

## The effects of schooling choices on the return to education

Deepraj Mukherjee

Department of Economics,  
Ambassador Crawford College of Business and Entrepreneurship,  
Kent State University, Kent, USA

Piya Chatterjee

LeBron James Family Foundation School of Education,  
Buchtel College of Arts and Sciences  
The University of Akron, Akron, USA

---

### Keywords

Schooling Choices, Return to Education, Demography, Human Capital, Wage Differential

---

### Abstract

*The present research looked at the return to schooling (Public and Private schools) in the census region of the East South-Central Division of the United States. The demographic diversity of this region is the primary reason for its selection. The functional form of return to schooling on education has the specified wage as the long-linear function of school attendance (in years), controlling for gender, race, work experience, occupation status, and educational attainment. We followed the extant literature to control a few socioeconomic variables. The article used 2020 sample data from Integrated Public Use Microdata Series (IPUMS) to estimate the parameter coefficients. The findings show private schools yield a higher return in terms of subsequent wage earnings compared to private schools. The results highlight a higher differential between private schools & public schools for Caucasian Americans over African Americans. Similarly, the differential is higher among females over males.*

---

### 1. Introduction

Education is often the single most important factor determining a person's economic and social achievement. The knowledge and skills children and youth acquire during school attendance will drive their employability, productivity, health, and well-being in the decades to come, and that will help ensure that their communities and nations thrive (Education for Global Development, The World Bank Group, 2011). In this context, return to education from public school and private school becomes immensely important as the schooling choice might determine the potential human capital accumulated by the nation. Public schools are generally the schools that are provided by state and federal funding. According to a report published by the National Center for Education Statistics (United States Department of Education, 2022) in 2002, in 1999–2000, approximately 27,000 private schools accounted for 24 percent of all schools in the US. 77 % of all private school students were white, compared to 63 % of all school students. The public school in 2020 constituted 66.3% of the school-going population while the private school counted for only 9% of the school-going population. The growing literature on the issue addresses the questions related to the quality of the educational institution that might affect the subsequent earnings. An excellent review is provided by McEwan (2000), who argues that, with a few exceptions, there is generally insufficient evidence to reach strong conclusions regarding such comparisons. Methodological difficulties found in this literature include the size and nature of the available samples of schools and students (e.g., small sample sizes, self-selection into public or private schools), as well as key student, family, school, and community variables that remain unmeasured but may be associated with both public versus private school attendance and student achievement.

Lack of academic scrutiny provided us the impetus to investigate the effect of attending public school and private school (i.e., the decision of school choices) on future wage earnings. We differed from the earlier studies in our choice of quality. Rather than looking into more school-specific details like the student-teacher ratio, the relative salary of teachers, and the length of school terms on the estimates of the

return to education (Card and Kreuger, 1992), we have considered a more simplistic qualitative indicator by considering the choice of schooling (public school versus private school) as the key quality parameter. Obtaining accurate estimates of payoff from these two types of schools would be important to the students and their parents and general societal welfare. The literature review is closely connected to the theoretical and empirical findings of college quality on earnings in the later years. US Central Division of East South Central is our chosen geographic region (TN, AL, MS, and KY) for the data sample. The objective behind the selection of this division is twofold: (1) we wanted to capture the returns to the schooling choices for relatively poorer states where income inequality is much higher than the states of other divisions, (2) diversified demographic composition especially the large presence of African American in this region motivated the selection. We admit in priority that the study typically relies on a single proxy variable for school quality. The wisdom of following such a path can be questioned, given that a single proxy likely measures school quality with substantial error. The remainder of the paper is organized as follows: The next section briefly describes the relevant literature on this subject, while section 3 discusses the data and methodology. Section 4 elaborates on the empirical results and inferences, while Section 5 and 6 concludes by delineating future research avenues. An appendix containing variable descriptions is also included at the end.

