Revista de Análisis Económico, Vol. 29, Nº 1, pp. 53-76 (Abril 2014)

THE DISTRIBUTIVE EFFECTS OF EDUCATION: AN UNCONDITIONAL QUANTILE REGRESSION APPROACH* **

EFECTOS DISTRIBUTIVOS DE LA EDUCACION: UN ENFOQUE DE REGRESIONES POR CUANTILES NO CONDICIONADOS

JAVIER ALEJO[a]
Universidad Nacional de La Plata and CEDLAS

MARIA FLORENCIA GABRIELLI[b]
Universidad Nacional de Cuyo and CONICET

WALTER SOSA-ESCUDERO[c]
Universidad de San Andrés and CONICET

Abstract

We use recent unconditional quantile regression methods (UQR) to study the distributive effects of education in Argentina. Standard methods usually focus on mean effects, or explore distributive effects by either making stringent modeling assumptions, and/or through counter- factual decompositions that require several temporal observations. An empirical case shows the flexibility and usefulness of UQR methods. Our application for the case of Argentina shows that education contributed positively to increased inequality in Argentina, mostly due to the effect of strongly heterogeneous effects of education on earnings.

* Corresponding author: Maria Florencia Gabrielli.
** We thank Leonardo Gasparini and the Centro de Estudios Distributivos Laborales y Sociales (CEDLAS) at National University of La Plata, for access to the data used in this paper. All errors and omissions are our responsibility.
[a] CEDLAS - Facultad de Ciencias Económicas - UNLP, 777 Calle 6, La Plata, Bs. As., Tel: tel/fax +54 -221- 423-6769. E-mail: javier.alejo@depeco.econo.unlp.edu.ar
[b] Facultad de Ciencias Económicas - UNCUYO, Centro Universitario, M5502JMA, Mendoza Argentina, tel/fax +54-261-413-5000 ext. 2454. E-mail: florgabrilli@gmail.com
[c] Departamento de Economía, UDESA, 284 Vito Dumas, Victoria, Argentina, tel/fax +54-11-4725-7000. E-mail: wsosa@udesa.edu.ar

Keywords: *Unconditional quantile regression, income inequality, education, Argentina.*

JEL Classification: *C21, I24, I31, D3.*

Resumen

A través de la utilización del método de regresiones por cuantiles no condicionados (UQR) estudiamos los efectos distributivos de la educación en Argentina. Los métodos estándar usualmente se focalizan en los efectos sobre la media, o exploran efectos distributivos, ya sea imponiendo supuestos restrictivos para la modelización y/o a través de descomposiciones contra factuales que requieren una cantidad importante de observaciones temporales. Desde un punto de vista empírico se ha demostrado la flexibilidad y utilidad de los métodos UQR. Nuestra aplicación empírica para el caso de Argentina muestra que la educación ha contribuido positivamente a aumentar la desigualdad en Argentina. En mayor medida, esto se debe al gran efecto heterogéneo que tiene la educación sobre los ingresos.

Palabras clave: *Regresión por cuantiles no condicionados, desigualdad del ingreso, educación, Argentina.*

Clasificación JEL: *C21, I24, I31, D3.*

1. INTRODUCTION

Fostering education is widely perceived as a powerful policy measure to improve the welfare of any society, through its direct impact on enhancing personal productivity and earnings, and as a way to equalize opportunities and, hopefully, improving the distribution of income.

The massive literature on returns to education focuses mostly on the first issue, that is, on the impact of education on expected earnings. In spite of the enormous methodological difficulties in providing clean, consistent estimates of this causal effect, there is wide agreement that education has an economically significant impact on expected personal earnings. See Card (2001) for a review of this literature.

The second issue –the distributive effects of increased education– has received less attention. The effects of education on distributional features, like earnings inequality, depend not only on the returns to education, but also on the initial distribution of education and other population characteristics, and on how changes in education are translated into changes in the distribution of incomes, and then on inequality.

The literature on quantile regressions that dates back to Buchinsky (1994)'s seminal study, reveals an important empirical result: increased education has the double effect of augmenting expected earnings while increasing the dispersion of earnings for each level of education. That is, more educated groups face a (conditional) distribution of earnings that is shifted to the right but also more disperse. Consequently, this literature suggests a dual effect of education that may imply a policy trade-off: more education increases both expected earnings and the within-group dispersion of earnings for each level of education. Martins and Pereira (2004) carefully documented this effect for a sample of 16 European countries, suggesting that this undesirable effect, through increased within-group earnings dispersion, may harm the role of education as a tool to improve income inequality.

Though quantile regression is a powerful tool to explore the effects of education on the conditional distribution of earnings, it is important to remark that the interest lies in the way education alters the unconditional distribution. The fact that education leads to a more disperse conditional distribution of earnings does not necessarily mean that the unconditional distribution would be more disperse. The latter can be seen as the product of the conditional distribution of earnings (on education) and the marginal distribution of education. Hence, the effect of increased education ultimately depends on both, the interaction between the conditional distribution of earnings, and the marginal distribution of education.

From this perspective, the quantile regression result that suggests increased within dispersion in the conditional distribution, albeit a very important one should be seen as an intermediate step, towards the final goal of assessing the impact of education on the unconditional distribution of earnings.

