JOURNAL OF APPLIED ECONOMETRICS
J. Appl. Econ. **32**: 218–223 (2017)
Published online 14 July 2015 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/jae.2477

REPLICATION SECTION

REPLICATION OF UNCONDITIONAL QUANTILE REGRESSIONS BY FIRPO, FORTIN AND LEMIEUX (2009)

BADI H. BALTAGI[a*] AND PALLAB KUMAR GHOSH[b]
[a] *Department of Economics and Center for Policy Research, Syracuse University, NY, USA*
[b] *Department of Economics, University of Oklahoma, Norman, OK, USA*

SUMMARY

This paper is a narrow replication of Firpo, Fortin and Lemieux (Unconditional quantile regressions. *Econometrica* 2009; 77(3): 953–973), who propose a new estimation method, called 'unconditional quantile regressions'. Using their empirical example, we confirm their results for the effects of unionization on US wage inequality during the period 1983–1985 for both conditional and unconditional quantile regressions. In addition, this paper applies the proposed estimation method to another paper by Autor, Katz and Kearney (Trends in US wage inequality: revising the revisionists. *Review of Economics and Statistics* 2008; 90(2): 300–323). The latter paper looks at measures of wage inequality in the US labor market data over the period 1963–2005. Copyright © 2015 John Wiley & Sons, Ltd.

Received 12 September 2014; Revised 19 January 2015

Supporting information may be found in the online version of this article.

1. INTRODUCTION

Firpo *et al.* (2009) introduce the concept of unconditional quantile regression method, which estimates the impact of changes in explanatory variables on the unconditional quantiles of the outcome variable. Unconditional quantile regressions are also known as RIF regressions since the regression method consists of running a regression of a transformation — the recentered influence function[1] of the outcome variable on the explanatory variables. An important difference between conditional quantile regressions (Koenker and Bassett, 1978) and this approach is that conditional quantiles do not average up to their unconditional population counterparts. As a result, the estimates obtained by running a quantile regression cannot be used to estimate the impact of an explanatory variable X on the corresponding unconditional quantile of the dependent variable Y.

* Correspondence to: Badi H. Baltagi, Department of Economics and Center for Policy Research, Syracuse University, NY, USA. E-mail: bbaltagi@maxwell.syr.edu
[1] Recentered influence function is defined in the next section.

The unconditional quantile regression approach has attracted much attention in the econometrics literature since Firpo *et al.* (2009) introduced the concept. The purpose of this replication is to investigate the robustness of the findings in Firpo *et al.* (2009). Oaxaca's (1973) decomposition method is widely used in labor economics. In the wage decomposition method literature, unconditional quantile regressions can be used to construct a counterfactual wage distribution function. Hence we also compare the decomposition of the US wage distribution function results as shown by Autor *et al.* (2008) with the results obtained using the unconditional quantile regressions approach.

2. UNCONDITIONAL PARTIAL EFFECTS

Firpo *et al.* (2009) consider the following model:

$$Y = h(X, \epsilon)$$

where X represents the set of independent variables and ϵ is the scalar unobserved error term. It is assumed that $h(\cdot)$ is strictly monotonic in ϵ. The unconditional partial effect (UPE) is defined as the small location shift in the distribution of a continuous variable X on the distributional statistic $\nu(F_Y)$. This can be expressed as

$$\alpha(\nu) = \int \frac{\mathrm{d}E\left[\mathrm{RIF}(Y,\nu)|X = x\right]}{\mathrm{d}x}\mathrm{d}F(x)$$

where $\mathrm{RIF}(Y, \nu)$ is the recentered influence function. When the distributional statistic ν is the τth quantile function, $q_\tau = \inf_q \{q : F_Y(q) \geq \tau\}$, the $\mathrm{RIF}(Y, q_\tau)$ can be represented as

$$\mathrm{RIF}(Y, q_\tau) = q_\tau + \frac{\tau - 1\{y \leq q_\tau\}}{f_Y(q_\tau)}$$

