Returns to Education and Wage Equations: a Dynamic Approach

Corrado Andini *

ABSTRACT

We study the impact of education on within-groups wage inequality using quantile-regression techniques and U.S. data for the period of 1980-1987. Our contribution consists of comparing estimates based on a standard Mincer equation with estimates based on a modified Mincer equation in which past earnings play the role of additional explanatory variable. We find that a dynamic model does not give the same answer as a static model regarding the impact of schooling on earnings dispersion, and provide an explanation for this result.

Keywords: Education, Wage Inequality, Quantile Regression.
JEL Classification: I21, J31, C29.

* *Assistant Professor, University of Madeira, Department of Management and Economics, Funchal, Portugal. Email: andini@uma.pt, Tel: +351 291 70 5053, Fax: +351 291 70 5040.*

Financial support by the European Commission (EDWIN Project, HPSE-CT-2002-00108) is gratefully acknowledged. I sincerely thank Pedro Telhado Pereira, Santiago Budría and the participants at the 6th Meeting of the EDWIN Project for their valuable comments. The usual disclaimer applies.

I. INTRODUCTION

The literature using a wage equation *a la* Mincer (1974) is rich in contributions. Most of existing studies share three common features:

- the estimated models have a static nature, not allowing for at least one lagged value of the dependent variable as additional regressor;
- the estimated coefficient of education is dependent on number and type of explanatory variables added to the standard Mincer equation[1];
- estimation is generally based on ordinary least squares, instrumental variables, random effects.

We attempt to do a step onwards from the current "state of the art" and explore returns to schooling in a dynamic framework. From an empirical point of view, this mainly involves keeping the autoregressive nature of earnings into account. In doing so, we also deal with the problem of choosing control-regressors by replacing the whole set of explanatory variables suitable to be added to a standard Mincer framework with one lagged value of earnings (this approach may be extended to more than one lag). Finally, following Buchinsky (1994) among others, we use quantile-regression techniques for studying the impact of education on within-groups wage inequality.

II. EMPIRICAL MODELS

We compare estimates of the return to schooling implied by the following two models:

$$\ln w_{it} = \alpha_\theta + \beta_\theta s_i + \delta_\theta z_{it} + \phi_\theta z_{it}^2 + \varepsilon_{\theta it} \tag{1}$$

$$\ln w_{it} = \gamma_\theta + \upsilon_\theta \ln w_{it-1} + \pi_\theta s_i + \chi_\theta z_{it} + \varsigma_\theta z_{it}^2 + \mu_{\theta it} \tag{2}$$

[1] See Martins and Pereira (2004).

where $\ln w$ represents the logarithm of the hourly wage, s is years of education, z is labor-market experience, θ goes from 0 to 1 and represents the wage-distribution quantile. As both models are linear in parameters, and we do not focus "on the case of iid innovations in which conditioning variables play the classical role of shifting the *location* of the conditional density of y_t [the autoregressive variable], but they have no effect on the conditional scale or shape" (Koenker and Xiao, 2004, p. 3), then we can apply standard quantile-regression techniques due to Koenker and Bassett (1978).

The return to schooling implied by model (1) is equal to β, the return to schooling implied by model (2) is equal to $\frac{\pi}{1-\upsilon}$. We may label the first as *static return*, the latter as *dynamic return*.

III. ESTIMATION RESULTS

We use data from the U.S. National Longitudinal Survey of Youth for the period of 1980-1987, as provided by Verbeek (2000). This data-set has been already used by Vella and Verbeek (1998) as well as, more recently, by Wooldridge (2005). The sample contains 4360 annual observations on 545 young male workers. We assume absence of participation issues typically arisen for women. Summary sample statistics for the selected variables are reported in Table 1.

As Table 2 suggests, our two models roughly provide the same answer regarding the impact of education on both median and mean earnings.

An argument in favor of a dynamic approach is that the variability of earnings at each decile (measured by the pseudo R-square) explained by model (2) varies from 25 to 39 percent, while the variability explained by model (1) varies from 5 to 10 percent. This is not surprising and, of course, consequence of adding an explanatory variable (lagged earnings) to the standard Mincer specification. Nevertheless, the outcome of an increased explained variability of wages remains a desirable outcome in many circumstances, including the analysis of the impact of education on within-groups earnings dispersion.

On this very last point, we may observe Figure 1 and notice that our two models give contradictory answers. The static model (1) tells us that education does not reduce earnings dispersion since the return to schooling lightly increases over the wage distribution, *while* the dynamic model (2) suggests that education may reduce wage dispersion.

