

NBER WORKING PAPER SERIES

EDUCATIONAL INEQUALITY

Yoshiaki Azuma
Herschel I. Grossman

Working Paper 8206
<http://www.nber.org/papers/w8206>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2001

We have received helpful comments from Oded Galor, Dmitriy Gershenson, Arne Kalleberg, Frank Levy, Juan Mendoza, Omer Moav, and participants in a seminar at Osaka Prefecture University and in the Brown Macro Lunch. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

© 2001 by Yoshiaki Azuma and Herschel I. Grossman. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Educational Inequality
Yoshiaki Azuma and Herschel I. Grossman
NBER Working Paper No. 8206
April 2001
JEL No. D3, J3, O3

ABSTRACT

This paper develops a theoretical model that relates changes in educational inequality to the combined effects of innovations that have increased the relative demand for more educated labor and innovations that have increased ability premiums. Under the assumption that in the long run individual decisions to become more educated equalize the lifetime earnings of more educated workers and comparable less educated workers, our model yields two novel implications: First, given the existence of ability premiums, an innovation in the relative demand for more educated labor increases educational inequality in the short run, but, *ceteris paribus*, would decrease educational inequality in the long run. Second, in the long run innovations that increase ability premiums cause educational inequality to be larger than otherwise. In applying our theory to recent changes in educational inequality in the United States, we suggest that increases in ability premiums are dampening the long-run response of the relative supply of more educated workers that otherwise would reverse previous increases in educational inequality.

Yoshiaki Azuma
Brown University

Herschel I. Grossman
Brown University,
Russell Sage Foundation
and NBER

This paper develops a theoretical model that relates changes in educational inequality to the combined effects of innovations that have increased the relative demand for more educated labor and innovations that have increased ability premiums. We define educational inequality to be the ratio of the average wage or salary of workers with more years of education to the average wage or salary of workers with fewer years of education.¹ Educational inequality comprises the relative earnings of efficiency units of more and less educated labor, which we assume are complements in production, and the average abilities of more and less educated workers, measured as the average number of efficiency units of more or less educated labor that a worker supplies. We define ability premiums to be differences between wages and salaries received by workers with the same years of education who are more or less able.

Under the assumption that in the long run individual decisions to become more educated equalize the lifetime earnings of more educated workers and comparable less educated workers, our model yields two novel implications: First, given the existence of ability premiums, an innovation in the relative demand for more educated labor increases educational inequality in the short run, but, *ceteris paribus*, would induce a large enough increase in the supply of more educated labor to decrease educational inequality in the long run. Second, in the long run innovations that increase ability premiums, by causing the supply of more educated workers to be smaller than otherwise, causes educational inequality to be larger than otherwise.

Looking at workers who have and have not attended college, Claudia Goldin and Lawrence Katz (2001) found that in the United States educational inequality exhibited a generally U-

¹Some authors refer to educational inequality as “the return to education” or “the education premium”. As defined by Peter Gottschalk (1997, page 29), “The college premium is captured by the coefficient on the dummy college variable in a standard regression explaining (the log of) weekly earnings; essentially, that coefficient shows how much more a college graduate earns than does a high school graduate holding other factors, such as experience, constant.”

shaped pattern over the twentieth century. Educational inequality apparently decreased from 1900 until 1950, but then increased from 1950 to the end of the century. Strikingly, Goldin and Katz estimate that “the relative earnings of the more-educated [in 1999] are similar to that which prevailed in the early twentieth century.”

This U-shaped pattern, however, was not smooth. Goldin and Katz found that educational inequality decreased in “two giant steps”, one in the years before and after 1920 and the other during the 1940s. David Autor, Katz, and Alan Krueger (1998) found that educational inequality increased from 1950 to 1970, but then decreased during the 1970s, as a prelude to a large and widely discussed increase during the 1980s. Looking at more recent data for workers who are college and high school graduates, Lawrence Mishel, Jared Bernstein, and John Schmitt (2001) find that educational inequality increased further over the 1990s, although at a slower rate than in the 1980s.

Ability premiums result from the advantage in production associated with being more able, combined with factors that determine differences in ability, such as the innate characteristics of individual workers and the quality of the educations that workers receive. Analyzing American data from 1963 to 1989, Chinhui Juhn, Kevin Murphy, and Brooks Pierce (1993) found that the residual inequality in wages or salaries that remains after accounting for years of education and experience increased steadily from the late 1960s through the 1980s. According to Mishel, Bernstein and Schmitt (2001) more recent data indicate that residual inequality peaked about 1995 and since then has declined. We can presume that these changes in residual inequality reflect, either wholly or in part, changes in ability premiums.² In addition, Richard Murnane, John Willett, and Frank Levy (1995) report that the dependence of wages or salaries on a measure of mathematical skill, which we can take to be a

²Some authors refer to residual inequality as “within-group inequality”. For simplicity, our theoretical analysis abstracts from inequality associated with experience. Juhn, Murphy, and Pierce (1993) report that their data also show an increase in the relative return to experience. Robert Topel (1997) indicates that this increase occurred steadily from the late 1960s until the early 1980s.

proxy for ability, increased between the 1970s and the 1980s especially for college-educated workers and also for less educated workers.³

Our framework for analyzing changes in educational inequality allows for innovations in three structural factors: One factor is the relative demand for more educated labor. In modeling relative demand we focus on technological innovations that increase the relative importance of more educated labor in production. We could, however, readily generalize the analysis to take into account the effects of capital accumulation, working through complementarity in production between physical capital and human capital, as well as innovations that increase the relative demand for education-intensive products, such as might result from the removal of barriers to international trade. These factors would be isomorphic to technological innovations in increasing the relative demand for more educated labor.⁴

The other structural factors that drive our analysis are advantages in production associated with differences in ability and differences in the quality of education that underlie differences in ability. Positive innovations in either of these factors directly cause increases in ability premiums, and, as we shall see, in the long run also have an indirect effect on educational inequality.

