their analysis. Their cross-tabulations of mean scores by type of institution - whether private, supported by Ministry of Education, by municipal or by provincial (changwat) governments; father's occupation and location are reproduced here.

The distribution of achievement in the tests follow the same pattern as school attendance. Apparently children in urban environment whose fathers work as government officials and as professionals learned more (of what the tests measured) than children of farmers. The latter scored about 60 percent of the mean score of urban elite children. There is also substantial differences in scores between changwat and Ministry of Education (MOE) schools with mean of 51 percent and 83 percent, respectively.

The distribution of MS 5 scores is likely to be subject to selection bias. A much smaller proportion of youth in rural areas and from poorer socio-economic background reached Maw Saw 5. The few to reach this level probably possessed a different set of traits from the typical student for them to be able to survive the disadvantages of their environment. Thus we find that the variation in mean score by region, father's occupation or by mother's education does not follow the same pattern as for grade III. Farmers' children did as well as professional children. Furthermore, there was not as much regional variation and

Table 5
Distribution of Students by High School Test Scores, by Region and Father's Occupation
1975

Occupation	Lower than 2.5	(1) 2.50-2.99	(2) 3.0-4.0	(1) + (2)
Professional	43.1	37.1	19.8	56.9
Administrative	37.1	44.0	18.9	62.9
Clerical	42.1	38.6	19.3	57.9
Sales	37.6	38.5	23.9	62.4
Farmers	46.9	35.0	18.1	53.1
Transportation	39.5	41.9	18.6	60.5
Craftsmen	41.7	43.7	14.6	58.3
Services	25.0	37.5	37.5	75.0
Laborers	49.0	31.9	19.1	51.0
Unclassified	46.2	39.1	14.7	53.8
Region				
Bangkok	42.1	38.1	19.8	57.9
Central	41.7	36.9	21.4	58.3
North	29.6	43.0	27.4	70.4
Northeast	42.2	37.8	20.0	57.8
South	47.3	42.7	10.0	52.7
EAst	34.4	34.4	31.2	65.6

Source: Nitongkorn and Vutisart's Chapter, Tables 15 and 17, pp. 48-49, 52.

Table 5.a
Mean Test Scores of Grade III Pupils by
Type of School, Region, Father's Occupation and
Urbanization of Location
1973

	Scores	Standard Deviation
Average for All	61.1	24.2
A. Type of School		
Private	78.2	20.9
Min. of Education	83.0	18.5
Principal	67.5	20.1
Changwat	50.6	20.8
B. Region		
Bangkok	77.3	23.1
Central	57.6	21.1
North	49.2	20.5
C. Father's Occupation		
Agriculture	49.7	20.0
Services	69.4	24.8
Merchant	73.7	22.5
Industrial Work	65.7	19.5
Government Officer	75.5	21.8
Professional	80.3	25.3
Others and no response	62.4	21.0
D. Urbanization		
City Municipality	80.8	21.0
Town Municipality	70.8	20.2
Village	53.2	21.0

Source: Nitongkorn and Vutisart, Tables 1-4, pp. 30-33.

Bangkok youth did only as well as other students. Southern region did best.

We find location to be a very strong explanatory variable of school attendance especially at higher levels. This is mainly because of the greater distance to schools of rural populations especially of those in the Northeast. Two supplementary tables are presented to show the effect of distance on schooling. Table 6 gives the distribution of students by distance to school. Is is shown that 63 percent of students were located within 3 kilometers of their schools. Only 16 percent of students were farther than six kilometers.

Table 7 gives a breakdown of total cost of schooling at various levels including transport and additional living expenses for students from municipal and non-municipal locations. Total cost is, as expected, higher the higher the level. Cost of transport and food taken out of the home formed a very large proportion of the total at every level (56 to 78 percent). Total cost was higher for non-municipal students.

The effect of financial constraint as determined by income and cost of schooling is shown in Table 8 which describes the reasons for not attending school given by those not enrolled. Eighty-three percent of the sample cited financial reason for their non-attendance in school. The percentage was higher the higher the schooling level.

Table 6
Percentage Distribution of Children and Youth Attending School by Type of School and Distance to School, 1975

Type of School	Total of Both Sexes	Less than 3 km.	3-6 km.	More than 6-10 km.	More than 10 km.	Unknown
Public School	52.00	33.58	9.39	4.27	4.43	.32
Private School	48.00	29.81	11.30	4.02	2.46	.39
Total:	100.00	63.39	20.69	8.29	6.89	.72

Source: NSO, The Children and Youth Survey, 1975, Table 7A.

Table 7

...n per Person by Level of Education Attended and Type of Expenditure

...VEL OF EDUCATION ATTENDED

...Equiv.	Secondary or Equivalent			University				
...er ...tary	Lower Secondary	Upper Secondary	Upper Higher Vocational	Academic	Technical Vocational	Teacher Training	Short Course Vocational	Unknown
...94	577	714	841	1,094	909	1,015	2,519	-
...38	1,809	2,463	2,698	4,118	3,315	2,573	2,262	-
...74	240	341	407	572	640	432	200	-
...07	206	285	494	1,442	971	1,063	358	-
...20	278	314	371	441	398	394	350	-
...64	105	117	291	407	656	224	200	-
...97	3,215	4,234	5,102	8,074	6,889	5,701	5,889	
...06	660	764	888	858	1,086	-	-	-
...18	2,044	2,812	3,103	3,608	4,304	-	1,560	431
...10	264	380	446	541	717	-	1,000	35
...61	857	1,274	2,710	2,865	4,427	-	2,182	128
...47	288	318	384	382	466	-	400	-
...74	99	124	217	377	821	-	-	-
...16	4,212	5,672	7,748	8,631	11,821		5,142	-

... and 9B.

