How Did the Increase in Economic Inequality between 1970 and 1990 Affect American Children's Educational Attainment?

Susan E. Mayer

University of Chicago

s-mayer@uchicago.edu

Disparities in hourly wages, annual earnings, and household income have all increased over the past generation in the United States. A considerable amount of research has tried to determine why income inequality grew over this period. (See Morris and Western, 1999 for a review of much of this research.) Far less research has been done on the consequences of inequality than on its causes. This paper estimates the effect of the growth in income inequality on mean educational attainment and on the disparity in educational attainment between rich and poor children. If income inequality contributes to inequality in educational attainment between children from rich and poor families, inequality in one generation will be perpetuated in the next generation. I also separate the effect of income inequality that is due to the nonlinear effect of a family's own income from the effect due to interpersonal interactions.

This paper focuses entirely on economic inequality. Other forms of inequality such as racial and gender inequality may be correlated with economic inequality but they can also exert their own independent effects on educational attainment. However the growth in economic inequality has been especially great over the last generation. Table 1 shows that the Gini coefficient of household income increased between 1970 and 1980 and increased even more between 1980 and 1990.[1] An alternative measure of inequality, the standard deviation of log income, also increased over this period.

Table 2 shows that the proportion of 25 to 29 year olds who have graduated high school increased during the 1970s but hardly increased at all during the 1980s. Any increase in the rate of high school graduation that occurred in the 1980s appears to have been largely due to an increase in persons getting a GED.[2] The trend for college enrollment and college graduation is similar to the trend for high school graduation: increases in the 1970s but little change in the 1980s.[3]

Economists often explain differences in educational attainment using a human capital model in which school enrollment is a function of the monetary returns to schooling. All else equal, as the return to additional schooling increases, enrollment should rise. The rate of return is determined by the ratio of the monetary and non-monetary costs to the monetary and non-monetary benefits of attending a year of school. The monetary costs include tuition and foregone wages. Non-monetary costs include the psychological cost (or benefit) of being in a classroom and doing academic work rather than working in the labor market, rearing children, watching television, and so on. Monetary costs to individual students also depend on the cost of

---

[1] The Gini coefficient is the proportion of the total area below the 45 degree line and the Lorenz curve, which plots the cumulative percentage of households against the cumulative percentage of income received by them. See Firebaugh (1999) for a good comparison of inequality measures across countries. See also Atkinson 1970 or 1983 for a discussion of the statistical differences among inequality measures.

[2] The percentage of sixteen to twenty-four year olds high school graduates who had a GED increased from 15.8 percent in 1980 to 17.8 percent in 1990.

[3] Welch (1999) notes that the proportion of men working full-time year around who have at least one year of college increased greatly after 1980 when the returns to schooling also increased greatly. Welch takes this as evidence that men responded to the increase in returns to schooling by getting more schooling and that the increase in schooling is therefore a benefit of the rise in inequality. However, the test of the response to the rise in the return to schooling is not the increase in schooling among workers since workers no doubt became more selected on their education as the returns to schooling increased. The test is the change in schooling for all members of young cohorts.

borrowing, how much parents contribute to the costs of schooling, and on how much schooling is subsidized by the state, local school districts, and colleges. When children are in elementary and secondary schools, the greatest cost of schooling is incurred by parents who choose to live in school districts with high tax rates.

Hypotheses about the Effect of Inequality

Income inequality can affect the costs and benefits of educational attainment in two main ways, which I label "nonlinearity" effects and "macro" effects.

Nonlinearity. Income inequality will affect the *mean* level of educational attainment only if the effect of household income on children's educational attainment is nonlinear. If the relationship between educational attainment and parental income is linear, then when the rich gain a dollar and the poor lose a dollar, the educational attainment of the rich will increase by exactly as much as the educational attainment of the poor decreases, leaving the mean unchanged.

However, suppose that a 1 percent increase in income generates the same absolute increment in educational attainment, regardless of whether income is initially high or low. If the relationship between income and schooling takes this semi-logarithmic form, then all else equal a costless redistribution from richer to the poorer households would increase children's mean educational attainment because shifting a dollar from the rich to the poor child would increase the education of poor children by a larger percentage than it would decrease the education of rich children.[4] Thus if the effect of family income is nonlinear all else equal growth in economic inequality would be associated with a decrease in mean educational attainment.

Macro Effects. In theory, income inequality can also affect educational attainment even when we compare households with the same income and needs. What I refer to as the "macro" effects of inequality occur when inequality affects people's behavior in ways that are independent of their household income.

Most hypothetical macro effects flow primarily from interpersonal comparisons between the rich and poor. Such comparisons can take several forms, which are likely to have different effects. One kind of social comparison involves "relative deprivation." Another involves "role models." A third possible macro effect of inequality results from the incentives that cause inequality. A fourth possible macro effect works through political behavior that results from interpersonal comparisons between the rich and poor.

Relative deprivation theory holds that high levels of inequality make the poor feel worse off, increasing their alienation and stress (Jencks and Mayer 1990). One version of this hypothesis is that children feel deprived when they cannot have the same material possessions as other children in their school or neighborhood. Another version is that relative deprivation makes poorer parents feel stressed and alienated, lowering their expectations for their children or reducing the quality of their parenting (McLoyd 1990). The relative deprivation model assumes that children or parents compare themselves to others who are better off while largely ignoring those who are worse off. If parents all compare themselves to the richest people in society, for

[4] Mean educational attainment might not increase when the rich get richer and even if it does, the increase might not be efficient. If when the rich get richer they "over-invest" in schooling and the poor "under-invest" the mean level of educational attainment might stay the same, but the efficiency of the investment would decline. For now I assume that no one over-invests in schooling.

example, they will feel poorer whenever the rich get richer. Their hopes and expectations for their children may then decline and their parenting skills may worsen.[5]

Note, however, that if either children or their parents mostly compare themselves to the poorest people rather than the richest, increases in inequality will make most people feel richer because the distance between them and the people at the bottom of the distribution will grow. If people mostly compare themselves to some real or imagined national average, increases in inequality will make the rich feel richer and the poor feel poorer. How this affects educational attainment or other outcomes will depend on whether the rich gain more than the poor lose.

The role model hypothesis holds that children model their behavior on the behavior of those around them. Role models can be either positive or negative, but most of the literature on this subject stresses the harm done by the absence of "positive" role models rather than the harm done by the presence of "negative" role models (Ogbu 1981, Wilson 1987). So, for example, most writers assume that attending a school where most other children want to go to college increases a child's own chances of going to college. Thus if increased inequality made some households richer and if this changed their children's behavior for the better, the change might have positive externalities for poorer children, making them act more like the rich. But an increase in inequality can also make the poor even poorer, which could increase the number of negative role models. If children are more influenced by negative than positive role models, increasing inequality could thus reduce the frequency of behavior that parents usually promote.

Economists usually believe that economic incentives are the best available tool for motivating individuals to behave in ways that society favors. Economic incentives inevitably

---

[5] Robert Merton (1957) introduced the idea of reference groups in his analysis of Stouffer's *American Soldier*. He proposed that people's aspirations are determined by the group to whom they compare themselves.

generate disparities in wages, hours of work, and household income. Insofar as these disparities involve differences in "trainable" skills (rather than motivation or innate ability) any inequality that results from the existence of such incentives should motivate the next generation to invest more heavily in its human capital. Incentives partly work through role modeling – children learn what they can expect from different behaviors by observing people who do different things. If the rewards for emulating a particular role model increase, more children will presumably emulate the model. Part of the growth in inequality in the United States between 1970 and 1990 was due to increased returns to schooling (Murphy and Welch 1992, 1993; Juhn et al. 1993). Because higher returns increase the incentive for children to stay in school, we would expect educational attainment to increase in response. I will separate these incentive effects from the effects of increased inequality among adults with the same amount of schooling.

Another kind of macro effect involves changes in political or social behavior. For example, inequality may affect voters' willingness to support redistributive policies, which could affect the tax rate (Perotti 1996, Alesina and Rodrik 1994). Likewise, high levels of inequality may encourage the rich to enroll their children in private schools, making them less interested in supporting public schools.

Because macro effects depend on social interactions, which may depend to some extent on the proximity of the rich to the poor, economic inequality may also affect the degree of economic segregation. Durlauf (1996) proposes a model in which growth in economic inequality leads to growth in economic segregation, because increasing inequality between the rich and poor means that they have less in common. An increase in economic segregation may then reduce poor children's well-being. Wilson's (1987) arguments about neighborhood effects suggest that economic segregation causes economic inequality rather than the other way around.