## 2. Literature Review

Morgan and Sirageldin (1968) and Link and Ratledge (1975) are the early studies that examined the relationship between return to schooling and wage earnings. But it was Jacob Mincer who published his seminal book *Schooling, Experience, and Earnings* which had a profound and lasting influence in the field of labor economics in 1974. The Mincer earnings function is the foundation of a voluminous literature in empirical labor economics. In the general version of Mincer's "human capital earnings function," log earnings are modeled as the sum of a linear function of years of education and a quadratic function of years of potential experience. The earning equation was specified as,

$$\ln wage = c + years\ of\ schooling + potential\ experience + potential\ experience^2$$

The popularity of Mincer's equation lies in the fact that it is based on a formal human capital investment model, and the specification is parsimonious and fits well with the data. In this regard, it might be argued that the key contribution of Mincer's earning specification was the introduction of potential experience (Age-Years of Schooling-6) as a standard regressor on the earning regression specification. The introduction of potential experience as opposed to age in the earnings equation is a prudent way of capturing both the shape of the age-earnings profile and the differential slope of the age-earnings profile across education groups. In addition, such a model asserts that conditional on years of potential experience; there is a single rate of return to education in the labor market. This

line of thinking provides the basis for the growing literature that attempts to estimate the causal effect of education on earnings (see David Card, 1999, for a detailed review).

Dale and Kreuger (2002) provide an excellent empirical analysis that utilizes information on the set of colleges to determine the effect of college quality on subsequent earnings. Contemporary research has supported the importance of school quality on future earnings. Bills and Klenow (2000) discussed the value of schooling being affected by the human capital of previous generations. They followed the mechanism that essentially measures the school quality in terms of the school's teacher quality. Manuelli and Seshadri (2005) and Erosa, Koreshkova, and Restuccia (2006) describe various mechanisms of school quality and how that mechanisms affect the measured role of schooling. However, empirical research on the determinants of school quality remains skeptical until now (Hanushek 1995, Hanushek 2002). The effect of family background on education by influencing the level of education individuals obtain is well documented and analyzed in the literature (Krishnan, 1996; Altonji & Dunn, 1996; Beach & Finne, 1988). Altonji & Dunn (1996) investigated the conjecture that the education slopes of wage equations are influenced by family background as measured by the father's and mother's education. Beach & Finne (1988) have analyzed a similar pattern of return. In a nutshell, the article analyzed the effects of parents' education on their son's educational attainment, which thus ends up increasing the importance of indirect and total effects of the family background variables on earnings. Lam & Schoeni (1993) have shown that the relationship between having a father with a university education and getting a 20% wage advantage

when compared with an illiterate father and a 9% wage advantage when compared to a father with 4 years of schooling, controlling the workers' own schooling and the schooling of other relatives respectively. Sahn & Alderman (1988) have pointed out that the wage offered in developing countries is influenced by other genetic and environmental influences captured in the wage of one's father. Thus, the significant impact of family background on earnings could mean that family background determines the quality of education, and individuals from a better family background are able to get better jobs through social capital

The partial cause of earnings differentials may also be a sector of employment.

Mann & Kapoor (1988) have explored that, on average, public sector workers are paid much higher wages than private and joint-sector workers. Rees & Shah (1995)

have reasoned that the private wage determination is subject to profit constraint, whereas,

the public sector wage determination is subject to an ultimate political constraint. Thus,

wages in the public sector are higher than in the private sector. Pritchett (1999)

highlighted the situation in which governments are taking resources away from nongovernmental activity in the form of taxes so as to pay additional workers whose

marginal product in the public sector is very low but are paid much higher wages than

workers in the private sector. Overall, even if it is widely accepted that schooling is positively correlated with earnings, it is not acceptable to understate the effects of other influential factors on earnings. This implies that the basic Mincerian approach to estimating return to education is applicable but should be extended further to control other latent variations which the present study has attempted.