where $f_Y(q_\tau)$ is the density function of Y evaluated at q_τ. Firpo *et al.* (2009) show that the unconditional quantile partial effect (UQPE) can be expressed as the weighted average of the conditional quantile partial effects (CQPE):

$$\mathrm{UQPE}(\tau) = E\left[\omega_\tau(X).\frac{\partial h(X, \epsilon_\tau(X))}{\partial x}\right] \equiv E\left[\omega_\tau(X).\mathrm{CQPE}(\zeta_\tau(x), x)\right]$$

where the weight function is defined as $\omega_\tau(X) \equiv f_{Y|X}(q_\tau|x)/f_Y(q_\tau)$ and the matching function $\zeta_\tau : X \to (0, 1)$ is given by

$$\zeta_\tau(x) \equiv \{s : Q_s[Y|X = x] = q_\tau\} = F_{Y|X}(q_\tau|X = x)$$

The matching function indicates when the unconditional quantile q_τ falls in the conditional distribution of Y.

3. EMPIRICAL APPLICATIONS

Firpo *et al.* (2009) estimate the impact of declining unionization on the US wage inequality by using a sample of 266,956 men from the 1983–1985 Outgoing Rotation Group supplement of the Current Population Survey. Table I replicates the Firpo *et al.* (2009) results by using the ordinary least squares (OLS), conditional quantile regressions and three specific types of unconditional quantile

DOI: 10.1002/jae

Table I. Comparing OLS, conditional quantile regressions (CQR), linear unconditional quantile regressions (UQR), logit unconditional quantile regressions (UQR) and nonparametric unconditional quantile regressions (NP UQR); 1983–1985 CPS data for men (Firpo *et al*., 2009, Table I replication)

		10th centile				50th centile				90th centile			
	OLS	CQR	Linear UQR	Logit UQR	NP UQR	CQR	Linear UQR	Logit UQR	NP UQR	CQR	Linear UQR	Logit UQR	NP UQR
Union status	0.179	0.288	0.199	0.306	0.230	0.195	0.338	0.316	0.323	0.088	−0.145	−0.121	−0.113
	(0.002)	(0.004)	(0.003)	(0.006)	(0.003)	(0.002)	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.005)	(0.005)
Non-white	−0.134	−0.139	−0.120	−0.109	−0.122	−0.134	−0.164	−0.157	−0.158	−0.120	−0.098	−0.116	−0.103
	(0.003)	(0.004)	(0.005)	(0.004)	(0.005)	(0.003)	(0.004)	(0.004)	(0.005)	(0.004)	(0.006)	(0.005)	(0.006)
Married	0.140	0.166	0.199	0.180	0.187	0.146	0.157	0.147	0.152	0.089	0.040	0.073	0.070
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
Education													
Elementary	−0.351	−0.279	−0.313	−0.270	−0.356	−0.374	−0.455	−0.403	−0.394	−0.357	−0.237	−0.377	−0.245
	(0.004)	(0.006)	(0.006)	(0.006)	(0.009)	(0.004)	(0.006)	(0.006)	(0.009)	(0.007)	(0.008)	(0.016)	(0.006)
HS dropout	−0.190	−0.127	−0.350	−0.197	−0.242	−0.205	−0.196	−0.219	−0.221	−0.227	−0.066	−0.193	−0.154
	(0.003)	(0.004)	(0.005)	(0.003)	(0.006)	(0.003)	(0.004)	(0.004)	(0.005)	(0.004)	(0.006)	(0.010)	(0.006)
Some college	0.133	0.058	0.059	0.046	0.045	0.133	0.177	0.170	0.167	0.172	0.152	0.188	0.213
	(0.002)	(0.003)	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.006)
College	0.406	0.252	0.202	0.266	0.197	0.414	0.462	0.437	0.416	0.548	0.588	0.467	0.661
	(0.003)	(0.005)	(0.004)	(0.005)	(0.004)	(0.004)	(0.00)	(0.005)	(0.004)	(0.006)	(0.006)	(0.006)	(0.009)
Postgraduate	0.478	0.287	0.142	0.221	0.168	0.482	0.518	0.506	0.473	0.668	0.857	0.565	0.899
	(0.003)	(0.006)	(0.005)	(0.008)	(0.004)	(0.005)	(0.005)	(0.006)	(0.005)	(0.006)	(0.007)	(0.006)	(0.012)
Constant	1.742	1.145	0.967			1.744	1.736			2.332	2.514		
	(0.004)	(0.288)	(0.006)			(0.005)	(0.005)			(0.006)	(0.008)		

