The Impacts of Educational Expansion and Schooling Inequality on Income Distribution

Hong-Yih Chu
Feng Chia University*

Most previous studies use (i) either schooling level or schooling inequality, but not both, in their assessment of education on income inequality, and/or (ii) cross-country data that have parametric implications because they assume the relationship between income inequality and independent variables is the same in all sample countries; and/or (iii) a single income inequality measure as a dependent variable, however, any single income inequality measure is incomplete Using the indirect quantile approach, we establish a time series model in which the aggregate variables such as schooling level and schooling dispersion could affect income inequality through a set of quantile income levels. This model not only avoids the limitations stated above but also makes the distribution data more useful, the analysis more flexible, and provides a high level of disaggregation.

Introduction

The importance of education for economic growth is well known. In addition, schooling is often seen as a potentially powerful social equalizer The policy of equal access to education is supported, therefore, when there is an interest in equalizing income distribution The effect of schooling on income inequality remains unclear and controversial. If the rate of return on s years of schooling is independent from s years, an increase in the level of schooling leads to more income inequality. If they are not independent, the direction of the increase in schooling level on income inequality can go either way (Fields, 1980).

Knight and Sabot (1983) also observe that there are two effects of educational expansion on income inequality. The composition effect, raising the earnings of those who are more educated, tends to increase income inequality. The wage compression effect, which follows the expansion of supply of educated labor relative to demand, goes in the opposite direction According to Blaug, Dougherty, and Psacharopoulos

* I would like to thank an editor and an anonymous referee of this journal for helpful and constructive comments The usual disclaimer applies.

0747-5535/00/1300/0039/$2 00

(1982), an equalization in the distribution of schooling produces two separate but opposite effects on earnings. First, it raises the mean earnings of the low-paid. Second, it makes the age-earnings profile steeper, thus increasing the variance of earnings. In summary, whether schooling level or schooling inequality is positively or negatively related to income inequality remains somewhat uncertain The question of what role schooling can and should play in income distribution, therefore, is still open to debate

Ram (1989) argues that (i) most previous studies use either schooling level or schooling inequality, but not both, in their assessment of education on income inequality; and (ii) most studies use cross-country data, which have parametric implications, because they assume the relationship between income inequality and independent variables is the same in all sample countries. In addition, most studies use a single income inequality measure as a dependent variable; however, any single income inequality measure is incomplete

For instance, the Gini index is an ambiguous measure when Lorenz curves intersect. The Gini index also gives more weight to transfer near the mode of the distribution than at the tails. The coefficient of variation gives equal weight to transfer through the whole distribution income. The limitation of the relative mean deviation is its insensitivity to transfer income among individuals within the same group (Kakwani, 1980). In addition, a rapid increase in average income may obscure the observation and increase the dispersion of income (Weisskoff, 1970). In order to put more weight on the lower ends of the distribution, Basmann and Slottje (1988) present a new measure—the B measure—which is particularly sensitive to transfer of income from the poorest income shares to the richer income shares. If any income group has zero income, however, the value of the B measure indicates perfect inequality

For the reasons stated above, the appropriate approach, therefore, should (i) ideally use a time series model that can avoid the limitations of a cross-country study, (ii) include both schooling level and dispersion as independent variables, and (iii) not use any single measure as a dependent variable

Using the indirect quantile approach (Beach, 1977), we establish a time series model in which the aggregate variables such as schooling level and schooling dispersion could affect income inequality through a set of quantile income levels This model not only avoids the above limitations, but also makes the distribution data more useful, the analysis more flexible, and provides a high level of disaggregation.

A Review of Some Empirical Studies

Due to limitations of data, empirical studies about the effects of schooling on income inequality are not plentiful. Little of the literature studies time series observations within the same country. Even the number of cross-country studies is relatively low (Ram, 1989). Research in the area of how schooling impacts income inequality can usefully be conceived of as falling into three groups. First, some center around the impact of the average level of schooling on income inequality. Second, some focus on the effect of schooling inequality on income distribution. Third, some studies investigate the effects of both schooling level and inequality on income inequality. We will discuss some examples in each of the three groups below.