I estimate the effects of economic inequality on economic segregation and the effect of economic segregation on educational outcomes elsewhere.[6]

I begin by estimating the effect of the change in economic inequality between 1970 and 1990 on children's educational attainment. I then try to separate the effects of the incentive provided by the increasing returns to schooling from the effects of inequality. Next I try to separate the nonlinearity effects from the macro effects of inequality. Finally, I test the hypothesis that the effect of inequality is different for rich and poor children. College attendance has also risen. Ellwood and Kane (1999) show that between the early 1980s and 1992, the proportion of children in the lowest income quartile who went on to some post-secondary schooling increased from 57 percent to 60 percent. But the proportion of children in the top quartile going on for some post secondary schooling increased from 81 to 90 percent. Thus the increase in college enrollment was greater among affluent than among low-income children suggesting that the growth in inequality may have had different effects on children from different family backgrounds.

Data and Models.

Most research on the consequences of inequality uses aggregate data. For example, inequality in a geographic area such as a nation or state predicts an aggregate-level outcome such as the mortality rate or the crime rate for the geographic area. On the other hand most research on educational attainment uses individual-level data to predict years of schooling from measures

[6] In the short run income inequality can also affect public financial support for education in ways that are more or less mechanical. Given a progressive tax rate, anything that makes the rich richer should increase the revenue available to pay for schooling. On average across states, state and local taxes are slightly progressive. For example, in 30 large cities in 30 different states in 1995 a family of four with an income of $25,000 paid an average of 8.2

of family background. Some research combines individual-level data with aggregate data to predict, say, the effect of neighborhood social composition on children's educational attainment, holding constant their family background. These models of "neighborhood effects" or "school effects" have many well-known estimation problems (Tienda 1991, Manski 1993, Moffitt 1999). Nonetheless, they do solve some of the problems associated with using exclusively aggregate data. In this paper I estimate multi-level models similar to those used to estimate neighborhood effects. I estimate the effect of state-level inequality on individual children's educational attainment, often holding the children's family characteristics constant. In this section I describe the models that I use and my attempts to avoid major sources of bias.

Theory provides little guidance as to which geographic unit is most relevant for educational attainment, and inequality may have different effects at different levels of aggregation. If school financing plays a crucial role in educational attainment, the relevant units of aggregation are the political jurisdictions that fund schools and universities. Nationwide, states and local school districts typically provide roughly equal amounts of money for elementary and secondary schooling, with relatively little coming from the federal government. State funds generally tend to equalize district funding, so the state pays a greater share for poor districts and a smaller share for rich districts. It follows that decisions about the degree of inequality in school district funding are primarily made at the state level. If household income inequality affects educational outcomes primarily by affecting taxpayer's inclination to fund education, the state may be the right level of aggregation.

percent of their income in state and local taxes while a family of the same size with an income of $75,000 paid on average 9.6 percent of their income in taxes (U.S. Bureau of the Census, 1997).

On the other hand, school districts also share in funding decisions, and they probably make the most significant decisions about policies that affect educational outcomes. Income inequality *within* school districts might affect both voters' inclination to support tax revenues for schools and other school policies. To further complicate matters, parents often choose their school district partly on the basis of who lives there. If the same parental characteristics that cause parents to choose one district over another also affect their children's educational outcomes, and if these parental characteristics are not measured accurately, the estimated effect of school district characteristics on educational attainment is likely to be biased. This form of selection bias is likely to be less important for estimating the effect of state-level characteristics on educational attainment because parents are relatively unlikely to move to a different state in order to improve their children's educational prospects.

Theories about the effect of income inequality that involve interpersonal relationships are ambiguous about the most relevant geographic unit, because it is not clear how individuals select the other people to whom they compare themselves. If children compare themselves to the people they see on television, the nation as a whole is probably the relevant comparison group. If that were true the only way to study the impact of inequality would be to make cross-national comparisons or comparisons across time. If children compare themselves mainly to their neighbors and classmates, a relatively small geographic area such as a neighborhood or school district may be more relevant than a larger unit. But, as I have noted, selection problems are likely to be more serious at the level of the neighborhood or school district. In this paper I use states as aggregation units. I estimate the effect of inequality in smaller units of aggregation elsewhere.

Data. The educational attainment of a state's residents affects inequality as well as vice versa. This is not likely to be a problem in the short run, because while economic inequality in a state can affect the probability that a teenager will graduate from high school or enter college, it takes some time for the high school graduation rate or the college entrance rate to affect the dispersion of household incomes in a state. Using economic inequality in a state to predict the educational attainment of the adults in the state would pose more serious problems, because the direction of the causal arrow is unclear. This problem is exacerbated by the fact that many adults no longer live in the state where they were raised, and the distribution of income in a state may affect the kinds of migrants who settle there.

To solve these problems I use data from the Panel Study of Income Dynamics (PSID) to estimate the effect of state economic inequality measured during adolescence on children's chances of completing high school, entering college, and completing four years of college. (A full description of the data and the variables appears in the Appendix.) I treat these three educational outcomes as distinct, because income inequality is likely to have different effects on enrollment choices at different ages and in different kinds of schools. For example, if inequality is mainly important because income has nonlinear effects on educational attainment, inequality should have a smaller effect on high school graduation than on enrolling in college or graduating college, because attending high school costs far less than attending college.[7]

[7] It is true that the quality of the high school a child attends depends partly on parental income. Thus if students are more likely to stay in high quality than low quality schools, nonlinearity effects might not be zero. Likewise, while foregone wages are lower for sixteen-year olds than for nineteen-year olds, staying in high school involves *some* foregone earnings, so there are likely to be *some* nonlinearity effects as a result. But nonlinearity effects are still likely to be smaller for high school graduation than for post-secondary schooling.

Incentive effects also might be greater for college entrance and graduation than for high school graduation. The greatest increase in returns to schooling after 1980 were for college graduates; the increase in the returns to high school graduation was quite small.

My PSID sample includes all respondents who were in the data set both when they were twelve to fourteen years old and when the educational outcome of interest was measured. I count respondents as having graduated high school if they reported that they had completed twelve or more years of schooling when they are twenty years old. The analysis of high school graduation therefore includes all respondents who were in the sample at both age twelve to fourteen and at age twenty. I count respondents as having attended college if they reported completing one or more years of college by the time they are twenty-three years old. I count them as having graduated college if they report having completed at least sixteen years of schooling by the time they are twenty-three years old. For these outcomes, therefore, respondents had to be in the PSID sample both at ages twelve to fourteen years and at age twenty-three. Thus the sample is somewhat smaller for the measures of post-secondary schooling than for high school graduation (N = 3240 versus 3504).

Because states have a small number of PSID respondents in any given year, I created most of my measures of state characteristics from the 1970 1 percent Public Use Micro Sample (PUMS) of census data and from the 1980 and 1990 5 percent PUMS. Values for a few other state characteristics come from published sources described below. Appendix Table A1 provides means and standard deviations for the variables used in this paper.

*The Measure of Inequality***.** Different measures of income inequality could have different effects on educational attainment. Table 3 shows the correlation among several alternative measures of state income inequality and individual educational attainment. Three findings are notable. First, all but one of the correlations between state income inequality and individual educational attainment is negative and all are small (none differs from zero by more than .061). Second, no one measure of income inequality is a consistently better predictor of educational outcomes than the others. Third, the correlations among the three "comprehensive" measures of inequality (the Gini, the SD of log income, and the 90-10 ratio) are very high (0.925 to 0.963). The correlation between inequality in the top half of the distribution (the 90-50 ratio) and in the bottom half of the distribution (the 50-10 ratio) is also very high (0.721). Thus it is not clear that any one measure of economic inequality is likely to be better than another at predicting educational attainment.

In this paper I use the Gini coefficient as a measure of inequality. I use Census data to calculate the Gini coefficient for each state in 1970, 1980, and 1990. I then use linear interpolation to create a Gini coefficient for each state in each inter-censal year for which I have PSID data.[8] I assign children the level of inequality in their state when they were fourteen years old. I predict high school graduation among children who were fourteen years old between 1971 and 1987. I predict college outcomes for those who were fourteen years old between 1971 and 1984.

The Gini coefficient is quite sensitive to how income is top coded. I use consistent top codes for all years. The procedure for doing this is described in the Appendix. I tested the sensitivity of my estimates to alternative measures of inequality by substituting the variance of log income for the Gini coefficient. The variance of log income is especially sensitive to income inequality at the bottom of the distribution. This is important because income in the tails is less

---

[8] Nationwide inequality increased somewhat more rapidly in the later part of the 1970s than in the earlier part of the decade. During the early 1980s the increase in inequality was nearly linear. Any deviation from the linear trend is a

accurately reported than income in the middle of the distribution.[9] None of the results changed when I used this measure of inequality so I do not show these results in this paper.

Table 1 shows that the Gini coefficient of household income increased from .361 in 1970 to .368 in 1980 to .381 in 1990. The range of the Gini coefficient across states is large. In 1970 the lowest Gini coefficient was .320 and the highest was .427. In 1990 the lowest was .337 and the highest was .438. The standard deviation of the Gini coefficient across states was .027 in 1970, .018, in 1980, and .022 in 1990.

