World Bank Reprint Series: Number 289

J. B. Knight and R. H. Sabot

Educational Expansion and the Kuznets Effect

Reprinted with permission from *The American Economic Review*, vol. 73, no. 5 (December 1983). pp. 1132–36.

World Bank Reprints

No. 253. Larry E. Westphal, "Fostering Technological Mastery by Means of Selective Infant-Industry Protection," *Trade, Stability, Technology, and Equity in Latin America*

No. 254. Gershon Feder, "On Exports and Economic Growth," *Journal of Development Economics*

No. 255. Mohan Munasinghe, "Third World Energy Policies: Demand Management and Conservation," *Energy Policy*

No. 256. Keith Marsden and Alan Roe, "The Political Economy of Foreign Aid: A World Bank Perspective," *Labour and Society*

No. 257. James A. Hanson, "Contractionary Devaluation, Substitution in Production and Consumption, and the Role of the Labor Market," *Journal of International Economics*

No. 258. Christiaan Grootaert, "The Conceptual Basis of Measures of Household Welfare and Their Implied Survey Data Requirements," *The Review of Income and Wealth*

No. 259. Guy Pfeffermann and Richard Webb, "Poverty and Income Distribution in Brazil," *The Review of Income and Wealth*

No. 260. Pradeep K. Mitra, "A Theory of Interlinked Rural Transactions," *Journal of Public Economics*

No. 261. David L. Lindauer and Richard H. Sabot, "The Public/Private Wage Differential in a Poor Urban Economy," *Journal of Development Economics*

No. 262. J. B. Knight and R. H. Sabot, "Labor Market Discrimination in a Poor Urban Economy," *Journal of Development Studies*

No. 263. Carl Dahlman and Larry Westphal, "Technical Effort in Industrial Development: An Interpretative Survey of Recent Research," *The Economics of New Technology in Developing Countries*

No. 264. Michael Bamberger, "The Role of Self-Help Housing in Low-Cost Shelter Programs for the Third World," *Built Environment*

No. 265. Bela Balassa, "The Adjustment Experience of Developing Economies after 1973," *IMF Conditionality*

No. 266. Bela Balassa, "Outward Orientation and Exchange Rate Policy in Developing Countries: The Turkish Experience," *The Middle East Journal*

No. 267. Dipak Mazumdar, "Segmented Labor Markets in LDCs," *American Economic Review*

No. 268. Stephen P. Heyneman and William A. Loxley, "The Effect of Primary-School Quality on Academic Achievement across Twenty-nine High- and Low-Income Countries," *The American Journal of Sociology*

No. 269. James R. Follain, Jr., Gill-Chin Lim, and Bertrand Renaud, "Housing Crowding in Developing Countries and Willingness to Pay for Additional Space: The Case of Korea," *Journal of Development Economics*

No. 270. Bela Balassa, "Policy Responses to External Shocks in Sub-Saharan African Countries," *Journal of Policy Modeling*

No. 271. Jaime de Melo and Sherman Robinson, "Trade Adjustment Policies and Income Distribution in Three Archetype Developing Economies," *Journal of Development Economics*

No. 272. J. B. Knight and R. H. Sabot, "The Role of the Firm in Wage Determination: An African Case Study," *Oxford Economic Papers*

No. 273. William G. Tyler, "The Anti-Export Bias in Commercial Policies and Export Performance: Some Evidence from Recent Brazilian Experience," *Weltwirtschaftliches Archiv*

Educational Expansion and the Kuznets Effect

By J. B. KNIGHT AND R. H. SABOT*

There is a widely accepted view that in the process of economic development, the inequality of income first increases and later decreases (see Montek Ahluwalia, 1976, and Ahluwalia et al., 1979). It is also recognized that the inequality of pay is an important component of total income inequality (see Alan Blinder, 1974; Henry Phelps Brown, 1977), and that education is a major determinant of pay. Whether the rapid expansion of education which has occurred in many countries has increased or decreased the inequality of pay is therefore of interest and of relevance to policy.