The step from conditional to unconditional distributive effects is not a trivial one, and only recently there are available specific statistical tools to study them. The still infant but rapidly growing literature on unconditional quantile regressions (Firpo, Fortin, and Lemieux (2009)) based on the concept of the recentered influence function, seems to provide a natural and important step towards this goal.

This paper uses unconditional quantile and recentered influence function regressions to explore the effects of increased education on unconditional income inequality. The final goal is to explore whether education indeed has an undesirable effect on inequality, as advanced by studies like that of Martins and Pereira (2004), focused on within group effects. Hence, this study complements quantile regression based results by extending the effect of education beyond that on the conditional distribution.

The case of Argentina provides relevant variability for this study, not found in other countries. During the nineties, sustained improvements in educational achievements occurred simultaneously with a dramatic deterioration in inequality. After the drastic crisis of 2002, both, education and inequality, improved. Consistently, Argentina is a case where both education and inequality moved markedly, providing important sampling variability to study the distributive effects of education.

The paper is organized as follows. Section 2 briefly describes the literature on conditional and unconditional quantile regression. In Section 3 we provide our empirical application to the case of Argentina. Finally Section 4 collects the main conclusions and lines for future research.

2. STANDARD, CONDITIONAL, AND UNCONDITIONAL QUANTILE REGRESSIONS

Standard regression models are useful tools when the interest lies in measuring the effect of a covariate on the expected value of the variable of interest. Only under very stringent assumptions such model can be used to extrapolate the effects to other aspects of the distribution of the variable of interest, such as its quantiles, its variance, or its level of inequality as measured by a standard index, like the Gini coefficient.

In our context, the goal is to measure the effect of changes in educational levels on the distribution of income. As a first step, and for analytic convenience and in accordance with the natural notion of a derivative, by movements in educational levels we mean small changes in the location of the distribution of education.

Our target will be some functional of the distribution of income other than the mean, like any quantile, the variance, or its Gini coefficient. In this sense, standard regression analysis focuses on the impact of education on one particular functional (the mean).

In a recent article, Firpo, Fortin, and Lemieux (2009) propose unconditional quantile regressions as a simple way to estimate the effects of increasing education on the quantiles of the unconditional distribution of a variable. In what follows we present the main ideas, and refer to these authors for further details.

Let Y be a random variable with cumulative distribution function (CDF) $F_Y(y)$, and let $\nu(F_Y)$ be any functional. For simplicity, we will focus on linear functional that can be expressed as

$$\nu\left(F_Y\right) = \int \psi\left(y\right) dF_Y\left(y\right) \tag{1}$$

for some function $\psi(y)$. For example, the mean, μ_Y, corresponds to $\psi(y) = y$. (Appendix A contains some details of how to construct the Influence Function).

Firpo, Fortin, and Lemieux (2009), define the 'recentered influence function' (RIF) as

$$RIF(y,\nu) \equiv IF(y,\nu) + \nu(F_Y) = \psi\left(y\right), \tag{2}$$

and, trivially,

$$E[RIF\left(y,\nu\right)] = \nu\left(F_Y\right) \tag{3}$$

This is an important step, since, it implies that any functional of interest can be expressed as an expected value.

In order to incorporate the effect of covariates, let X be a vector of random variables. Note that, using the law of iterated expectations,

$$\nu\left(F_Y\right) = \int RIF(y,\nu)\ dF_Y\ \left(y\right) = \int E[RIF(Y,\nu) \mid X = x] dF_X\ \left(x\right) \tag{4}$$

where $F_X(x)$ is the marginal CDF of X. Suppose the distribution of X changes as a small location shift, and let $\alpha(\nu)$ be the vector of partial effects of moving each coordinate of X separately as a location shift. Assume also that the conditional distribution of Y given X stays constant. Then, Firpo, Fortin, and Lemieux (2009) show that the 'unconditional partial effect on' $\nu(F)$ of altering the CDF of X in such way is given by:

$$\alpha(\nu) = \int \frac{dE\left[RIF(y,\nu)|X = x\right]}{dx} dF_x(x) \tag{5}$$

In words, this means that the partial effects of altering shifting the CDF of X to the right (marginally) can be recovered by simple regression methods, that is, by regressing the RIF of Y with respect of the functional of interest, on the vector X (the 'RIF regression'), compute the marginal effects, and then integrate over the values of X, as in standard regression analysis.

A relevant application for our case corresponds to the effects of X on the unconditional quantiles of Y. Let now $\nu(F_Y) = q_\tau$ denote the τ–th quantile of $F_Y(\cdot)$. Its recentered influence function can be shown to be given by (Firpo, Fortin, and Lemieux (2009)).

$$\begin{aligned} RIF(y,q_\tau) = q_\tau + IF(y,q_\tau) \quad &= q_\tau + \frac{\tau - I(y \le q_\tau)}{f_Y(q_\tau)} \\ &= \frac{I(y \le q_\tau)}{f_Y(q_\tau)} + q_\tau - \frac{1-\tau}{f_Y(q_\tau)} \\ &= c_{1,\tau} I(y > q_\tau) + c_{2,\tau} \end{aligned} \tag{6}$$

where $c_{1,\tau} \equiv 1 / f_Y(q_\tau)$ and $c_{2,\tau} \equiv q_\tau - c_{1,\tau} \cdot (1 - \tau)$. Therefore

$$\begin{aligned} E[RIF(y,q_\tau) \mid X = x] \quad &= c_{1,\tau} E[I(Y > q_\tau) \mid X = x] + c_{2,\tau} \\ &= c_{1,\tau} \Pr[I(Y > q_\tau) \mid X = x] + c_{2,\tau} \end{aligned} \tag{7}$$

This last expression is the unconditional quantile regression, that is, a regression model that links the expected value of the quantiles (as measured by the RIF) to covariates. Particular specifications on $Pr[Y>q_\tau \mid X = x]$ lead to alternative regressions. We use linear probability model for estimation (see Appendix B for further details).

