I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12237

Education and Risk Compensation in Wages: A Quantile Regression Approach

José Vieira
Carolina Constância
João Teixeira

MARCH 2019

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12237

Education and Risk Compensation in Wages: A Quantile Regression Approach

José Vieira
University of the Azores, Centre for Applied Economic Studies of the Atlantic and IZA

Carolina Constância
University of the Azores

João Teixeira
University of the Azores and Centre for Applied Economic Studies of the Atlantic

MARCH 2019

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Education and Risk Compensation in Wages: A Quantile Regression Approach

This paper examines the effect of wage variation on individual wages. The results reveal that wage variation by educational classifications positively affects wages, while the skewness has a negative effect. As has been referred in previous literature on the issue, both results are consistent with the notion of wage compensation for risk-averse workers. However, our results show that the impact of wage-variation on wages is not reasonably described by a single parameter for all individuals. Such an effect is heterogeneous and varies throughout the conditional wage distribution. Indeed, the positive effect of dispersion increases, and the negative effect of skewness decreases, as we move up on the conditional wage distribution. Apparently, those at the upper end of the conditional wage distribution have both higher risk-aversion and higher affection for skewness.

JEL Classification: J31, J33

Keywords: risk compensation, education, wage formation, quantile regression

Corresponding author:
José Vieira
School of Business and Economics
University of the Azores
Rua Mãe de Deus
9500-801 Ponta Delgada
Portugal
E-mail: jose.ac.vieira@uac.pt

I. Introduction

Theoretically, risk-averse individuals dislike risk measured by the variance of the income distribution and thus demand a higher expected income for undertaking a higher risk (Arrow 1965; Pratt 1964). Such a prediction is well-established within the portfolio theory. Furthermore, Tsiang (1972), relying on requirements of the utility function for a risk-averse individual presented by Arrow (1965), shows that a decreasing absolute risk aversion implies a preference for income skewness. This risk-averse agent's preference for skewness has received attention in the literature regarding gambling (Goleck and Tomarkin 1998) and labor markets (Hartog and Vijverberg 2007).

With regard to the labor market, some authors have examined how wages respond to wage-variation (King 1974; Johnson 1977; Feinberg 1981; Moore 1995; McGoldrick and Robst 1996; Hartog et al. 2003; Hartog and Vijverberg 2007; Diaz-Serrano, Hartog, and Nielsen 2008; Nalg 2012). Most of the literature on the issue uses a Mincer-type human capital earnings function extended measures of wage variation (risk) and skewness. Formal derivation of this type of wage equation for a risk-averse individual, which predicts that individuals want to be compensated for a higher variance but are willing to accept lower wages as skewness increases, can be found in Hartog and Vijverberg (2007) and Nalg (2012).

Regarding the existing empirical work, the measures of risk and skewness have been calculated within occupational cells (e.g. King, 1974; Johnson, 1977; McGoldrick and Robst, 1996; Hartog et al., 2003; Nalg, 2012), educational classifications (Diaz-Serrano, Hartog, and Nielsen 2008) or a mix of occupational and educational cells (Hartog and Vijverberg 2007). In any case, the empirical results consistently reveal a wage compensation for risk and a payment for skewness. That is, workers dislike the variance of wage distribution but appreciate small probabilities of receiving very high wages (Hartog et al, 2003; Hartog and Vijverberg, 2007).

The extended wage function has usually been estimated by ordinary least squares (e.g. King, 1974; Johnson, 1977; McGoldrick and Robst, 1996; Hartog et al., 2003). As suggested by Hartog et al., (2003), it would be interesting to pursue with further panel data estimates. Moore (1995) used panel data and found no effect of risk on wages. However, Nagl (2012) used panel data for Germany and confirmed a positive coefficient associated to wages dispersion and a negative coefficient associated to skewness through a fixed effects estimation method.