If economic inequality were a random accident, we could compare educational attainment in states with high and low levels of inequality, assume that all else was more or less equal across states, and treat observed differences in educational outcomes as a byproduct of economic inequality. The effect of inequality (*Gini*) in state *s* and year *t* on the educational attainment (*E*) of individual *i* would then be given by the value of β_g in the following model:

$$E_{\mathrm{ist}} = \beta_0 + \beta_g Gini_{\mathrm{st}} + \varepsilon_{\mathrm{ist}} \qquad (1)$$

where $\varepsilon_{\mathrm{ist}}$ is the usual random error term.

Measures of Other State Characteristics. States vary in many ways besides their level of economic inequality. Some of these differences are associated with both economic inequality and with educational outcomes. In this paper I try to estimate what would happen to educational attainment as a result of an exogenous change in economic inequality. An exogenous increase in inequality might result from a technological innovation that changed the skill needs of employers

source of measurement error in the inequality measure and thus it probably biases the coefficient of the Gini coefficient towards zero.

[9] Except in the tails of the distribution, American households get most of their income regular sources such as wages or Social Security. Households report their income from these sources fairly accurately. Households are the tails of the distribution get more income from irregular sources, and such income is less accurately reported. In addition, the incentive to mis-report income is greater at the tails of the distribution.

and therefore resulted in the wage premium increasing for some skills and decreasing for others. In response to such a technological "shock," states might differ in how much inequality increased, depending on the skill distribution in the state, the available mechanisms for increasing high-premium skills, the generosity of their social programs, the "culture" of the state, and many other factors. To estimate the effect of such a change in inequality, one must control all the exogenous determinants of inequality that affect educational attainment. To do this I first control a set of exogenous state-level determinants of inequality ($\boldsymbol{X}'$):

$$E_{ist} = \beta_0 + \beta_g Gini_{st} + \boldsymbol{\beta}_{\mathbf{x}} \boldsymbol{X}_{st}' + \varepsilon_{ist} \qquad (2)$$

A large body of research suggests that across countries racial and ethnic diversity, weak central governments, weak labor unions, highly decentralized wage bargaining, weak left-wing political parties, and the distribution of human capital affect inequality. Some of these same variables might play a role in explaining why some American states have more economic inequality than others, but no one has developed a strong theory of why states differ in their level of inequality.

I use Census data to compute the percent of state residents who were African American and the percent who were Hispanic in each year. I control both variables. Inequality could affect the racial composition of the state as well as vice versa. This could happen, for example, if white families moved to states with high levels of inequality because they thought they would have more opportunity in those states, while black families moved to states with low levels of inequality because they thought they would have better opportunities in those states. In this case a state's level of inequality would "cause" its racial composition, and controlling racial and ethnic composition of states would downwardly bias the estimated effect of inequality. But in practice inequality cannot have much effect on a state's racial composition, because the correlation between the percent black in a state in 1970 and 1980 is 0.98 and the correlation between 1980 and 1990 is again 0.98. The inter-year correlations for Hispanics are equally high. I therefore treat the percent of state residents who are black and the percent who are Hispanic as exogenous.

A state's average household income is also negatively correlated with inequality, and mean household income could obviously affect children's educational attainment. The correlation between mean household income and the Gini coefficient was -0.724 in 1970, -0.425 in 1980, and –0.559 in 1990. Like racial composition, inequality can affect mean income as well as the other way around. Barro (1999) provides empirical evidence that while inequality reduces economic growth in poor countries, it has a small positive effect on economic growth in rich countries. States in the US are all rich by international standards, so we might expect more unequal states to grow slightly faster. But it is not at all clear that the (unknown) factors that generate a correlation between inequality and subsequent economic growth at the national level would also apply to the economies of US states. Because inequality and mean income are correlated, and because it seems more likely that state income levels affect inequality than vice versa in US states, I control state mean income.

I also control the state unemployment rate measured at the same time as inequality is measured. Fluctuations in the unemployment rate are mainly attributable to short-term fluctuations in the business cycle and do not contribute much to the level of inequality in the state. The correlation between state unemployment rate and the Gini coefficient is -.047. However, among states with the same mean income, those with high levels of unemployment are likely to have more inequality because unemployment reduces the income of some state residents. Some research suggests that local unemployment rates are associated with greater college enrollment (Betts and McFarland 1995, Kane 1994).[10]

As a second way to control potentially relevant omitted variables, I include region fixed effects to control characteristics of the region that remain unchanged over the period of observation. These characteristics could include such things as climate, political history, and culture. I use the four Census regions for the Northeast, South, Midwest and West. An alternative strategy would be to control state fixed effects. Such a model would be equivalent to estimating the within-state effect of a change in inequality. This strategy has the advantage of controlling all invariant characteristics of states. However, it has three important disadvantages. First, it can magnify measurement error in independent variables, including the measure of inequality, which would downwardly bias the estimated effects. Second, if the lag structure of the model is not correctly specified, this too can result in downwardly biased estimates of inequality. Third, and perhaps most serious, including state fixed effects greatly reduces the degrees of freedom available to estimate the model, which in turn increases the standard errors of the estimates. I experimented with both state fixed effects and smaller regions. I discuss models with state fixed effects below.

I also control year fixed effects to account for the secular national trend in educational outcomes. With both region and year fixed effects, all variation in inequality derives from a combination of changes in inequality within states and differences in equality among states in the same region. Thus the model becomes

$$E_{\text{ist}} = \beta_0 + \beta_g Gini_{\text{st}} + \boldsymbol{\beta}_{\mathbf{x}} \boldsymbol{X}_{st}' + \gamma_r + \gamma_t + \varepsilon_{\text{ist}} \qquad (3)$$

---

[10] Other earlier research finds little effect of local labor market conditions on college enrollment (Manski and Wise 1983, Grubb 1988).

where γ_r is a set of dummies for the four regions and γ_t is a set of dummies for years. In this model $\boldsymbol{X'}$ represents exogenous state characteristics that have changed over time described above.

The third way I control unobserved state characteristics is to create an instrument for inequality. One important cause of increased economic inequality over the last 30 years has been changes in the returns to skills. These changes have increased relative wages in some industries and decreased relative wages in others, resulting in more inequality. I use this fact to create an instrument for predicting changes in economic inequality that are arguably determined by national economic forces and thus exogenous with respect to other changes at the state level. The reasoning for the instrument is that if a technological "shock" raised wages in some industries and not others, and if the state's industrial mix cannot respond quickly, the industrial mix when the shock occurred should be a good exogenous predictor of the change in inequality over the short to medium run. The industrial mix at the time of the shock would presumably affect educational attainment at that time, but it would affect subsequent changes in educational attainment only through its affect on inequality, including its effect on returns to schooling.

I first measure the industrial mix of states in 1970 by assigning each worker between the ages of 25 and 64 in a state the national mean for the earnings of workers in the same three-digit industry. I then calculate a new variable that assigns each worker in 1970 the average 1980 wage for his industry. I repeat this using average wages for 1990. I calculate the standard deviations of the 1970, 1980, and 1990 measures in each state. This measure, which I call "industrial mix," is the amount of inter-industry income inequality we would expect to find in the state if the industrial mix had not changed between 1970 and 1990. The R^2 when I regress the change in the

Gini coefficient on the change in industry mix is .365. I use this instrument in a two-stage least squares model.

I also control state mean returns to schooling. Imagine two states where inequality grows but mean income does not change. In state A inequality within education groups increases, perhaps because employers increase the pay of workers with non-cognitive skills not usually learned in school. In state B inequality grows because the returns to schooling increase. In both states the rich get richer and the poor get poorer. If the effects of inequality are due entirely to nonlinearity or relative deprivation, the change in children's educational outcomes should be the same in both states, because these effects do not depend on the returns to schooling. If the effect of inequality on educational attainment is due entirely to the incentive effect of increased returns to schooling, educational attainment should increase in state B but not in state A.

My measure of returns to schooling is the average effect of an extra year of schooling on log wages in a given state and year, estimated for workers aged 18 to 65 using the following model:

$$\ln W_{is} = \beta_0 + \beta_s S_{is} + \varepsilon_{is} \qquad (4)$$

where $\ln W_{is}$ is log wages for individual i in state s, and S is the individual's schooling. Thus β_s is the percentage increase in wages due to an additional year of schooling. β_s averages .062, with a standard deviation of 0.011. This means that wages rise by an average of 6.4 percent for each extra year of schooling.[11]

[11] I experimented with over 12 different measures of returns to schooling, using different age groups, different functional forms and separating returns for men and women. I selected the measure that increased R^2 the most in when added to the estimation model. The measure I use also corresponds best to economic theory about the functional form of returns to schooling and produces an estimated return to a year of schooling that is consistent with previous research on the returns to schooling (Winship and Korenman 1999, Mayer and Knutson, Ceci 1991).