Among the first group, Blaug, Dougherty, and Psacharopoulos (1982) use the standard earnings function of human capital theory and claim that raising the school leaving age increases income inequality Mohan and Sabot (1988) claim that the compression effect dominates the composition effect; therefore, educational expansion will reduce income inequality. Psacharopoulos and Steier (1988) find that an increase in the schooling level can significantly reduce income inequality in terms of the variance of the log earnings Ram (1989) argues that the effect of the mean schooling level of the labor force on the Gini index is small. Regressing income share to private rate of return to education level and other variables. Tilak (1989) concludes that educational expansion tends to reduce income inequality.

Turning to the second group of analyses, Caniglia (1988) declares that there is a positive relationship between schooling inequality and income inequality. In another study, Chiswick (1974) concludes that the coefficient of schooling inequality is not significant.

In the third group of studies, Lam and Levison (1992) and Park (1996) conclude that the rising schooling level associated with the steady declines in schooling dispersion will have beneficial effects on income inequality Plotnick (1982) finds that either increases in schooling level or schooling dispersion tends to worsen the income distribution Winegarden (1979) deduces that higher average levels of schooling exert an equalizing effect, and that schooling dispersion plays a larger role in generating income inequality than previous studies have revealed. With a model similar to Winegarden's specification and more recent data, Ram (1984) asserts that there are no significant effects of schooling level and dispersion on income inequality.

In summary, as claimed by Ram (1989), even if schooling inequality is positively related to income inequality, the effect of changes in schooling inequality on income distribution remains somewhat uncertain Predictions about the effect of higher schooling level are even more ambiguous.

Economists have been interested in how economic gain is distributed in society, i.e , who gains relatively more from educational expansion and who suffers less from the widening of schooling dispersion. In addition, to what extent does income inequality respond to change in schooling level and dispersion? The model presented in the following section investigates these issues.

The Model

The indirect quantile approach aims to link the schooling effect of income inequality to a set of quantile income levels. The first step establishes a simple reduced-form model in which schooling affects individual quantile income. The second step shows the measures of income concentration in terms of these income quantile.

The quintiles selected are nine income deciles y(1), y(2), … , y(9). The i[th] decile represents a real income point such that 10*i percent number of the population have real incomes less than or equal to this level and 10(10-i) percent have real incomes that exceed it. We should note that this breakdown is relative to the real income position and not the real income level of a particular group.

Each laborer's income is divided into two components: labor and non-labor income.

(1) $y(i) \equiv ly(i) + ny(i)$.

Term ly(i) is real income received from employment, while ny(i) is real non-labor income.

Labor income is a function of schooling level and schooling dispersion. The first term on the right side, therefore, can be expressed as

(2) $ly(i) = f(s(i), VS)$,

where s(i) represents s years of schooling of the i[th] decile income. It should be emphasized at this point that everyone with income y(i) is not assumed to have the same schooling years, but only that the schooling in equation (2) represents the average over the members of this quantile-specific group. Symbol VS represents the schooling dispersion of the population. Equation (1) can then be represented as

(3) $y(i) = f(s(i), VS, ny(i))$.

Two variables on the right side, s(i) and ny(i), are not observable. We assume that there is a unique set of values for the intercept and the slope and the relationship between the subgroup sum (or mean) and overall sum (or mean) is in a linear fashion.[1] The relationship can be shown as

(4) $S(i) = \delta_s(i) + \alpha_s(i)S + u_s(i)$

$$ny(i) = \delta_{ny}(i) + \alpha_{ny}(i)NY + u_{ny}(i)$$

where.

S = Average schooling years of employed labor,
NY – Average real non-labor income of the population;
u(i) – A disturbance term that account for the disagreement.

