Web Appendix to
"Estimating Marginal and Average Returns to Education"

Pedro Carneiro, James J. Heckman and Edward Vytlacil

July 10, 2009

Abstract

Pedro Carneiro
Department of Economics
University College London
Gower Street
London WC1E 6BT
United Kingdom
p.carneiro@ucl.ac.uk
+44-20-76795824

James Heckman
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637
United States
jjh@uchicago.edu
+1-773-702-0634

Edward Vytlacil
Department of Economics
Yale University
Box 208281
New Haven, CT 06520-8281
United States
edward.vytlacil@yale.edu
+1-203-436-3994

A A Model of Schooling Choice

Consider a standard model of schooling choice. Let $Y_1(t)$ be the earnings of the schooled at experience level t while $Y_0(t)$ is the earnings of the unschooled at experience level t. Assuming that schooling takes one period in which earnings are foregone, a person takes schooling if

$$\frac{1}{(1+r)}\sum_{t=0}^{\infty}\frac{Y_1(t)}{(1+r)^t} - \sum_{t=0}^{\infty}\frac{Y_0(t)}{(1+r)^t} - C^* \geq 0,$$

where C^* is direct cost which may include psychic cost components, r is the discount rate, and lifetimes are assumed to be infinite to simplify the algebra. This is the prototypical discrete choice model applied to human capital investments.[1] We follow Mincer (1974) and assume that earnings profiles in logs are parallel in experience across schooling levels. Thus $Y_1(t) = Y_1 e(t)$ and $Y_0(t) = Y_0 e(t)$, where $e(t)$ is a post-school experience growth factor. Think of "1" as college and "0" as high school.

The agent attends school if

$$\left(\frac{1}{(1+r)}Y_1 - Y_0\right)\sum_{t=0}^{\infty}\frac{e(t)}{(1+r)^t} \geq C^*.$$

Let $K = \sum_{t=0}^{\infty}\frac{e(t)}{(1+r)^t}$ and absorb K into C^* so $\widetilde{C} = \frac{C^*}{K}$. Define discount factor $\gamma = \frac{1}{(1+r)}$. Using growth rate g to relate potential earnings in the two schooling choices we may write $Y_1 = (1+g)Y_0$ (in our empirical model, $\beta = \ln(1+g)$). Thus the decision to attend school $(S = 1)$ is made if

$$Y_0[\gamma(1+g) - 1] \geq \widetilde{C}.$$

This is equivalent to

$$\beta \geq \ln(1 + \frac{\widetilde{C}}{Y_0}) + \ln(1+r).$$

[1] This formulation makes it clear that we are analyzing *ex post* returns, as is conventional in the schooling literature. For an analysis of *ex ante* and *ex post* returns, see Carneiro, Hansen, and Heckman (2003) and Cunha, Heckman, and Navarro (2005).

For $r \approx 0$ and $\frac{\widetilde{C}}{Y_0} \approx 0$, we may write the decision rule as $S = 1$ if

$$\beta \geq r + \frac{\widetilde{C}}{Y_0}. \tag{1}$$

Ceteris paribus, a higher r or $\widetilde{C}$ lowers the likelihood that $S = 1$. As long as $g > r$ (so $\gamma(1+g) - 1 > 0$), a higher Y_0 implies a higher absolute return to college and leads people to attend college. Assuming that direct costs are zero ($\widetilde{C} = 0$) and that the only cost is the opportunity cost of foregone income, the marginal return for those indifferent between going to school and facing interest rate r is $E(\beta \mid \beta = r)$. In the empirical analysis of this paper, we introduce variables Z that shift costs and discount factors ($\widetilde{C} = \widetilde{C}(Z)$, $r = r(Z)$). We use $C(Z) = r(Z) + \frac{\widetilde{C}(Z)}{Y_0}$ in the text.

B Description of the Data

We restrict the NLSY sample to white males.[2] We define participation in college as having attended some college or having completed more than 12 grades in school. The wage variable that is used is an average of deflated (to 1983) non-missing hourly wages reported in 1989, 1990, 1991, 1992 and 1993. We delete all wage observations that are below 1 or above 100. Experience is actual work experience in weeks (we divide it by 52 to express it as a fraction of a year) accumulated from 1979 to 1991 (annual weeks worked are imputed to be zero if they are missing in any given year). The remaining variables that we include in the X and Z vectors are mother's years of schooling, number of siblings, urban residence at 14, schooling corrected AFQT, dummies indicating the year of birth, the presence of a four-year college in the county of residence at age 14 (from Kling, 2001),[3] local average earnings in the county of residence at 17 and local unemployment rate in state of residence at age 17, and

[2]For a description of the NLSY 1979, see Bureau of Labor Statistics (2001).