### 3. Data and Methodology

We have used the US Census Bureau data for the year 2020. The data are collected from Integrated Public Use Microdata Series: Minneapolis, MN: Minnesota Population Center <http://usa.ipums.org/usa/>. There are two distinct advantages of using IPUMS data for our analysis. First, the convenience of this data, the data provides codes and names for all samples in an easy-to-use format. Samples include demographic measures, educational, occupational, and all work indicators. We tailored our data filter to answer the questions using the desired variables. Second, the 5% sample of this database provides a large sample base crucial to capture the differential effects of certain variables on the outcome variable. However, there is a notable disadvantage. Using IPUMS data implies that we are restricted only to the cross-sectional dimensions of the data. In our sample, we extracted approximately 100,000 numbers of observations and further cleaned the data to maintain our sample size of 47,268 numbers of observations.

We have constructed the dummy variables (to obtain a binary specification) for the independent variables. However, a couple of such variables need special attention from the readers. From the data, we have divided the educational attainment records of the observations into four categories, namely observations with less than a high school diploma, with a high school diploma, with some associate degree, and with baccalaureate or higher degree. We have chosen the less-than-high-school diploma category as the reference category. Specifically, we have constructed three dummies in the other three categories. Similarly, we have categorized the race variable into three parts: White Americans, Black Americans, and other races, and have used the 'other races' category as the reference to construct the respective dummies for the remaining two categories. We have constructed two race dummies, namely *Racew* and *Raceb*. *Racew* =1 for whites and =0 otherwise; *Raceb* =1 for blacks and =0 otherwise. To capture the effect of family backgrounds for different observations, we have used two of the constructed variables from the IPUMS data, namely the variable 'mother's location in the household and the 'father's location in the household' variable. Such variables are specified as,

0 = Absence of parents in the same household when the observation was a child.

1 = Presence of parents in the same household when the observation was a child.

### 4. Empirical Results

This section consists of three distinct but related forms of empirical analysis. First, we report the descriptive statistics for the model specified in equation (1). Second, we present the estimated coefficients

of the relevant parameters using OLS regression as baseline specifications. Third, we show the independent causal relations of the explanatory variables on the outcome variable to counter any potential endogeneity. Such an analysis is the best alternative without any valid instrument. Table 1 summarizes the mean wage income for White Americans and African Americans. While the mean wage income is higher for private school attendees than for public school attendees, the wage differential in terms of return to education in these two schools is higher for African Americans than for white Americans. The descriptive statistics for the relevant parameters are given in Table 1 and Table 2.

Table 1. *Mean Earnings by School Types for Two Dominant Races*

Race	Number of Observations	Mean Income (Public School)	Mean Income (Private School)	Mean Income
White	35469	12126.86	13983.64	12450.64
Black	10043	11744.29	15645.16	12253.12

Table 2. *Descriptive Statistics for the relevant parameters*

Variable	Mean	Standard Deviation
Log of Wage	8.646	1.367
School type	0.834	0.371
Potential experience	11.257	9.037
Less than High School	0.236	0.257
High school diploma	0.130	0.337
Some college	0.485	0.499
College degree and higher	0.146	0.353
Mom location	0.620	0.782
Pop location	0.270	0.519
Marital Status	0.250	0.433
Sex	1.540	0.499
Number of Siblings	0.102	0.303
Language	0.069	0.253

Initially, regression was conducted for the entire sample as presented. Table 3. Table 4 and Table 5 summarize the regression estimates for Caucasian Americans and African Americans respectively while in Tables 6 and 7 we ran separate regressions gender-wise: i.e., for the male observations and female observations separately. For all the specifications adjusted  $R^2$  is reasonably high. Standardized coefficients or beta coefficients are reported for each variable. It represents the change in the log of wage earnings that results from a change of one standard deviation in one of the independent variables. The standardized coefficient helps us determine the relative weights of each independent variable on the outcome variable. We have conducted the multicollinearity test for all the variables in each specification and reported the respective variance Inflation factor (from here on referred to as VIF). For all the variables, the reported VIF measures indicate that there is no multicollinearity except for the potential

experience variable and its quadratic function. However, the findings suggest that VIF for them is much below 30 (potential threat level) i.e., the collinearity is very moderate and can safely be ignored. Table 3 shows the estimated parameters' estimated coefficients and expected signs.