Our objective in this note is to show the change in wage dispersion in response to increases in the relative supply of educated workers in two low-income countries. We also measure the relative contributions to that change of its two components: the effect of the educational expansion on the educational composition of the labor force (holding the educational structure of wages constant), and the resultant compression of that structure (holding composition constant). We base our analysis on three precisely comparable surveys of wage employees conducted in Tanzania in 1971 and 1980, and in Kenya in 1980.[1]

The change in the educational composition of the labor force itself has an effect on inequality. Whether it raises or lowers inequality, *ceteris paribus*, depends on the relative sizes of the different educational categories, their relative mean wages, and their relative wage dispersions. In a two-group model, a transfer of workers from the low- to the high-education (and wage) group can be shown to raise the variance (or log variance) until the high-education group reaches a certain proportion of the total, the precise proportion depending on the difference in the means and the variances of the two groups. Sherman Robinson (p. 438) has shown for workers in the jth group, that where $\bar{x}_j$ is mean earnings, p_j is proportion of workers, σ_j is the variance of earnings, $j = 1, 2$ = educated and uneducated group, respectively, and σ is the variance of earnings in the population, then inequality (as measured by σ) rises to a maximum where p_1 is equal to p_1^*:

$$(1) \quad p_1^* = \left(\sigma_1^2 - \sigma_2^2\right)/2(\bar{x}_1 - \bar{x}_2) + 1/2.$$

This implies that $p^* \gtrless \frac{1}{2}$ according as $\sigma_1^2 \gtrless \sigma_2^2$. If the variance of wages of the educated group exceeds that of the uneducated ($\sigma_1^2 > \sigma_2^2$), inequality reaches a peak after more than one-half the labor force have become educated. The condition $\sigma_1^2 > \sigma_2^2$ also implies that p_1^* is larger the smaller the difference in mean wages, that is, inequality peaks later in the process of educational expansion the lower is the premium on education. However, if the variance is higher for the uneducated group, this result is reversed.

This was the basis of Simon Kuznets' hypothesis: the transfer of people between sectors at different income levels initially raises inequality as more people acquire high income, but eventually lowers it as fewer low-income people remain; if the expanding sector has more inequality, the peaking of aggregate inequality is delayed.

Kuznets concentrated on the "composition effect" and did not incorporate any resulting "compression effect." Indeed, he suggested (p. 8) that urban-rural income differentials (with which he was concerned) were likely to increase with economic development. However, from the competitive market prediction

*Institute of Economics and Statistics, University of Oxford, Manor Road, Oxford, England, and Development Research Department, World Bank, 1818 H Street, NW, Washington, D.C. 20433.

[1]Each survey was an establishment-based stratified random sample survey of manufacturing employees in the capital city. The 1980 surveys were carried out by a team including the authors as part of a World Bank project on the economic consequences of educational expansion; the 1971 survey was carried out by one of the authors

that the returns to a factor decrease as its relative supply increases, we expect the coefficient on education in an earnings function to decline as education expands relative to other factors, that is, the premium on education falls as supply increases relative to demand. The narrowing of the educational structure of wages should, *ceteris paribus*, reduce inequality. This result can be formalized by differentiation of the well-known identity

$$(2) \qquad \sigma^2 = p_1\sigma_1^2 + (1-p_1)\sigma_2^2 + p_1(1-p_1)(\bar{x}_1 - \bar{x}_2)^2$$

to derive:

$$(3) \quad \partial\sigma^2/\partial p_1 = (\sigma_1^2 - \sigma_2^2) + (1-2p_1) \times (\bar{x}_1 - \bar{x}_2)^2 + 2p_1(1-p_1)(\bar{x}_1 - \bar{x}_2) \times (\partial\bar{x}_1/\partial p_1 - \partial\bar{x}_2/\partial p_1).$$

Since $\bar{x}_1 > \bar{x}_2$, $\partial\bar{x}_1/\partial p_1 < 0$ and $\partial\bar{x}_2/\partial p_1 > 0$, the third term, which shows the effect of educational expansion on the wage structure, is negative. The effect of educational expansion on inequality is therefore the net outcome of two potentially countervailing tendencies, and no a priori prediction can be made about its sign.