Note: Robust standard errors (OLS) and bootstrapped standard errors (200 replications) for CQR, linear UQR, logit UQR and NP UQR are given in parentheses. All regressions include a set of dummies for labor market experience categories. Following Firpo *et al.* (2009) for nonlinear RIF regressions, we do not report the constant term.

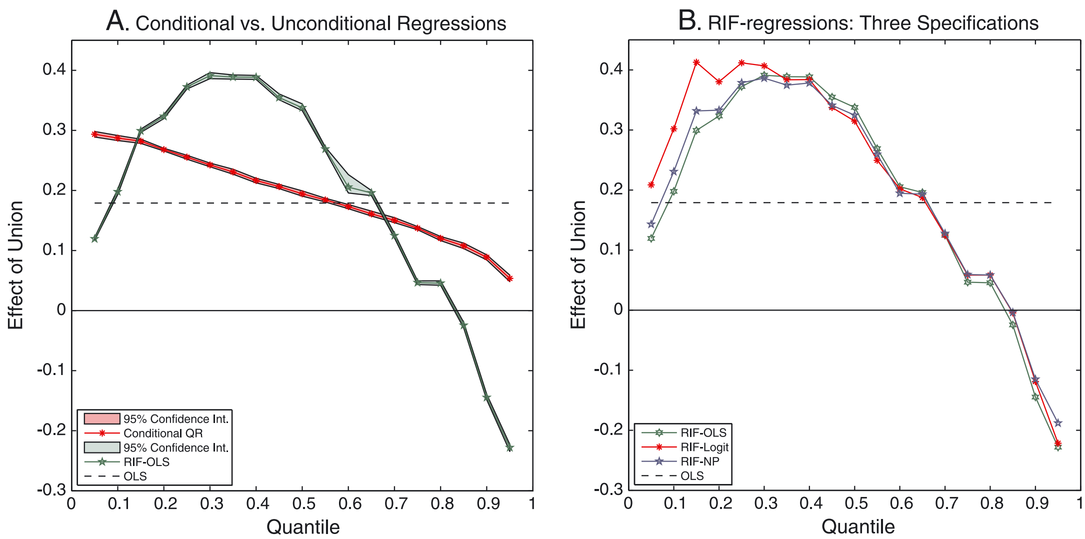

Figure 1. Unconditional and conditional quantile regression estimates of the effect of union status on log wages. This figure is available in color online at wileyonlinelibrary.com/journal/jae

regressions, namely RIF-OLS, RIF-logit and RIF-nonparametrics. All these regressions include a set of dummy variables for different experience groups and the standard errors are calculated using 200 bootstrap replications.

Table I confirms the results of Firpo *et al.* (2009). However, the unconditional quantile regressions estimates are slightly different because we use the Botev *et al.* (2010) diffusion method[2] to obtain the optimal bandwidth parameters for estimating the kernel density functions. In contrast, Firpo *et al.* (2009) arbitrarily choose the bandwidth parameter, $\tilde{h}_{\text{opt}} = 0.06$ for all quantiles. Despite this, the replication results are qualitatively the same as shown in Figure 1. The unconditional quantile partial effect increases from the 5th quantile to the 35th quantile and reaches a maximum around 0.4, then monotonically decreases until -0.2 at the 95th quantile. The conditional quantile partial effect declines monotonically from about 0.3 at the 5th quantile to 0.05 at the 95th quantile.