Table 3. *Standardized Regression Estimates*

Variable	Standardized Coefficients Beta	Sig. level (P values)	Variance Factor	Inflation
(Constant)	6.902	.000		
School type	0.123	.000	1.015	
Potential experience	0.156	.000	13.969	
Potential experience <sup>2</sup>	-0.02	.000	12.106	
High school diploma	.156	.000	1.028	
Some college	0.304	.000	1.026	
College degree and higher	0.325	.000	1.409	
Mom location	0.003	0.418	1.831	
Pop location	0.004	0.215	1.865	
Marital Status	0.044	.000	1.411	
Sex	-0.083	.000	1.014	
Racew (Controlling for white)	0.021	0.015	5.992	
Raceb (Controlling for Black)	-0.004	0.507	5.891	
Number of Siblings	-0.001	0.687	1.133	
Language	-.008	.015	1.108	
<b>Number of Observation</b>	<b>47265</b>			
<b>Adjusted R<sup>2</sup></b>	<b>0.436</b>			

The parameter school type potential experience and different educational attainment categories are statistically significant. As expected for the race variable (controlling for whites and blacks), *Racew* is positive, *Raceb* is negative, and both are statistically significant. Intuitively, the differential effect of being black on the outcome variable is negative with respect to the reference category. Two indicator variables, namely the number of siblings and the language spoken in the household, has their expected signs. The increase in the number of siblings is negatively associated with wage earnings. Similarly, the variable 'language spoken at the household' is negative, indicating that for the households whose native language is not English, the differential effect of it on the log earnings is negative with respect to the English-speaking households on log earnings.

Table 4. *Regression Estimates for Sub-Sample (Caucasian)*

Variable	Standardized Coefficients Beta	Sig. level (p values)	Variance factor	Inflation
(Constant)	7.048	.000		
School type	0.115	.000	1.014	
Potential experience	0.157	.000	15.088	

Potential experience <sup>2</sup>	-0.002	.000	12.681
High school diploma	.163	.000	1.392
Some college	0.320	.000	2.098
College degree and higher	0.339	.000	2.068
Mom location	0.018	0.011	3.293
Pop location	0.005	0.459	2.885
Marital Status	0.061	.000	1.584
Sex	-0.095	.000	1.012
Racew (Controlling for white)	-	-	-
Raceb (Controlling for Black)	-	-	-
Number of Siblings	0.001	0.848	1.001
Language	-	-	-
<b>Number of Observation</b>	<b>35,457</b>		
<b>Adjusted R<sup>2</sup></b>	<b>0.454</b>		

Table 5. *Regression Estimates for African Americans*

Variable	Standardized Coefficients Beta	Sig. level (p values)	Variance Inflation factor
(Constant)	6.762	.000	
School type	0.137	.000	1.033
Potential experience	0.170	.000	12.333
Potential experience <sup>2</sup>	-0.003	.000	11.301
High school diploma	.135	.000	1.623
Some college	0.258	.000	2.105
College degree and higher	0.249	.000	1.773
Mom location	0.006	0.451	1.002
Pop location	0.006	0.412	1.002
Marital Status	0.021	.000	1.209
Sex	-0.035	.000	1.030
Racew (Controlling for white)	-	-	-
Raceb (Controlling for Black)	-	-	-
Number of Siblings	-0.011	0.153	1.027
Language	-	-	-
<b>Number of Observation</b>	<b>10,042</b>		
<b>Adjusted R<sup>2</sup></b>	<b>0.384</b>		

Tables 4 and 5 summarize the regression results when the same set of regressions are run for the Caucasians and the African Americans separately. All the coefficients have their expected sign, and the adjusted R<sup>2</sup> is reasonably high. The wage differential between private and public-school goers among



blacks is higher than the same for whites. Family background variables that determine whether the parents were in the same household during the schooling years also have their expected signs.