In our view, a possible explanation of the above contradiction is that the hypothesis underlying the so-called model (1), that is:

$$(3) \qquad \mathrm{Quant}_{\theta}(\varepsilon_{\theta it} | s_{it}, z_{it}, z_{it}^{2}) = 0 \quad \text{for each } \theta \ ,$$

does not necessarily hold for each θ, resulting in inconsistent estimates at some θ. Conditioning on past earnings, as in model (2), allows both attenuating the omitted-variable bias and estimating the *total* return to schooling by means of simple algebra[2]. Alternatively, if we are interested in estimating the *total* return to schooling using model (1) and we want to attenuate the omitted-variable bias, then we should be very careful in selecting as additional regressors for model (1) a set of variables *not* being education-dependent covariates (as a matter of example, we should not use industry dummies[3]). This, however, may be an uneasy exercise as previously argued by Martins and Pereira (2004).

IV. CONCLUSION

[2] Martins and Pereira (2004) provide an exhaustive discussion of the concept of *total* return to schooling.

[3] Jobs in same industries require more years of education than in other industries.

In a recent *Applied Economics* article, Martins and Pereira (2004) argued in favor of a simple Mincer specification for estimating the *total* return to schooling. This approach, however, incurs the risk of providing inconsistent estimates because of omitted-variable problems, particularly when using quantile-regression techniques that imply stronger assumptions of correct model-specification than ordinary least squares. We propose an alternative method for the estimation of the *total* return to schooling when longitudinal data are available. The introduction of past earnings as additional explanatory variable increases the explained variability of wages and reduces the risk of inconsistency *without* implying any additional difficulty for the issue of recovering the *total* return to education.

REFERENCES

Buchinsky, M. (1994) Change in the U.S. Wage Structure 1963-1987: Application of Quantile Regression, *Econometrica*, 62(2), pp. 405-458.

Koenker, R. and Bassett, G. (1978) Regression Quantiles, *Econometrica*, 46(1), pp. 33-50.

Koenker, R. and Xiao, Z. (2004) Quantile Autoregression, unpublished manuscript.

Martins, P.S. and Pereira, P.T. (2004) Returns to Education and Wage Equations, *Applied Economics*, 36(6), pp. 525-531.

Mincer, J. (1974) *Schooling, Experience and Earnings*, Cambridge, National Bureau of Economic Research.

Vella, F. and Verbeek, M. (1998) Whose Wages Do Unions Raise? A Dynamic Model of Unionism and Wage Rate Determination for Young Men, *Journal of Applied Econometrics*, 13, pp. 163-183.

Verbeek, M. (2000) *A Guide to Modern Econometrics*, Chichester, John Wiley & Sons.

Wooldridge, J. M. (2005) Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Effects, *Journal of Applied Econometrics*, 20(1), pp. 39-54.

Table 1. Descriptive Statistics

Variable	Obs.	Mean	Std. Dev.	Min	Max
Logarithm of hourly wage	4360	1.64	0.53	–3.57	4.05
Years of schooling	4360	11.76	1.74	3.00	16.00
Experience	4360	6.51	2.82	0.00	18.00
Age	4360	24.28	2.77	17.00	30.00

Table 2. Estimation Results

Distribution decile	$\hat{\upsilon}$	$\hat{\pi}$	$\hat{\beta}$	$\frac{\hat{\pi}}{1-\hat{\upsilon}}$
0.1	0.7591	0.0404	0.0884	0.1677
	(0.0215)	(0.0069)	(0.0090)	
0.2	0.8036	0.0292	0.0954	0.1486
	(0.0121)	(0.0121)	(0.0062)	
0.3	0.8201	0.0220	0.0954	0.1222
	(0.0094)	(0.0033)	(0.0055)	
0.4	0.8231	0.0214	0.1004	0.1209
	(0.0076)	(0.0026)	(0.0052)	
0.5	0.7911	0.0216	0.1036	0.1033
	(0.0081)	(0.0027)	(0.0047)	
0.6	0.7660	0.0219	0.1066	0.0935
	(0.0089)	(0.0028)	(0.0047)	
0.7	0.7017	0.0224	0.1070	0.0750
	(0.0113)	(0.0032)	(0.0054)	
0.8	0.6298	0.0326	0.1058	0.0880
	(0.0160)	(0.0041)	(0.0056)	
0.9	0.4789	0.0511	0.1072	0.0980
	(0.0287)	(0.0061)	(0.0063)	
OLS	0.5786	0.0447	0.1021	0.1061
	(0.0274)	(0.0050)	(0.0044)	

Standard errors in parentheses. Estimates of υ's, π's and β's are all significant at 1 percent level.

Figure 1. Returns to education at different deciles of the wage distribution