The main purpose of this letter is to shed further light on this literature through a quantile regression analysis. Indeed, a few authors have shown that the impact on the mean of the conditional wage distribution overlooks important features of the wage formation process (Hartog et al., 2001; McGuiness and Bennett, 2007; Prieto-Rodriguez, Barros, and Vieira 2008).

The remainder of the letter is organized as follows. The next section describes the model and the estimation method. Section 3 includes the data. Section 4 examines the estimation results. Finally, Section 5 concludes.

II. The Empirical Model

The Wage Equation

As in previous literature (e.g., McGoldrick and Robst, 1996; Hartog et al., 2003) we proceed in two stages. Firstly, we estimate by ordinary least squares a standard human capital wage function of the type:

$$\ln w_i = \beta_0 + \beta_1' X_i + u_i \qquad i=1,\ldots,N \qquad (1)$$

where w_i denotes X_i is a vector of explanatory variables.

Once estimated, in this first equation we take the antilog of the residuals given by e_i=exp($\hat{u}_i$), where $\hat{u}_i$ denotes the estimated residuals in equation (1), which are collected within educational classifications. For each classification j, we calculate the measures of risk (R_j) and skewness (S_j) as follows:

$$R_j = \frac{\sum_i (e_{ij} - \bar{e}_j)^2}{N_j - 1} \tag{2}$$

$$S_j = \frac{N_j}{(N_j - 1)(N_j - 2)} \frac{\sum_i (e_{ij} - \bar{e}_j)^3}{sd_j^3} \tag{3}$$

where sd_j denotes the standard deviation of e_i for the educational classification j.

These values are assigned to each individual in the sample in accordance with the educational classification to which he/she belongs. Then, as done by McGoldrick and Robst (1996) and Hartog et al. (2003), in a second step the following final equation is estimated:

$$\ln w_i = \alpha_0 + \alpha_1' X_i + \theta_1 R_{ij} + \theta_2 S_{ij} + v_i \tag{4}$$

If individuals dislike risk and enjoy wage skewness, then the estimated values for θ_1 and θ_2 must be positive and negative, respectively.

The Estimation Method

Equation (2) is estimated by ordinary least squares (OLS) as in most of the literature, which evaluates only the impact of wage variation and skewness on the mean of the conditional wage distribution. It is also estimated using a quantile regression approach (QR).

Let us consider $M_i = \alpha_0 + \alpha_1' X_i + \theta_1 R_{ij} + \theta_2 S_{ij}$ and $y_i = \ln w_i$. Following Koenker and Basset (1978), the θth quantile regression is the solution to the problem:

$$\min_{\beta \in R^k} \left[\sum_{(i:y_i \geq M_i)} \theta |y_i - M_i| + \sum_{(i:y_i < M_i)} (1-\theta) |y_i - M_i| \right], \ \theta \in (0,1) \qquad (5)$$

This can also be written as:

$$\min_{\beta \in R^k} \sum_i \rho_\theta (y_i - M_i) \qquad (6)$$

where $\rho_\theta(\varepsilon)$ is the so-called check-function defined as

$$\rho_\theta(\varepsilon) = \begin{cases} \theta\varepsilon & if \ \ \varepsilon \geq 0 \\ (\theta - 1)\varepsilon & if \ \ \varepsilon < 0 \end{cases} \qquad (7)$$

By variation in θ, different quantiles can be estimated. Therefore, by changing θ from 0 to 1, we trace the entire distribution of wages conditional on the explanatory variables. Since M_i is formulated as a linear function of parameters, the minimization problem can be solved by linear programming methods (Koenker and Basset 1978; Koenker and Hallock, 2001). For this purpose, we use the Stata statistical software package, which relies on the linear programming algorithm suggested by Armstong, Frome, and Kung (1979). Standard errors are obtained by bootstrap methods.