In practice, in a first step the RIF is estimated by replacing all unknown quantities by their observable counterparts. In this case, unknown quantities are qτ and $f_Y(q_\tau)$, which are estimated by the sample τ–*th* quantile of Y, and a standard nonparametric density estimator (e.g. kernel), respectively. The second stage regresses the estimated RIF on x using a standard OLS estimator.

Some remarks on this strategy are the following. First, the linear probability assumption may sound restrictive. Replacing it by a standard *probit* or *logit* specification can be easily implemented. Nevertheless, the empirical results of Firpo, Fortin, and Lemieux (2009) indicate that results are almost indistinguishable of those using the linear probability model, much in accordance to the recent literature that favors it in light of its conceptual and computational advantages, as clearly advocated by Angrist and Pischke (2008)[1]. Second, (asymptotic) inference in the second stage must accommodate the fact that q_τ and $f_Y\ (q_\tau)$ are estimated in a first stage. This is discussed in detail in Firpo, Fortin, and Lemieux (2009). Finally, RIF regressions for other functional of interest can be derived. For example, if the functional of interest is the mean, then the RIF of Y for the mean is simply y, then, as expected, the RIF regression is the standard regression. In our case, we will be interested in the RIF regression for the Gini coefficient, derived in Firpo, Fortin, and Lemieux (2009). See Appendix C for further details.

Finally, it is relevant to compare unconditional quantile regression with standard quantile regressions, as defined originally by Koenker and Basset (1978). The linear quantile regression model specifies

$$Q_{Y|X}(x,\tau) = x'\beta(\tau) \tag{8}$$

where $Q_{Y|X}(\tau)$ denotes the τ–*th* quantile of the conditional distribution of Y given $X = x$. Consequently

$$\beta(\tau) = \frac{\partial Q_{Y|X}(x,\tau)}{\partial x} \tag{9}$$

that is, the elements of $\beta(\tau)$ measure the effect of altering the components of x marginally, on the τ–*th* quantile of the conditional distribution of Y on X. In this model, $\beta(\tau)$ is understood as a non-specified function of τ, hence its semiparametric nature.

In this context, the standard result (mentioned in the Introduction) that for the case of education, $\beta(\tau)$ is a positive and monotonically increasing function means that increasing education impacts more in higher quantiles of the conditional distribution of income, that is, by increasing education, all conditional quantiles move up, but at an increasing rate along quantiles. This effect is clearly and naturally captured by quantile regressions, as documented by Martins and Pereira (2004).

Nevertheless, the effect on the unconditional distribution (the subject of interest of distributive analysis) requires ‘averaging’ these effects according to the levels of education observed in the sample. In intuitive terms, if the distribution of Y can be thought as factored by its conditional distribution given X, and the marginal distribution of X, then inequality in Y represents the interaction of the inequality in X and the way Y is affected by X. Conditional quantile regressions can be seen as modeling the second channel, whereas unconditional quantile regressions integrate both. For

[1] The use of OLS as a regression method is just a lineal approximation to the true model and thus it is not clear how to measure the goodness of fit.

example, and as seen in the empirical part of this paper, the observed unequalizing effect in the conditional quantile regression might be enhanced if it takes place over an already unequal distribution of education, or dampened if increases in education result in a more equal distribution of education.

3. EXPLORING THE DISTRIBUTIVE EFFECTS OF EDUCATION: ARGENTINA 1992-2008

The analysis is based on micro data from Argentina's Permanent Household Survey (EPH) for years 1992, 1998, and 2008, for all regions available in the period under analysis. The cities included are: Greater La Plata, Greater Santa Fe, Greater Paraná, Comodoro Rivadavia - Rada Tilly, Greater Córdoba, Neuquén - Plottier Santiago del Estero - La Banda, Jujuy - Palpalá, Rio Gallegos, Salta, San Luis - El Chorrillo, Greater San Juan, Santa Rosa - Toay, Ushuaia - Rio Grande, Buenos Aires City and Greater Buenos Aires. The sample considered is composed of men between 15 and 65 years old. Income is defined as the salary obtained in all occupations measured in pesos as of December 2008.

Inequality, poverty and other aspects of the distribution of income changed dramatically in the last twenty years. Even though the nineties started with a period of sustained GDP growth, the same decade witnessed a monotonic increase in inequality and poverty. The drastic crisis experienced by Argentina in 2002 led to historic records in these measures. After that, a period of recovery followed, and inequality and poverty decreased at a monotonic rate, reaching, in 2008, levels similar to those observed at the beginning of the nineties. The three periods chosen for the analysis (1992, 1998 and 2008) are representative of this behavior. For example, the Gini coefficient of hourly wages (see Table 3.1) started in 40.5, increased to 44 in 1998, and after 2001 a period of sustained decline started and reached 39.8 in 2008. See Gasparini and Cruces (2009) and Sosa Escudero and Petralia (2012) for a complete description of these evolutions.