Substituting (4) into equation (3), the equation becomes

(5) $y(i) = X\beta(i) + v(i)$,

where:

y(i) – A column of vectors of n observations on the i[th] group,

[1] In equation (4) there are disturbance terms that account for the disagreement. The author is obliged to an anonymous referee for his guidance of the model specification

x = A n*4 matrix (four variables are S, VS, NY, and constant);
$\beta(i)$ = The coefficient; and
v(i) = A n-dimensional vector of the random term [2]

By the same token, we can specify an income equation for the real mean μ of the distribution.

$$(6)\quad \mu = X\beta_{\cdot} + v_{\mu}\,.$$

One can now set out the overall income distribution for the nine decile equations and associate with the mean income equation (6) Combining these ten seemingly unrelated equations, we can write the matrix notation as,

$$(7)\quad y = X\beta + V\,.$$

The dependent variable y that is used in this analysis is log income. The motivation for choosing log income is that it provides a better way to interpret the link between the variance of schooling and the log variance of incomes. More generally, it decomposes the link between schooling and inequality into a component driven by schooling level and a component driven by the dispersion in schooling. Thus far, what we have done is to express a set of income deciles in terms of a set of aggregate economic variables. Further constraints are needed in the model.

The coefficients to be estimated are subjected to a set of adding-up constraints. The income decile, y, represents a real income point such that j percent of the population have real incomes less than or equal to this level and 100 - j percent have real incomes that exceed it. $g(y_i)$ is the underlying density function for the income distribution It can be seen that if the schooling level were to shift up by one percentage point, the resulting distribution of the shift among the income groups in the distribution would have to satisfy the constraint (for the case of schooling level):

$$(8)\quad \int_{j=0}^{100}\frac{\partial y_j g(y_j)}{\partial s}\,dy = \frac{\partial \mu}{\partial s}$$

The left side of equation (8) can then be written as

$$(.10)\sum_{i=1}^{9}\frac{\partial y_{i0}}{\partial s}$$

Because we only have nine decile equations in this study, the weights sum to .90 rather than to unity. Therefore, the right side should be written as

[2] The absolute values of the normal statistics for the disturbance terms range from 0269 to 1 3319, and the normal distribution of the disturbance term cannot be rejected I would like to thank an anonymous referee for bringing to my attention this important character

$$(.90)\frac{\partial\mu}{\partial s}$$

We have thus established a time series model in which the aggregate variables such as schooling level and schooling dispersion could affect income inequality through a set of quantile income levels. This model allows one to see where the strongest impacts fall in the income distribution and to uncover the underlying factors affecting the basic structure of income concentration.

Estimation of the Model

Because many macroeconomic time-series variables contain unit roots, we use the augmented Dickey-Fuller (ADF) test as recommended by Schwert (1989) to avoid spurious results. The ADF test shows that dependent variables y_i and S are integrated of order one I(1), whereas VS and PI are integrated as I(0) Therefore, we use the first difference for dependent variables y_i and S in our model. The model, consisting of ten seemingly unrelated equations (7) with the adding-up constraints equation (8) and Tai-

Table 1—Regression Estimate of Schooling, Schooling Dispersion and Non-labor Income from 1966-1995

	Constant	Δ Schooling	Schooling Dispersion	Non-labor Income
Δy1	04489 (2)	19436 (2 7)	- 00046 (- 05)	00024 (3)
Δy2	08708 (4)	15691 (1 9)	- 00188 (- 17)	- 00040 (- 5)
Δy3	11647 (6)	15328 (2 0)	- 00308 (- 31)	- 00082 (-1 0)
Δy4	12699 (6)	17540 (2 4)	- 00374 (- 38)	- 00076 (-1 0)
Δy5	12979 (6)	15331 (2 1)	- 00373 (- 38)	- 00062 (- 8)
Δy6	34048 (1 7)	14246 (1 9)	- 01335 (-1 4)	- 00106 (-1 3)
Δy7	41434 (1 9)	18733 (2 4)	- 01713 (-1 6)	- 00108 (-1 3)
Δy8	35899 (1 8)	15278 (2 1)	- 01418 (-1 5)	- 00122 (-1 6)
Δy9	30442 (1 5)	12626 (1 7)	- 01170 (-1 2)	- 00076 (-1 0)
Δμ	21318 (1 1)	16023 (2 3)	- 00769 (- 85)	- 00072 (-1 0)