...enditure

D

Teacher Training	Short Course Vocational	Unknown
1,323	600	-
3,220	1,320	-
687	20	-
1,246	149	-
540	200	-
361	500	-
7,377	2,789	
-	792	-
-	1,800	132
-	-	50
-	2,500	-
-	-	100
-	1,500	-
	6,592	

Table 7

Average Annual Expenditure on Education per Person by Level of Education Attended and Type of Expe

(cont'd)

	LEVEL OF EDUCATIONAL ATTENDE							
	Elementary & Kindergarten or Equiv.			Secondary or Equivalent			University	
	Kinder-garten	Lower Elementary	Upper Elementary	Lower Secondary	Upper Secondary	Upper Higher Vocational	Academic	Technical Vocational
NON-MUNICIPAL								
Public School								
1. Transportation	300	418	526	824	1,481	1,930	1,071	2,236
2. Food taken outside home	514	386	649	1,255	2,545	2,584	3,932	2,907
3. Book, Materials and equipment	33	64	145	250	370	376	1,076	842
4. School Fees	106	61	63	187	205	309	761	1,070
5. Uniform	109	110	199	296	399	409	622	582
6. Other	-	36	57	113	147	254	1,392	300
7. Total	1,062	1,075	1,639	2,925	5,147	5,862	8,854	7,937
Private School								
1. Transportation	661	601	547	651	896	1,020	970	-
2. Food taken outside home	782	730	890	1,123	1,689	1,751	2,629	-
3. Book, Materials and equipment	54	106	184	252	305	385	276	-
4. School Fees	426	378	421	468	806	2,638	3,504	-
5. Uniform	146	176	233	301	289	328	347	-
6. Other	137	42	68	99	95	262	250	-
7. Total	2,206	2,033	2,343	2,894	4,080	6,384	7,976	

Source: NSO, The Children and Youth Survey, 1975, Tables 9A and 9B.

Table 8
Percentage Distribution of Children Not Attending School by Age and Reasons for Not Attending School, 1975

Reasons	Total for Both Sexes	7-9	10-14	15-19	20-24
1. Sickness, disability/ mental handicap	1.38	.002	.24	.62	.51
2. No financial support	82.75	.12	6.09	29.34	47.18
3. Others (no interest, inadequate qualifications, etc.)	15.87	.07	1.43	6.19	8.18
Total:	100.00	.19	7.76	36.15	55.87

Source: NSO, Children and Youth Survey, 1975, Table 5A.

The above data provided some salient information on the degree of inequality of educational opportunity. We are led to believe that financial constraints explain most of the inequality and how it gets worse the higher the education level.

Section 3. Constrained Optimization in Education Choice

The model used here was first developed by Tan and Danao in 1976 for the PREPF[1/] project. Education is treated as a human capital and conventional investment decision criterion is applied to the choice of pursuing alternative types of education including zero education. Two major constraints are built into the model - financial and informational.

Following Hicks (1973) we consider choice among relevant alternative capital processes rather than of alternative capital stocks, or capital expenditures. Education has a peculiar process of capital build-up. A capital process is a flow of inputs and outputs over a time period. Each process is evaluated by taking its capitalized value, i.e., discounting all input and output flows. No distinction is made between capital and current expenditures. What matters in the valuation of capital is the timing of expenditure and output. A given plant and equipment may be built with different time paths and once built, the flow of inputs and output can still vary. The same physical plant and equipment may involve different capital processes. Using the concept of capital process frees capital choices from the more limited choices of capital goods. It might also rid us of the two-Cambridge capital controversy of what capital is and how it is to be valued.

Assuming a number of proceeses, each process j has a net

[1/] PREPF is Population, Resources, Education in the Philippine Future.

worth computed from its flow of inputs, C_{jt} and outputs, R_{jt} over T periods:

$$NW_j = \sum_{t=0}^{T-1} \frac{R_{jt} - C_{jt}}{(1 + r)^t}, \quad j = 1, 2, \ldots, n \qquad (1)$$

where r, the rate of interest, is assumed constant over time $t = 0, 1, 2, \ldots, T - 1$. At some rate of interest and other cost, we obtain a ranking of the processes 1 to n. The ranking order may change with changes in the cost of capital or discount rate so that a process with large capital expenditures and early pay-off or early flow of receipts would become more profitable than one with a longer spread of capital inputs and later output flow at high rates of interest. Once the NW_j's are estimated choice is easily made with the objective of maximizing NW_j. Constraints on choice may be imposed such as capital rationing and mutually exclusive processes. Net worth is to be maximized subject to whatever constraints apply.

Application of Capital Theory to Education.

The concept of human capital has been accepted without much question. In just over a decade since Schultz' revival in 1960 of earlier works on education as human capital, works on economics of education have assumed a major bulk of the economics literature. However capital choice theory may be applied to education only at great simplification of education options. Education is not clearly observable. It can be defined more meaningfully in a philosophical sense than as an observable capital variable. As a capital good counted at a point in time, it is a whole range of knowledge acquired by an individual consisting of all scientific, linguistic and artistic information absorbed; the discipline to reason, analyze information, and make correct judgment; and the ability to search for new information and create new knowledge. Special skills used in production of goods and services such as engineering, bookkeeping and surgery have also been included in the definition of education. This multi-faceted nature of education as a capital makes it difficult to categorize and to measure. There is no accurate measure of its many components. For this reason, very rough categorization has been used in planning and empirical works.

The usual classification used is by type of schooling such as the various grade levels, the various fields of specialization in college, and formal or non-formal education, the latter including skills training and informal campaigns. These, it is to be noted, are extremely rough categorization since it does not distinguish levels of actual knowledge. Despite this problem we proceed in the conventional manner.

Most of the schooling types of capital are additions to basic, i.e., primary education capital. Let us take the case of medical education. An investor in this capital would have to acquire elementary, high school, a few years of general education in college before being able to enroll in medicine itself. The capital is acquired in a certain sequence. In fact each type of schooling capital could be considered as the sum of a sequence of schooling, or sequences of investment, i.e., completion of grade 1, grade 2, ... pth year medical college so that medical schooling capital, K^p_{mp}, valued at time p can be written as

$$K^p_{mp} = K^p_{mp-1} + I_p = \sum_{t=1}^{p} I_t (1 + r)^{p-t} \qquad (2)$$

since in general

$$K^t_t = K^t_{t-1} + I_t \qquad (2')$$

If capital is valued at time zero

$$K^0_{mp} = \sum_{t=1}^{p} \frac{I_t}{(1 + r)^t} \qquad (3)$$

where I_t is investment at time t, $t = 1, 2, ..., p$. Time t in education is generally in annual units.

Any change in desired level of schooling capital would be made through a similar sequence of investment. The observable flow of investment is the expenditure in each grade leading to the desired schooling. If the desired number of medical degrees increases at time t everything else constant, investment or enrollment in preparatory schooling for medicine will also increase. Necessarily, there will be a lagged relation between changes in desired schooling capital and its acquisition, the lag depending on the length of the program.