Changes in education were also dramatic in the period under analysis. Schooling, as measured by years of education increased from 9.9 in 1992 to 10.8 in 2008, as can be seen in Table 3.1. A clearer picture is obtained when looking at educational levels. For example, the proportion of individuals whose maximum level of education is complete primary dropped from 30.3% in 1992 to 19.4% in 2008. Similarly, the same proportion for complete high school rose from 16.4% to 22.4% during the same period. Educational levels increased monotonically, with most of the action taking place in the center of the distribution (around complete high school). If we split the sample by age groups (16 to 24, 25 to 40 and 41 to 64 years old) and educational level, a simple way of analyzing the evolution of wage inequality is through a decomposition of the Theil index by group. Table 3.2 shows that between 1992 and 2008 wage inequality between groups declined but within the interior of each group the inequality increased, thus giving a small increase in overall inequality. Moreover, looking at the changes by period, the main component in the change of inequality both during the increase in inequality in the nineties and the decline in the first decade after year 2000 was

TABLE 3.1

SUMMARY STATISTICS OF SURVEY DATA. ARGENTINA 1992-2008
SAMPLE: MEN BETWEEN 16 AND 64 YEARS OLD

Year	Variable					
	Hourly wage					
	Mean	Gini	Quantile 0.10	Median	Quantile 0.90	Range 90-10
1992	11.5	40.5	4.1	8.2	21.6	17.5
1998	12.6	44.0	3.8	8.6	25.7	21.9
2008	11.4	39.8	3.6	8.7	21.1	17.5

Year	Age					
	Mean	Std. Dev.	Quantile 0.10	Median	Quantile 0.90	Range 90-10
1992	36.6	13.1	20	36	56	36
1998	37.4	12.3	22	36	55	33
2008	36.5	13.6	19	35	56	37

Year	Year of education					
	Mean	Gini	Quantile 0.10	Median	Quantile 0.90	Range 90-10
1992	9.9	21.1	7	10	15	8
1998	10.0	20.9	7	10	16	9
2008	10.8	19.2	7	12	16	9

Year	Educational Level					
	Primary incomplete	Primary complete	Dropouts	Highschool	College incomplete	College complete
1992	9.3%	30.3%	24.5%	16.4%	11.4%	8.1%
1998	7.6%	27.4%	24.7%	18.0%	11.4%	10.7%
2008	6.9%	19.4%	23.8%	22.4%	14.8%	12.7%

Year	Region					
	GBA	Pampa	Cuyo	NOA	Patagonia	Total
1992	65.9%	20.7%	3.5%	6.2%	3.7%	100%
1998	73.8%	14.4%	3.2%	5.2%	3.4%	100%
2008	70.7%	16.6%	3.3%	6.0%	3.4%	100%

Source: Own calculations based on EPH (INDEC).

basically led by the change in inequality 'within' (roughly 80% in the 90's and 65% in the 00's). These changes of the overall distribution can be more drastically appreciated in Figure 3.1, which shows the estimated densities of education for the three periods.

TABLE 3.2

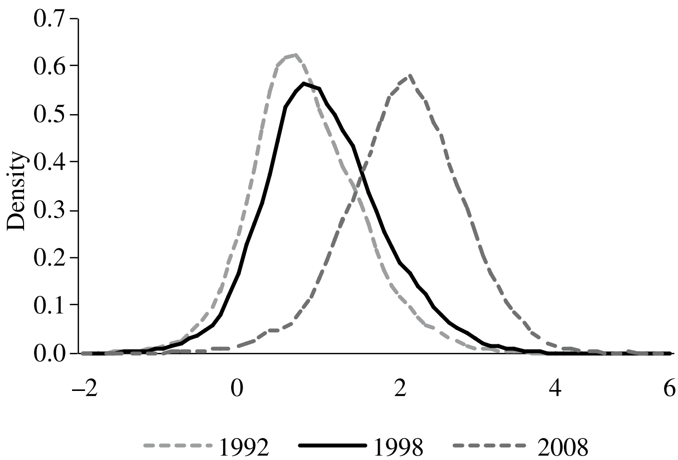
THEIL GROUP DECOMPOSITION
SAMPLE: MEN BETWEEN 16 AND 64 YEARS OLD

	Overall	Within	Between
Level			
1992	31.76	29.98	1.78
1998	36.83	34.11	2.72
2008	32.52	31.32	1.20
Annual change			
92-98	0.85	0.69	0.16
98-08	–0.43	–0.28	–0.15
92-08	0.05	0.08	0.04

Source: Own calculations based on EPH (INDEC).

FIGURE 3.1

WAGE MARGINAL DISTRIBUTION. ARGENTINA 1992 -2008
SAMPLE: MEN BETWEEN 16 AND 64 YEARS OLD



Source: Own calculations based on EPH (INDEC).