Table 6. *Regression Estimates for Males*

Variable	Standardized Coefficients Beta	Sig. level ( <i>p</i> values)	Variance Factor	Inflation
(Constant)	6.472	.000		
School type	0.123	.000	1.021	
Potential experience	0.172	.000	14.282	
Potential experience <sup>2</sup>	-0.003	.000	12.021	
High school diploma	.165	.000	1.337	
Some college	0.304	.000	1.681	
College degree and higher	0.287	.000	1.805	
Mom location	0.002	0.645	1.015	
Pop location	0.007	0.138	1.012	
Marital Status	0.080	.000	1.550	
Sex	-	-	-	
Racew (Controlling for white)	0.031	0.015	5.098	
Raceb (Controlling for Black)	-0.019	0.770	4.964	
Number of Siblings	0.005	0.376	1.099	
Language	-.013	0.014	1.167	
<b>Number of Observation</b>	<b>21861</b>			
<b>Adjusted R<sup>2</sup></b>	<b>0.484</b>			

Table 7. *Regression Estimates for females.*

Variable	Standardized Coefficients Beta	Sig. level ( <i>p</i> values)	Variance Factor	Inflation
(Constant)	6.563	.000		
School type	0.127	.000	1.019	
Potential experience	0.147	.000	14.141	
Potential experience <sup>2</sup>	-0.002	.000	12.326	
High school diploma	0.150	.000	1.538	
Some college	0.311	.000	2.273	
College degree and higher	0.362	.000	2.101	
Mom location	0.013	0.645	0.104	
Pop location	0.017	0.138	0.020	
Marital Status	-0.051	.004	1.407	
Sex	-	-	-	
Racew (Controlling for white)	0.009	0.486	6.905	
Raceb (Controlling for Black)	0.019	0.634	6.823	

Number of Siblings	-0.050	0.380	1.212
Language	-0.006	0.270	1.167
<b>Number of Observation</b>	<b>25405</b>		
<b>Adjusted R<sup>2</sup></b>	<b>0.398</b>		

Table 6 and Table 7 summarize the regression results when the same set of regressions in terms of sex, i.e., separately for males and females. The wage differential from attending private with respect to public schools is marginally higher for females (0.4 points) than for males. Notably, for females, married women earn less than unmarried women, and the estimated coefficient for the relevant variable has a negative sign.

## 5. Discussions

Early studies (Morgan & Sirageldin, 1968; Link & Ratledge, 1975) only examined the relationship between return to schooling and wage earnings. Jacob Mincer further extended such studies in the field of empirical labor economics through his seminal works on schooling, experience, and earnings. The present study extended the basic Mincerian approach to estimating return to education by controlling for other latent variables of schooling on return to education, which could possibly be a persuasive study in the field of labor economics. Gary Becker's pioneering research on human capital launched a large and active industry, estimating causal effects and returns to schooling. However, the earlier results often do not report the causal effect. The main result obtained with the help of the extended Mincerian earnings function indicates that people educated in a private school earn more than those educated in a public school. In addition, the study indicates that the wage differential between private school goes and public-school goes among African Americans is higher than the wage differential for the same among White Americans.

## 6. Policy Implications, limitations, and future research avenues

The findings of the study indicate sustaining income inequality among blacks and can be an interesting research question to analyze. In addition, in terms of the gender issue, the wage differential is higher for females, which might delineate possible research avenues in the literature on family, feminist, and educational research economics. The methodology has certain limitations, as the estimations are restricted to only cross-sectional observations. There is possible endogeneity of certain explanatory variables. However, the study can be an important addition to the extant literature as it addresses a different proxy for school quality in terms of choice of schooling. In the future, the authors would like to construct balanced panel data and check for the robustness of the results.

