10-15-2000

The Effect of Educational and Residential Characteristics on the Private Return to Education

Dwight Adamson
South Dakota State University

James Swenson

The Effect of Educational and Residential Characteristics on the Private Return to Education

By

Dwight W. Adamson and James A. Swanson*

Economics Staff Paper 00-11

October 2000

*Respectfully, Department of Economics, South Dakota State University, Brookings, SD 57007, (O) 605.688.4105, (F) 605.626.3022, bill_adamson@sdstate.edu; Department of Economics, Central Missouri State University, Warrensburg, MO 64093, (O) 660.543.4246, (F) 660.543.8580, Swanson@cmsu1.cmsu.edu

Twenty-five copies of this document were made by the Economics Department at a cost of $1.40 per copy.

ABSTRACT:

This paper attempts to investigate the individual wage rate of return to college education. Over the last twenty years, the return to college education has increased dramatically. We propose to investigate how the return to college differs across educational and residential characteristics. We first consider differential returns to an associate, bachelor, advanced, and professional degrees and what characteristics of an individual's college education influence the rate of return to different degrees. Secondly, we focus on the variation in the return to college education in different areas of residence by comparing the return to education across regions and metropolitan and nonmetropolitan areas. These issues will be explored using the National Longitudinal Survey of Youth data set. In our preliminary analysis, we explored the differential impact of educational characteristics on wages across regions of residence, on the location where the college was attended, and on wages metropolitan versus nonmetropolitan workers. Our initial findings suggest that returns to education do vary with college location and area worked, and are greater for metro workers.

I. INTRODUCTION

The link between education and earnings is one of the best documented relationships in labor economics. While there continues to be a debate over why more educated persons earn more, the evidence strongly suggests that average earnings increase with more education. This relationship has been documented for many decades, yet it has become the focus of recent studies partly because of the increase in returns to higher education and the increase in income inequality that appears to have taken place in the United States in the past few decades (see, for example, Gottschalk and Smeeding, 1997, and Ashenfelter and Rouse, 1999).

Besides the recent increase in the returns to education, there has also been greater variation in these returns. A recent study found that the income difference of a male with some college in the 90^{th} percentile of income from one in the 10^{th} percentile grew by over $35,000 between 1972 and 1995 (Hoxby and Terry, 1999). This same study found that part of the explanation for this wider variation in college earnings arose because of differences in characteristics of the colleges attended.

Other studies in this area suggest that higher earnings are associated with attending colleges with higher average tuition (Dale and Krueger, 1999). An earlier study had found some evidence of differences in earnings based upon the quality of the college, although this study attributed more of the variation in earnings to the choices of majors and grade point average (James, et. al., 1989).

In this paper we are using a large data source that not only allows us to test the relationship between earnings and college quality, but also to investigate if there are regional differences in the returns to quality. The next section discusses the theoretical wage equation. The third section discusses the data and the empirical model. The fourth section provides the empirical results, and the final section presents conclusions and areas for future research.

II. THE THEORETICAL WAGE EQUATION

A typical human capital wage equation is modeled in which the primary determinates of an individual's wage are his or her human capital accumulation, their ability, and their labor market experience. Prior research has shown, however, that additional factors must be included to control for other demographic factors such as marital status and health. In addition, characteristics of the occupation of the worker and the industry he or she works in will impact on their earnings. Finally, as discussed in the previous section, the quality of the college attended and the location of employment will have an impact on earnings.

This discussion leads to the following wage equation:

$$w_{it} = f(X_i, I_i, L_i, Q_i, t)$$

where w_{it} is worker i's wage in year t, **X** is a vector of time variant and invariant measures of the worker's human capital, demographic characteristics, and job characteristics, **I** is a vector of time variant indicators for the worker's industry of employment, **L** is a vector that indicates the working location, **Q** is a vector of factors reflecting college quality, and t is a vector of time period effects to control for factors that have a common influence.

III. DATA AND EMPIRICAL MODEL

The primary data source is the 1988-1993 NLSY. The NLSY is a longitudinal survey began in 1979 with a survey of 12,686 young males and females between the ages of 14 and 22 (Center for Human Resource Research, 1997). A key feature of the NLSY sample is that the retention rate has been in the neighborhood of 90%. Budgetary constraints forced the elimination of 878 members of the military subsample in 1985 and 1,643 members of the disadvantaged white subsample were eliminated in 1991. Including the disadvantaged white subsample, 10,465 were surveyed in 1988 and 9,011 were surveyed in 1993. Observations were then deleted for the self employed and for members of the military as well as cases where an individual's observation was unavailable for all six years between 1988 and 1993. Likewise, observations were also omitted when the respondent's hourly wage was less than \$1.50 or over \$250. Finally, individuals were deleted if there were missing data for the variables of interest The total sample size is 25,284 over the entire sample period.