A special feature of the formal schooling process is that there is a fairly rigid annual sequence of capital build-up from kindergarten up to the university level. The curricula for elementary up to sophomore college are usually for general education. Specialization takes place beyond sophomore in college and in sub-professional vocational-technical training. Intensity of specialization in a field increases as one moves from first to second and to third degree programs. The latter is a very important feature of formal schooling as it is provided currently. The heavy content of general education up to first degree programs permits much flexibility in labor adjustment to changes in skill requirements. One can easily move up the sequence or change fields of specialization while still in school or after joining the labor force. All these changes can be accomplished in a fairly short time. A shift of specialization in college will take less than four years and may be completed in just one year for related fields. A shift in graduate degree fields may take as much as the full length of a graduate program. Pursuit of medical or a Ph.D. degree takes about four additional years from a first degree. Nevertheless these are short gestation periods relative to worklife, or the lifetime.

Because of the sequential nature of formal schooling investment, an important benefit of attaining a certain level of schooling is the value of being able to pursue succeeding levels. A high school education is very valuable in this sense as it allows one to pursue a number of post-secondary education. In contrast an elementary education permits one to go on only to the next higher level of general education. Another feature of this capital is that it cannot be destroyed except by obsolescence or depreciation resulting from its idleness. Memory dims with time elapsed since last perception or learning of an object, a fact or a theory. On the other hand, one can build on a given stock of knowledge after leaving school. The ability to learn is itself an output of education.

Men are born with differing innate characteristics. There is unequal distribution of mathematical, artistic, linguistic, and even physical abilities. The varying historical and cultural backgrounds of families develop dissimilar values and attitudes. There is a distribution of these characteristics among the population of a nation though there may be common relative strength in any one or in a few characteristics viz other nationalities. The cost to an individial, psychic as well as monetary of acquiring a certain category of education, depends on his innate characteristics given school-related cost. Let us call this personal cost. Personal cost varies depending on the degree of matching of innate characteristics and those suited or required by the education category pursued. A mismatch will require of the student longer and more intensive input in his studies to overcome his poor ability in the subject. It might also mean psychic cost in the form of smaller satisfaction from the education pursued. The personal cost of pursuing a Ph.D. degree in Physics may be prohibitive for someone who has very poor innate ability in this field and who enjoys and is inclined to music or the stage.

Because of differences in innate abilities, attitudes and values, an individual will not be indifferent to education options that give equal monetary returns. Or an increase in the relative monetary returns to an option would not attract everybody to undertake that education process. We would expect instead an upward sloping enrollment of students in a given option. Monetary return has to compensate the personal cost of undertaking an education process in which students have weak abilities and inappropriate attitude. At some point the supply may turn vertical as personal cost becomes prohibitive because of utter lack of talent and dislike for the education of the marginal population. For this reason the supply elasticity is expected to decrease with intensity of specialization and for fields that require special abilities like the arts. There are many areas of education where supply is fairly elastic. Everyone qualifies for the lowest education levels. In fact many countries have compulsory elementary or even high school enrollment for children of the corresponding schooling ages starting from age six or seven years old. Most first degree college programs containing many courses of general interest should also have fairly elastic supply. First degree graduates of teacher education, business, psychology, liberal arts and even engineering have been more prone to shifting occupations from those which correspond to their majors.

Monetary cost of providing education varies by level and by field of specialization. In general cost increases with level. Higher levels seem to have larger scale economies. These have lead to the establishment of relatively large colleges and universities and their location in population centers. Some fields require more capital stock per pupil. Quality of instruction also depends on level of expenditures for teachers, laboratory and library.

The monetary cost to the student includes cost of instruction charged to him, foregone income and marginal living expenses including transportation in going to school. Geographic access to schools differs among students. There is a concentration of colleges and universities in large cities while primary schools are provided in all towns and large villages. High schools are not yet universally accesssible though many larger towns have them. The degree of dispersal of schools determines the distribution of distance among population groups. In most cases provincial students must bear a higher schooling cost than city students because of distance of their homes to schools. For this reason cost of distance tends to be higher the higher the level.

The distribution of schools among geographic areas and the distribution of ability and aptitudes are basic factors determining the supply function of students in each educational program. The more specialized it is, the steeper the supply curve. The derivation of the supply curve is discussed below using the above constrained optimization model. Psychic and distance costs have to be included in the cost parameter used in the optimization problem.

Section 3. Constrained Optimization Model and the Supply of Graduates

In the model we assume families to be maximizing the network of their children's education subject to two constraints, cost and ability. (Net worth is used equivalently with returns to education.)

$$\text{Max } NW_j^i = \sum_{t=1}^{j} \frac{R_t^{ij} - C_t^{ij}}{(1+r)^t} - \sum_{t=1}^{e} \frac{R_t^{i,e} - C_t^{i,e}}{(1+r)^t} \qquad (4)$$

st.

$$C_t^{ij} \leq B_t^i, \quad C_t^{ie} \leq B_t^i$$

for each year of schooling $t = 1, 2 \ldots j$ corresponding to age 7, 8 ... j + 6. For those desiring college

$$A^i \geq \overline{A}^j, \quad A^i \geq \overline{A}^e$$

where R is expected benefit, C is total cost for each year t in pursuit of education, j, of child i. The budget B_t for each child of corresponding age a is a function of family income, Y_f.

$$B_t^i \equiv B_a^i = b_a(Y_f^i) \qquad (5)$$

$\overline{A}$ is the minimum ability required for education e, j. Ability A is found to be strongly influenced by family income also. These constraints determine the set of alternatives which are relevant to children of given backgrounds and abilities.

Consider a population of children of a certain age range with its distribution by family income, distance to school facilities and abilities. Array the children by their schooling cost and map this distribution to the budgets for schooling. Children of age a can go to school so long as $C_a \leq B_a$. From this mapping we identify the children who can and those who cannot pursue each education alternative corresponding to their age a. The whole set of alternatives over all schooling ages of a child may also be obtained for given values of expected family income. We find that a given distribution of income, distance to school and abilities generate a distribution of sets of alternatives. The richer and brighter a child is, the larger the set facing him. The poor bright child in a distant location may face a set that is not significantly larger than that of his dull counterpart. On the other hand, superior quality of school and home environment in which rich children are brought up can so offset inferior inherent ability as to break the ability constraint for their higher education.

Market adjustment to relative rates of return to education will be made within the defined sets of relevant alternatives. A decreasing proportion of the population is expected to respond to positive returns to higher education, in particular, the more costly college programs. For this reason, disequilibrium in rates of return between costly and inexpensive program may be expected to persist or to be stable. In fact, this stable disequilibrium situation is frequently observed in many LDCs including Thailand and the Philippines.