Owing to the expansion of educational enrollments over time in Tanzania, and the faster rate of expansion at the secondary school level in Kenya, our three samples show sharp differences in the educational attainment of the manufacturing sector labor force (Table 1).[2] The differences are most marked at the secondary forms 1–4 level. The occupational structure is very similar in the three samples. Given our skill-based occupational classification, this suggests that the composition of demand for skills is also similar. The resultant differences in the education-occupation matrix are therefore likely to be due to differences in the relative supply of educated labor, which is lowest in Tanzania 1971 and highest in Kenya 1980.[3]

TABLE 1—MANUFACTURING SECTOR: THE DISTRIBUTION OF EMPLOYEES BY LEVEL OF EDUCATIONAL ATTAINMENT IN SIMPLE EARNINGS FUNCTIONS

	Tanzania 1971	Tanzania 1980	Kenya 1980
No Education	24.4	11.9	5.1
	(−0.236)	(−0.254)	(−0.202)
Primary Standards			
1–4	27.6	18.6	12.8
	(−0.150)	(−0.124)	(−0.176)
5–8	36.2	48.7	46.3
Secondary Forms 1–4	10.2	16.6	33.3
	(0.748)	(0.502)	(0.256)
Post-Form 4	1.6	4.2	2.6
	(1.199)	(0.968)	(0.932)

Notes: The distribution is the percentage of the column total. The coefficients on the education dummy variables are show in parentheses. The coefficients are in relation to the base subcategory primary standards 5–8, with log earnings as the dependent variable. All the coefficients are significantly different from zero at the 5 percent level.

An earnings function which includes the educational categories among the independent variables[4] and log earnings as the dependent variable, shows high and significant (gross) returns to education in all three samples (also Table 1). The coefficients for post-primary education (with primary standards 5–8 as the base subcategory) are higher in Tanzania 1971 than in Tanzania 1980, and higher in Tanzania 1980 than in Kenya 1980. Again, the differences are most marked at the secondary forms 1–4 level. This compression of the educational structure of wages with educational expansion is consistent with the competitive market prediction.

To measure the effect of educational expansion on the inequality of pay, and the relative contributions of the compression effect and the composition effect, we proceed

[2]As Kenya and Tanzania inherited a common educational system, the same educational categories have the same meaning and involve the same number of years of schooling in the two countries.

[3]The samples might be thought of as three points in a time-series, with Kenya 1980 representing Tanzania some years hence.

[4]The other independent variables are years of wage employment experience and its square, and dummy variables representing race, sex, employment status, and possession of formal training.

as follows.[5] First, we estimate an earnings function

$$(4) \qquad w = a + \sum_i b_i e_i + \sum_j c_j z_j + u,$$

where w is log earnings, e_i is a set of dummy variables representing different educational categories, z_j is a set of independent variables, u is an error term, and a, b_i, c_j are parameters.

The estimated earnings function is used to predict the wage of each individual worker ($\hat{w}$) from his set of characteristics; the inequality of predicted earnings is then measured. Second, we simulate the compression effect of increasing the relative supply of educated labor. The wages of workers in each sample are predicted using the education coefficients estimated for the other samples instead of the actual coefficients. For instance, where the subscript k denotes the Kenya 1980 and t the Tanzania 1980 sample, we predict the wages of the Kenya sample using the Tanzania education coefficients:

$$(5) \qquad \hat{w}_k = a_k + \sum_i b_{it} e_{ik} + \sum_j c_{jk} z_{jk}.$$

Third, we simulate the composition effect of increasing relative supply. Instead of the actual educational composition of each sample, we assume in turn the composition of the other samples. This involves reweighting the observations. For example, where α_{ik} represents the proportion of workers in each educational category e_i in the Kenya sample, and α_{it} the proportion in the Tanzania 1980 sample, in applying to the Kenya sample the Tanzania 1980 composition, we use the set of weights α_{it}/α_{ik}.[6] In this case the simulation involves a decrease in the proportion of workers with secondary education and an increase in the proportion with primary education. The former observations are therefore given a weight less than unity, and the latter a weight greater, in generating the assumed distribution. Fourth, we combine the composition effect and its consequent compression effect in a single simulation. The inequality of the simulated wage distribution is in each case measured.

The results of the first exercise, based on equation (4), indicate that the inequality of predicted wages is greatest in Tanzania 1971 and least in Kenya 1980: the log variances are .230, .128, and .097. Inequality is less the greater the relative supply of educated labor. But how much of these differences is due to the compression effect, to the composition effect, and/or to differences in the samples which are unconnected with education?