In light of these results, it is natural to explore the interaction between changes in the distribution of education along with those in the distribution of income. As a motivation, in Table 3.3 we show (log) wage dispersion by group of educational level. As can be seen, inequality is increasing in educational level, although we control for the effect in the means of the remaining covariates. Therefore, it is clear that studying the wage distribution along with education only through the conditional mean (OLS) gives an incomplete picture of the distributive effects of education. Gasparini, Marchionni, and Sosa Escudero (2004) explore this link using a microeconometric decomposition frame-work, and conclude that education had equalizing effect in the period 1989-1992, and an unequalizing effect for 1992-1998. Bustelo (2004) adopts the approach of Mata and Machado (2005), that estimates a conditional quantile regression model from which, through simulations, a counterfactual unconditional distribution is obtained, and finds that that an increase in education is associated with a decrease in poverty, and a small unequalizing effect in the period 1992-2001, with a stronger effect for higher levels of education. Alejo (2006) explores the statistical significance of all these results. With Chilean data, some exercises along these lines can be found in Behrman (2011) and Ruiz Tagle (2007)[2].

TABLE 3.3

VARIANCE OF LOG-WAGES, BY EDUCATION LEVEL
SAMPLE: MEN BETWEEN 16 AND 64 YEARS OLD

	Primary incomplete	Primary complete	Secondary incomplete	Secondary complete	College incomplete	College complete
Without controls						
1992	0.328	0.299	0.362	0.420	0.461	0.554
1998	0.417	0.375	0.401	0.463	0.424	0.564
2008	0.454	0.421	0.476	0.412	0.408	0.494
Controling for remaining covariates (1)						
1992	0.322	0.263	0.313	0.355	0.403	0.505
1998	0.400	0.343	0.341	0.400	0.368	0.486
2008	0.406	0.383	0.415	0.375	0.371	0.467

Source: Own calculations based on EPH (INDEC).
Note:
(1) Log-wages net from the effect of other regressors (OLS). Covariates are: age, marital status and regional *dummies*.

As discussed in the previous section, we estimate RIF regressions for several unconditional quantiles, using a linear probability specification. We also estimate a

2 We thank a referee for pointing out this connection.

RIF regression for the Gini coefficient. We use the usual covariates in standard Mincer equations, namely: age, years of education, marital status and dummy variables to control for regional effects.

As a previous step, Table 3.4 presents a conditional quantile regression analysis for quantiles ranging from 0.1 to 0.9. The last column of this Table presents results of a standard OLS regression. Tables 3.5 present results based on unconditional quantile regressions, and the last columns show results for the RIF regression of the Gini coefficient. For convenience, estimated coefficients for these two tables are represented in the first row of graphics in Figure 3.2.

Consider the first graph of Figure 3.2, which represents the estimated coefficients of years education, for the conditional and unconditional quantile regressions in Tables 3.4 and 3.5, for 1992. The horizontal line represents the 'mean' effect associated to the standard OLS estimator; 0.084, in this case. Were education set exogenously, this implies that an extra year of education led to an increase of around 8.4% in expected wages. The solid line with triangles represent conditional quantile regression estimates, and the solid line (with no ticks), represents estimates for unconditional quantiles.

A first interesting fact is that, consistently with most previous results, effects are heterogeneous and increasing along the quantiles. CQR results suggest that effects range from 0.063 for the first decile to 0.095 to the 9th decile of the conditional distribution of wages. As stressed in the Introduction, this result must be interpreted carefully. It only suggests that after controlling for the covariates, all quantiles of the conditional distribution increase when education is enhanced, but at an increasing rate for the higher the quantiles. A common difficulty associated with the interpretation of these results is that the top (bottom) of the conditional distribution does not coincide with the top (bottom) of its unconditional counterpart. That is, the positive and heterogeneous CQR effects do not imply that education has a stronger effect for the, say, rich, but for the conditionally rich, that is, after controlling for all covariates. Consequently, in the CQR it is difficult to see if this unequalizing effect translates into the unconditional distribution of incomes, hence the usefulness of the UQR approach that studies this effect directly on the distribution of income.

Interestingly, UQR results show an even more pronounced heterogeneous behavior, the effect ranging from 0.046 to 0.140. UQR results are more directly interpretable since now they truly suggest that the effect of education is stronger for the rich. Differences between the CQR and the UQR approach might be due to the fact that the originally unequalizing effect of the CQR is further enhanced by applying it to the already unequal (and markedly asymmetric) distribution of education of 1992. As stressed in the previous sections, and unlike CQR, UQR integrates the heterogeneous effects on the conditional distribution with the existing levels of education, leading to an enhanced heterogeneous effect.

Finally, RIF regressions results for the Gini coefficient (last column of Table 3.5) are interesting. First, in order to obtain comparable results, the regression is estimated using levels of wages, not logs as in standard Mincer equations. Hence, results suggest