Consider the following supply curves of graduates of different programs, A, B, C, which have increasing ability requirements. Gross return of A is in relation to B, that of B in relation to C. More specifically,

$$GR_B = \sum_{t=1}^{B} \frac{R_{Bt} - C_{Bt}}{(1+r)^t} - \sum_{t=1}^{A} \frac{R_{At} - C_{At}}{(1+r)^t} \quad \text{and}$$

$$GR_C = \sum_{t=1}^{C} \frac{R_{Ct} - C_{Ct}}{(1+r)^t} - \sum_{t=1}^{B} \frac{R_{Bt} - C_{Bt}}{(1+r)^t}$$

R_t is expected monetary benefits and C_t is cost of tuition, books and supplies at time t, excluding pesonal cost.

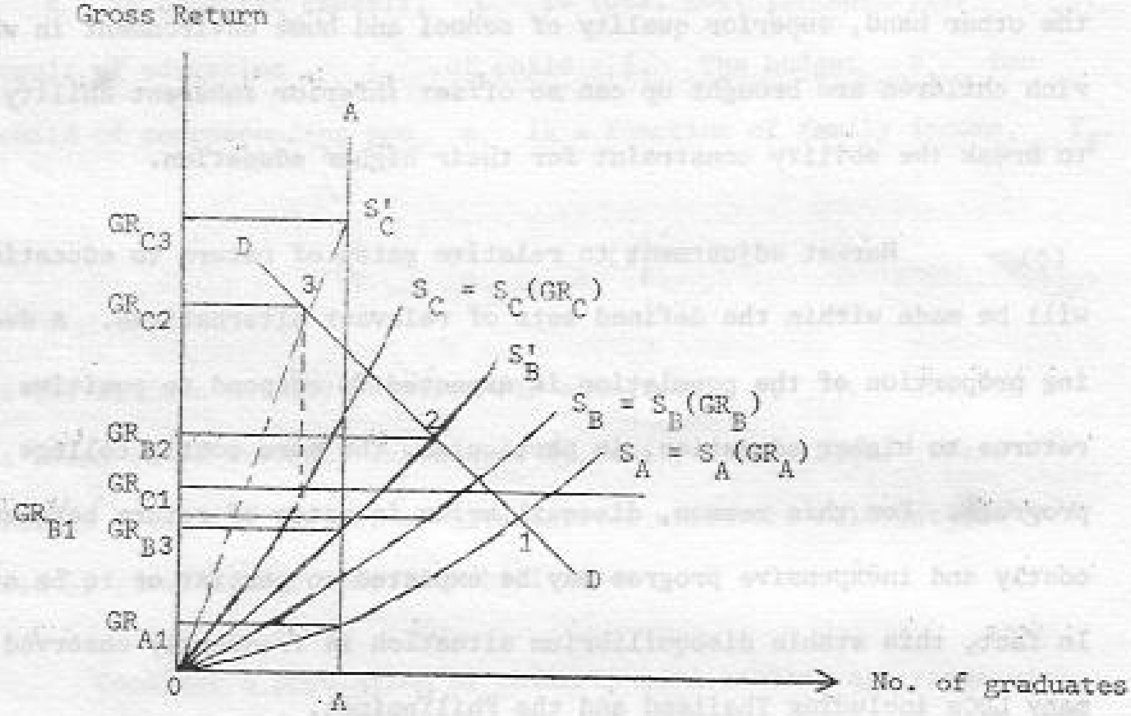

Chart 1

The S curves are the supply curves that would make marginal net returns (net of personal and distance cost) zero. As ability increases, the supply curve becomes less elastic. Differential in gross return must cover increasing personal cost of pursuing more stringent programs. Financial constraints should make the more costly higher education C programs even less elastic than if these were absent. This constraint is reflected in shifts to S'_C and S'_B. It is seen that a positive return to C over B education attracts less students than the same return to B relative to A. Take a demand curve DD. Operating at the S' curves would result in a disequilibrium in the market. Those able to meet the financial constraint for C will be earning positive net returns, say $GR_{C2} - GR_{C1}$ for C program and $GR_{B2} - GR_{B1}$ for B program.

The chart also shows that at a given level of gross return, the supply is smaller the higher or more specialized the education. A substantial gross return differential as indicated by the intersection of the vertical supply line AA and the respective S' curves is needed in order for there to be an equal number of graduates pursuing each program. The return differentials required for equal response of students are $GR_{C3} - GR_{B3}$ and $GR_{B3} - GR_{A1}$. The quantity supplied in each program is not positively related to the relative gross return. Moreover, an upward sloping aggregate supply cannot be meaningfully drawn.

We underscore the implications of the budget constraints on decision and test a hypothesis of determination of school attendance based mainly on this constraint. Irrespective of what may be the relative gross returns to schooling of different levels but provided they are not negative for the higher levels, we may argue that school attendance at each level will be mainly a function of variables affecting the financial constraint: family income, distance to school and level and type of program. As earlier explained, the longer the distance to school the larger are transport and additional living expenses.

Tuition, other fees and other expenses, in general, increase with level. Distance also increases with level since there are fewer higher educational institutions most of which are located in central cities. The financial constraint thus becomes more stringent with level of schooling.

Other variables that might influence family decision are father's education and number of sibling. There is a tendency for families to preserve its socio-economic position so that children are directed to the education and occupation that are at least as prestigious as the father's.

Given the above reasoning we argue that the probability of a child of a given age to attend a schooling level corresponding to this age is taken to be a function of all the above mentioned explanatory variables. And since the process of education is sequential, we assume a conditional probability function as follows:

$$P^{j}_{ia} = P^{j}_{a}(Y_{fi}, E_{fi}, L_{fi}, N_{i})P^{j-1}_{ia-1} \qquad (6)$$

$$P^{j-1}_{ia-1} = P^{j}_{a-1}(Y_{fi}, E_{fi}, L_{fi}, N_{i})P^{j-2}_{ia-2}$$

If a child did not complete the preceeding level, the probability of his attending the next level is zero. The probability of his pursuing other succeedingly higher levels is also zero at a given point in time. The family may decide to let him complete the previous level at a later time. This allows him subsequently to pursue next levels.

Section 4. Empirical Test

Equation (6) above applies to each child i of age a. For testing purposes this equation simplifies to

$$P_{ia} = P_a \ (Y_{fi}, E_{fi}, L_{fi}, N_i) \tag{7}$$

The equation estimates the probability of school attendance by all children in each group, i.e., those who completed and those who did not complete previous levels. It therefore estimates the product of p^a and p^{a-1} corresponding to level j and $j-1$.