The results of the second exercise are shown in Figure 1. On the vertical axis is the log variance, and the horizontal axis shows which set of education coefficients (panel a), which educational composition (panel b), and which set of education coefficients and educational composition (panel c) is being assumed. The unsimulated but predicted log variances for the three samples are circled in each panel. A line connecting the points in panel a indicates the effect on the log variance for one sample of altering the education coefficients to equal those estimated for the other samples. The line marked T71, for instance, indicates the log variance when the education coefficients of the Tanzania 1971, Tanzania 1980, and Kenya 1980 samples are in turn applied to the Tanzania 1971 function. As expected, each of the lines is downward sloping: the decline in the premium on education, considered in isolation, consistently reduces inequality.

The effect of altering the educational composition of a sample while assuming that the coefficients on education remain unchanged is shown by the connecting lines in panel b. The line labeled T71, for instance, shows how inequality varies as the educational compositions of the Tanzania 1980 and Kenya 1980 samples are applied to the Tanzania 1971 earnings function and sample. In all but one of the six cases, inequality rises with the simulated expansion of education.

Panel c shows the effect on the inequality of wages in each sample of simulating both the educational composition of another sample and its accompanying education coefficients. The vertical distance between points therefore represents the differences in inequality among the samples which are attrib-

[5] The method is an extension of that used by Blinder to decompose the inequality of an income distribution; it is discussed in our paper with Jere Behrman (1983).

[6] The weighting procedure implies that the total number of observations is left unchanged.

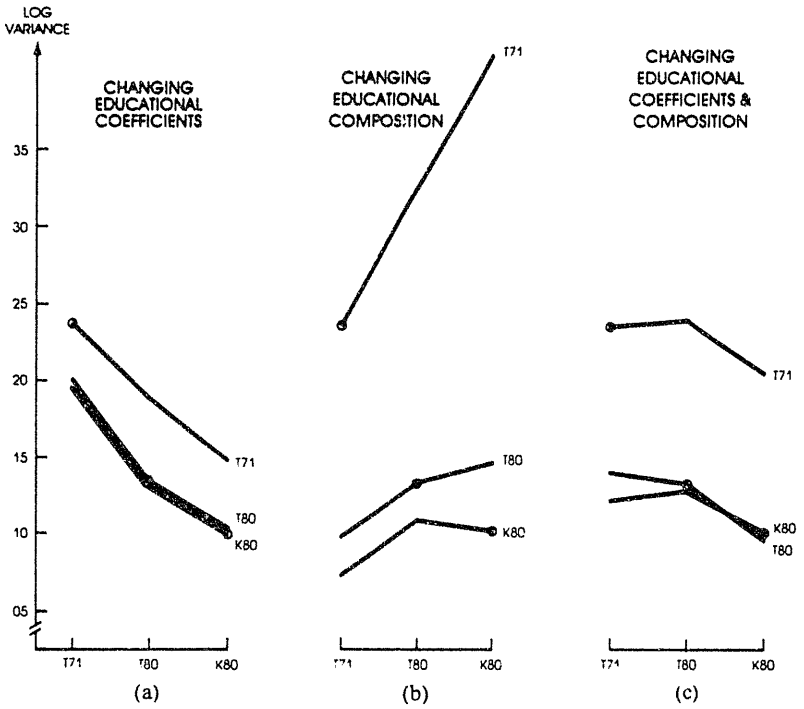

FIGURE 1

utable to differences in the noneducational characteristics and coefficients of the samples. As these differences are irrelevant to the analysis, it is the shape of each line which is important. In their affect on inequality, educational expansion and the consequent compression of the educational wage structure in Tanzania cancel each other out: the log variance rises very slightly in two cases and falls very slightly in the other. However, the move from the Tanzania 1980 to the Kenya 1980 educational composition and coefficients reduces inequality in every case.[7]

There remains a puzzle. We argued above for the two-group case that, if the variance of wages of the educated exceeds that of the uneducated, the variance of the sample peaks when the proportion of workers in the educated group exceeds one-half; and that this proportion is greater the smaller the difference in group mean wages. Panel b, however, suggests that the reverse is the case. Comparing the educational composition of Tanzania and Kenya in 1980, inequality actually falls when the simulation is conducted on the Kenya earnings function, that is, the one for which educational wage differences are smallest. However, there is no necessary inconsistency here. First, it may not be the case that the inequality of earnings of the more educated is greater than that of the less educated. Second, the empirical analysis involves five educational categories, whereas the theoretical prediction is based on two.