III. The Data

The data used were drawn from *Quadros de Pessoal*, and the sample includes 132,755 observations for Portugal. All firms with wage earners must complete a standardized questionnaire for the Portuguese Department of Labour. The data includes information on workers, such as age, education, tenure with the firm, gender, monthly hours of work and monthly wages. The data also includes information on firm size measured by the number of employees, firm age, regions and industry affiliation. Hourly wages were calculated by dividing monthly wages by monthly hours of work. Years of education were calculated by imputing the normal number of years of education to complete the corresponding type of education. The data include 99 educational classifications for each of which variance and skewness can be calculated following the procedure mentioned in Section II. These educational classifications range from primary to higher education. Moreover, within each level they are divided into sub-levels and, within these, by areas of study (e.g. higher education, bachelor's degree, computer science).

IV. Results

We first we estimate the equation (1) by ordinary least squares. The dependent variable is the logarithm of hourly wages and the explanatory variables include years of education, age and age squared, years of tenure with the firm, gender, firm size, firm age, regions, and industries. Then, we estimate equation (4) that, in addition to these covariate, also includes measures of risk and skewness, as indicated in Section II. These measures were calculated separately for each of the 99 educational classifications from the estimated residuals of equation (1), and assigned to each individual based on the classification to which he belongs.

Estimated results are depicted in Figure 1 and Figure 2 and included in Table 1. The findings confirm previous studies: the coefficient of risk is positive, and the coefficient of skewness is negative. Moreover, this is valid not only for the impact of these covariates on the mean of the conditional wage distribution (OLS estimates) but throughout the entire distribution (QR estimates).

However, the effects are far from being the same across different conditional quantiles. The coefficient associated to risk rises as we move up on the conditional distribution, and the coefficient associated to skewness decreases. The impact of an additional unit of risk is 4.2 times higher at the .95 quantile than at the 0.05 quantile.

The coefficient of skewness is 7.3 times higher at the .95 quantile than at the .005 quantile. A test of equality of coefficients rejects the null hypotheses that the effects are equal across the conditional wage distribution (Table 2).

The analysis of the effect on the conditional mean overlooks important aspects of those covariates to wage formation. These results uncover a considerable amount of unobserved heterogeneity on the effect of wage variation and skewness on individual wages. Such heterogeneity might be related to unobserved differences in risk aversion and skewness affection across individual workers. As we move up on the conditional wage distribution, workers demand higher compensation for each additional unit of risk. On the other hand, workers' payment for skewness drops (that is, workers at higher quantiles are willing to pay less for small probabilities of receiving very high wages).

V. Conclusion

We have provided evidence that evaluating at the mean overlooks important aspects of the effect of wage-variation on wages. Quantile regression analysis uncovers important features of the contribution of risk and skewness for wage formation. The outcome of wage-variation on wages is not reasonably described by a single parameter for all individuals. The results suggest that those at the upper (lower) end of the conditional wage distribution require a higher (lower) compensation for an additional amount of risk. In other words, as we move up in the conditional wage distribution risk-aversion apparently increases. Moreover, workers at the upper (lower) end of the conditional wage distribution have a higher (lower) affection for skewness.

References

Armstong, R. D., E. L. Frome, and D. S. Kung. 1979. "Algorithm 79-01: A Revised Simplex Algorithm for the Absolute Deviation Curve Fitting Problem." *Communications in Statistics, Simulation and Computation* B8 (2): 175-190.

Arrow, K. 1965. *Aspects of the Theory of Risk-Bearing*. Helsinki: Yrjö Jahnssonin Säätiö.

Diaz-Serrano, L., J. Hartog, and H. S. Nielsen. 2008. "Compensation Wage Differentials for Schooling Risk in Denmark." *Scandinavian Journal of Economics* 110 (4): 711-731.

Feinberg, R. M. 1981. "Earnings-risk as a Compensation Differential." *Southern Economic Journal* 48 (1): 156-163. doi: 10.2307/1058607.