Innovations in any of the three structural factors that drive our analysis can be a concomitant of technological progress and economic growth, but are conceptually distinct from

³Eric Gould, Omer Moav, and Bruce Weinberg (2000) point out that since the late 1960s inequality among less educated workers has increased as much as inequality among more educated workers. This observation suggests that an increased ability premium does not account completely for increased inequality among less educated workers. Appealing to the distinction introduced by Steven Davis and John Haltiwanger (1991) between inequality within plants and inequality among plants, Gould, Moav, and Weinberg associate increased inequality among less educated workers with “increasing variance of technological implementation across industries”.

⁴Robert Baldwin and Glen Cain(1997) claim that technological innovations have been the main force in increasing educational equality, but James Harrigan and Rita Balaban (1999) report that both capital accumulation and a decrease in the price of traded goods also have contributed.

technological progress and economic growth. For example, in a standard production function exhibiting constant returns to scale an increase in the relative importance of more educated labor in production could be associated either with an increase or with a decrease in total factor productivity. As another example, increases in differences in the quality of education could result either from improvements in higher quality education or, as some people think has been happening, deterioration in lower quality education. Also, although some authors argue that the relative importance of more educated labor and/or the advantage associated with high ability increase with the rate of technological progress,⁵ we think that distinguishing the effects of the innovations that drive our analysis from the effects of technological progress provides a useful clarification.

Our analysis shows that the determination of educational inequality is complex. In the short run, in which the number of more educated workers is given, an increase in the relative demand for more educated labor causes an increase in educational inequality. In the long run, however, this short-run increase in educational inequality causes a larger fraction of workers to choose to become more educated. Specifically, our analysis implies that, given that ability premiums exist, an increase in the relative demand for more educated labor would induce a long-run increase in the relative supply of more educated labor that by itself would be large enough more than to reverse the short-run increase in educational inequality.

But, our analysis also implies that educational inequality is positively related to ability premiums. Most interestingly, in our model the larger are ability premiums the smaller is the fraction of workers that must choose to become more educated in order to equalize the lifetime earnings of more educated workers and comparable less educated workers. Consequently, in the long run innovations that cause increases in ability premiums also cause the relative supply of more educated labor to be smaller than otherwise. As a result increases in ability

⁵For various versions of this story, see, for example, Oded Galor and Daniel Tsiddon (1997), Francesco Caselli (1999), Huw Lloyd-Ellis (1999), Philippe Aghion, Peter Howitt, and Giovanni Violante (1999), Galor and Moav (2000), and Gould, Moav, and Weinberg (2000).

premiums in the long run causes educational inequality to be larger than otherwise. In applying our theory to recent changes in educational inequality in the United States, we suggest that increases in ability premiums have dampened the long-run response of the relative supply of more educated workers that otherwise would reverse previous increases in educational inequality.

In related papers that focus on educational inequality, Daron Acemoglu (1998, 2000), picking up on the observation that the relative number of more educated workers increased unusually rapidly in the 1910s and 1920s and again in the 1970s, argues that these changes in the pattern of supply were at least partly exogenous, and he accepts that their contemporaneous effect was to decrease educational inequality. Acemoglu assumes, however, that the larger number of more educated workers caused the development of more education-intensive production technologies. He suggests that in subsequent decades this endogenous response of technology more than reversed the initial decrease in educational inequality.

The analysis in the present paper is complementary in that it treats the relative demand for more educated labor as exogenous and the relative supply of more educated labor as endogenous. Together Acemoglu's analysis and the present analysis imply that either an exogenous increase in the relative number of more educated workers or an exogenous increase in the relative demand for more educated labor could lead to an endogenous cycle in educational inequality. More importantly, the present paper adds an analysis of the combined effects on educational inequality of innovations that change the relative demand for more educated labor and innovations that affect ability premiums.

1. Educational Inequality in the Short Run

Assume that more educated labor and less educated labor perform complementary functions in the production process. Specifically, assume that the output per period of a representative firm in a representative industry, denoted by Y , is a Cobb-Douglas function of

inputs of more educated labor and less educated labor, as in

$$(1) \quad Y = L_m^\sigma L_\ell^{1-\sigma}, \quad \sigma \in (0, 1),$$

where L_m and L_ℓ denote the numbers of efficiency units of more educated labor and less educated labor that the firm employs per period. This formulation implies that more educated labor and less educated labor differ qualitatively. The parameter σ measures the relative importance of more educated labor in production.⁶

Let \tilde{L}_m and \tilde{L}_ℓ denote the quantities supplied per firm per period of efficiency units of more educated labor and less educated labor, and let w_m and w_ℓ represent the earnings per period of an efficiency unit of more educated labor and less educated labor. Assume that the firm takes the earnings of an efficiency unit of labor as given and demands quantities of each type of labor such that the marginal product of an efficiency unit of labor equals the earnings of an efficiency unit of labor. Calculating marginal products from equation (1), and using the market-clearing conditions that \tilde{L}_m equals L_m and \tilde{L}_ℓ equals L_ℓ , we find that

$$(2) \quad \frac{w_m}{w_\ell} = \frac{\sigma}{1-\sigma} \frac{\tilde{L}_\ell}{\tilde{L}_m}.$$

Equation (2) shows that the relative earnings per period of efficiency units of more and less educated labor depend negatively on the relative supply of more educated labor and positively on the relative importance of more educated labor in production.