I estimate the effect of state returns to schooling (R_{st}) when a child was fourteen years old on his or her eventual educational attainment assuming that the child decides on his or her level of schooling based on the information available before the decision is made. Thus the model becomes:

$$E_{ist} = \beta_0 + \beta_g Gini_{st} + \beta_r R_{st} + \boldsymbol{\beta}_{\mathbf{x}} \mathbf{X'}_{st} + \gamma_r + \gamma_t + \varepsilon_{ist} \qquad (5)$$

This model makes sense if returns to schooling are an exogenous determinant of both educational attainment and inequality. Then β_g is the effect of an exogenous increase in inequality that is not related to schooling. Such an exogenous increase could occur because of a state's "luck" in its industrial mix when a technological shock increases wages in some industries but not others, or for other reasons.

However, the return to schooling may depend partly on the level of inequality. States with few colleges and universities may, for example, have higher returns to schooling because they offer fewer opportunities for post-secondary schooling, reducing the supply of highly educated people in the state and increasing their wage premium. Absent migration, states with less post-secondary schooling opportunities will then have more inequality. This would also presumably increase the incentive to go to college. Over time states would then reach an equilibrium in which they provided the "right" supply of college graduates to meet the state's demand for college graduates. This implies that over the long run we would expect little variation across states in returns to schooling. The standard deviation for the returns to schooling in my sample is small (.011, making the coefficient of variation .177).[12]

[12] This also implies that the opportunity for post-secondary schooling in a state is a potentially important omitted variable. However, the ratio of the number of currently enrolled post-secondary students in a state to the number of eighteen to twenty-four year olds in the state is very weakly related to returns to schooling (r = .001).

Separating Nonlinearity from Macro Effects. To estimate the role of nonlinearity, we first must determine the approximate functional form of the relationship between parental income and children's educational attainment. I average household cash income over the three years when a child was age twelve, thirteen and fourteen. Table 4 shows that any of three nonlinear functions of this measure of household income is a better predictor of educational outcomes than a linear variable. The cube root of income yields a slightly better fit than log income for both graduating high school and enrolling in college, but the difference between the Chi-square values is small. Not surprisingly, adding dummies for the poorest and richest income decile yield the best predictions of all three educational outcomes.[13] These data do not allow one to choose among the non-linear specifications. I use log income as a measure of nonlinearity because of its theoretical appeal and ease of interpretation. To estimate the effect of nonlinearity I estimate the following equation:

$$E_{ist} = \beta_0 + \beta_Y \ln Y_{ist} + \beta_g Gini_{st} + \boldsymbol{\beta}_{\mathbf{x}} \boldsymbol{X'}_{st} + \gamma_r + \gamma_t + \varepsilon_{ist} \qquad (6)$$

In this equation $\beta_Y \ln Y_{ist}$ captures the effect of household income. In this model β_g captures the remaining "macro" effects of inequality, which operate independently of a household's own income. If β_g is negative in equation 3 and zero in equation 4, the entire cost of inequality is traceable to nonlinearity effects.[14]

Most models of neighborhood or school effects begin with an individual-level model of an outcome such as high school graduation, then add an aggregate-level variable for the

---

[13] Surprisingly, the effect of being in the poorest decile is positive and the effect of being in the richest decile is negative for all educational outcomes. This suggests that poor children get more schooling than we expect from the log function and rich children get less.

[14] If log household income is not the right functional form, the estimate of β_g could be biased. However, the three functional forms that I test produce very similar estimates of the non-linearity in the effect of parental income on educational attainment, so this is not likely to be a large source of bias.

neighborhood or school social composition. For example, a reduced form model of the effect of school social composition on high school graduation might predict high school graduation from family background characteristics such as parental education and income and a school-level variable such as the mean level of family income among students (Crane 1991a, 1991b, Evans et al. 1992, Mayer 1991). In such models family background variables are intended to control for the fact that parents select the schools their children attend and the neighborhoods in which they live. This form of selection bias is unlikely to be a problem for a state-level model. But to see if parental selection across states with different levels of inequality is a problem I control two other measures of family background, namely race and parental education.

Economic inequality can affect tax revenues, and tax revenues can affect both expenditures for schooling and other child inputs that help children succeed in school (such as higher quality or lower cost health care). To see if the macro effect of inequality operates through these mechanisms, I include a measure of tax revenue per person for each state in each year. I also include a measure of state expenditures on primary and secondary schooling per pupil. I could not find a consistent measure of either state expenditures on post-secondary schooling or the costs of college to students in different states.

In all models the standard errors are corrected for the fact that individuals are clustered in states and years.

Results.

Model 1 in Table 5 shows the effect of the Gini coefficient on educational outcomes controlling no characteristics of states or families. These results are from a probit model. The cell entries are partial derivatives evaluated at the mean of the distribution. For an average child in a state with an average Gini coefficient, the probability of graduating high school is 0.832. Living in a state with a Gini coefficient one standard deviation (.02) above the national average is associated with a decline in the high school graduation rate from 0.832 to 0.820. This is a small effect with a large confidence interval. The association between income inequality and a student's chances of entering or completing college is also negative but very weak.

Model 2 in Table 5 controls region and year fixed effects. In this model the negative effect of inequality on graduating high school is smaller, and its effect on college entrance and graduation becomes positive. This model suggests that an average child's probability of enrolling in college increases by .040 when inequality increases by a standard deviation. The mean probability of enrolling in college is .427, so this is a 9.4 percent increase. This effect is statistically significant at the .05 level. The fact that controlling fixed effects changes the apparent effect of the Gini coefficient so much shows that unmeasured differences among regions are important correlates of both the Gini coefficient and educational attainment.

Model 3 in Table 5 controls a state's mean income and unemployment rate, the percent of state residents who are African American, and the percent who are Hispanic in addition to region and year fixed effects. With these controls the effect of inequality on graduating high school becomes positive (but with a large confidence interval). The positive effect of inequality on enrolling in college and graduating college becomes even larger. This is because richer states have less inequality than poorer states ($r = -.441$), and states with a lot of economic inequality also have large African American and Hispanic population.

Table 6 repeats Models 2 and 3 from Table 5 using OLS. OLS results are easier to interpret than probit results, and are comparable to the two-stage least squares model that I use later. The OLS and probit models yield qualitatively similar results. Both the probit and OLS models show that with year and region fixed effects the effect of state income inequality is small, negative, and statistically insignificant for high school graduation; moderate, positive, and statistically significant for college entrance; and fairly small, positive, and statistically insignificant for college graduation.

Both OLS and probit models suggest that controlling states' mean income, racial composition, and ethnic composition increases the positive effect of the Gini coefficient. If omitting mean state income and the racial and ethnic composition of states results in downwardly biased estimates of the effect of inequality on educational attainment, omitting other exogenous state characteristics may do the same. Table 7 shows two-stage-least-squares (2SLS) estimates with industrial mix as an instrument for the Gini coefficient. The first Model in Table 7 is the same as Model 2 in Table 6. In this model the effect of the Gini coefficient on high school graduation is large, negative, and statistically significant. The effect of the Gini coefficient on graduating college is large and positive. The effect of the Gini coefficient on enrolling in college is small and statistically insignificant. The fact that the 2SLS results differ from the OLS results suggests that there may be important omitted variables in the OLS model. For the remainder of this paper I mainly discuss the 2SLS results. Below I describe the sensitivity of these results to various changes in the models.

As noted above economic inequality grew partly because of an increase in the returns to schooling. The positive effect of inequality on college going could be partly attributable to rising returns to schooling.[15] Model 2 in Table 7 controls the state-specific returns to schooling in the year a child was fourteen years old. Not surprisingly, students in states with high returns to

[15] Rising returns to schooling is not the main source of inequality growth. The within-education group variance of income rose almost as fast as the between-group variance of income, and educational attainment only accounts for 15-20 percent of the variance in income initially.

schooling are more likely than other students to graduate high school, enroll in college, and graduate college. These results suggest that a 1 percentage point increase in the state's return to schooling increases the chances of graduating high school by 4 percentage points, the chances of going to college by 6.5 percentage points, and the chances of graduating college by 2.3 percentage points. If the return to schooling were entirely exogenous, Model 2 in Table 7 would provide a better estimate of the "true" effect of inequality on educational attainment than earlier tables did. It suggests that a one standard deviation increase in inequality reduces children's chances of graduating high school by .072 (8.8 percent) but increases their chances of graduating college by .079 (39 percent). The effect of the Gini coefficient on enrolling in college is relatively small and statistically insignificant.

Although Model 2 in Table 7 is in principle be the reduced form estimate of the effect of the Gini coefficient, the combination of growing inequality and rising returns to schooling may have caused people to exaggerate the benefits of additional schooling. In particular, people in states where returns to schooling did not rise much but inequality grew substantially may have stayed in school because they (or their parents) knew that the labor market was increasingly competitive, had heard that going to college was the best way to protect themselves from economic stringency, and either did not know that the market for college graduates was soft in their own state or were willing to move elsewhere if they could not get the kind of job they wanted in their own state. Thus the effect of the Gini coefficient in Table 7 may be greater than it would be if there had been no increase in the returns to schooling. Nonetheless these results suggest that not all the positive effect of inequality is due to the incentive provided by the rising returns to schooling.