TABLE 3.4

MARGINAL EFFECTS ON CONDITIONAL WAGE DISTRIBUTION QUANTILE REGRESSION - MEN BETWEEN 16 AND 64 YEARS OLD

Argentina 1992

	q(0.10)	q(0.20)	q(0.30)	q(0.40)	q(0.50)	q(0.60)	q(0.70)	q(0.80)	q(0.90)	Mean
Age	0.005	0.007	0.008	0.009	0.010	0.011	0.012	0.015	0.015	0.010
	(3.29)**	(6.67)**	(7.07)**	(11.59)**	(10.65)**	(11.79)**	(11.19)**	(12.10)**	(9.00)**	(20.33)**
Years of education	0.063	0.065	0.071	0.076	0.079	0.086	0.089	0.093	0.095	0.084
	(16.46)**	(24.55)**	(22.87)**	(33.80)**	(28.00)**	(28.84)**	(25.06)**	(22.21)**	(16.09)**	(59.12)**
Married	0.154	0.173	0.175	0.157	0.181	0.168	0.196	0.159	0.153	0.180
	(3.88)**	(6.62)**	(5.84)**	(7.38)**	(6.96)**	(6.24)**	(6.23)**	(4.47)**	(3.19)**	(13.52)**
Region 2 (Pampa)	–0.208	–0.214	–0.208	–0.227	–0.239	–0.253	–0.265	–0.274	–0.292	–0.241
	(7.62)**	(11.52)**	(9.60)**	(14.86)**	(12.75)**	(13.07)**	(11.71)**	(10.82)**	(8.51)**	(16.53)**
Region 3 (Cuyo)	–0.472	–0.427	–0.424	–0.446	–0.470	–0.464	–0.468	–0.472	–0.529	–0.461
	(13.62)**	(18.43)**	(15.88)**	(23.85)**	(20.59)**	(19.83)**	(17.29)**	(15.64)**	(13.11)**	(14.27)**
Region 4 (NOA)	–0.476	–0.409	–0.397	–0.402	–0.423	–0.419	–0.421	–0.429	–0.462	–0.430
	(16.95)**	(21.60)**	(18.03)**	(25.91)**	(22.21)**	(21.33)**	(18.32)**	(16.63)**	(13.27)**	(17.02)**
Region 5 (Patagonia)	–0.031	0.055	0.093	0.112	0.126	0.138	0.146	0.133	0.109	0.088
	(1.08)	(2.84)**	(4.13)**	(7.12)**	(6.54)**	(7.00)**	(6.38)**	(5.26)**	(3.21)**	(2.86)**
Constant	–0.525	–0.411	–0.372	–0.314	–0.258	–0.220	–0.160	–0.065	0.200	–0.256
	(7.86)**	(9.22)**	(7.05)**	(8.30)**	(5.51)**	(4.52)**	(2.78)**	–0.97	(2.11)*	(10.99)**
Sample size	12,196	12,196	12,196	12,196	12,196	12,196	12,196	12,196	12,196	12,196

Argentina 1998

	q(0.10)	q(0.20)	q(0.30)	q(0.40)	q(0.50)	q(0.60)	q(0.70)	q(0.80)	q(0.90)	Mean
Age	0.004	0.007	0.011	0.014	0.015	0.016	0.018	0.020	0.021	0.013
	(2.71)**	(5.75)**	(9.76)**	(14.20)**	(14.69)**	(16.40)**	(19.68)**	(16.16)**	(12.90)**	(24.84)**
Years of education	0.091	0.090	0.096	0.100	0.100	0.102	0.107	0.111	0.113	0.104
	(23.77)**	(29.63)**	(32.78)**	(37.58)**	(34.87)**	(34.93)**	(35.89)**	(26.70)**	(19.05)**	(66.76)**
Married	0.204	0.162	0.124	0.106	0.098	0.122	0.128	0.121	0.066	0.116
	(5.37)**	(5.44)**	(4.41)**	(4.23)**	(3.78)**	(4.78)**	(5.16)**	(3.60)**	(1.41)	(8.21)**
Region 2 (Pampa)	–0.226	–0.208	–0.221	–0.217	–0.228	–0.239	–0.235	–0.243	–0.252	–0.230
	(7.11)**	(8.22)**	(9.40)**	(10.48)**	(10.65)**	(11.41)**	(11.63)**	(9.07)**	(6.69)**	(13.61)**
Region 3 (Cuyo)	–0.370	–0.325	–0.339	–0.337	–0.359	–0.378	–0.395	–0.424	–0.382	–0.361
	(10.66)**	(11.95)**	(13.24)**	(14.96)**	(15.39)**	(16.64)**	(18.10)**	(14.67)**	(9.32)**	(10.73)**
Region 4 (NOA)	–0.529	–0.471	–0.458	–0.456	–0.462	–0.482	–0.474	–0.475	–0.472	–0.475
	(15.64)**	(17.80)**	(18.52)**	(20.93)**	(20.53)**	(21.92)**	(22.38)**	(16.97)**	(12.11)**	(17.88)**
Region 5 (Patagonia)	0.013	0.026	0.025	0.052	0.062	0.091	0.091	0.111	0.148	0.064
	(0.39)	(1.01)	(1.05)	(2.47)*	(2.89)**	(4.38)**	(4.62)**	(4.31)**	(4.16)**	(1.99)*
Constant	–0.693	–0.495	–0.485	–0.468	–0.347	–0.286	–0.277	–0.179	0.036	–0.359
	(9.42)**	(8.87)**	(9.40)**	(10.29)**	(7.32)**	(6.11)**	(5.96)**	(2.84)**	(0.40)	(14.09)**
Sample size	11,228	11,228	11,228	11,228	11,228	11,228	11,228	11,228	11,228	11,228