⁶The Cobb-Douglas function in equation (1) is a special case of an aggregate production function that exhibits constant elasticity of substitution between more educated labor and less educated labor, as in $Y = [\sigma L_m^\rho + (1-\sigma) L_\ell^\rho]^{1/\rho}$, $\rho < 1$, $\rho \neq 0$. Equation (1) obtains in the limit as the parameter ρ goes to zero. If we were to replace equation (1) with the general CES function, then we would have to replace subsequent equations that are derived using equation (1) with more complicated equations involving ρ . These equations are given in the mathematical appendix. As we can see, as long as ρ is not a large negative number, none of our qualitative conclusions, which we draw mainly from equations (11) and (13), are changed. Autor, Katz, and Krueger (1998) cite various estimates of the elasticity of substitution between workers who are or are not college educated, all of which imply that ρ is in the neighborhood of 1/3.

A worker's wage or salary equals the product of the number of efficiency units of labor that he (or she) supplies per period and the earnings per period of an efficiency unit of his type of labor. We assume that all more educated workers perform the same function in the production process, and that all less educated workers perform the same complementary function in the production process. But, the number of efficiency units of labor that a worker supplies per period depends on his ability. In other words, differences in ability result in quantitative differences among more educated workers and among less educated workers.

For simplicity, we assume that a worker can have either high ability, or ordinary ability, or low ability, and that a worker's ability is readily discernable. We also assume that workers with either high ability or ordinary ability are capable of becoming more educated, whereas workers with low ability are not educable beyond a basic education, and that workers with high ability can realize their advantage over workers with ordinary ability only by becoming more educated. As we shall see, this last property implies that, if any workers with ordinary ability choose to become more educated, then all workers with high ability choose to become more educated.

Abstracting from differences in experience, let α denote the number of efficiency units of labor that a more educated worker who has high ability supplies per period, and let β denote the number of efficiency units of labor that a more educated worker who has ordinary ability supplies per period, where $\beta < \alpha$. Accordingly, we have $W_{mh} = \alpha w_m$ and $W_{mo} = \beta w_m$, where W_{mh} denotes the wage or salary of a more educated worker with high ability, and W_{mo} denotes the wage or salary of a more educated worker with ordinary ability. The ratio W_{mh}/W_{mo} , which equals the ratio α/β , measures the ability premium for more educated workers.⁷ These ratios reflect the advantage in production associated with being a more

⁷Because in our model more educated workers have either high ability or ordinary ability and less educated workers have either ordinary ability or low ability, ability premiums do not depend on the numbers of more or less educated workers. If, alternatively, we were to model ability as being continuously distributed, then an increase in the fraction of workers who choose to become more educated would imply an increase in the

educated worker with high ability rather than ordinary ability, combined with factors that determine the differences among workers, such as their innate characteristics and the quality of the educations that they receive.

Again abstracting from differences in experience, let γ denote the number of efficiency units of labor that a less educated worker who has ordinary ability supplies per period, and let δ denote the number of efficiency units of labor that a less educated worker who has low ability supplies per period, where $\delta < \gamma$. Thus, we have $W_{\ell o} = \gamma w_{\ell}$, and $W_{\ell \ell} = \delta w_{\ell}$, where $W_{\ell o}$ denotes the wage or salary of a less educated worker with ordinary ability, and $W_{\ell \ell}$ denotes the wage or salary of a less educated worker with low ability. The ratio $W_{\ell o}/W_{\ell \ell}$, which equals the ratio γ/δ , measures the ability premium for less educated workers.

Let N_{mh} denote the number of more educated workers who have high ability, and let N_{mo} denote the number of more educated workers who have ordinary ability. Also, let $N_{\ell o}$ denote the number of less educated workers who have ordinary ability, and let $N_{\ell \ell}$ denote the number of less educated workers who have low ability. In the short run all of these numbers are given, whereas in the long run some of them depend on the decisions of workers to become more educated. Given α , β , γ , and δ , the relative supply of efficiency units of more educated labor is related to the numbers of workers of each type according to

$$(3) \quad \frac{\tilde{L}_{\ell}}{\tilde{L}_m} = \frac{\gamma N_{\ell o} + \delta N_{\ell \ell}}{\alpha N_{mh} + \beta N_{mo}}.$$

Let W_m denote the average wage or salary per period of more educated workers, and let W_{ℓ} denote the average wage or salary per period of less educated workers. To calculate W_m we divide the aggregate earnings per period of more educated workers, $(\alpha N_{mh} + \beta N_{mo})w_m$, ratio of the wage or salary of the most able more educated worker to the wage or salary of the least able more educated worker and an decrease in the ratio of the wage or salary of the most able less educated worker to the wage or salary of the least able less educated worker. If ability premiums were defined as these ratios, then ability premiums would depend on the identities of the least able more educated worker and the most able less educated worker, hence, on the numbers of more or less educated workers.