The dependent variable used in the regression analysis is the usual hourly wage taken as usual weekly earnings divided by usual weekly hours. Assuming a linear form the following equation can be written:

$$w_{it} = \beta_0 + \beta_1 X_i + \beta_2 I_i + \beta_3 Q_i + \beta_4 t + e_{it}.$$

In the equation β_0 is the constant, β_1 - β_4 are coefficient vectors, and t is a vector of year dummies. The e term represents the typical regression error with mean zero and variance σ_e^2. The **X** vector contains the standard human capital, demographic, and workplace characteristic variables that have been used in other micro data studies. They include tenure with the current employer and actual experience in linear and quadratic form; a vector of indicator dummies for years of educational attainment; a dummy for minority and gender; a union member indicator; a dummy for health status that affects work or pay; an indicator of part-time employment, an indicator for larger employers with 1000 or more employees; and the respondent's Armed Forces Qualifications Test (AFQT) score, which was administered to most of the NLSY sample in 1980. There are also eight occupation dummy variables with the low-skilled labor occupation being the omitted category so that wages are measured relative to laborers. Labor market differences by region will be captured by regional dummies for the Midwest, South, and West, with the Northeast being the omitted group.

The **I** vector is a set of nine indicator variables for the worker's industry of employment. The omitted industry is manufacturing, so that all the other industry coefficients measure *ceteris paribus* industry wages relative to manufacturing. One characteristic of the industry variables is they control for various unmeasured aspects of the worker's industry, which may capture some of the effects of the human capital and other variables. For example, some high-paying industries also employ a well-educated workforce, suggesting that some of education's effect may be captured by the industry dummy coefficient.

The Q vector contains the information on college quality. At the present time, we use the region where the college was located as our measure of quality. The regions are the New England region (MA, ME, VT, NH, CT, and RI), the middle Atlantic region (MATIREG: NY, NJ, and PA), the south Atlantic region (SATLREG: WV, VA, DC, DE, MD, NC, SC, GA, and FL), the east south central region (ESCREG: KY, TN, MS, and AL), the west south central region (WSCREG: TX, OK, AR, and LA), the east north central region (ENCREG: MI, OH, IN, IL, and WI), the west north central region (WNCREG: MN, IA, MO, KS, NE, ND, and SD), the mountain region (CO, AZ, NM, ID, NV, UT, and WY), and the pacific region (PACREG: CA, WA, OR, AK, and HI). The New England region is excluded with dummy variables for the remaining seven regions.

The means and standard deviations for all of the variables used in the study are reported in Table II in the appendix. Note, the sample includes people who attended college in all of the fifty states except Arkansas and Delaware.

IV: EMPIRICAL RESULTS

The results of the estimation of the empirical model are reported in Table I. The semi-log function form was estimated by a maximum likelihood procedure using the SAS statistical package. Four separate versions of the empirical model were estimated. Model 1 contains only the traditional human capital and demographic variables along with dummy variables for region of employment (the North East is the base case) and dummy variables for each year after 1988. As expected, the ability variable (AFQT81) is positive and significant, and both experience (EXPACT and EXPSQ) and tenure (TEN and TENSQ) have positive but declining effects on wages. Married (MAR) workers earn more while those with health problems earn less, although the latter variable is not significant in any of the models. Part-time (PT) workers earn less (although only significant in the first and third versions of the model), and there is a primum for working in a metropolitan area (METRO). Women (FEMALE) consistently earn less than males, while Asians (ASIAN) consistently earn more than whites. The other minority variables (BLACK, HISPANIC, and NATIVE) are not significant or have different signs across models. Finally, the dummies for the different years (T89 - T93) are all positive and significant.

As predicted by human capital theory, high school graduates (HSDEG) earn less than those with some college (the base case), while those with college degrees (AADEG, BACDEG, ADVDEG, and PRODEG) all earn more. Finally, the region where employed has a significant effect with all other regions (MWEST, SOUTH, and WEST) earning less than the north east region.