In fact, there is a monotonic increasing relationship between the log variance of earnings and educational level in both Kenya and Tanzania 1980, and when each sample is divided into two educational categories instead of five, again the log variance of earnings of the "educated" (postprimary schooling) exceeds that of the "uneducated" (primary schooling or none) in each case. For the Kenya and Tanzania 1980 two-group earnings functions, Table 2 reproduces the simulation analysis conducted in panel b, viz, it indicates the effect on inequality of varying educational composition so as to equal that in the other two samples. For both functions, an increase in the proportion of

[7]Results precisely corresponding to those of Figure 1 were obtained with the Gini coefficient as the measure of earnings inequality in place of the log variance.

TABLE 2—THE COMPOSITION EFFECT IN THE TWO-GROUP CASE: PARAMETERS FOR THE ESTIMATION OF PEAK INEQUALITY, AND SIMULATIONS OF INEQUALITY

	Tanzania 1970	Tanzania 1980	Kenya 1980
σ_1^2	.467	.367	.363
σ_2^2	.187	.154	.240
$\bar{x}_1$	6.709	7.043	6.852
$\bar{x}_2$	5.570	6.389	6.597
p_1	.120	.201	.359
p_1^*	.608	.748	1.446
$\delta\sigma^2/\delta p_1$	1.266	.469	.141
σ_{K80}^2	.064	.074	.082
σ_{T80}^2	.084	.110	.141

Notes: p_1^* is obtained from equation (1), and $\delta\sigma^2/\delta p_1$ from equation (3) (in the absence of compression effects). The inequality of predicted wages, used in the simulations (σ_{K80}^2 and σ_{T80}^2), is much less than actual inequality because the predictive equation distinguishes only two educational categories.

educated workers in the total (line p_1 in the table) raises the log variance of predicted wages (lines σ_{K80}^2 and σ_{T80}^2, where the subscript indicates the earnings function being used). It is no longer the case that inequality declines when the educational composition is changed from that of Tanzania 1980 to that of Kenya 1980 using the Kenya earnings function. These results are to be expected because the actual value of p_1 in the two-group case falls well short of the proportion yielding peak inequality in each sample (p_1^* in Table 2). Indeed, in the case of the Kenya earnings function, the difference in log means is so small in relation to the difference in log variances that the peak is never reached ($p_1^* > 1$).[8] The implications of this exercise are two-fold. First, our evidence and method do not contradict the predictions of the two-group model. Second, the predictions of the two-group model do not necessarily carry over into the more realistic multigroup case.

The Kuznets effect has normally been considered in relation to the transfer of population from low-income rural to high-income urban areas. Such a transfer is not expected to produce a narrowing of the income gap, either because of labor market imperfections or because the accumulation of human and nonhuman capital in the urban areas tends to widen the gap. We have shown, however, that the expansion of the supply of educated labor relative to the demand has a powerful compressing effect on the intraurban educational structure of wages. The composition effect of educational expansion can indeed raise intraurban inequality, but the consequent compression effect outweighs it: relative educational expansion reduces inequality. Since this process occurs within the relatively expanding high-income, urban sector, it is hastening the arrival at the point beyond which economic growth is associated with a reduction in overall inequality.

[8] The flatter slope of the curve based on the Kenya sample ($\delta\sigma^2/\sigma p_1$ in Table 2) does not, therefore, indicate that the peak is nearby.

REFERENCES

Ahluwalia, Montek S., "Inequality, Poverty and Development," *Journal of Development Economics*, December 1976, *3*, 307–42.

_____, **Carter, Nicholas G. and Chenery, Hollis B.,** "Growth and Poverty in Developing Countries," *Journal of Development Economics*, September 1979, *6*, 299–341.

Behrman, Jere R., Knight, John B. and Sabot, Richard H., "A Simulation Alternative to the Comparative R^2 Approach to Decomposing Inequality," *Oxford Bulletin of Economics and Statistics*, August 1983, *45*, 307–12.