The estimated equations are

Δ y(i) = Cons+ Δ Schooling + Schooling dispersion + Non-labor income, i=1. ,9, μ
y(i) = The decile income
Δ y(i) = Change in the decile income
S = Average years of schooling per employed unit of labor
Δ Schooling = Change in average years of schooling per employed unit of labor
Schooling Dispersion = Variance of schooling years of the employed
Non-labor Income = Average real income from property and entrepreneurship per employed unit of labor The system R-square - 6764

Data sources The decile income data y(i) can be found in the *Report on the Survey of Family Income and Expenditures* (Table 6, p 23), Taiwan, R O C , 1996. the Department of Budget, Accounting and Statistics The schooling level per employed person and schooling dispersion of the employed are calculated from data of the *Yearbook of Manpower Survey Statistics, Taiwan Area* (Table 11, pp 34-37), the Department of Budget, Accounting and Statistics, R O C , 1996 The data for income from property and entrepreneurship can be found in the *Report of National Income* (Table 3, pp 188-192) the Department of Budget, Accounting and Statistics, R O C , 1996 wan data from 1966 to 1995, has been estimated for the real income distribution of households and is presented in Table 1 with t-ratios in parentheses [3] In order to test whether there is a difference between the unrestricted and restricted equations, the likelihood ratio test is used. The hypotheses are.

- $H_0 : \beta_u = \beta_r$, where u = unrestricted, r = restricted.
- $H_1 : \beta_u \neq \beta_r$.

The likelihood ratio test is $\phi = \sigma_u^2 / \sigma_r^2$, -2ln* = -2[lnDETu(S-matrix) - lnDETr(S-matrix)] = 32.72, and pr($\chi^2_{[illegible]} < c$) ≈ 26.51. At a significance level of .95, we would reject the null hypothesis

The regression results obtained will now be used to analyze how changes in schooling level and schooling dispersion affect individual income deciles The partial elasticities (evaluated at the mean), which are computed from Table 1, are presented in Table 2. As Table 2 shows, the schooling level elasticity is positively inelastic at the top decile income and positively elastic at the remaining decile income levels All the schooling dispersion elasticities are negatively inelastic. Non-labor income elasticity is positively inelastic at the first decile income and negatively inelastic at the remaining decile income levels.

Table 2—Estimated Decile Income Elasticities of Schooling, Schooling Dispersion and Non-labor Income from 1966-1995

	y1	y2	y3	y4	y5	y6	y7	y8	y9	μ
S	1 4689	1 1859	1 1585	1 3257	1 1587	1 0767	1 4158	1 1547	9543	1 2110
VS	- 0096	- 0390	- 0638	- 0776	- 0773	- 2766	- 3549	- 2940	- 2425	- 1595
NY	0027	- 0046	- 0095	- 0088	- 0072	- 0122	- 0125	- 0141	- 0088	- 0083

S = Schooling years,
VS = Schooling dispersion,
NY = Non-labor income

The primary concern of this paper has been to analyze the schooling impact on income concentration. The three most basic inequality measures–relative median income, relative mean income, and income-shared behavior–are examined. The relative median and mean income figures, which are computed from Table 2, $y_i/y5$ and y_i/μ, are presented in Table 3 and Table 4, respectively.

Table 3—Relative Median Income Elasticities of Schooling, Schooling Dispersion and Non-labor Income from 1966-1995

	y1/y5	y2/y5	y3/y5	y4/y5	y5/y5	y6/y5	y7/y5	y8/y5	y9/y5
S	3103	0272	- 0002	1670	0	- 0820	2571	- 0040	- 2045
VS	0677	0383	0135	- 0003	0	- 1993	- 2776	- 2167	- 1652
NY	0099	0025	- 0024	- 0017	0	- 0050	- 0054	- 0070	- 0017

S = Schooling years,
VS = Schooling dispersion.
NY = Non-labor income

[3] Individual t tests for each parameter are printed, but for nonlinear models one should be cautious in drawing any inferences from t tests

Table 4— Relative Mean Income Elasticities of Schooling, Schooling Dispersion and Non-labor Income from 1966-1995