The existing unionization literature suggests that the marginal effects of union status are positive throughout the wage distribution function. However, all three specifications of the unconditional quantile regressions show that the UQPE of declining unionization are negative in the upper tail of the wage distribution. Firpo *et al.* (2009) also show that the UQPE can be represented as a weighted average of the CQPE. Figure 1 shows that the UQPE are significantly different from CQPE, especially in the upper tail of the wage distribution. This creates a puzzle as to why the estimated UQPE of union status are negative in the upper tail of the wage distribution.

The puzzle can be solved by using the definition of the CQPE and UQPE. The conditional quantile partial effect measures the impact of union status only for union workers. This is known as the within effect of unionization. Unlike CQPE, the unconditional quantile regression estimates the marginal effect of unionization for all workers at the τth quantile of the wage distribution. In fact, the unconditional quantile partial effect estimates both the within and between effects of unionization. The between effect is the impact of change in the unionization rate on the non-union workers. Following similar arguments to those in Firpo *et al.* (2009), the between effect of unionization can be negative especially for skilled workers, and in this empirical application the between effect dominates the positive within effect, causing the UQPE to be negative in the upper tail of the distribution.

[2] The Matlab code used to implement the Botev *et al.* (2010) diffusion method was written by Zdravko Botev.

DOI: 10.1002/jae

Table II. 100 × observed and composition constant changes in overall and residual wage inequality measures using the counterfactual decomposition method (replication of Autor *et al.* (2008, Table 4) using unconditional quantile regressions)

		Residual inequality			Overall inequality		
		1973–1989	1989–2005	1973–2005	1973–1989	1989–2005	1973–2005
Δ 90/50							
Males	Observed	4.8 [4.4]	4.8 [4.0]	9.6 [8.4]	10.3 [10.2]	15.4 [14.2]	25.7 [24.4]
	1973 X's	2.5 [4.0]	1.5 [2.8]	4.0 [6.8]	10.4 [11.2]	16.5 [9.2]	27.0 [20.4]
	1989 X's	3.2 [3.6]	2.4 [2.6]	5.6 [6.2]	10.8 [8.8]	18.0 [13.5]	28.8 [22.3]
	2005 X's	3.5 [2.8]	3.2 [2.9]	6.7 [5.7]	11.5 [7.9]	14.5 [13.0]	26.1 [20.9]
Females	Observed	7.2 [8.2]	4.4 [4.8]	11.5 [12.9]	11.1 [11.3]	13.0 [9.8]	24.0 [21.1]
	1973 X's	6.7 [6.8]	−3.2 [−0.9]	3.5 [5.9]	11.5 [11.8]	10.6 [9.9]	22.1 [21.7]
	1989 X's	7.3 [7.6]	−2.1 [1.9]	5.2 [9.5]	12.1 [13.2]	12.5 [10.5]	24.6 [23.7]
	2005 X's	5.9 [8.0]	0.2 [4.0]	6.1 [12.0]	12.9 [10.7]	12.3 [10.1]	25.3 [20.8]
Δ 50/10							
Males	Observed	7.3 [5.7]	−1.4 [−1.3]	5.9 [4.4]	7.6 [8.1]	−2.4 [2.1]	5.2 [10.2]
	1973 X's	4.6 [3.9]	−5.8 [−7.1]	−1.2 [−3.2]	14.3 [8.7]	−11.5 [−12.6]	2.8 [−3.9]
	1989 X's	4.4 [6.3]	−4.5 [−4.0]	−0.1 [2.3]	11.0 [5.2]	0.8 [−6.4]	11.8 [−1.2]
	2005 X's	7.7 [4.9]	−3.6 [−3.5]	4.0 [1.4]	15.1 [15.7]	−7.9 [−8.4]	7.2 [7.3]
Females	Observed	8.1 [8.8]	0.2 [0.4]	8.3 [9.2]	13.3 [14.4]	0.4 [2.8]	13.7 [17.3]
	1973 X's	3.4 [4.4]	−2.9 [−4.5]	0.5 [−0.1]	11.4 [4.9]	1.3 [−8.4]	12.7 [−3.5]
	1989 X's	3.8 [5.7]	−3.4 [−3.9]	0.4 [1.8]	10.4 [8.5]	−3.3 [−7.9]	7.1 [0.6]
	2005 X's	2.8 [6.2]	−0.9 [−2.5]	1.9 [3.7]	10.7 [8.3]	−1.2 [−0.1]	9.5 [8.2]
Δ 90/10							
Males	Observed	12.2 [10.1]	3.4 [2.8]	15.3 [12.8]	17.9 [18.3]	13.0 [16.4]	30.9 [34.6]
	1973 X's	7.1 [7.9]	−4.3 [−4.3]	4.9 [3.6]	24.7 [19.9]	12.9 [−3.4]	33.2 [16.5]
	1989 X's	7.7 [8.0]	−2.1 [−1.5]	6.2 [6.5]	21.8 [13.7]	14.0 [7.0]	35.8 [20.7]
	2005 X's	11.2 [7.7]	−0.4 [−0.7]	7.6 [7.0]	26.6 [23.4]	13.5 [4.6]	34.6 [28.0]
Females	Observed	15.2 [16.9]	4.6 [5.2]	19.8 [22.1]	24.3 [25.7]	13.4 [12.7]	37.7 [38.4]
	1973 X's	10.1 [11.2]	−6.1 [−5.4]	3.9 [5.8]	23.0 [16.7]	11.9 [1.5]	34.8 [18.2]
	1989 X's	11.1 [13.3]	−5.6 [−2.0]	5.5 [11.3]	22.5 [21.7]	9.2 [2.6]	31.7 [24.3]
	2005 X's	8.7 [14.2]	−0.6 [1.4]	8.1 [15.6]	23.6 [18.9]	11.2 [9.9]	34.8 [28.8]