I now turn to the question of *why* inequality seems to have these effects. Table 8 tries to separate the effect of inequality on households' own incomes from the macro effects of inequality. In this model the thought experiment is what would happen to two children with the same family income (averaged over several years) in states with the same mean income but different levels of inequality.

To understand the nonlinearity effect, imagine that children's educational attainment is a linear function of the logarithm of income, so that poor children benefit more than rich children from additional income. If we estimate the effect of state-level inequality without controlling household income, inequality will have a negative effect on educational attainment simply because individual income has a nonlinear effect on educational attainment. Once we control a household's own income, however, the effects of income inequality will be reduced to zero unless inequality also has "macro" effects that operate independently of a household's own income. But since the effect of inequality on college attendance is *positive* in most of the models presented so far, it appears that the net benefits of inequality outweigh the costs, including the costs of nonlinearity. Thus, when I control household income, the positive effect of inequality should *increase* in magnitude.

Comparing Model 2 in Table 7 to Table 8 shows that this is just what happens. Controlling log household income reduces the negative effect of the Gini coefficient on high school graduation and increases its positive effect on enrolling in college and graduating college. This implies that negative effects of inequality result from non-linearity in the relationship between educational attainment and parental income (or its correlates). Because the effect of the Gini coefficient on college graduation remains large and statistically significant, these results imply that the "macro" benefits of inequality out-weight the nonlinearity costs.

Not surprisingly, the effect of family income on all three outcomes is large. In addition, controlling family income reduces the effect of the returns to schooling, especially its effect on graduating college. The state returns to schooling are positively correlated with parental income because parents' income partly depends on the returns to schooling. The effect of higher returns to schooling remains important for high school graduation and college enrollment. This is what we would expect if children graduate high school and enroll in college because they hope to earn more in the future, but the high monetary and non-monetary costs of college drive out children from low-income families.

If the level of inequality in a state is correlated with other family characteristics independent of income, then omitting these characteristics could bias the estimated effect of inequality on children's outcomes. Table 9 tests this hypothesis by controlling parents' education and race. These controls hardly change the effect of the Gini coefficient on any of the measures of educational attainment. Thus selective migration is not likely to be an important source of bias, since other family background characteristics that affect migration patterns are likely to be correlated with parents' race and education.

Controlling parental education and child's race reduces the effect of household income by between a quarter and a third, depending on the outcome. This is because the apparent effect of household income on children's educational attainment is partly due to parental characteristics that both increase the parents' income and improve their children's educational attainment (Mayer 1997). These omitted variables result in upwardly biased estimates of the effect of family income. Thus the estimates of the nonlinearity effect of income inequality is not entirely due to income per se.

To see if the macro effect of inequality affects educational attainment through its effects on states' tax revenues or state expenditures on primary and secondary schools, I control both in Table 10. Both measures are positively correlated with the three schooling outcomes, although the correlation is very weak. They are also positively correlated with state mean income and negatively correlated the Gini coefficient. The effect of both tax revenues and school expenditures is small, positive, and insignificant for all three outcomes. It is not surprising that these are insignificant because they are highly correlated (r = .73). Controlling these increases the negative effect of inequality on high school graduation. This suggests that higher tax revenue and school spending partly ameliorate the effect of inequality on high school graduation. Controlling tax revenue and school spending slightly raises the positive effect of inequality on college going. This means that the net effect of these two factors is to decrease college going.

Macro effects of inequality involve social comparisons and "relative income." If we compare children with the same absolute income, those in high-inequality states have lower incomes relative to the rich and higher incomes relative to the poor than those in low-inequality states. If people mainly compare themselves to the rich, the relative income hypothesis implies that households at any given income level will feel poorer when inequality rises. This could reduce children's educational attainment. If people mainly compare themselves to the poor, however, households at any given income level will feel richer when inequality rises. This could increase children's educational attainment. The effect of the Gini coefficient on high school graduation is negative in Table 10 but its effect on college graduation is positive. This could be because social processes affect college-bound students differently from students likely to only graduate high school.

Poor Children. Table 11 shows results separately for "high" and "low" income children. I divide the sample at the median income. "High-income" children are thus those whose household income is in the top half of the income distribution. "Low-income" children are those whose household income is in the bottom half of the income distribution. I use this division of the sample to preserve enough high and low-income cases for a meaningful analysis. Dividing the sample reduces the t-statistics for all coefficients. Nonetheless, because there is good reason to expect that neither inequality nor other state characteristics affect high and low income children in the same way, these fully-interacted models are instructive even though the division of the sample is somewhat arbitrary.

The first Model in Table 11 shows that the effect of the Gini coefficient on high school graduation is positive but very small and statistically insignificant for high-income children. Its effect is large, negative, and statistically significant for low-income children. In contrast, the effect of the Gini coefficient on graduating college is large, positive, and statistically significant for high-income children but negative, small and statistically insignificant for low-income children. The effect of the Gini coefficient on enrolling in college is also positive for high-income children and negative for low-income children, although the t-statistic is relatively small. Thus high-income children appear to benefit from inequality while low-income children appear to be hurt by inequality.[16]

The second model in Table 11 controls parental income, race, and education and the third model adds state tax revenues and state school expenditures per child. These controls do not

[16] An alternative to dividing the sample might be to simply add an interaction for the Gini coefficient and a child's own household income. That would in principle allow one to determine the income level at which inequality has the greatest benefit or cost. However, as it turns out household income interacts with several variables in the model. Ignoring these other interactions produces misleading estimates of the effect of inequality at various income levels.

change the basic story. The fact that the macro effects of inequality appear to reduce low-income children's educational attainment and increase high income children's educational attainment suggests that parents or children may mainly compare themselves to the middle of the income distribution. If people mostly compare themselves to the middle of the income distribution, all else equal as inequality grows those below the mean become relatively poorer while those above the mean become relatively richer. In this case the poor should be worse off and the rich better off even if their own income does not change.

Robustness. Most of the results that I have discussed so far are based on two-stage-least-squares models that use industrial mix as an instrument for inequality in an effort to control unobserved state characteristics. I rely on this model because of the evidence of omitted variable bias. Because there is always uncertainty about the validity of any specific instrument, I examine the robustness of these results using several other models. I compare the reduced form model in Table 7 (model 2) with the same model using OLS rather than 2SLS, with a 2SLS model with state rather than region fixed effects, and with an OLS model with state fixed effects.

Table 12 shows these comparisons. In all models the effect of income inequality on college graduation is large and positive. Its effect is smaller in models with region fixed effects than in models with state fixed effects. It is smallest in the OLS model with region fixed effects. The effect of income inequality on enrolling in college is positive in all models except one. The only model in which the effect of the Gini coefficient on enrolling in college is statistically significant at even the .10 level is the OLS model with region fixed effects. The effect of the Gini coefficient on high school graduation is small, usually negative and never statistically significant at the .05 level. Its effect is largest in the 2SLS model with region fixed effects.

One could include interaction terms for all variables, which would avoid the arbitrary division of the sample. But

The conclusions in this paper are based on the 2SLS model with region fixed effects. The robustness tests suggest that my conclusions could understate the beneficial effects of inequality on college graduation and overstate the negative effect of inequality on high school graduation. Regardless of the model, the benefits of inequality accrue to the most affluent half of the income distribution.

Conclusions

Taken as a whole the results in this paper suggest five conclusions. First, the growth in inequality since 1970 did not increase high school graduation and probably decreased it. The reduced form model in Table 7 suggest that a one standard deviation increase in inequality results in a decline in high school graduation of about 10 percent.

Second, the growth in inequality since 1970 may have increased college enrollment by a small amount. The reduced form model in Table 7 suggest that a one standard deviation increase in inequality results in a 2.7 percentage point increase in college enrollment (6 percent), but the estimate has a large confidence interval and is not robust.

Third, the growth in inequality since 1970 increased college graduation considerably. The reduced form estimate in Table 7 suggests that a standard deviation increase in inequality results in a 7.4 percentage point (40 percent) increase in college graduation. These results seem to be quite robust.

Fourth, the effect of inequality is largely due to what I have called macro effects, not to the non-linear relationship between parental income and children's outcomes, or to the incentive provided by increasing returns to schooling.

such models are difficult to interpret and in this case provide results similar to those presented in the paper.

Fifth, low-income children did not benefit from increases in income inequality. The growth in inequality probably reduced high school graduation among low-income children. The increase in college graduation due to the growth in inequality was confined to high-income children. This is not primarily because the growth in inequality increased the income of rich children but not poor children. The differential effect of inequality on high and low-income children remains even when their own family income is controlled. This suggests that if relative income is the mechanism through which these marco effects operate, children's or parent's reference for income comparisons may be roughly the middle of the income distribution.