by the number of more educated workers, $N_{mh} + N_{mo}$. To calculate W_ℓ we divide the aggregate earnings per period of less educated workers, $(\gamma N_{\ell o} + \delta N_{\ell \ell})w_\ell$, by the number of less educated workers, $N_{\ell o} + N_{\ell \ell}$. Alternatively, we can characterize W_m as the product of the earnings of an efficiency unit of more educated labor, w_m , and the average ability of more educated workers, $(\alpha N_{mh} + \beta N_{mo})/(N_{mh} + N_{mo})$, and similarly for W_ℓ .⁸

The ratio W_m/W_ℓ measures educational inequality. Calculating W_m and W_ℓ and dividing we obtain

$$(4) \quad \frac{W_m}{W_\ell} = \frac{\alpha N_{mh} + \beta N_{mo}}{N_{mh} + N_{mo}} \frac{N_{\ell o} + N_{\ell \ell}}{\gamma N_{\ell o} + \delta N_{\ell \ell}} \frac{w_m}{w_\ell}.$$

Equation (4) shows how educational inequality comprises the relative earnings per period of efficiency units of more and less educated labor and the average abilities of more and less educated workers.

Substituting equations (2) and (3), which determine the ratio w_m/w_ℓ , into equation (4), we obtain

$$(5) \quad \frac{W_m}{W_\ell} = \frac{\sigma}{1 - \sigma} \frac{N_{\ell o} + N_{\ell \ell}}{N_{mh} + N_{mo}}.$$

Equation (5) implies that educational inequality depends positively on the number of less educated workers relative to the number of more educated workers. Topel (1997) reports that the data for many countries are consistent with this implication. In addition, equation (5) implies that, given the relative number of more educated workers, which is fixed in the short run, a technological innovation that increases the parameter σ , the relative importance of more educated labor, increases educational inequality.⁹

⁸These calculations correspond to the standard way of calculating the average wage or salary of more and less educated workers from panel data. For more detailed explanations, see Katz and Murphy (1992) and Mishel, Bernstein, and Schmitt (2001).

⁹According to equation (5) educational inequality in the short run does not depend on α , β , γ , or δ . Given the earnings per efficiency unit of more educated labor, an innovation that increases either α

2. Educational Inequality in the Long Run

To analyze educational inequality in the long run, consider a constant population of workers per firm, normalized to one, with constant fractions of workers that have high ability, ordinary ability, and low ability. Let H , O , and L , respectively, denote these fractions, where $H + O + L = 1$.

Low ability can be either innate or a result of the quality of the basic education that a worker receives. By assumption workers with low ability are not educable beyond a basic education. In contrast, workers with either ordinary ability or high ability can choose to become more educated. Ordinary ability or high ability can be either innate or a result of a combination of the quality of the basic education that a worker receives and the quality of the advanced education that is available to a worker.

Let M denote the fraction of workers with ordinary ability that chooses to become more educated. The fraction $1 - M$ remains less educated. For ability premiums to be observed both for more educated workers and for less educated workers, both M and $1 - M$ must be positive. Hence, workers with ordinary ability must be indifferent between becoming more educated and remaining less educated.

Remember that workers with high ability can realize their advantage over workers with ordinary ability only by becoming more educated. Given this assumption, and assuming that a worker chooses to become more educated or to remain less educated according to which choice yields higher lifetime earnings for him,¹⁰ with workers with ordinary ability being or β would increase the average wage or salary of a more educated worker. But, the increased supply of efficiency units of more educated labor also would decrease the relative earnings per efficiency unit of more educated labor. With a Cobb-Douglas production function these two effects on educational inequality are exactly offsetting. With the general CES production function educational inequality in the short run would be either positively or negatively related to both α and β as ρ is positive or negative. A similar analysis applies to γ and δ . See the mathematical appendix.

¹⁰This assumption abstracts from liquidity constraints on the ability to finance education. Galor and Moav (2000) point out that, if technological progress relaxes liquidity constraints on the ability to finance

indifferent between becoming more educated or not, all workers with high ability choose to become more educated.

We define the long run to be a steady state in which M is the same in all age cohorts. This definition of the long run implies that σ has stabilized at a constant value. Realization of the long run, which requires convergence to the steady state associated with a constant value of σ , can require as much as enough time for an age cohort to pass completely through a life cycle.

Assume that each worker is active for T periods and that to become more educated a worker must spend τ periods in school rather than in the work force, where $\tau < T$. Thus, in the long run $H\tau/T$ young workers who have high ability and $OM\tau/T$ young workers who have ordinary ability are in school. Accordingly, the number of workers of each type who are in the work force is

$$\begin{aligned}
 N_{mh}^* &= H(T - \tau)/T, \\
 N_{mo}^* &= OM(T - \tau)/T, \\
 N_{lo}^* &= O(1 - M), \quad \text{and} \\
 N_{\ell\ell}^* &= L,
 \end{aligned}
 \tag{6}$$

where the $*$ denotes a value that obtains in the long run.