Note: The bracketed terms represent Autor *et al.* (2008) Table 4 values.

Next, we apply the unconditional quantile regression method to the empirical application by Autor *et al.* (2008) in order to explain the trends in the US wage inequality over the period 1973–2005.[3] Table II presents the overall and residual wage inequality over the period 1973–2005. In order to decompose the changes in the wage distribution function into price and composition effects, Autor *et al.* (2008) use a modified DiNardo *et al.* (1996) decomposition method. In order to construct counterfactual wages, we use the unconditional quantile regressions of Firpo *et al.* (2009).

The results in Table II show that, holding the labor force composition constant at its 1973, 1989 and 2005 levels, the price effect is the dominating factor for the changes in residual, as well as over all wage inequality in almost all cases considered, for both men and women. Thus the unconditional quantile regression decomposition method gives qualitatively similar results to the the modified DiNardo *et al.* (1996) decomposition method proposed by Autor *et al.* (2008).

[3] See the descriptive statistics in Table 1 of Autor *et al.* (2008).

REFERENCES

Autor D, Katz L, Kearney M. 2008. Trends in the U.S. wage inequality: revising the revisionists. *Review of Economics and Statistics* **90**: 300–323.

Botev Z, Grotowski J, Kroese D. 2010. Kernel density estimation via diffusion. *Annals of Statistics* **38**(5): 2916–2957.

DiNardo J, Fortin N, Lemieux T. 1996. Labor market institution and the distribution of wages, 1973–1992: a semiparametric approach. *Econometrica* **64**: 1001–1044.

Firpo S, Fortin N, Lemieux T. 2009. Unconditional quantile regressions. *Econometrica* **77**: 953–973.

Koenker R, Bassett G. 1978. Regression quantiles. *Econometrica* **46**: 33–50.

Oaxaca R. 1973. Male-female wage differentials in urban labor markets. *International Economic Review* **14**: 693–709.

DOI: 10.1002/jae