[3]The distance variable we use is the one used in Kling (2001), available at the *Journal of Business and Economics Statistics* website.

in 1991. Permanent local earnings and unemployment are computed by location of residence at 17 (county for wages, state for unemployment), by averaging values of (deflated) local labor market variables between 1973 and 2000. County earnings correspond to the average wage per job in the county constructed using data from the Bureau of Economic Analysis, deflated to 2000. The state unemployment rate data come from the BLS website. However, from the BLS website it is not possible to get state unemployment data for all states for all the 1970s. Data are available for all states from 1976 on, and for 29 states for 1973, 1974 and 1975. Therefore for some of the individuals we have to assign them the unemployment rate in the state of residence in 1976 (which will correspond to age 19 for those born in 1957 and age 18 for those born in 1958). County and state of residence at 17 are not available for everyone in the NLSY, but only for the cohorts born in 1962, 1963 and 1964 (age 17 in 1979, 1980 and 1981). However, county and state of residence at age 14 is available for most respondents. Therefore, we impute location at 17 to be equal to location at 14 for cohorts born between 1957 and 1962 unless location at 14 is missing, in which case we use location in 1979 for the imputation. The NLSY79 has an oversample of poor whites which we exclude from this analysis. We also exclude the military sample. To remove the effect of schooling on AFQT we implement the procedure of Hansen, Heckman, and Mullen (2004). See the estimates reported in Table A1.

References

Bureau of Labor Statistics (2001). *NLS Handbook 2001: The National Longitudinal Surveys.* Washington, DC: U.S. Department of Labor.

Carneiro, P., K. Hansen, and J. J. Heckman (2003, May). Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice. *International Economic Review 44*(2), 361–422.

Cunha, F., J. J. Heckman, and S. Navarro (2005, April). Separating uncertainty from heterogeneity in life cycle earnings, The 2004 Hicks Lecture. *Oxford Economic Papers 57*(2), 191–261.

Hansen, K. T., J. J. Heckman, and K. J. Mullen (2004, July–August). The effect of schooling and ability on achievement test scores. *Journal of Econometrics 121*(1-2), 39–98.

Kling, J. R. (2001, July). Interpreting instrumental variables estimates of the returns to schooling. *Journal of Business and Economic Statistics 19*(3), 358–364.

Mincer, J. (1974). *Schooling, Experience and Earnings.* New York: Columbia University Press for National Bureau of Economic Research.

Table A1
Regression of AFQT on Schooling at Test Date and Completed Schooling

Schooling at Test Date	Coefficient
9	12.6802
	(1.5105)
10	16.9406
	(1.5158)
11	22.0232
	(1.5354)
12	23.1203
	(1.4901)
13 to 15	26.6032
	(1.7298)
16 or greater	29.0213
	(2.1278)

Note: These are coefficients of the AFQT score on schooling at test date and complete schooling:

$$\text{AFQT} = \delta_0 + \Sigma_{ST} D_{ST} \delta_{ST} + \Sigma_{SC} D_{SC} \delta_{SC} + \eta$$

D_{ST} are dummy variables, one for each level of schooling at test date and δ_{ST} are the coefficients on these variables. D_{SC} are dummy variables, one for each level of completed schooling and δ_{SC} are the coefficients on these variables. The omitted category in the table is "less or equal to eight years of schooling."