Substituting equations (6) into equation (5) we obtain

$$\frac{W_m^*}{W_\ell^*} = \frac{\sigma}{1 - \sigma} \frac{O(1 - M) + L}{H + OM} \frac{T}{T - \tau}.
 \tag{7}$$

Equation (7) relates educational inequality in the long run to the fractions of the workers who have high, ordinary, or low ability, the fraction of workers with ordinary ability that chooses to become more educated, the fraction of his active life that a more educated worker spends in the work force, and the relative importance of more educated labor in production.

education, then technological progress can be associated with a decrease in educational inequality. Gould, Moav, and Weinberg (2000) introduce an additional precautionary motive for becoming more educated.

The final problem is to determine the fraction of workers with ordinary ability that choose to become more educated. A critical element in our analysis is that in the long run this fraction responds positively to short-run increases in the relative earnings of more educated workers. Topel (1997) suggests that the facts are consistent with such a positive response.

To analyze this response, let E_{mo}^* denote the lifetime earnings of a more educated worker with ordinary ability in the long run, and let $E_{\ell o}^*$ denote the lifetime earnings of a less educated worker with ordinary ability in the long run. Assume that the cost of becoming more educated includes both the earnings foregone while becoming more educated and an out-of-pocket cost for tuition and other expenses. Let c , which for simplicity we assume to be the same for more educated workers with high ability and with ordinary ability, denote the ratio of the out-of-pocket cost of becoming more educated to the gross earnings of a more educated worker. Thus, we have

$$(8) \quad \begin{aligned} E_{mo}^* &= (1 - c) (T - \tau) W_{mo}^* \quad \text{and} \\ E_{\ell o}^* &= T W_{\ell o}^*. \end{aligned}$$

As we have noted, for ability premiums to be observed both for more educated workers and for less educated workers, workers with ordinary ability must be indifferent between becoming more educated and remaining less educated. This property of indifference implies that E_{mo}^* must equal $E_{\ell o}^*$. Equating E_{mo}^* and $E_{\ell o}^*$ we find that in the long run the fraction of the workers with ordinary ability that chooses to become more educated is such that the wage or salary of a more educated worker with ordinary ability relative to the wage or salary of a less educated worker with ordinary ability satisfies

$$(9) \quad \frac{W_{mo}^*}{W_{\ell o}^*} = \frac{\beta w_m^*}{\gamma w_\ell^*} = \frac{1}{1 - c} \frac{T}{T - \tau}.$$

According to equation (9) neither $W_{mo}^*/W_{\ell o}^*$ nor w_m^*/w_ℓ^* depend on σ . Thus, equation (9) has the following implication:

In response to an innovation that increases the relative importance of more educated labor, in the long run an increase in the fraction of workers with ordinary ability that chooses to become more educated reverses the short-run increase in earnings per efficiency unit of more educated labor relative to earnings per efficiency unit of less educated labor, and, hence, also reverses the short-run increase in the average wage or salary of a more educated worker with ordinary ability relative to the average wage or salary of a less educated worker with ordinary ability.

Earnings per efficiency unit of more educated labor and less educated labor also must satisfy market-clearing conditions. Substituting equation (3) into equation (2), and using equation (6), we find that market clearing implies that

$$(10) \quad \frac{w_m^*}{w_\ell^*} = \frac{\sigma}{1-\sigma} \frac{\gamma N_{\ell o}^* + \delta N_{\ell \ell}^*}{\alpha N_{m h}^* + \beta N_{m o}^*} = \frac{\sigma}{1-\sigma} \frac{\gamma O(1-M) + \delta L}{\alpha H + \beta O M} \frac{T}{T-\tau}.$$

Solving equations (9) and (10) for M , we obtain

$$(11) \quad M = \frac{(1-c) \frac{\sigma}{1-\sigma} \left(1 + \frac{\delta L}{\gamma O}\right) - \frac{\alpha H}{\beta O}}{(1-c) \frac{\sigma}{1-\sigma} + 1}.$$

We assume that the values of the parameters are such that M , as given by equation (11), is a positive fraction. Equation (11) shows that M is not only positively related to σ , an effect that already was implicit in equation (9), but that M is also negatively related to the ratios α/β and γ/δ .

This important result obtains because ability premiums affect the average wage or salary of a more educated worker with ordinary ability relative to the average wage or salary of a less educated worker with ordinary ability. Specifically, we can understand why M depends negatively on α/β as follows: (An analogous explanation applies to γ/δ .) Given $N_{m o}^*$ and $N_{\ell o}^*$, the larger is α the larger would be the relative supply of more educated labor and, hence, from equation (10), the smaller would be the ratio $\beta w_m^*/\gamma w_\ell^*$ that satisfies

the market-clearing conditions. But, the ratio $\beta w_m^*/\gamma w_\ell^*$ also must satisfy the equality $E_{\ell o}^* = E_{mo}^*$, and the implied equation (9), which does not involve α . Consequently, equation (10) implies that, with a larger value of α , N_{mo}^* must be smaller and $N_{\ell o}^*$ must larger. Both a smaller value of N_{mo}^* and a larger value of $N_{\ell o}^*$ imply a smaller value of M .

Also, given N_{mo}^* and $N_{\ell o}^*$, the smaller is β the smaller is the relative supply of more educated labor. But, given w_m^* , the smaller is β the smaller is βw_m^* . It is easy to see that, with a Cobb-Douglas production function, the latter effect dominates. Consequently, to satisfy the equality $E_{\ell o}^* = E_{mo}^*$, the smaller is β again the smaller must be M .