The second model includes all of the variables in the first model, and it includes the industry and occupational dummy variables. Unionized employers (UNION) pay higher wages. Workers in agriculture (AGIND), wholesale (WHOSALE), retail (RETAIL), finance and insurance (FININS), (PROBUSER), personal service (PERSERV), and public administration (PUBICADM) have lower wage than manufacturing. Mining (MINING) and transportation and utilities (TRASUTIL) have higher wages.

Managers and professionals (MANGPROF), technical (TECHOCC), and sales (SALES) workers earn more then low-skill workers, while the coefficients for the remaining occupations are not consistently significant. Finally, as has been found in several studies, large firms (LAREMP) pay higher wages.

Models three and four include the location where the person attended college. In model three, these variables are included while the industry and occupational variables are excluded. In the fourth model, all of the variables are included. Since the results differ little, the discussion will be for the fourth model.

The excluded region was the New England region which contains most of the Ivy League schools. Thus, it was expected that people who attended college outside of New England would have, ceteris paribus, lower wages. For the most part, just the opposite was true. In fact, except for the mountain and west north central regions, all of the coefficients were positive and significant. While college location is not necessarily the ideal measure of differences in quality, the fact that most of these variables were significant tends to support the idea of variable returns to higher education depending upon characteristics of the college attended.

V. CONCLUSIONS AND AREAS OF FURTHER RESEARCH

Recent research on the returns to higher education have found not only a higher return in the last two decades, but also a growing inequality of earnings. Studies have suggested that at least part of the rising inequality may be due to increased variation in college quality. The findings of this paper tend to support this conclusion. However, better measures of quality need to be incorporated into the model. This is the direction we intend to take with this paper.

References

ADAMSON, DWIGHT W., CLARK, DAVID E. AND PARTRIDGE, MARK D., "Are Worker's Wages Driven by National or Local Factors?" mimeo, November 1999.

ASHENFELTER, ORLEY AND ROUSE, CECILIA, "Schooling, Intelligence, and Income in America: Cracks in the Ball Curve, *NBER Working Paper, #* 6902, January 1999.

DALE, STACY BERG AND KRUEGER, ALAN B., "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables," *NBER Working Paper, #* 7322, August 1999.

GOTTSCHALK, PETER AND SMEEDING, TIMOTHY, "Cross-National Comparisons of Earnings and Income Inequality," *Journal of Economic Literature,* Vol. XXXV (June 1997), pp. 633-687.

HOXBY, CAROLINE M. AND TERRY, BRIDGET, "Explaining Rising Income and Wage Inequality among the College-Educated," *NBER Working Paper, #* 6873, January 1999.

JAMES, ESTELLE, ET AL. "College Quality and Future Earnings: Where should You Send Your Child to College?" *American Economic Review*, May 1989, pp. 247-252.

Table I: Model Estimates*

Variable	Model 1	Model 2	Model 3	Model 4
INTERCEPT	1.5700 (59.14)	1.617 (62.18)	1.5799 (58.63)	1.6267 (61.58)
AFQT81	0.0046 (20.76)	0.0041 (19.87)	0.0041 (18.02)	0.0037 (17.39)
EXPACT	0.0446 (11.75)	0.0389 (10.54)	0.0446 (11.79)	0.0389 (10.58)
EXPSQ	-0.0006 (-3.24)	-0.0004 (-2.23)	-0.0006 (-3.13)	-0.0004 (-2.14)
TEN	0.0368 (20.09)	0.0353 (19.53)	0.0366 (20.02)	0.0351 (19.45)
TENSQ	-0.0021 (-13.09)	-0.0020 (-12.83)	-0.0020 (-12.87)	-0.0020 (-12.61)
MAR	0.0295 (5.35)	0.0268 (4.97)	0.0301 (5.47)	0.0273 (5.07)
HSDEG	-0.0326 (-2.58)	-0.0323 (-2.72)	-0.0370 (-2.95)	-0.0368 (-3.11)
AADEG	0.0978 (5.84)	0.0884 (5.52)	0.0769 (4.54)	0.0687 (4.24)
BACDEG	0.1758 (11.57)	0.1627 (11.18)	0.1541 (9.94)	0.1424 (9.58)
ADVDEG	0.1691 (7.83)	0.1657 (7.91)	0.1607 (7.46)	0.1574 (7.53)
PRODEG	0.3533 (6.57)	0.3353 (6.49)	0.3404 (6.37)	0.3231 (6.29)
HEACOND	-0.0083 (-0.83)	-0.0102 (-1.03)	-0.0088 (-0.89)	-0.0106 (-1.08)
MWEST	-0.1443 (-10.14)	-0.1441 (-10.77)	-0.1495 (-9.02)	-0.1504 (-9.57)
SOUTH	-0.1803 (-13.88)	-0.1712 (-13.95)	-0.1924 (-13.18)	-0.1828 (-13.16)
WEST	-0.0757 (-4.96)	-0.0649 (-4.52)	-0.0832 (-4.69)	-0.0700 (-4.16)
METRO	0.0992 (10.95)	0.0973 (11.10)	0.0957 (10.59)	0.0939 (10.74)
PT	-0.0290 (-4.19)	-0.0102 (-1.48)	-0.0302 (-4.36)	-0.0111 (-1.61)
FEMALE	-0.1457 (-14.70)	-0.1337 (-14.06)	-0.1543 (-15.66)	-0.1404 (-14.86)
BLACK	0.0366 (2.76)	0.0117 (0.94)	0.0205 (1.55)	-0.0025 (-0.021)
ASIAN	0.1386 (2.52)	0.1276 (2.50)	0.1145 (2.10)	0.1058 (2.10)