It is also important to emphasize that state-level inequality explains very little of the variation in children's educational attainment. The Gini coefficient alone explains hardly any of the difference in children's educational attainment. Models that include only state-level variables explain less than 5 percent of the variation in any measure of children's educational attainment. Family background characteristics are far more important to children's educational attainment than the inequality level of the state in which a child lives. This paper focuses on inequality at the level of the state. Inequality at smaller levels of aggregation could be either more or less important than inequality at the state level. Inequality at the national level may also be more important than inequality in a state.

If true, the results in this paper present a problem. The increase in income inequality between 1970 and 1990 appears to have in increased mean educational attainment, which presumably will contribute to economic growth. However, only the most affluent half of the income distribution benefited from the growth in inequality. This is likely to contribute to inequality in the next generation. This suggests that it is important to find ways to reduce the potentially harmful effects of inequality on low-income children.

References

Alessina A. and D. Rodrik. 1994. “Distributive Politics and Economic Growth” *Quarterly Journal of Economics,* 109: 465-490.

Atkinson, Anthony B. 1970. “On the Measurement of Inequality.” *Journal of Economic Theory*. Vol. 2.

Atkinson, Anthony B. 1983. *The Economics of Inequality.* Oxford: Oxford University Press.

Barro, Robert J. 1999. “Inequality, Growth, and Investment.” Harvard: unpublished manuscript.

Betts, Julian and L. McFarland. 1995. “Safe Ports in a Storm: The Impact of Labor Market Conditions on Community College Enrollment.” *Journal of Human Resources.* 30(4): 741-765.

Ceci, Stephen J. 1991. “How Much Does Schooling Influence General Intelligence and Its Cognitive Components? A Reassessment of the Evidence.” *Developmental Psychology,* 27(5):703-722.

Crane, Jonathan. 1991a. “Effects of Neighborhoods on Dropping Out of School and Teenage Childbearing” in Christopher Jencks and Paul Peterson (Eds) *The Urban Underclass.* Washington, D.C.: Brookings Institution Press.

Crane, Jonathon. 1991b. “The Epidemic Theory of Ghettos and Neighborhood Effects on Dropping Out and Teenage Childbearing.” *American Journal of Sociology* 96(5):1226-1259.

Durlauf, Steven. 1996. “A Theory of Persistent Income Inequality.” *Journal of Economic Growth.* 1:75-93.

Ellwood, David A. and Thomas J. Kane. 1999. “Who Is Getting A College Education? Family Background and the Growing Gaps in Enrollment.” Unpublished manuscript, Harvard University.

Evans, William, William Oates and R.M. Schwab. 1992. “Measuring Peer Group Effects: A Study of Teenage Behavior.” *Journal of Political Economy* 100: 966-91.

Firebaugh, Glenn. 1999. “Empirics of World Income Inequality.” *American Journal of Sociology.* 104(6):1597-1630.

Grubb, N. W. 1988. Vocationalizing Higher Education: The Causes of Enrollment and Completion in Public Two-Year Colleges 1970-1980. *Economics of Education Review* 7(3): 301-19.

Jencks, Christopher and Susan E. Mayer. 1990. “The Social Consequences of Growing Up in a Poor Neighborhood.” in Laurence Lynn and Michael McGeary, editors. *Inner- City Poverty in the United States*. Washington: National Academy Press.

Juhn, Chinhui, Kevin Murphy and Brooks Pierce. 1993. “Wage Inequality and the Rise in Returns to Skill. *Journal of Political Economy.* 101(3):410-42.

Kane, Thomas. 1994. “College Entry by Blacks since 1970: The Role of College Costs, Family Background and the Returns to Education.” *Journal of Political Economy.* 102(5):878-911.

Manski, Charles F. 1993. “Identification of Endogenous Social Effects: The Reflection Problem.” *Review of Economic Studies* 60(3): 531-542.

Manski, Charles and David Wise. 1983. *College Choice in America.* Cambridge, MA: Harvard University Press.

Mayer, Susan E., 1991. “How Much Does A High School’s Racial and Socioeconomic Mix Affect Graduation and Teen Fertility Rates?” in Christopher Jencks and Paul Peterson (Eds) *The Urban Underclass.* Washington, D.C.: Brookings Institution Press.

Mayer, Susan E. 1997. *What Money Can’t Buy: Family Income and Children’s Life Chances*. Cambridge, MA: Harvard University Press.

Mayer, Susan E. and David Knutson. 1999. “Does the Timing of Schooling Effect How Much Children Learn?” in *Earning and Learning: How Schools Matter* by Susan E. Mayer and Paul E. Peterson. Washington, D.C.: Brookings Institution Press.

McLoyd, Vonnie. 1990. "The Impact of Economic Hardship on Black Families and Children: Psychological Distress, Parenting and Socioemotional Development". *Child Development*, 61: 311-346.

Merton, Robert. 1957. *Social Theory and Social Structure.* Glencoe: The Free Press of Glencoe.

Morris, Martina and Bruce Western, 1999. “Inequality in Earnings at the Close of the Twentieth Century.” *American Review of Sociology*. 25:623-57.

Moffitt, Robert. 1999. “Policy Interventions, Low-Level Equilibria and Social Interactions” Johns Hopkins University, Unpublished Manuscript.

Murphy, Kevin and Finis Welch. 1993. “Industial Change and the Rising Important of Skill” in *Uneven Tides* edited by Sheldon Danziger and Peter Gottschalk. New York: Russell Sage Foundation.

Murphy, Kevin and Finish Welch. 1992. “The Structure of Wages.” *Quarterly Journal of Economics* 107 (February): 215-236.

Ogbu, John. 1981. "Origins of Human Competence: A Cultural-Ecological Perspective." *Child Development*, 52:413-429.

Perotti, Roberto. 1996. "Growth, Income Distribution, and Democracy: What the Data Say" *Journal of Economic Growth*, 1: 149-187.

Tienda, Marta. 1991. "Poor People and Poor Places: Deciphering Neighborhood Effects on Poverty Outcomes." In *Macro-Micro Linkages in Sociology* edited by J. Haber. Newberry, CA: Sage.

U. S. Bureau of the Census. 1998. *Statistical Abstract of the United States.* Washington, D.C.: U.S. Government Printing Office.

Welch, Finis. 1999. "In Defense of Inequality." *American Economic Review*. 89(2): 1-17.

Wilson, William Julius. 1987. *The Truly Disadvantaged.* Chicago: University of Chicago Press.

Winship, Christopher and Sanders Korenman. 1999. "Economic Success and the Evolution of Schooling' in *Earning and Learning: How Schools Matter* by Susan E, Mayer and Paul E. Peterson. Washington, D.C.: Brookings Institution Press.

Appendix
Description of the Data and Variables

PSID Data

The PSID sample comes from the 1993 wave of the PSID. It includes individuals who were in the sample when they were ten to fourteen years old and when the outcome was measured. High school graduation was measured when children were twenty years old. Thus the high school graduation sample includes all people in the Individual File of the PSID who were ages 20 through 37 in 1993 and who are not missing data on independent variables. College enrollment and graduation was measured when children were twenty-four years old. Thus the college sample includes all people who were ages 23 through 37 in 1993 and who are not missing data on independent variables. The PSID initially oversampled the low-income, non-Hispanic population of the United States. We weighted the observations in our sample to account for this sample design.

The variables in both samples were constructed by pooling across the 26 currently available waves of the PSID Family File (years 1968 through 1993). We assigned values to each individual based on that individual's age rather than a particular year. For example, we average family income when children were aged twelve to fourteen. Thus it was averaged over 1985 to 1987 for children born in 1973 and over 1990 to 1992 for children born in 1978. Following is the description of the variables created with PSID data. The means and standard deviations of all variables are in Table A1. The correlations among most variables are in Table A2

High School Graduate: A dummy variable set equal to 1 if the individual had either earned a GED or had 12 years of education by age 20, 0 otherwise.

Enrolled in College: A dummy equal to 1 if the individual had completed 13 years of education by age 23, 0 otherwise.

Completed College: A dummy equal to 1 if the individual had completed 16 years of education by age 23, 0 otherwise.

Log Family Income: The child's family income was averaged over the three years when the child was age twelve through fourteen. All income values were coded to1998 dollars using the CPI-U-X1 index for inflation. We then calculated the natural logarithm of the averaged value.

Parental Education: This is the highest year of schooling completed by the mother when the child was aged fourteen unless that value was missing. If this was missing, we used the mother's education when the child was age thirteen. If this value was missing, we used the value when the child was age twelve. If that is missing, we used the value at age eleven. If all of these values were missing, then we assigned the father's education when the child was age fourteen.

Black: A dummy variable set equal to 1 if the child was African American, 0 otherwise.