Y = family income in thousands of baht

R_r = regions dummy, $r = 1, 2 \ldots 5$

1 is North

2 is Northeast

3 is Central

4 is Bangkok metropolis

5 is South

C_L = urban/rural dummy, $L = 0, 1, 2$

0 is for village

1 is for metropolis

2 is for municipal

E_e = father's education, $e = 0, 1, 2, \ldots 5, 6\text{-}7, 8\text{-}9$

0 is zero schooling

1 is Pratum 1-4

2 is completed Pratum 4

3 is completed Pratum 7

4 is completed Maw Saw 3

5 is completed Maw Saw 5

6-7 is completed college and teacher education

8-9 is technical and other training

Rel_q = relation of youth to head dummy,* $q = 1, 2 \ldots 9$

1 is spouse of head

2 is unmarried son/daughter

3 is married son/daughter

4 is son/daughter in-law

5 is nephew/niece

6 is parents of head

7 is relatives

8 is other dependent

9 is servant and employee

N is number of sibling of schooling age 6-24, a continuous variable

---

*Used in the analysis are relationships coded as 2, 5, 7 and 8.

We grouped the sample population by age ranges 7-10, 11-13, 14-18 and 19-24 which correspond to Thailand's schooling stages – Pratum 1-4 (primary level), Pratum 5-7 (upper elementary), Maw-Saw 1-5 (high school) and post secondary. The latter includes teacher and technical training and university education.

Both regression and logit methods were used in testing the hypothesis. Logit fits the model better as it directly gives an estimate of the probability of attendance in school. Note that the dependent variable is dichotomous, i.e., attendance or non-attendance in school assuming a value of 1 or zero and therefore has a non-normal distribution. This violates an important assumption of least squares estimate rendering interpretation of the estimated regression coefficients and the regression statistics unclear. (Please see the technical note on this in the appendix.) Nevertheless, regression analysis was applied for two reasons. First, we wanted to compare the results of the two methods to see whether the less expensive (computer timewise) regression estimates approximate closely enough the more suitable but expensive logit model. Some recent studies involving dichotomous dependent variables relied on regression estimates (Encarnación and Canlas, 1976; Canlas and Razak, 1979). Acceptance of their results would depend in part on our findings.

A second reason is expediency. Logit analysis takes more than ten times computer time for the equivalent regression estimate. We used regression to obtain the "best" independent variable selection in terms of significance and R^2s. The "best" selection was used for the logit analysis. We did not try to investigate the validity of this approach.

We tested the function by regression including all independent variables on each sample group and for a sub-set of independent variables after eliminating those that contributed insignificant value to the R^2. (Please see Appendix table for results of the complete regression runs.) Presented below are the results that we consider to provide the best estimate of the function. Included variables are family income, fathers' education, number of sibling, region and municipal/village location category. Though we cannot interpret the coefficients in terms of probability values we take their significance, sign and movement as rough estimates of the relationship. The regression results seem to be very satisfactory. Most of the regression parameters are significant and of the expected sign. Furthermore, they moved in the expected direction as we go from the youngest age group to the oldest.

First let us look at the constants. They declined from .786 to .262 as age (or level of schooling) increases from 7-11 to 19-24. This pattern reflects in part society's attitude to the different levels of schooling. The lower levels are regarded as a basic need required of every citizen for his orderly and satisfactory participation in all types of social interaction. Hence the large constant. As we move up the schooling ladder, the role of education becomes more specialized and there is no longer a common or equal desire for each level or type. Demand partly depends on expected net monetary benefits, partly on matching of perceived ability and required qualification. For these reasons the value of the constant tends to fall as schooling level increases. On the other hand, mainly because of financial constraints, the influence of cost-related factors tended to rise with level of schooling. Recall that schooling cost increases with level. Hence, the larger absolute value of the coefficient of family income, location, number of sibling and fathers' education. Fathers' education influences more strongly decision to pursue higher levels for two other reasons. One is its positive effect on home education which enhances inherent ability and therefore scholastic performance that is recognized in schools and in the selection of applicants for college or high school. Another influence is on education and occupational identity developed in children. Since lower education levels are regarded as basic for everybody, the influence of identity factor is likely to be weak, hence father's education.

The coefficients of the independent variables obtained by logit were practically all significant at the one percent level, similar to the results obtained by OLS. However, their values and movement as we go up the education ladder were very different. The sign and value of the intercepts were also very different. The influence of the various variables as reflected in the Beta coefficients increased with age range but up to 14-18 age range only. Then their values dropped. In contrast the movement of the regression coefficient was upward throughout.

The value of the probability of attendance in school for each age range can be estimated from the logit Beta coefficients for different values or categories of the independent variables. We limited our exercise to the extreme values of the variables to obtain the range of value of the probability for each age range. We take as the extreme values for number of youth to be one and five and for income B5,000 and B100,000. The probability, P, of the dependent variable, attending or not attending school with value of 1 or 0, taking on value of 1 is

$$P/_{Y=1} = \frac{1}{1 + e^{-X'B}} = \frac{e^{X'_B}}{1 + e^{X'B}}$$

X and B are vectors of independent variables and the β coefficients respectively. We obtain the following val0es of the probability of attending school.

Table 9
Results of Linear Regressions of
Attending or not Attending School

	6-10	7-10	11-13	14-18	19-24
Constant	.7180	.7862	.8170	.5361	.262
Family income	.0004**	.0003**	.0005**	.0007**	-.0000
Region 2 (NE)	-.0581**	-.0353	-.1069**	-.0154	-.0342**
Region 3 (Central)	-.0008	-.0016**	-.0128	.0606**	-.0268*
Region 4 (Bangkok)	.0184	.0302**	.0296**	.0987**	.0815**
Region 5 (South)	-.0693**	-.0381**	.0096	.0949**	.0090
Municipal	.0806**	.0390**	.0610**	.2050**	.1521**
Non-municipal	-.0397**	-.0339**	-.0951**	-.1721**	-.0355**
Number of youth	.0060**	.0031*	-.0069**	-.0143**	-.0021
FE_2 (Father's Education)	.0152*	.0276**	.0193**	.0310**	.0258**
FE_3	.0939**	.0749**	.1365**	.2260**	.1271**
FE_4	.1358**	.0994**	.1022**	.2740**	.2006**
FE_5	.1574**	.1270**	.0758**	.2743**	.2267**
FE_{6-7}	.1899**	.1205**	.1134**	.3108**	.3156**
FE_{8-9}	.1981**	.1181**	.1170**	.2417**	.2828**
$\bar{R}^2$	.055	.036	.098	.240	.132

Table 10
Range of Values of Probability of Attending School by Age
1975

	AGE RANGE 7-10	11-13	14-18	19-24
1. Lowest				
$X'\beta$	.912	.137	-1.787	-1.374
$P/_{Y=1}$	.7134	.5342	.1435	.2020
2. Highest				
$X'\beta$	4.030	4.807	3.358	.257
$P/_{Y=1}$	.9825	.9919	.9664	.5639