Variable	Model 1	Model 2	Model 3	Model 4
HISPANIC	0.0316 (2.01)	0.0184 (1.26)	0.0251 (1.61)	0.0131 (0.90)
NATIVE	0.0285 (0.99)	0.0131 (0.49)	0.0242 (0.85)	0.0089 (0.34)
T89	0.0322 (5.85)	0.0316 (5.83)	0.0314 (5.71)	0.0309 (5.70)
T90	0.0673 (10.74)	0.0670 (10.97)	0.0658 (10.53)	0.0657 (10.78)
T91	0.0813 (11.08)	0.0825 (11.66)	0.0786 (10.76)	0.0802 (11.37)
T92	0.0939 (10.90)	0.0946 (11.49)	0.0907 (10.59)	0.0918 (11.21)
T93	0.1113 (11.09)	0.1158 (12.05)	0.0907 (10.76)	0.1121 (11.73)
UNION		0.0739 (12.46)		0.0729 (12.31)
AGIND		-0.1282 (-6.81)		-0.1292 (-6.87)
MINING		0.0692 (2.54)		0.0729 (2.65)
TRASUTIL		0.0410 (4.19)		0.0392 (4.01)
WHOSALE		-0.0440 (-4.08)		-0.0442 (-4.41)
RETAIL		-0.1527 (-19.38)		-0.1527 (-19.41)
FININS		-0.0224 (-2.02)		-0.0218 (-1.97)
PROBUSER		-0.0619 (-8.91)		-0.0635 (-9.16)
PERSERV		-0.1396 (-10.85)		-0.1409 (-10.96)
PUBICADM		-0.0140 (1.32)		0.0120 (1.13)
MANGPROF		0.0603 (7.27)		0.1581 (7.01)
TECHOCC		0.0729 (6.71)		0.0710 (6.54)
SALES		0.0245 (2.53)		0.0230 (2.38)
CLEROCC		0.0114 (1.38)		0.0096 (1.16)
SERVOCC		-0.0064 (-0.78)		-0.0067 (-0.81)

Variable	Model 1	Model 2	Model 3	Model 4
NROCC		-0.0212 (-1.35)		-0.0198 (-1.26)
CRAFTOCC		0.0180 (2.13)		0.0195 (1.16)
OPEROCC		-0.0013 (-0.16)		0.0006 (0.08)
LAREMP		0.0376 (8.38)		0.0377 (8.42)
PACREG			0.1213 (6.22)	0.1071 (5.78)
MOUNTREG			-0.0876 (-3.27)	-0.0859 (-3.38)
WSCREG			0.0681 (3.51)	0.0616 (3.36)
ESCREG			0.0699 (2.29)	0.0608 (2.10)
WNCREG			0.0093 (0.36)	0.0056 (0.23)
ENCREG			0.0812 (4.67)	0.0807 (4.89)
SATLREG			0.1001 (6.34)	0.0939 (6.28)
MATLREG			0.0693 (3.61)	0.0642 (3.53)
OBSERVATIONS	25,284	25,284	25,284	25,284
-2 Res Log Likelihood	9420.297	8583.403	9348.787	8515.515

* *t-statistics* in parentheses.