Goleck, J., and M. Tomarkin. 1998. "Bettors love skewness, not risk, at the horse track." *Journal of Political Economy* 106 (1): 205-225. doi: 10.1086/250007.

Hartog, J., and W. Vijverberg. 2007. "On Compensation for Risk Aversion and Skewness Affection in Wages." *Labour Economics* 14 (6): 938-956. doi:10.1016/j.labeco.2007.08.002.

Hartog, J., E. Plug, L. Diaz-Serrano, and J. Vieira. 2003. "Risk Compensation in Wages – A Replication." *Empirical Economics* 28 (3): 639-647. doi: 10.1007/s001810200150

Hartog. J., P. T. Pereira, and J. A. C. Vieira. 2001. "Changing Returns to Education in Portugal during the 1980s and Early 1990s: OLS and Quantile Regression Estimators." *Applied Economics* 33 (8): 1021-1037. doi: 10.1080/00036840122679.

Johnson, W. R. 1977. "Uncertainty and the Distribution of Earnings". In *The Distribution of Economic Well-Being*, edited by F. Fuster, 379-396. Cambridge: National Bureau of Economic Research, Inc.

King, A. G. 1974. "Occupational Choice, Risk Aversion and Wealth." *Industrial and Labor Relations Review* 27 (4): 586-196. doi: 10.2307/2522049.

Koenker, R., and G. Basset. 1978. "Regression Quantiles." *Econometric*a 46 (1): 33-50. doi: 10.2307/1913643.

Koenker, R., and K. F. Hallock. 2001. "Quantile Regression." *Journal of Economic Perspectives* 15 (4): 143-156. doi: 10.1257/jep.15.4.143.

McGoldrick, K., and J. Robst, J. 1996. "The Effect of Worker Mobility on Compensating Wages for Earnings Risk." *Applied Economics* 28 (2): 221-232. doi:10.1080/000368496328858.

McGuiness, S., J. Bennett. 2007. "Overeducation in the Graduate Labour Market: A Quantile Regression Approach." *Economics of Education Review* 26 (5): 521-531. doi: 10.1016/j.econedurev.2005.12.003.

Moore, M. 1995 "Unions, Employment Risks, and Market Provision of Employment Risk Differentials." *Journal of Risk and Uncertainty* 10 (1): 57-70. doi: 10.1007/BF01211528.

Pratt, J. 1964. "Risk Aversion in the Small and in the Large." *Econometrica* 32 (1/2): 122-136. doi:10.2307/1913738.

Prieto-Rodriguez, J., C. P. Barros,, and J. A. C. Vieira,. 2008. "What a Quantile Approach can Tell us About Returns to Education in Europe." *Education Economics* 16 (4): 391-410. doi:10.1080/09645290801977033.

Schweri, J., Hartog, J., and Wolter, S. 2011. "Do Students Expect Compensation for Wage Risk?." *Economics of Education Review*, 30 (2): 215-227. doi:10.1016/j.econedurev.2010.12.001.

Tsiang, S. C. 1972. “The Rationale for Mean-Standard Deviation Analysis, Skewness Preference and the Demand for Money.” *American Economic Review* 63 (2): 354-371.

Table 1. Quantile regression (QR) and ordinary least squares (OLS) estimates for risk and skewness