Census Data

Most of the state level variables used in this analysis were created from the 1970, 1980, and 1990 Public Use Microdata Sample (PUMS) from the U.S. Census. In 1980 and 1990 we used the full 5 percent samples. In 1970 we use the 1 sample because that is what is available.

Because state-level data is attached to individual cases, the state-level means and standard deviations in tables are approximately weighted by the state population.

Household Income. Household income was computed by summing the components of income for each person in a household. Using components of person's income rather than person's total income increases the detail available at the upper tail of the distribution by avoiding Census Bureau top-coding of person's total income. To limit the detrimental effect on comparability of changes in the Census Bureau's top-coding of income components, we created uniform income components and top-codes that we used in all years. There were six comparable income categories for all years: Wages/Salary, Non-farm Self-employment, Farm Self-employment, Social Security, Public Assistance, and Other. Each category was top-coded by reassigning values above the lowest 99th percentile of positive values among the years to the median of all values across years that lie above that lowest 99th percentile. The same was done for negative values using the highest 1st percentile as the cutoff. All dollars are adjusted to1998 dollars using CPI-U-X1.

The resulting components are then summed to get household income. Our state-level measures of income, including income inequality, were then calculated from the resulting household incomes, both at the household level and the person level. Persons in group quarters were excluded from all calculations.

Percent Black and Percent Hispanic. I estimate these variables using 1970, 1980, and 1990 PUMS data and then use linear interpolation to assign values for the state in the year when the child was fourteen years old.

Industrial Mix. The industry instruments are the state-level standard deviations across 3-digit industries of the mean person's weekly wages within each industry, weighted by the number of currently working persons age 25-64 in each industry (the "industry mix"). Weekly wages were made comparable following the same procedure as for each income component (see above), with real values above $2906.28 set to $3877.82. The mean weekly wages were calculated for each industry and the state-level person-weighted standard deviation of these means for each year became the industry instruments.

The correlation between industrial mix and the Gini coefficient is .561. As with the Gini coefficient the correlation between industrial mix and the measures of educational attainment is small (-.033 with high school graduation, .020 with enrolling in college and .033 with graduating from college). In the first stage the effect of industrial mix on the Gini coefficient is always large and statistically significant. When I regress the measures of educational attainment on the Gini coefficient and other controls and industrial mix, industrial mix has a small and statistically insignificant effect.

Returns to Schooling. This variable is the calculated by regressing log wages on years of schooling for each state in each year. Thus it represents the state average percent increase in wages for each additional year of schooling. I experimented with numerous alternative measures of returns to schooling. I created an alternative variable based on the relationship between household income and schooling. An individual might evaluate the potential returns to schooling in terms of his or her own potential wages. But individuals may also think of both the potential wages and the potential spouse that a level of schooling can "earn." I also created several ratios to compare the wages and household income for different levels of schooling. These include the ratio for less than high school and high school graduates, high school graduates to those with some college and high school graduates to those who graduated college. Finally I created age-specific returns to schooling because children might assess their own potential return to schooling by comparing incomes among young people or by comparing incomes among their parent's generation. I selected the variable with the largest effect on educational attainment, log returns to schooling.

Other State Level Variables

State tax revenue per capita: This variable is equal to the total state tax collection divided by the population of the state when the individual was 14. The source for this information is U.S. Bureau of the Census, *Statistical Abstract of the United States* annual volumes *1970* (91st edition.) through 1990 (110th Edition) Washington, DC, various years.

Elementary and secondary public school expenditures per capita: This variable is equal to the state's elementary and secondary public school expenditures divided by the state population of individuals ages five through seventeen when the individual was 14. State population data for 5-17 year olds is from the U.S. Census Bureau's web page. Elementary and secondary public school expenditures are from U.S. Bureau of the Census, *Statistical Abstract of the United States: 1970* (91st edition.) annual volumes *1970* (91st edition.) through 1990 (110th Edition), Washington, DC, various years.

Table A1 Means and Standard Deviations

Variable	Mean	Standard Deviation
High School Graduate	0.832	0.374
Enrolled in College	.427	.494
Graduated College	.182	.386
State Mean Household Income in 1998 dollars	36.728	4.526
Percent Black	11.460	7.744
Percent Hispanic	4.971	6.170
Child's Race is Black	0.155	0.362
Gini Coefficient of Household Income	0.402	0.018
Returns to Schooling	0.062	0.011
Unemployment Rate	8.396	2.573
Log Household Income in 1998 Dollars	10.759	0.657
Parent's Education	11.431	2.661
State Tax Revenue Per Capita in 1998 Dollars	998.207	217.846
State School Expenditures in 1998 Dollars	2934.891	780.330
College Enrollment	0.258	0.054

Source: PSID data described in Appendix.
Notes: These means and standard deviations are based on the 3,504 cases in the sample for models predicting high school graduation. The sample for college outcomes is 3,240 cases so the means differ slightly.

Table A2 Correlations among Variables

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. High School Graduate	1.000												
2. Enrolled in College	.022	1.00											
3. Graduated College	.031	.547	1.00										
4. Mean Household Income	0.055	0.101	0.068	1.000									
5. Percent Black	-0.044	-0.089		-0.305	1.000								
6. Percent Hispanic	0.085	0.075	0.042	0.369	-0.135	1.000							
7. Child's Race is Black	-0.114	-0.116	-0.137	-0.982	0.330	-0.005	1.000						
8. Gini Coefficient	-0.030	-0.013	-0.022	-0.492	0.506	0.188	0.156	1.000					
9. Returns to Schooling	0.030	0.036	0.030	0.148	0.394	0.334	0.104	0.447	1.000				
10. Log Household Income	0.256	0.333	0.269	0.263	-0.208	0.120	-0.374	-0.184	-0.609	1.000			
11. Parent's Education	0.271	0.347	0.276	0.257	-0.213	0.092	-0.229	-0.113	0.030	0.425	1.000		
12. State Tax Revenue	0.040	0.090	0.063	0.576	-0.196	0.283	-0.031	-0.027	0.192	0.166	0.167	1.000	
13. Expenditures on Schools	0.064	0.133	0.103	0.716	-0.263	0.217	-0.096	-0.118	0.279	0.182	0.251	0.738	1.000

Table 1 Trends in the Income Distribution in the United States

	Year			
Inequality Measure	1970	1980	1990	Ratio 1990/1970
Gini Coefficient				
Census[a]	.361	.368	.381	1.055
CPS[b]	.391	.404	.431	1.102
SD Log Income	.789	.822	.856	1.085

Notes: Income is household cash income adjusted to 1998 dollars with the CPI-U-X1.
a. Author's calculations from Census data described in the Appendix.

b. U.S. Bureau of the Census, Current Population Reports, P60-193.

Table 2 Trends in Percent of Twenty-five to Twenty-nine Year Olds with Various Levels of Educational Attainment

Year	Completed High School or Equivalent[a]	Completed Some College[b]	Completed at Least Four Years of College[a]
1970	75.4	44[c]	16.4
1975	83.1	50	21.9
1980	85.4	52	22.5
1985	86.1	51	22.2
1990	85.7	52	23.2
1993	86.7	59	23.7
1995	86.9	62	24.7

[a] National Center for Educational Statistics, *Digest of Educational Statistics, 1996* Table 8.
[b] Trends in the Well-being of America's Youth, 1998, Table EA1.6
[c] Data for 1971.

Table 3 Correlation between Measures of State Inequality and Educational Outcomes

	High School Graduate	SD Log Income	Gini	50-10 Ratio	90-10 Ratio	90-50 Ratio	Entered College	College Graduate
HS Graduate	1.00							
SD Log Income	-.037	1.00						
Gini	-.038	.925	1.00					
50-10 ratio	-.061	.918	.861	1.00				
90-10 ratio	-.046	.963	.948	.966	1.00			
90-50 ratio	-.010	.872	.944	.721	.873	1.00		
Entered College	.022	-.025	-.023	-.034	-.024	.006	1.00	
College Graduate	.031	-.003	-.019	-.003	-.008	-.011	.543	1.00
Mean	.855	.825	.368	3.543	7.429	2.088	.429	.180
SD	.352	.047	.020	.334	1.030	.108	.495	.384

Source: PSID samples described in Appendix.

.

Table 4 Effect of Various Functional Forms of Household Income on Educational Attainment

	Chi- Square for Functional Form			
Dependent Variable	Linear	Logarithm	Cube Root	Logarithm Plus Deciles
High School Graduate	174.52	214.69	218.26	219.28
Entered College	386.94	401.35	417.50	425.81
College Graduate	252.06	273.21	278.61	290.99

Source: PSID data described in the Appendix.
Notes: The "Logarithm Plus Deciles" predicts educational attainment from log household income, a variable equal to one if the child was in the highest income decile and another variable equal to one if the child was in the poorest income decile.