Table 11
Value of Beta in Logit

Independent Variable	AGE RANGE 7-10	11-13	14-18	19-24
Constant	1.193	1.462	- .472	-1.042
Family income	.009** (39.82)	.011** (43.35)	.006** (85.12)	- .0003 (0.144)
Region 2 (Northeast)	- .229** (8.70)	- .620** (54.69)	- .130* (4.47)	- .115** (3.01)
Region 3 (Central)	- .053 (0.37)	- .132 (2.01)	.251** (15.99)	- .084** (2.18)
Region 4 (Bangkok)	.294** (9.67)	.358** (11.64)	.434** (51.69)	.221** (6.19)
Region 5 (South)	- .283** (8.34)	.099 (0.74)	.445** (32.74)	.027 (0.554)
Municipal city	.253 (11.36)	.446** (28.88)	.867** (284.90)	.454** (12.99)
Non-municipal	- .137 (3.42)	- .405** (26.14)	- .800** (178.38)	- .185** (4.74)
Number of youth	.008 (0.30)	- .071** (20.39)	- .083** (66.49)	- .006 (1.01)
Father's Education 2	.200** (13.10)	.156** (6.77)	.151** (13.67)	.796** (3.29)
Father's Education 3	.603** (13.55)	1.696** (39.18)	1.149** (121.84)	.361** (6.27)
Father's Education 4	.967** (48.78)	1.175** (41.66)	1.542** (290.20)	.563** (13.41)
Father's Education 5	2.020** (11.84)	.848** (3.90)	1.747** (63.34)	.576** (6.16)
Father's Education 6-7	1.382** (19.31)	1.512** (16.85)	2.012** (110.94)	.660** (10.29)
Father's Education 8-9	1.380** (19.33)	1.602** (14.33)	1.335** (58.06)	.672** (8.

* .05 significance level.

** .01 significance level (chi-square values with one degree of freedom) are in parenthesis except for age 19-24 where asymptotic t-values

To be noted from Table 10 are the rapid drop of the value of the probability of school attendance as schooling level increases and its wide range for the different socio-economic and location classes. For the possibly worst-off children the probability drops successively from .7134 to .5342, to .1435, to .2020 for ages 7-10, 11-13, 14-18 and 19-24. The corresponding figures for the possibly "best-off" children are .9825, .9919, .9664 and .5639.

Attendance rate at each age range can be predicted for different population groups and from this the distribution of formal schooling. Attendance rate is determined by the set of characteristics of each group. The model tells us further the importance of financial variables on schooling decision and provides us with an inter-generational link in the acquisition of education capital. An increase in a group's income or a change in location would have a permanent impact on all its future generations. The link is through future fathers' education and their subsequent income. Finally, the model explains why distribution tends to be more unequal as level of schooling rises. The policy implications seem to be obvious since financial factors are not difficult to change.

Appendix A

Table A.1
Regression Results on Alternative Selection of
Independent Variables, Age 7-10

Independent Variables	REGRESSION COEFFICIENTS					
Intercept	0.0807	0.7854	0.7899	0.7850	0.7862	0.7838
Income	0.0004 (6.1568)	0.0004 (6.0977)	0.0003 (4.6406)	0.0004 (5.0815)	0.0003 (4.5598)	0.0003 (5.0162)
Reg 2	-0.0445 (-4.3451)	-0.0447 (-4.3670)	-0.0350 (-3.3993)	-	-0.0353 (-3.4276)	-
Reg 3	0.0046 (0.4165)	0.0041 (0.3718)	-0.0010 (-0.0930)	-	-0.0016 (-0.1400)	-
Reg 4	0.0522 (4.8310)	0.0519 (4.8053)	0.0305 (2.7662)	-	0.0302 (2.7373)	-
Reg 5	-0.0327 (-2.4800)	-0.0324 (-2.4595)	-0.0383 (-2.9097)	-	-0.0381 (-2.8911)	-
C1	-	-	0.0388 (4.0643)	0.0426 (4.5103)	0.0390 (4.0851)	0.0428 (4.5314)
C0	-	-	-0.0342 (-3.3435)	-0.0510 (-5.1378)	-0.0339 (-3.3128)	-0.0507 (-5.1075)
N Youth	0.0029 (1.6666)	0.0031 (1.7543)	0.0029 (1.6700)	0.0029 (1.6332)	0.0031 (1.7665)	0.0030 (1.7144)
E 2	0.0249 (3.0635)	0.0278 (3.6362)	0.0244 (3.0119)	0.0211 (2.6078)	0.0276 (3.6201)	0.0238 (3.1276)
E 3	0.0886 (4.9342)	0.0908 (5.0984)	0.0724 (4.0250)	0.0695 (3.8644)	0.0749 (4.1963)	0.0717 (4.0128)
E 4	0.1189 (8.6941)	0.1219 (9.3172)	0.0961 (7.0873)	0.0915 (6.7537)	0.0994 (7.4886)	0.0943 (7.1132)
E 5	0.1442 (4.7208)	0.1477 (4.8644)	0.1231 (4.0302)	0.1242 (4.0639)	0.1270 (4.1834)	0.1275 (4.1971)
E 6-7	0.1339 (6.0064)	0.1370 (6.2075)	0.1172 (5.2573)	0.1172 (5.2541)	0.1205 (5.4581)	0.1200 (5.4304)
E 8-9	0.1331 (5.7265)	0.1358 (5.8850)	0.1152 (4.9560)	0.1113 (4.7774)	0.1181 (5.1151)	0.1137 (4.9163)
Rel 2	-0.0114 (-0.1602)	-	0.0009 (0.0129)	0.0029 (0.0411)	-	-
Rel 5-7	-0.0209 (-0.2922)	-	-0.0096 (-0.1349)	-0.0061 (-0.0852)	-	-

Table A.2
Regression Results on Alternative Selection of
Independent Variables, Age 11-13