	Risk			Skewness		
Quantile	Coef.	Std. Error		Coef.	Std. Error	
0.05	0.00997	0.00040	***	-0.00048	0.00001	***
0.10	0.01290	0.00038	***	-0.00066	0.00001	***
0.15	0.01421	0.00038	***	-0.00077	0.00001	***
0.20	0.01550	0.00038	***	-0.00088	0.00001	***
0.25	0.01769	0.00039	***	-0.00099	0.00001	***
0.30	0.01903	0.00040	***	-0.00111	0.00001	***
0.35	0.02091	0.00038	***	-0.00125	0.00001	***
0.40	0.02250	0.00042	***	-0.00139	0.00001	***
0.45	0.02448	0.00044	***	-0.00155	0.00001	***
0.50	0.02611	0.00048	***	-0.00172	0.00001	***
0.55	0.02772	0.00048	***	-0.00188	0.00001	***
0.60	0.02931	0.00053	***	-0.00205	0.00002	***
0.65	0.03049	0.00056	***	-0.00223	0.00002	***
0.70	0.03206	0.00063	***	-0.00245	0.00002	***
0.75	0.03380	0.00066	***	-0.00265	0.00002	***
0.80	0.03576	0.00078	***	-0.00287	0.00002	***
0.85	0.03733	0.00097	***	-0.00307	0.00003	***
0.90	0.03845	0.00113	***	-0.00326	0.00004	***
0.95	0.04183	0.00144	***	-0.00350	0.00005	***
OLS	0.02737	0.00051	***	-0.00201	0.00001	***

*** Significant at the 1% level. Each regression includes also controls for years of education, age, age squared, tenure with the firm, gender, firm size, firm age, regions and industries. The dependent variable is the logarithm of hourly wages.

Table 2. Interquantile regression (for specific comparisons)

Quantiles	Risk Coef.	Risk Std. Error		Skewness Coef.	Skewness Std. Error	
0.25 - 0.05	0.00772	0.00053	***	-0.00051	0.00001	***
0.50 - 0.25	0.00842	0.00052	***	-0.00073	0.00002	***
0.75 - 0.50	0.00769	0.00069	***	-0.00093	0.00002	***
0.95 - 0.75	0.00803	0.00113	***	-0.00085	0.00004	***

*** The difference is statistically significant at the 1% level. Bootstrap standard errors with 100 repetitions.

Figure 1. Ordinary least squares (OLS) and quantile regression (QR) estimates for the effect of risk on wages

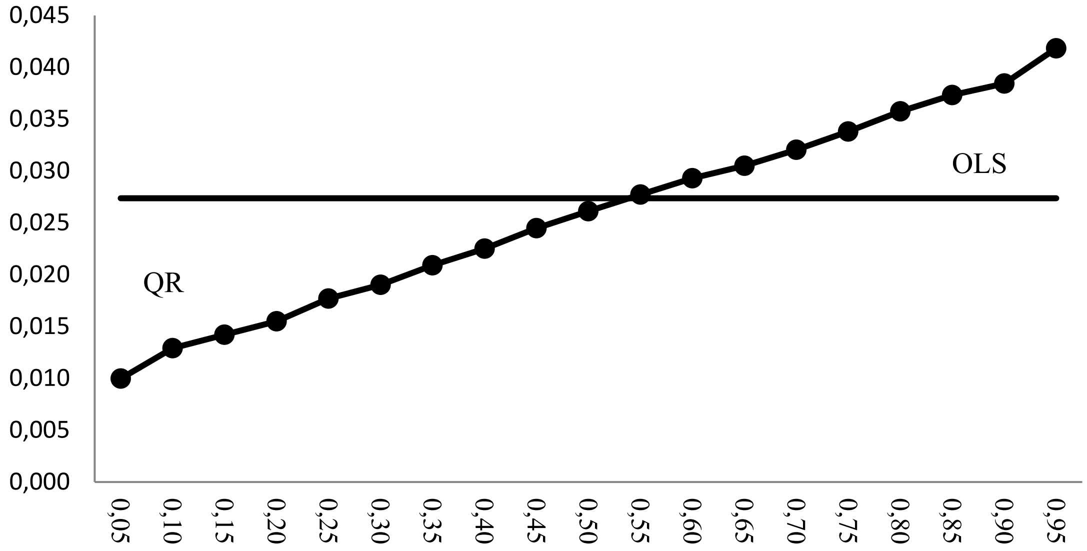

Figure 2. Ordinary least squares (OLS) and quantile regression (QR) estimates for the effect of skewness on wages