	$y1/\mu$	$y2/\mu$	$y3/\mu$	$y4/\mu$	$y5/\mu$	$y6/\mu$	$y7/\mu$	$y8/\mu$	$y9/\mu$
S	2580	-0251	-0525	1147	-0523	-1343	2048	-0563	-2568
VS	1499	1205	0957	0819	0822	-1171	-1955	-1345	-0831
NY	0111	0037	-0012	-0005	0012	-0039	-0042	-0058	-0005

S = Schooling years,
VS = Schooling dispersion,
NY = Non-labor income

The elasticity of a ratio is equal to the difference in the elasticities. Positive values below the median (mean) income represent a reduction of income inequality, and positive values above the median (mean) represent an increase in income inequality. Because the positive values below the median (mean) income represent the ratio yi/y5 (yi / μ) toward unity and the positive values above median (mean), the upper income levels will move further away from the median (mean) of the distribution.

As Table 3 shows, it is the income groups, y1, y2, y4, and y7 that gain as schooling level increases. With an increase in schooling dispersion, the relative positions of the bottom three have improved, while the remaining worsen. Looking at Table 4, the increase in schooling level tends to improve the relative positions of y1, y4, and y7 Under an increase in schooling dispersion, the top four deciles worsen.

Another basic inequality measure is a set of income share statistics corresponding to the different ten-percentage groups in the population. The income shares can be derived in terms of the elasticities of the relative mean income from Table 4. Formulas for those estimates are detailed in Beach (1977).[4] Table 5, which is computed from Table 4, shows that the relative positions of y1, y2, y4, y5, y7, and y8 have improved as schooling level increases. When schooling dispersion increases, the top four income shares shrink and the remaining shares expand.

[4] If R(i) represents the i[th] relative mean income figure, $R(i) = y(i)/\mu$, then the elasticity is $\eta_{R(i),x} = \eta_{y(i),x} - \eta_{\mu,x}$, where x stands for schooling level, schooling dispersion and non-labor income The income share elasticity therefore can be shown as

$$\eta_{I(i),x} = \frac{y(i-1)}{y(i-1)+y(i)}\eta_{R(i-1),x} + \frac{y(i)}{y(i-1)-y(i)}\eta_{R(i),x}$$

The results in Table 5 have been computed from the equation above The Gini inequality estimate can be computed as

$$\eta_{g,x} = -.2\sum_{i=1}^{9}\frac{\partial B(i)}{\partial x}\frac{x}{g}; B(i) = \sum_{j}^{i} I(i)$$

from which the figures in Table 6 have been calculated

Table 5— Income Share Elasticities of Schooling, Schooling Dispersion and Non-labor Income from 1966-1995

	Lowest 10% Share	Second 10% Share	Third 10% Share	Fourth 10% Share	Fifth 10% Share	Sixth 10% Share	Seventh 10% Share	Eighth 10% Share	Ninth 10% Share
S	2580	1145	- 0393	0349	0273	- 0953	0446	0656	- 1659
VS	1499	1350	1076	0884	0820	- 0224	- 1584	- 1630	- 1064
NY	0111	0073	0012	- 0008	0004	- 0015	- 0040	- 0050	- 0029

S = Schooling years,
VS = Schooling dispersion.
NY = Non-labor income

Given the above disaggregate results, it is not necessary to run equations to obtain aggregate results The aggregate inequality estimates are contained in Table 6. Table 6 shows the overall inequality that confirms the disaggregate results. The negative value indicates the aggregate income distribution measured as Gini becomes equal. For instance, we can see the elasticity of the schooling level to the Gini index is -.0789. This negative value implies that the increase in schooling level will tend to reduce the inequality

Table 6— Real Elasticities of Gini Coefficient

$$e_{g,s} = -.0789, \quad e_{g,vs} = -.0913, \quad e_{g,ny} = -.0029$$

G = Gini index,
S = Schooling years,
VS = Schooling dispersion,
NY = Non-labor income

Conclusion

The purpose of this paper has been to consider the impact of schooling level and schooling dispersion on income distribution for the period 1966-1995 in Taiwan. The results show that either increases in schooling level or schooling dispersion tend to improve the income distribution. Educational expansion can reduce income inequality in Taiwan. The policy to subsidize publicly supported colleges to buy educational equality, therefore, may be an appropriate one in Taiwan.