Blinder, Alan S., *Toward an Economic Theory of Income Distribution*, Cambridge: MIT Press, 1974.

Kuznets, Simon, "Economic Growth and Income Inequality," *American Economic Review*, March 1955, *45*, 1–28.

Phelps Brown, Henry, *The Inequality of Pay*, Oxford: Oxford University Press, 1977.

Robinson, Sherman, "A Note on the U Hypothesis Relating Income Inequality and Economic Development," *American Economic Review*, June 1976, *66*, 437–40.

World Bank, "Kenya and Tanzania Surveys of Wage, Employment and Education," RPO 672-01, Washington, 1980.

No. 274. Ron Duncan and Ernst Lutz, "Penetration of Industrial Country Markets by Agricultural Products from Developing Countries," *World Development*

No. 275. Malcolm D. Bale, "Food Prospects in the Developing Countries: A Qualified Optimistic View," *The American Economic Review* (with Ronald C. Duncan) and "World Agricultural Trade and Food Security: Emerging Patterns and Policy Directions," *Wisconsin International Law Journal* (with V. Roy Southworth)

No. 276. Sweder van Wijnbergen, "Interest Rate Management in LDCs," *Journal of Monetary Economics*

No. 277. Oli Havrylyshyn and Iradj Alikhani, "Is There Cause for Export Optmism? An Inquiry into the Existence of a Second Generation of Successful Exporters," *Weltwirtschaftliches Archiv*

No. 278. Oli Havrylyshyn and Martin Wolf, "Recent Trends in Trade among Developing Countries," *European Economic Review*

No. 279. Nancy Birdsall, "Fertility and Economic Change in Eighteenth and Nineteenth Century Europe: A Comment," *Population and Development Review*

No. 280. Walter Schaefer-Kehnert and John D. Von Pischke, "Agricultural Credit Policy in Developing Countries," translated from *Handbuch der Landwirtschaft und Ernährung in den Entwicklungsländern* (includes original German text)

No. 281. Bela Balassa, "Trade Policy in Mexico," *World Development*

No. 281a. Bela Balassa, "La política de comercio exterior de México," *Comercio Exterior*

No. 282. Clive Bell and Shantayanan Devarajan, "Shadow Prices for Project Evaluation under Alternative Macroeconomic Specifications," *The Quarterly Journal of Economics*

No. 283. Anne O. Krueger, "Trade Policies in Developing Countries," *Handbook of International Economics*

No. 284. Anne O. Krueger and Baran Tuncer, "An Empirical Test of the Infant Industry Argument," *American Economic Review*

No. 285. Bela Balassa, "Economic Policies in Portugal," *Economia*

No. 286. F. Bourguignon, G. Michel, and D. Miquel, "Short-run Rigidities and Long-run Adjustments in a Computable General Equilibrium Model of Income Distribution and Development," *Journal of Development Economics*

No. 287. Michael A. Cohen, "The Challenge of Replicability: Toward a New Paradigm for Urban Shelter in Developing Countries," *Regional Development Dialogue*

No. 288. Hollis B. Chenery, "Interaction between Theory and Observation in Development," *World Development*

THE WORLD BANK

Headquarters
1818 H Street, N.W.
Washington, D.C. 20433, U.S.A.

Telephone: (202) 477-1234
Telex: WUI 64145 WORLDBANK
RCA 248423 WORLDBK
Cable address: INTBAFRAD
WASHINGTONDC

European Office:
66, avenue d'Iéna
75116 Paris, France

Telephone: (1) 723-54.21
Telex: 842-620628

Tokyo Office:
Kokusai Building
1-1, Marunouchi 3-chome
Chiyoda-ku, Tokyo 100, Japan

Telephone: (03) 214-5001
Telex: 781-26838

The full range of World Bank publications, both free and for sale, is described in the *World Bank Catalog of Publications*, and of the continuing research program of the World Bank, in *World Bank Research Program: Abstracts of Current Studies*. The most recent edition of each is available without charge from:

PUBLICATIONS DISTRIBUTION UNIT
THE WORLD BANK
1818 H STREET, N.W.
WASHINGTON, D.C. 20433
U.S.A.

ISSN 0253-2131