In addition, with a Cobb-Douglas production function, the effects of α and $1/\beta$ on M are equal, as are the effects of γ and $1/\delta$.¹¹ To see these results algebraically, observe that, after multiplying by β/γ , equation (10) becomes

$$(12) \quad \frac{W_{mo}^*}{W_{\ell o}^*} = \frac{\sigma}{1-\sigma} \frac{O(1-M) + (\delta/\gamma)L}{(\alpha/\beta)H + OM} \frac{T}{T-\tau}.$$

Equation (12) implies that, ceteris paribus, the larger is either α/β or γ/δ the smaller is the ratio $W_{mo}^*/W_{\ell o}^*$ for any given value of M . Thus, to satisfy the equality $E_{\ell o}^* = E_{mo}^*$ and equation (9), the larger is either α/β or γ/δ the smaller must be M .

In sum, equation (11) has the following implication:

Whereas an innovation that increases the relative importance of more educated labor causes a larger fraction of the workers with ordinary ability to choose to become more educated, innovations that increase ability premiums cause a smaller fraction of workers with ordinary ability to choose to become more educated.

The net change in the fraction of workers that chooses to become more educated depends on

¹¹With the general CES production function M would be either more or less sensitive to $1/\beta$ than to α as ρ is positive or negative. But, as long as ρ is not a large negative number, M still would be negatively related to both α and $1/\beta$. The same effect would carry over to the relation between W_m^*/W_ℓ^* and α and $1/\beta$. Analogous results apply to γ/δ . See the mathematical appendix.

the net effect of innovations that increases the relative importance of more educated labor and innovations that increase ability premiums.¹²

We turn now to educational inequality in the long run, W_m^*/W_ℓ^* . Substituting equation (11) into equation (7) we obtain a solution for W_m^*/W_ℓ^* as a function of the exogenous variables,

$$(13) \quad \frac{W_m^*}{W_\ell^*} = \frac{(1-c) \frac{\sigma}{1-\sigma} L\left(1 - \frac{\delta}{\gamma}\right) + H\left(\frac{\alpha}{\beta} - 1\right) + 1}{(1-c) \left[1 - L\left(1 - \frac{\delta}{\gamma}\right)\right] - \frac{1-\sigma}{\sigma} H\left(\frac{\alpha}{\beta} - 1\right)} \frac{T}{T - \tau}.$$

Given that α is larger than β and that γ is larger than δ , equation (13) implies that W_m^*/W_ℓ^* is smaller the larger is σ . In the limit as the ratios α/β and γ/δ approach one, W_m^*/W_ℓ^* becomes independent of σ . Thus, equation (13) has the following implication:

If and only if ability premiums exist, an innovation that increases educational inequality in the short run causes a decrease in educational inequality in the long run.

We can explain this result as follows: As we have seen from equation (9), an increase in σ in the long run causes sufficiently more workers with ordinary ability to become more educated to offset the short-run effect of σ on the relative earnings per efficiency unit of more educated labor. The addition of more workers with ordinary ability to the total

¹²Equation (11) also tells us that M does not depend on either T or τ . This result obtains because the larger is $T/(T - \tau)$ the larger is $E_{m_o}^*$ relative to $E_{\ell_o}^*$ for a given ratio $W_{m_o}^*/W_{\ell_o}^*$, but also the smaller is the ratio $W_{m_o}^*/W_{\ell_o}^*$ that satisfies market-clearing conditions. Given the assumed Cobb-Douglas production function, these two effects are offsetting. Hence, the value of M that satisfies the equality $E_{\ell_o}^* = E_{m_o}^*$ is independent of $T/(T - \tau)$. With the general CES production function M would be either positively or negatively related to T as ρ is positive or negative. If ρ is positive, then the conclusions about the effect of $T/(T - \tau)$ on W_m^*/W_ℓ^* that we draw from equation (13) would be reinforced. See Sebnem Kalemli-Ozcan, Harl Ryder, and David Weil (2000) for a model in which the longer that a worker expects to live the larger is the number of years of schooling that he chooses.

number of more educated workers also decreases the average number of efficiency units of more educated labor that more educated workers supply and increases the average number of efficiency units of less educated labor that less educated workers supply. The result is that w_m^*/w_ℓ^* is unchanged and that W_m^*/W_ℓ^* is decreased.

Equation (13) also implies that W_m^*/W_ℓ^* is larger the larger are the ratios α/β and γ/δ . Thus, equation (13) also has the following implication:

Innovations that increase ability premiums cause educational inequality in the long run to be larger than it otherwise would be.

This result obtains because, as we have seen from equation (11), an increase in either α/β or γ/δ causes a decrease in M .

From equation (13) we also see that W_m^*/W_ℓ^* is larger the larger is either c or $T/(T-\tau)$. The effect of c on educational inequality obtains because, from equation (11), the larger the out-of-pocket cost the smaller is the fraction of the workers with ordinary ability that chooses to become more educated. The effect of $T/(T-\tau)$ on educational inequality obtains because, from equation (7), for any value of M , the smaller is the fraction of his active life that a more educated worker spends in the work force the smaller is the relative supply of more educated labor, and the larger is W_m^*/W_ℓ^* .