References

1 Basmann, R L, and D J Slottje, "A New Index of Income Inequality," *Economic Letters*, 24 (1988), pp 384-389

2 Beach, C M, "Cyclical Sensitivity of Aggregate Income Inequality," *Review of Economics and Statistics* (1977), pp 56-66

3 Blaug, M, C Dougherty, and G Psacharopoulos, "The Distribution of Schooling and the Distribution of Earnings Raising the School Leaving Age in 1972," *The Manchester School* (1982), pp 24-40

4 Caniglia, A S, "Inequality in Education Attainment and the Distribution of Income," *Review of Business*, 10, no 2 (1988), pp 5-10

5 Chiswick, Barry, *Income Inequality* (New York, National Bureau of Economic Research, 1974)

6 Fields, G S , "Education and Income Distribution in Developing Countries A Review of the Literature," in T King (ed), *Education and Income* (Washington D C World Bank Staff Working Paper No 402, 1980)

7 Kakwani, Nanak C , *Income Inequality and Poverty Methods of Estimation and Policy Applications* (Oxford University Press, Oxford) (1980)

8 Knight, J B , and R H Sabot, "Educational Expansion and the Kuznets Effect," *American Economic Review*, 73, no 5 (1983), pp 1132-36

9 Lam, David, and Deborah Levison, "Declining Inequality in Schooling in Brazil and its Effects on Inequality in Earnings," *Journal of Development Economics*, 37, nos 1 and 2 (1992), pp 199-225

10 Mohan, R and R Sabot, "Educational Expansion and the Inequality of Pay Colombia 1973-78," *Oxford Bulletin of Economics and Statistics*, 50, no 2 (1988), pp 175-182

11 Park, K H , "Expansion and Educational Inequality on Income Distribution," *Economics of Education Review*, 15, no 1 (1996), pp 51-58

12 Plotnick, R D , "Trends in Male Earnings Inequality," *Southern Economic Journal*, 48, no 3 (1982), pp 724-732

13. Psacharopoulus, George, and F Steier, "Education and the Labor Market in Venezuela, 1975-1984," *Economics of Education Review*, 7, no 3 (1988), pp 321-332

14 Ram, Rati, "Population Increase, Economic Growth, Educational Inequality, and Income Distribution Some Recent Evidence," *Journal of Development Economics*, 14 (1984), pp 419-428

15 Ram, Rati, "Can Educational Expansion Reduce Income Inequality in Less-Developed Countries?" *Economics of Education Review*, 8, No 2 (1989), pp 185-195

16 Schwert, G W , " Test for Unit Roots A Monte Carlo Investigation," *Journal of Business and Economic Statistics*, 7 (1989), pp 147-159

17 Tilak, Jandhyala B G , "Rate of Return to Education and Income Distribution," *The Economist*, 137, no 4 (1988), pp 454-465

18 Weisskoff, R , "Income Distribution and Economic Growth in Puerto Rico, Argentina, and Mexico," *Review of Income and Wealth* (1970), pp 303-332

19 Winegarden, Calvin R , "Schooling and Income Distribution Evidence from International Data," *Economica*, 46 (1979), pp 83-87

Appendix: Data Sources

Variable	Source
y(i) (Decile Income)	*Report on the Survey of Family Income and Expenditures, the Department of Budget, Accounting and Statistics, R.O C. 1996* (Table 6, Upper Limit of Each Tenth by Disposable Income Decile, p. 23.) These data were divided by the price index.
S and VS (Schooling Years and Schooling Dispersion)	Calculated from the data of the *Yearbook of Manpower Survey Statistics, Taiwan Area, the Department of Budget, Accounting and Statistics, R O C 1996* (Table 11, Educational Attainment of Employed Person, pp. 34-37).
NY (Non-labor Income)	*Report of National Income, the Department of Budget, Accounting and Statistics, R.O.C. 1996* (Table 3, Distribution of National Income, pp. 188-192) These data were divided by the price index.