Independent Variables	REGRESSION COEFFICIENTS					
Intercept	0.5393	0.8049	0.5256	0.5073	0.8246	0.8181
Income	0.0007 (8.4062)	0.0007 (8.3688)	0.0004 (5.2454)	0.0006 (6.2330)	0.0004 (5.2220)	0.0006 (6.1996)
Reg 2	-0.1269 (-10.6863)	-0.1270 (-10.6808)	-0.1065 (-9.0183)	–	-0.1069 (-9.0413)	–
Reg 3	-0.0037 (-0.2960)	-0.0026 (-0.2057)	-0.0140 (-1.1283)	–	-0.0128 (-1.0315)	–
Reg 4	0.0741 (6.1338)	0.0740 (6.1187)	0.0292 (2.3948)	–	0.0296 (2.4247)	–
Reg 5	0.0233 (1.5710)	0.232 (1.5618)	0.0095 (0.6462)	–	0.0096 (0.6559)	–
C1	–	–	0.0627 (5.8432)	0.0691 (6.4853)	0.0610 (5.6788)	0.0672 (6.3018)
C0	–	–	-0.0955) (- 8.1969)	-0.1309 (-11.5173)	-0.0951 (- 8.1474)	-0.1308 (-11.4886)
N Youth	-0.0062 (-2.9530)	-0.0063 (-3.0064)	-0.0070 (- 3.3790)	-0.0078 (- 3.7898)	-0.0069 (- 3.3836)	-0.0078 (- 3.7638)
E 2	0.0166 (1.8851)	0.0146 (1.7052)	0.0206 (2.3746)	0.0073 (0.8328)	0.0193 (2.2806)	0.0065 (0.7688)
E 3	0.1696 (8.5087)	0.1683 (8.4338)	0.1374 (6.9525)	0.1208 (6.0697)	0.1365 (6.9018)	0.1200 (6.0224)
E 4	0.1459 (9.4242)	0.1434 (9.2839)	0.1038 (6.7085)	0.0867 (5.5711)	0.1022 (6.6184)	0.0853 (5.4952)
E 5	0.1181 (3.2723)	0.1143 (3.1620)	0.0792 (2.2204)	0.0697 (1.9367)	0.0758 (2.1194)	0.0661 (1.8351)
E 6-7	0.1476 (5.8824)	0.1383 (5.5214)	0.1230 (4.9613)	0.1106 (4.4280)	0.1134 (4.5794)	0.1010 (4.0420)
E 8-9	0.1591 (5.5600)	0.1544 (5.3987)	0.1209 (4.2705)	0.1072 (3.7540)	0.1170 (4.1302)	0.1034 (3.6235)
Rel 2	0.2634 (5.2717)	–	0.2977 (6.0340)	0.3102 (6.2329)	–	–
Rel 5-7	0.2748 (5.4189)	–	0.3055 (6.1010)	0.3145 (6.2269)	–	–

Table A.3
Regression Results on Alternative Selection of
Independent Variables, Age 14-18

Independent Variables	REGRESSION COEFFICIENTS					
Intercept	0.1671	0.4125	0.0654	0.0893	0.3897	0.4351
Income	0.0012	0.0011	0.0007	0.0008	0.0007	0.0007
	(16.6221)	(16.0596(	(10.7090)	(11.2777)(	9.9985)	(10.5833)
Reg 2	-0.0582	-0.0591	-0.0125	-	-0.0154	-
	(- 4.8486)	(- 4.9160)	(- 1.0920)		(- 1.3353)	
Reg 3	0.0817	0.0848	0.0564	-	0.0606	-
	(6.4410)	(6.6656)	(4.6782)		(4.9995)	
Reg 4	0.2059	0.2087	0.0928	-	0.0987	-
	(17.4918)	(17.6885)	(8.0937)		(8.5658)	
Reg 5	0.1465	0.1464	0.0953	-	0.0949	-
	(9.4852)	(9.4599)	(6.4843)		(6.4316)	
C1	-	-	0.2108	0.2249	0.2050	0.2194
			(20.0142)	(21.6315)(	19.3857)	(21.0029)
CO	-	-	-0.1759	-0.2023	-0.1721	-0.2009
			(-15.0737)	(-17.7692)(	-14.6757)	(-17.5474)
N Youth	-0.0160	-0.0153	-0.0157	-0.0150	-0.0143	-0.0135
	(- 7.7982)	(- 7.5370)	(- 8.1126)	(- 7.7873)(	- 7.3931)	(- 7.0138)
EDH 2	0.0192	0.0186	0.0302	0.0249	0.0310	0.0253
	(2.2528)	(2.1808)	(3.7323)	(3.0749)(	3.8333)	(3.1204)
E 3	0.3041	0.2968	0.2364	0.2298	0.2260	0.2182
	(15.9341)	(15.5506)	(13.0275)	(12.6162)(	12.4129)	(11.9347)
E 4	0.3865	0.3762	0.2866	0.2749	0.2740	0.2604
	(26.3810)	(25.7111)	(20.3581)	(19.5050)(	19.4096)	(18.4300)
E 5	0.3588	0.3530	0.2857	0.2911	0.2743	0.2801
	(12.3052)	(12.1182)	(10.3222)	(10.4928)(	9.8902)	(10.0690)
E 6-7	0.3960	0.3856	0.3282	0.3221	0.3108	0.3032
	(15.9075)	(15.5369)	(13.8877)	(13.5861)(	13.1446)	(12.7826)
E 8-9	0.3320	0.3190	0.2623	0.2536	0.2417	0.2312
	(11.7189)	(11.2806)	(9.7573)	(9.3977)(	8.9772)	(8.5509)
Rel 2	0.2505	-	0.3365	0.3561	-	-
	(9.0156)		(12.7457)	(13.4621)		
Rel 5-7	0.2493	-	0.3091	0.3288	-	-
	(8.6342)		(11.2888)	(11.9751)		

Table A.4
Regression Results on Alternative Selection of
Independent Variables, Age 19-24