## References

- Altonji, J., & Dunn, T. (1996). "The effects of family characteristics on the Return to Education" *The Review of Economics and Statistics*, 78 (4), 692-704.
- Bils, M., and P. J. Klenow (2000): "Does Schooling Cause Growth?" *The American Economic Review*, 90(5), 1160-1183.
- Bryman, A. and D. Cramer (1990). *Quantitative data analysis for social scientists*, pp. 246-251.
- Card, David (1999). "The Causal Effect of Education on Earnings" in Orley Ashenfelter and David Card, editors, *Handbook of Labor Economics*, North Holland: Amsterdam and New York.
- Erosa, A., T. Koleshikova, and D. Restuccia (2006): "On the Aggregate and Distributional Implications of Productivity Differences across Countries," Working Paper.
- Griliches, Zvi (1979). "Sibling Models and Data in Economics: Beginnings of a Survey" *The Journal of Political Economy*, Vol. 87, No. 5, Part 2: Education and Income Distribution, pp. S37-S64.
- Hanushek, E. A. (1995): "Interpreting Recent Research on Schooling in Developing Countries," *The World Bank Research Observer*, 10(2), 227-246.
- (2002): "Publicly Provided Education," in *Handbook of Public Economics*, ed. by A. J. Auerbach, and M. Feldstein, vol. 4, chap. 30, pp. 2045-2141. Elsevier Science, North-Holland Publishers.



- Heckman, James & Polachek, S. (1974). "Empirical Evidence on the Functional Form of the Earnings-Schooling Relationship." *Journal of the American Statistical Association*, Vol. 69, No. 346, pp. 350-354
- Krishnan, P. (1996). "Family background, education and employment in urban Ethiopia" *Oxford Bulletin of Economics and Statistics*, 58 (1), 167-183.
- Link, Charles R., & Edward C. Ratledge. (1975). "Social Returns to Quantity and Quality of Education: A Further Statement." *Journal of Human Resources* 10(1),78-89.
- Manuelli, R. E., and A. Seshadri (2005): "Human Capital and the Wealth of Nations," Working Paper.
- Morgan, J., I Sirageldin. (1968). "A Note on the Quality Dimension in Education." *The Journal of Political Economy*, Vol. 76, No. 5, 1069-1077
- Mann, P., & Kapoor, B. (1988). "Earnings differentials between public, private and joint sector in Punjab (India)" *The Journal of Development Studies*, 25 (1), 97-111.
- McEwan, P.J. (2000). "The Potential Impact of Large-Scale Voucher Programs" *Review of Educational Research*, 70(2): 103-149.
- Mincer, J. (1958). "Investment in human capital and personal income distribution" *The Journal of Political Economy*, 66 (4), 281-302.
- Mincer, J. (1974). "Schooling, experience, and earnings" New York: National Bureau of Economic Research.
- King, E. (2011, January 28). *Education is Fundamental to Development and Growth*. World Bank Blogs. Retrieved February 15, 2023, from <https://blogs.worldbank.org/education/education-is-fundamental-to-development-and-growth#:~:text=It%20is%20the%20knowledge%20and,their%20communities%20and%20nations%20thrive>.
- NCES (2020). "Racial Ethnic Enrollment in Public Schools." National Center for Education Statistics. Retrieved on February 8<sup>th</sup>, 2023, from <https://nces.ed.gov/programs/coe/indicator/cge/racial-ethnic-enrollment>
- NCES (1997). Findings from the condition of Education 1997 - National Center for Education Statistics, No. 12 - Public and Private Schools How do they differ? Retrieved February 15, 2023, from <https://nces.ed.gov/pubs97/97983.pdf>

## Appendix

### **Variable Description**

**Wage and salary income (lnwage):** Reports each respondent's total pre-tax wage and salary income - that is, money received as an employee - for the previous year. The censuses collected information on such income during the preceding calendar year; for the ACS and the PRCS, the reference period was the past 12 months. Log of wage is derived from taking the logarithm of this variable.

**Mother's location in the household (Mom Location):** this is a constructed variable that indicates whether the person's mother lived in the same household when the subject was a child.

**Father's location in the household (Pop Location):** is a constructed variable that indicates whether the person's father lived in the same household when the subject was a child.

**High School Diploma:** classifies high school graduates according to their degree earned.

**Some College:** Occupational associate degree or Academic associate degree

**College Degree and Higher:** Bachelor's /Master's /Professional/Doctorate

**Marital Status (Marstatus):** Marstatus gives each person's current marital status.

Categories:

- Married:

- 1) Spouse Present
- 2) Spouse Absent

- Unmarried:

- 1) Widowed
- 2) Divorced
- 3) Separated
- 4) Single