Table A2
Maximum Likelihood Estimates of the Normal Switching Regression Model

	$\mu_1(X)$	$\mu_0(X)$	$\mu_Z(X)$
CONTROLS (X):			
Years of Experience	0.0799	0.0541	
	(0.0179)	(0.0202)	
Years of Experience Squared	-0.0035	0.0004	
	(0.0012)	(0.0011)	
Local Log Earnings in 1991	0.8259	0.5279	
	(0.1080)	(0.1068)	
Local Unemployment in 1991	0.0038	-0.0032	
	(0.0154)	(0.0147)	
Corrected AFQT	0.1328	0.0563	3.1990
	(0.0444)	(0.0290)	(2.7991)
Corrected AFQT Squared	0.0550	-0.0469	0.2031
	(0.0185)	(0.0174)	(0.0382)
Mother's Years of Schooling	-0.0095	-0.0201	-1.8849
	(0.0461)	(0.0300)	(1.1197)
Mother's Years of Schooling Squared	0.0015	0.0010	0.0094
	(0.0017)	(0.0014)	(0.0039)
Number of Siblings	-0.0119	0.0033	-4.0161
	(0.0278)	(0.0173)	(1.3345)
Number of Siblings Squared	0.0003	-0.0005	0.0023
	(0.0036)	(0.0017)	(0.0067)
Urban Residence at 14	0.0562	0.0089	0.1113
	(0.0404)	(0.0307)	(0.0858)
"Permanent" Local Log Earnings at 17	0.5713	12.2602	-51.3135
	(7.0701)	(7.2801)	(17.3948)
"Permanent" Local Log Earnings at 17 Squared	-0.0270	-0.5898	2.5234
	(0.3446)	(0.3557)	(0.8490)
"Permanent" State Unemployment Rate at 17	0.1436	0.0716	0.2345
	(0.1688)	(0.1354)	(0.3821)
"Permanent" State Unemployment Rate at 17 Squared	-0.0133	-0.0059	-0.0167
	(0.0132)	(0.0104)	(0.0292)
INSTRUMENTS (Z):			
Presence of a College at 14			0.8694
			(0.4931)
* AFQT			0.1270
			(0.0881)
* Mother's Education			-0.0669
			(0.0388)
* Number of Siblings			0.0301
			(0.0401)
Local Log Earnings at 17			-4.2099
			(1.4644)
* AFQT			-0.2508
			(0.2737)
* Mother's Education			0.1942
			(0.1102)
* Number of Siblings			0.3843
			(0.1318)
Local Unemployment Rate at 17			0.3358
(in %)			(0.1323)
* AFQT			-0.0015
			(0.0241)
* Mother's Education			-0.0240
			(0.0101)
* Number of Siblings			-0.0061
			(0.0109)

Note: This table reports the maximum likelihood estimates of the coefficients of the variables in the selection and outcome equations for the normal selection model. The model also includes cohort dummies which are not reported.

Table A3
Average Derivatives for the Wage Equation - Estimates from the Partially Linear Model

	$\mu_0(X)$	$\mu_1(X) - \mu_0(X)$
Years of Experience	0.0459	-0.0598
	(0.0012)***	(0.0210)***
Local Log Earnings in 1991	0.0417	0.7390
	(0.1627)**	(0.2981)**
Local Unemployment in 1991	0.0030	0.0030
	(0.0246)	(0.0443)
Corrected AFQT	-0.0456	0.2505
	(0.1112)	(0.2246)
Mother's Years of Schooling	-0.0016	0.0074
	(0.0212)	(0.0341)
Number of Siblings	0.0183	-0.0284
	(0.0166)	(0.0305)
Urban Residence at 14	0.0220	-0.0010
	(0.0530)	(0.0968)
"Permanent" Local Log Earnings at 17	0.1681	-0.3552
	(0.1647)	(0.2590)
"Permanent" State Unemployment Rate at 17	0.0080	-0.0515
	(0.0302)	(0.0559)

Note: *** Significant at 1%. ** Significant at 5%. * Significant at 10%. This table reports the coefficients on the variables in the outcome equations for the semi-parametric selection model. These coefficients are obtained using Robinson's (1989) partially linear regression method. The model also includes cohort dummies which are not reported. Standard Errors are Bootstrapped (250 Replications).

Table A4 - OLS and IV Estimates of the Return to a Year of College

Two Stage Least Squares						
	OLS	IV				
		College	Earnings	Unemployment	All	P
Return to College	0.0726	-0.0059	0.2169	0.2149	0.1597	0.0226
	(0.0070)	(0.0866)	(0.0709)	(0.1339)	(0.0499)	(0.0362)
Limited Information Maximum Likelihood						
Return to College	0.0726	0.0576	0.1736	0.1582	0.1233	0.0672
	(0.0070)	(0.0727)	(0.0789)	(0.1898)	(0.0517)	(0.0480)

Note: This table reports OLS and IV estimates of the return to college attendance (a dummy variable that is equal to 1 if an individual has ever attended college and equal to 0 if he has never attended college but has graduated from high school). The coefficient on college is divided by 4 to reflect the difference in years of schooling between those with and without college. Instruments are the presence of a college in the county of residence at 14, local earnings and local unemployment in the area of residence at 17 (interacted with AFQT, mother's education and number of siblings). Standard Errors are bootstrapped (250 replications).