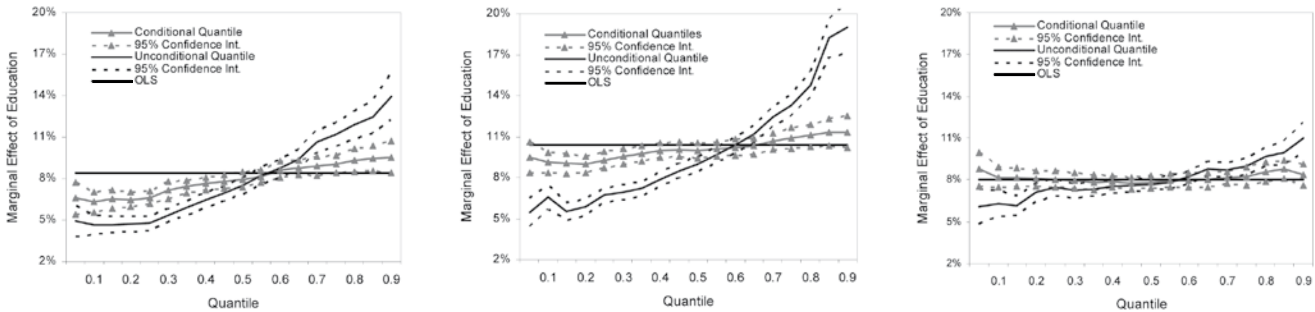
Argentina 2008

	q(0.10)	q(0.20)	q(0.30)	q(0.40)	q(0.50)	q(0.60)	q(0.70)	q(0.80)	q(0.90)	Mean
Age	0.005	0.006	0.008	0.008	0.010	0.011	0.013	0.014	0.014	0.009
	(3.84)**	(6.49)**	(8.39)**	(12.78)**	(13.93)**	(14.27)**	(21.44)**	(13.80)**	(14.48)**	(20.31)**
Years of education	0.082	0.081	0.080	0.079	0.078	0.080	0.082	0.086	0.084	0.080
	(22.06)**	(28.21)**	(29.39)**	(38.77)**	(34.52)**	(30.96)**	(39.96)**	(23.50)**	(21.50)**	(57.67)**
Married	0.183	0.123	0.093	0.096	0.080	0.060	0.056	0.055	0.027	0.102
	(5.19)**	(4.76)**	(3.89)**	(5.50)**	(4.19)**	(2.88)**	(3.47)**	(1.96)*	(0.92)	(8.40)**
Region 2 (Pampa)	–0.163	–0.102	–0.084	–0.085	–0.101	–0.103	–0.089	–0.125	–0.118	–0.109
	(5.23)**	(4.44)**	(3.94)**	(5.42)**	(5.95)**	(5.53)**	(6.19)**	(5.16)**	(4.64)**	(7.51)**
Region 3 (Cuyo)	–0.386	–0.331	–0.294	–0.312	–0.308	–0.324	–0.311	–0.306	–0.338	–0.336
	(9.99)**	(11.28)**	(10.77)**	(15.35)**	(13.92)**	(13.35)**	(16.81)**	(9.65)**	(10.65)**	(11.18)**
Region 4 (NOA)	–0.723	–0.619	–0.554	–0.558	–0.536	–0.500	–0.475	–0.465	–0.452	–0.543
	(20.96)**	(24.31)**	(23.35)**	(32.35)**	(28.66)**	(24.42)**	(30.44)**	(17.43)**	(16.36)**	(23.46)**
Region 5 (Patagonia)	0.252	0.318	0.348	0.334	0.361	0.365	0.383	0.407	0.434	0.351
	(7.41)**	(12.60)**	(14.98)**	(19.92)**	(20.03)**	(18.70)**	(26.09)**	(16.46)**	(16.95)**	(12.07)**
Constant	0.270	0.537	0.693	0.835	0.930	1.028	1.085	1.182	1.431	0.896
	(3.71)**	(10.07)**	(14.17)**	(23.31)**	(23.76)**	(23.82)**	(32.41)**	(20.51)**	(23.59)**	(37.09)**
Sample size	14,580	14,580	14,580	14,580	14,580	14,580	14,580	14,580	14,580	14,580

Source: Own calculations based on EPH (INDEC).
Note: Absolute value of t statistics in parentheses, * significant at 5%; ** significant at 1%.