3. The Observed Pattern of Educational Inequality

Over the past few centuries life expectancy at all ages has steadily increased. A concomitant of this steady increase in life expectancy has been a steady increase in the fraction of his active life that a more educated worker can spend in the work force. In addition, from the middle of the nineteenth century until at least the middle of the twentieth century expansion of publicly-financed higher education steadily decreased the relative out-of-pocket cost of becoming more educated. In terms of our model, there has been a secular increase in $(T-\tau)/T$ and a secular decrease in c . Equation (13) implies that the secular increase in the relative supply of more educated labor resulting from both the secular increase in $(T-\tau)/T$

and the secular decrease in c tends to decrease educational inequality. Goldin and Katz (2001) suggest that over the first half of the twentieth century the increased relative supply of more educated labor seems to have been the dominant influence on educational inequality, despite technological innovations that increased the relative importance of more educated labor in production.

Now, suppose that during 1950s and 1960s and again during the 1980s and 1990s innovations that increased the relative demand for more educated labor were large enough to outweigh the effect of continuing increases in the relative supply of more educated labor. In our analysis we have modeled innovations that increased the relative demand for more educated labor by increases in σ . These innovations could account for the observed deviations during these decades from the secular trend of decreasing educational inequality.¹³

Our analysis implies, however, that increases in σ do not cause permanent increases in educational inequality. On the contrary, our analysis implies that an induced long-run increase in the relative supply of more educated labor more than reverses any short-run increase in educational inequality caused by an increase in σ . Thus, our model predicts that, to the extent that an increase in educational inequality reflects the short-run effect of an increase in σ , educational inequality will tend to return to its long-run downward trend. We might speculate that the slowing of the rate of increase in educational inequality in the 1990s compared to the 1980s indicates the beginning of this process.

But, our analysis also suggests that there is more to the story than increases in the relative demand for more educated labor. As we have noted, the data tell us that ability premiums increased steadily from the late 1960s until the mid 1990s. We suggested that these increases in ability premiums resulted from a combination of increases in the advantage in production associated with being more able and increased differences in the quality of education. In our

¹³Autor, Katz, and Krueger (1998) speculate that a temporary slowing in the rate of increase in the relative supply of more educated labor, resulting from exogenous demographic factors, also contributed to the increases in educational inequality during the 1980s and 1990s.

model ability premiums affect the average wage or salary of a more educated worker with ordinary ability relative to the average wage or salary of a less educated worker with ordinary ability. This effect has the critical implication that increases in ability premiums cause a smaller fraction of workers with ordinary ability to choose to become more educated. Hence, in the long run increases in ability premiums cause educational inequality to be larger.

This implication suggests an additional reason for the increase in educational inequality during the 1980s and 1990s. More importantly, to the extent that increases in ability premiums are contributing in the long run to educational inequality, the theory implies that, although induced increases in the relative supply of more educated labor may moderate educational inequality, the increases in educational inequality during the 1980s and 1990s are unlikely to be soon reversed.

4. Summary

This paper has analyzed a model in which in the short run an innovation that increases the relative demand for more educated labor increases both the earnings of an efficiency unit of more educated labor relative to the earnings of an efficiency unit of less educated labor and the average wage or salary of a more educated worker relative to the average wage or salary of a less educated worker. Thus, in the short run such an innovation increases educational inequality.

But, in the long run the relative supply of more educated labor is endogenous, and an increase in the relative demand for more educated labor causes an increase in the fraction of workers that chooses to become more educated. By itself, this educational response would reverse the short-run increase in the relative earnings of an efficiency unit of more educated labor and, if ability premiums exist, would more than reverse the short-run increase in educational inequality.

The model also allowed for the effects of innovations that increase ability premiums. Most interestingly, we found that such innovations cause a smaller fraction of workers to choose

to become more educated than otherwise. As a result, in the long run increases in ability premiums result in more educational inequality than otherwise. In applying our theory to recent developments, we suggested that increases in ability premiums have dampened the long-run response of the relative supply of more educated workers that otherwise would be reversing previous increases in educational inequality.

References

- Acemoglu, Daron. "Technical Change, Inequality, and the Labor Market." National Bureau of Economic Research (Cambridge MA) Working Paper No. 7800, July 2000.
- Acemoglu, Daron. "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality." *Quarterly Journal of Economics*, November 1998, *113*(3), pp. 1055-1089.
- Aghion, Philippe, Howitt, Peter, and Violante, Giovanni L. "General Purpose Technology and Within-Group Inequality." unpublished, December 1999.
- Autor, David H., Katz, Lawrence F., and Krueger, Alan B. "Computing Inequality: Have Computers Changed the Labor Market?" *Quarterly Journal of Economics*, November 1998, *113*(3), pp. 1169-1213.
- Baldwin, Robert E. and Cain, Glen G. "Shifts in U.S. Relative Wages: The Role of Trade, Technology and Factor Endowments." National Bureau of Economic Research (Cambridge MA) Working Paper No. 5934, February 1997.
- Caselli, Francesco, "Technological Revolutions." *American Economic Review*, March 1999, *89*(1), pp. 78-102.
- Davis, Steven J. and Haltiwanger, John. "Wage Dispersion Between and Within U.S. Manufacturing Plants, 1963-86." *Brookings Papers on Economic Activity: Microeconomics*, 1991, pp. 115-180.
- Galor, Oded and Moav, Omer. "Ability-Biased Technological Transition, Wage Inequality, and Economic Growth." *Quarterly Journal of Economics*, May 2000, *115*(2), pp. 469-497.
- Galor, Oded and Tsiddon, Daniel. "Technological Progress, Mobility, and Economic Growth." *American Economic Review*, June 1997, *87*(3), pp. 363-381.
- Goldin, Claudia and Katz, Lawrence F. "Decreasing and then Increasing Inequality in America: A Tale of Two Half Centuries," in Finis Welch, ed., *The Causes and Consequences of Increasing Inequality*, Chicago: University of Chicago Press, 2001.