Table 5 Probit Estimates of the Effect of the Gini Coefficient on Educational Attainment

Variable	High School Graduate	Entered College	College Graduate
Model 1			
Gini Coefficient	-.619	-.347	-.559
	(-1.48)	(-.560)	(-1.130)
Pseudo R^2	.001	.001	.001
Model 2 (with Region and Year Fixed Effects)			
Gini Coefficient	-.245	2.006	.702
	(-.460)	(2.480)	(1.270)
Pseudo R^2	.018	.019	.021
Model 3 (with Region and Year Fixed Effects)			
Gini Coefficient	.221	3.404	1.024
	(.270)	(2.900)	(1.150)
Mean Income	.003	.012	.005
	(.960)	(2.320)	(1.350)
Percent Black	-.001	-.004	-.004
	(-.810)	(-2.360)	(-1.710)
Percent Hispanic	-.001	-.001	.001
	(-.850)	(-.180)	(.720)
Unemployment Rate	-.012	-.010	-.007
	(-3.620)	(-1.920)	(-1.760)
Pseudo R^2	.019	.023	.025

Source: PSID samples described in the Appendix.
Notes: Coefficients are the partial derivative evaluated at the mean of the distribution. Z-statistics are in parenthesis.

Table 6 OLS Estimates of the Effect of the Gini Coefficient on Educational Attainment with Region and Year Fixed Effects

Model Variable	High School Graduate	Entered College	College Graduate
Model 1			
Gini Coefficient	-.268	1.951	.764
	(-.474)	(2.542)	(1.245)
R^2	.010	.026	.018
Model 2			
Gini	.123	3.307	1.142
Coefficient	(.147)	(2.937)	(1.349)
Mean Income	.003	.012	.005
	(.863)	(2.526)	(1.462)
Percent Black	-.001	-.005	-.003
	(-.729)	(-2.360)	(-1.723)
Percent Hispanic	-.002	-.000	.001
	(-.758)	(-.158)	(.681)
Unemployment	-.012	-.010	-.007
Rate	(-3.725)	(-2.019)	(-1.862)
R^2	.017	.031	.024

Source: PSID samples described in Appendix.
Notes: Models control region and year fixed effects. T-statistics are in parentheses and are corrected for clustering in state-year cells.

Table 7 2SLS Estimates of the Effect of the Gini Coefficient on Educational Attainment Controlling the State Returns to Schooling and State Unemployment Rate

	Model 1			Model 2		
Variables	High School Graduate	Entered College	College Graduate	High School Graduate	Entered College	College Graduate
Gini Coefficient	-4.164 (-1.946)	1.5529 (.527)	3.602 (1.774)	-3.637 (-1.789)	2.534 (.930)	3.972 (2.080)
Mean Income	-.012 (-1.746)	.000 (.015)	.010 (1.591)	-.010 (-1.599)	.006 (.651)	.011 (1.806)
Percent Black	.003 (1.150)	-.002 (-.820)	-.003 (-1.862	.000 (.190)	-.006 (-2.399)	-.005 (-2.438)
Percent Hispanic	.005 (1.155)	.005 (1.105)	.001 (.245)	.001 (.368)	-.003 (-.647)	-.003 (-.849)
Unemployment Rate	-.012 (-3.737)	-.010 (-2.053	-.006 (-1.785)	-.011 (-3.381)	-.008 (-1.621)	-.006 (-1.575)
Returns to Schooling				4.087 (2.288)	6.477 (2.669)	2.385 (1.243)
R^2	.006	.029	.020	.011	.034	.021

Source: PSID sample described in the Appendix.
Notes: All models include region and year fixed effects. T-statistics are in parentheses and are corrected for clustering in state-year cells.

Table 8 2SLS Estimates of the Effect of the Gini Coefficient on Educational Attainment Controlling Household Income

Variables	High School Graduate	Entered College	College Graduate
Gini Coefficient	-3.086	3.589	4.639
	(-1.537)	(1.451)	(2.518)
Mean Income	-.013	.002	.009
	(-2.023)	(.197)	(1.400)
Percent Black	.002	-.002	-.003
	(1.130)	(-.958)	(-1.319)
Percent Hispanic	.001	-.004	-.003
	(.283)	(-.876)	(-1.054)
Log Household Income	.144	.254	.160
	(9.768)	(15.580)	(11.418)
Returns to Schooling	2.711	3.613	.576
	(1.533)	(2.314)	(.309)
Unemployment Rate	-.008	-.004	-.004
	(-2.471)	(-.946)	(-1.032)
R^2	.070	.126	.077

Source: PSID data described in the Appendix.
Notes: All models include region and year fixed effects. T-statistics are in parentheses and are corrected for clustering in state-year cells.

Table 9 2SLS Estimates of the Effect of the Gini Coefficient on Educational Attainment Controlling Parental Education and Child's Race

Variable	High School Graduate	Entered College	College Graduate
Gini Coefficient	-3.080 (-1.538)	3.686 (1.500)	4.460 (2.478)
Mean Income	-.014 (-2.154)	.001 (.077)	.008 (1.250)
Percent Black	.003 (1.719)	-.001 (-.400)	-.001 (-.271)
Percent Hispanic	.001 (.424)	-.004 (-.881)	-.003 (-.975)
Returns to Schooling	2.384 (1.341)	3.081 (1.356)	.380 (.212)
Unemployment Rate	-.007 (-1.948)	-.002 (-.433)	-.002 (-.679)
Log Household Income	.095 (6.176)	.184 (9.668)	.104 (7.279)
Parent's Years of Schooling	.030 (7.798)	.044 9.816	.028 (8.214)
Child is Black	-.014 (-.555)	.051 1.562	-.035 (-1.856)
R^2	.105	.172	.109

Source: PSID sample described in the Appendix.
Notes: All models include region and year fixed effects. T-statistics are in parentheses and are corrected for clustering in state-year cells.

Table 10 2SLS Estimate of the Effect the Gini Coefficient Controlling State Tax Revenues and School Expenditures on Educational Attainment

Variable	High School Graduate	Entered College	College Graduate
Gini Coefficient	-3.885 (-1.713)	3.024 (1.061)	4.070 (1.866)
Mean Income	-.019 (-2.298)	-.004 (-.340)	.005 (.577)
Percent Black	.00 (1.862)	-.001 (-.259)	-.000 (-.114)
Percent Hispanic	.002 (.608)	-.003 (-.672)	-.002 (-.758)
Returns to Schooling	2.960 (1.560)	3.353 (1.393)	.434 (.220)
Log Household Income	.093 (6.068)	.184 (9.557)	.104 (7.129)
Parent's Years of Schooling	.030 (7.823)	.044 9.656	.027 (8.096)
Child is Black	-.016 (-.623)	.052 1.583	-.034 (-1.793)
Tax Revenue per Person/1,000	.018 (.271)	-.048 (-.507)	-.058 (-.827)
State School Expenditures per Child/1,000	.037 (1.327)	.049 (1.045)	.035 (.993)
R^2	.103	.173	.112

Source: PSID sample described in the Appendix.
Notes: All models include region and year fixed effects. T-statistics are in parentheses and are corrected for clustering in state-year cells.
.

Table 11 2SLS Estimate of the Effect of the Gini Coefficient on Educational Attainment for High and Low-Income Children

Variables	High School Graduate		Entered College		College Graduate	
	High	Low	High	Low	High	Low
Model 1: controls mean income, racial and ethnic composition, returns to schooling, unemployment rate	1.623 (.434)	-4.167 (-1.675)	9.641 (1.657)	-.141 (-.059)	15.198 (3.077)	-1.030 (-.674)
Model 2: adds log household income, parent's education and child's race	2.002 (.531)	-4.489 (-1.780)	10.894 (1.949)	.464 (.192)	16.379 (3.376)	-1.011 (-.663)
Model 3: adds states tax revenue and school expenditures	.997 (.223)	-4.619 (-1.663)	12.843 (1.597)	-.393 (-.152)	19.457 (2.690)	-1.544 (-.926)

Source: PSID sample described in the Appendix.
Notes: All models include region and year fixed effects. T-statistics are in parentheses and are corrected for clustering in state-year cells.

Table 12 Comparison of Models of the Effect of the Gini Coefficient on Educational Attainment

Models	High School Graduation	Entered College	Graduated College
Model 1: OLS with Region Fixed Effects	-.050 (-.057)	2.826 (2.530)	3.408 (1.908)
Model 2: OLS with State Fixed Effects	-1.634 (-.685)	-.700 (-.155)	5.086 (1.631)
Model 3: 2SLS with Region Fixed Effects	-3.637 (-1.789)	2.534 (.980)	3.972 (2.080)
Model 4: 2SLS with State Fixed Effects	1.484 (.243)	11.655 (1.502)	16.093 (3.164)

Source: PSID sample described in the Appendix.
Notes: All models control state mean income, state racial and ethnic composition, state returns to schooling, state unemployment rate and year fixed effects.