Independent Variables	REGRESSION COEFFICIENTS					
Intercept	0.0607	0.1660	-0.0479	-0.0649	0.0972	0.1030
Income	0.0000	0.0002	0.0000	0.0001	-0.0000	0.0000
	(3.0133)	(2.5304)	(0.3623)	(0.8243)	(- 0.2008)	(0.2242)
Reg 2	-0.0476	-0.0496	-0.0306	-	-0.0342	-
	(- 3.2148)	(- 3.3500)	(- 2.0858)		(- 2.3273)	
Reg 3	-0.0236	-0.0231	-0.0281	-	-0.0268	-
	(- 1.5423)	(- 1.5068)	(- 1.8646)		(- 1.7747)	
Reg 4	0.1346	0.1367	0.0773	-	0.0815	-
	(9.9737)	(10.1230)	(5.6635)		(5.9660)	
Reg 5	0.0446	0.0431	0.0118	-	0.0090	-
	(2.3182)	(2.2464)	(0.6211)		(0.4717)	
C1	-	-	0.1567	0.1823	0.1521	0.1782
			(11.7715)	(13.9036)	(11.4406)	(13.6015)
C0	-	-	-0.0381	-0.0564	-0.0355	-0.0554
			(- 2.5193)	(- 3.7716)	(- 2.3404)	(- 3.6966)
Youth	-0.0040	-0.0039	-0.0027	-0.0021	-0.0021	-0.0014
	(- 1.8583)	(- 1.8152)	(- 1.2775)	(- 0.9853)	(- 0.9806)	(- 0.6778)
E 2	0.0178	0.0162	0.0285	0.0244	0.0258	0.0209
	(1.8495)	(1.6890)	(3.0115)	(2.5699)	(2.7223)	(2.2014)
E 3	0.1537	0.1509	0.1331	0.1347	0.1271	0.1279
	(7.1557)	(7.0485)	(6.2869)	(6.3300)	(6.0139)	(6.0200)
E 4	0.2483	0.2415	0.2122	0.2075	0.2006	0.1937
	(15.3223)	(15.1225)	(13.2027)	(12.8702)	(12.6157)	(12.1482)
E 5	0.2510	0.2480	0.2351	0.2604	0.2267	0.2526
	(8.4495)	(8.3865)	(8.0439)	(8.8971)	(7.7800)	(8.6515)
E 6-7	0.3457	0.3420	0.3262	0.3355	0.3156	0.3243
	(14.5063)	(14.5469)	(13.8953)	(14.2609)	(13.6002)	(13.9349)
E 8-9	0.3226	0.3163	0.2991	0.0351	0.2828	0.2872
	(12.1829)	(12.1959)	(11.4655)	(11.6370)	(11.0463)	(11.1579)
Rel 2	0.1054	-	0.1481	0.1699	-	-
	(4.4762)		(6.3605)	(7.3042)		
Rel 5-7	0.1126	-	0.1379	0.1607	-	-
	(4.5602)		(5.6689(	(6.5977)		

Technical Note on Logistic Model

by

Fe Lisondra

In multiple linear regression model,

$$y_i = \beta X_i + \mu_i \text{ a/}$$

y_i is normally distributed for fixed X_i, in which the following assumptions are made:

1. $E(\mu_i) = 0$

2. $E(\mu_i \mu_j) = \sigma^2 \qquad i = j$

$\qquad\qquad = 0 \qquad i \neq j$

In actual practice, it is not always reasonable to assume y_i to be normally distributed. Although X_i in some cases may not be normal, i.e., when some of its components are dummy variables, y_i for fixed X_i could be assumed normal and the variance-covariance matrix for y_i given X_i does not depend upon X_i. But in situations where the dependent variable y_i is dichotomous 0 or 1, the ordinary least squares method will yield $E(\mu_i) = 0$ but with $var(y_i) = Var(\mu_i) = X_i' \beta (1 - X_i' \beta)$ since y_i is a Bernoulli

---

a/ y_i denotes the dependent variable for the i^{th} observation. $X_i = (X_{i1}, X_{i2}, X_{i3}, \ldots, X_{ip})$ is the vector of p independent variables for the i^{th} observation. $\beta = (\beta_1, \beta_2, \ldots, \beta_p)$ denotes the vector of regression parameters. μ_i denotes the i^{th} uncorrelated disturbance term.

random variable. It is clear that the resulting error variance is not constant for all observations. Thus, regression of y_i on X_i is heteroscedastic and should not be estimated using the ordinary least squares since it violates one of the basic assumptions generally made in a linear model. This heteroscedasticity (unless necessary corrections or transformations are made) will generate inefficient estimators of β. Standard errors of the sample regression coefficients would be therefore incorrect and as a result, tests of significance and confidence intervals for regression coefficients may be seriously misleading. Furthermore, in OLS, estimators of $X_i'\beta$ can have any numerical value despite the fact that $E(y_i) = X_i'\beta$ and $0 \leq y_i \leq 1$, $0 \leq X_i'\beta \leq 1$. This means that y_i being a probability, rules out the linear model because y_i may not be bounded by 0 or 1.

The logistic model provides an appropriate analysis of binary response data. The model in a logistic cumulative distribution function form,

$$P_{y_{i=1}} = F(X_i'\beta) = (1 + \exp(-X_i'\beta))^{-1} \text{ b/}$$

---

b/ This nonlinear function represents the relationship between the probability of attending school $P_{y_{i=1}}$ and the socio-economic and demographic characteristics represented by vector X_i. β represents the vector of regression parameters.

has a curve similar to the cumulative curve of the normal distribution. Its likelihood function is

$$L = \prod_{i=1}^{n} \left[\frac{1}{1 + \exp(-X_i \beta)} \right]^{y_i} \left[1 - \frac{1}{1 + \exp(-X_i \beta)} \right]^{1-y_i}$$

$$= \frac{(\exp \beta' \sum_{i=1}^{n} X_i y_i)}{\prod_{i=1}^{n} \left[1 + \exp(X_i \beta) \right]}$$

where the maximum likelihood estimator[c/] of β is obtained by differentiating the logarithm of the likelihood function, setting the result equal to 0 and solving for β.

This method using the logistic cdf in solving regressions problems with qualitative dependent variable is called logit analysis. The chi-square statistic for testing the hypothesis that a parameter is zero is calculated by computing the square of the parameters estimate divided by its standard error.

[c/] In this paper, the maximum likelihood estimates (MLE) were computed using the Newton-Raphson method.

Bibliography

Canlas, Dante B. and Mohd Razak. "Education and the Labor Force Participation of Married Women: West Malaysia 1970," Discussion Paper 7910. Quezon City: IEDR, School of Economics, University of the Philippines (July).

Chatterjie and Price. Regression Analysis by Example, Wiley, New York, 1977, 117-120.

Encarnación, Jose, Jr., and Dante B. Canlas, "Income, Education, Fertility and Employment: Philippines 1973," Discussion Paper 7628. Quezon City: IEDR, School of Economics, University of the Philippines (December).

Hicks, John R. Capital and Time, Oxford, 1973.

Jorgenson, Dale W., "Capital Theory and Investment Behavior," American Economic Review, Papers and Proceedings, May 1963.

Nerlove and Press, Univariate and Multivariate Log-linear and Logistic Models, Rand, 1973.

Pernia, Ernesto, "An Intergenerational and Sequential Analysis of Migration Decision: The Philippines," The Philippine Review of Business and Economics, March 1969.

Press, James, Applied Multivariate Analysis, Holt, Rinehart and Winston, New York, 1972, 265-268.

Wesolowsky, Multiple Regression and Analysis of Variance, Wiley, New York, 1976, 126-128.

Source:
National Statistical Office of Thailand, 1975 Survey of Youth, Bangkok, Thailand.