- Gottschalk, Peter. "Inequality, Income Growth, and Mobility: The Basic Facts." *Journal of Economic Perspectives*, Spring 1997, 11(2), pp. 21-40.
- Gould, Eric D., Moav, Omer, and Weinberg, Bruce A. "Precautionary Demand for Education, Inequality, and Technological Progress." unpublished, July 2000.
- Harrigan, James and Balaban, Rita. "U.S. Wages in General Equilibrium: The Effects of Prices, Technology, and Factor Supplies, 1963-1991." National Bureau of Economic Research (Cambridge MA) Working Paper No. 6981, February 1999.
- Juhn, Chinhui, Murphy, Kevin M., and Pierce, Brooks. "Wage Inequality and the Rise in Returns to Skill." *Journal of Political Economy*, June 1993, 101(3), pp. 410-442.
- Kalemli-Ozcan, Sebnem, Ryder, Harl E., and Weil, David N. "Mortality Decline, Human Capital Investment, and Economic Growth." *Journal of Development Economics*, June 2000, 62(1), pp.1-23.
- Katz, Lawrence and Murphy, Kevin. "Changes in Relative Wages, 1963-1987: Supply and Demand Factors." *Quarterly Journal of Economics*, February 1992, 107(1), pp. 35-78.
- Lloyd-Ellis, Huw. "Endogenous Technological Change and Wage Inequality." *American Economics Review*, March 1999, 89(1), pp. 47-77.
- Mishel, Lawrence, Bernstein, Jared, and Schmitt, John. *The State of Working America 2000-2001*, Ithaca: Cornell University Press, 2001.
- Murnane, Richard, Willett, John, and Levy, Frank. "The Growing Importance of Cognitive Skills in Wage Determination." *The Review of Economics and Statistics*, May 1995, 77(2), pp. 251-266.
- Topel, Robert H. "Factor Proportions and Relative Wages: The Supply-Side Determinants of Wage Inequality." *Journal of Economic Perspectives*, Spring 1997, 11(2), pp. 55-74.

Mathematical Appendix: General CES Production Function

With a general CES production function, equations (1), (5), (11), and (13) would become the following:

$$(1') \quad Y = [\sigma L_m^\rho + (1 - \sigma) L_\ell^\rho]^{1/\rho}, \quad \rho < 1, \quad \rho \neq 0.$$

$$(5') \quad \frac{W_m}{W_\ell} = \frac{\sigma}{1 - \sigma} \frac{N_{\ell o} + N_{\ell \ell}}{N_{mh} + N_{mo}} \left(\frac{\alpha N_{mh} + \beta N_{mo}}{\gamma N_{\ell o} + \delta N_{\ell \ell}} \right)^\rho.$$

$$(11') \quad M = \frac{\left[(1 - c) \frac{\sigma}{1 - \sigma} \right]^{\frac{1}{1 - \rho}} \left(1 + \frac{\delta}{\gamma} \frac{L}{O} \right) - \frac{\alpha}{\beta} \frac{H}{O} \left(\frac{\gamma}{\beta} \frac{T}{T - \tau} \right)^{\frac{\rho}{1 - \rho}}}{\left[(1 - c) \frac{\sigma}{1 - \sigma} \right]^{\frac{1}{1 - \rho}} + \left(\frac{\gamma}{\beta} \frac{T}{T - \tau} \right)^{\frac{\rho}{1 - \rho}}}.$$

$$(13') \quad \frac{W_m^*}{W_\ell^*} = \frac{\sigma}{1 - \sigma} \left[\frac{(1 - c) \frac{\sigma}{1 - \sigma}}{\frac{\gamma}{\beta} \frac{T}{T - \tau}} \right]^{\frac{\rho}{1 - \rho}} \frac{T}{T - \tau} \cdot$$

$$\frac{\left[(1 - c) \frac{\sigma}{1 - \sigma} \right]^{\frac{1}{1 - \rho}} L \left(1 - \frac{\delta}{\gamma} \right) + \left[H \left(\frac{\alpha}{\beta} - 1 \right) + 1 \right] \left(\frac{\gamma}{\beta} \frac{T}{T - \tau} \right)^{\frac{\rho}{1 - \rho}}}{\left[(1 - c) \frac{\sigma}{1 - \sigma} \right]^{\frac{1}{1 - \rho}} \left[1 - L \left(1 - \frac{\delta}{\gamma} \right) \right] - H \left(\frac{\alpha}{\beta} - 1 \right) \left(\frac{\gamma}{\beta} \frac{T}{T - \tau} \right)^{\frac{\rho}{1 - \rho}}}.$$