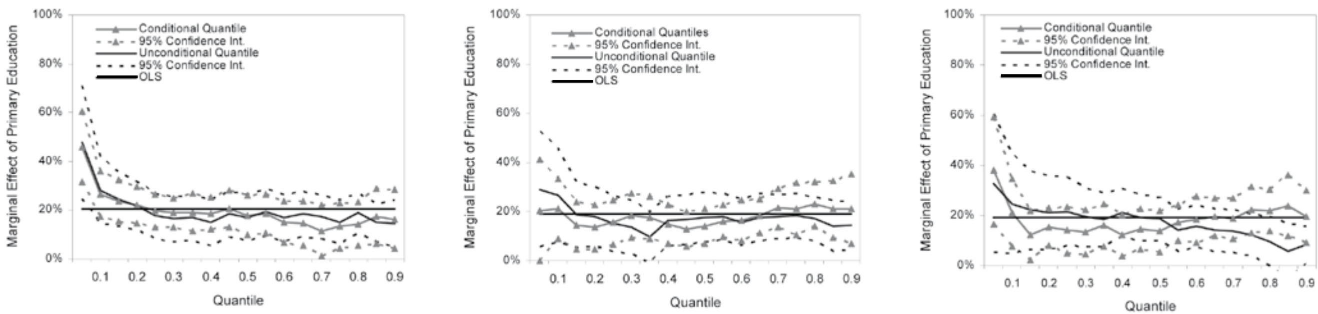
FIGURE 3.2

MARGINAL EFFECTS OF EDUCATION ON UNCONDITIONAL WAGE DISTRIBUTION - MEN BETWEEN 16 AND 65 YEARS OLD



EPH 1992 EPH 1998 EPH 2008

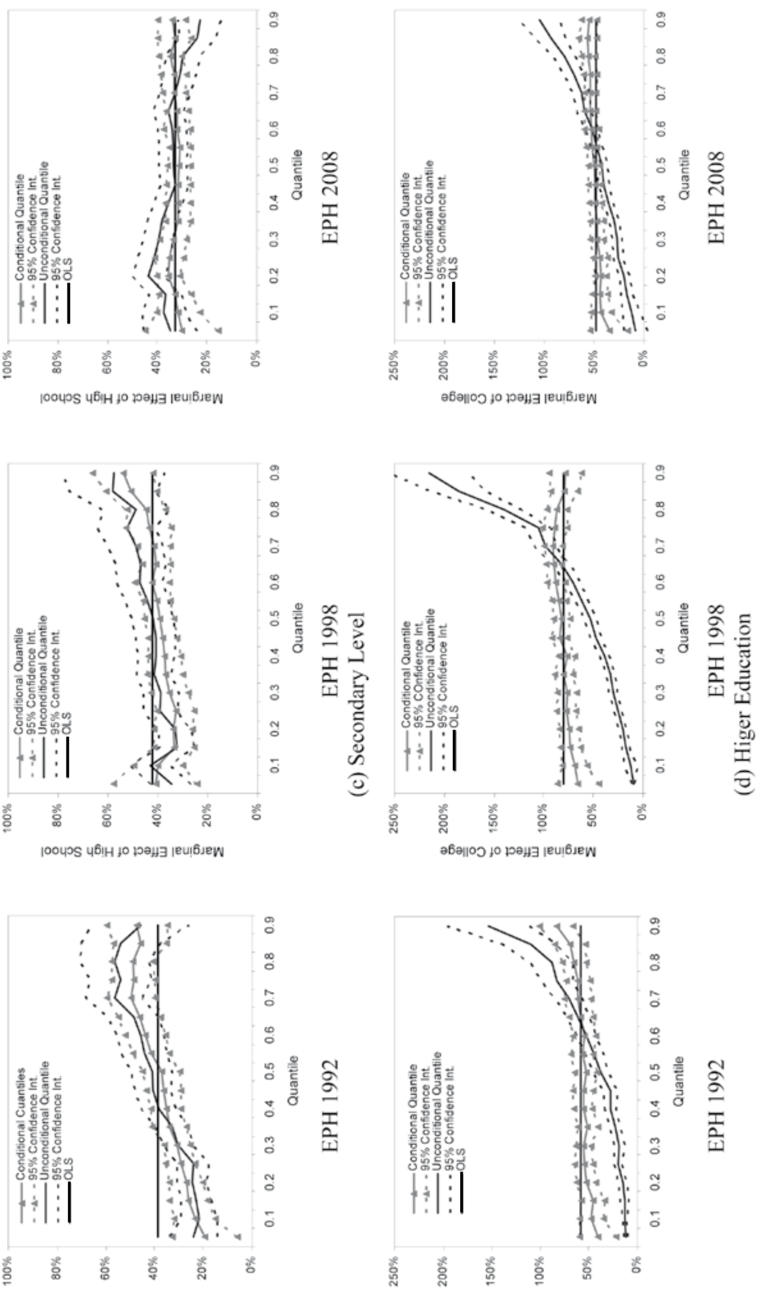
(a) Years of education



EPH 1992 EPH 1998 EPH 2008

(b) Primary Level

Figure 3.2 (continuation)



that shifting the distribution of education marginally, leads to an unequalizing increase of 1.83 points in the Gini coefficient[3]. It is interesting to remark that, qualitatively and quantitatively, these results are in agreement with those found by alternative methods (Gasparini, Marchionni, and Sosa Escudero (2004) and Bustelo (2004)). A major advantage of UQR is that the result is obtained using a single cross section for a given period, like in standard Mincer analysis, and unlike microeconometric decompositions that require either two periods or the construction of counterfactual distributions by simulation as in Mata and Machado (2005).

We then explore these effects for the remaining two periods (1998 and 2002). First, in 1998 all effects increase, for example, the mean effect moves from 0.084 in 1992 to 0.104 in 1998. Interestingly, the CQR results, though still positive and increasing along the quantiles of the conditional distribution, are less disperse, now with a difference of less than 0.02 points between the 0.1 and the 0.9 decile, suggesting a decreasing unequalizing effect. On the contrary, UQR results are markedly more heterogeneous, ranging from 0.066 to 0.19 along the quantiles of the unconditional distribution of wages, suggesting a strong unequalizing effect of education through this channel. This coincides with the beginning of the worst part of the performance in inequality in the period under analysis. This effect is further confirmed by the corresponding coefficient for education in the RIF regression for the Gini coefficient, which now leads to an increase of almost 2 points. It is important to remark that beyond the qualitative or statistical relevance of this figure, in economic terms, 2 points along the Gini coefficient of Argentina is an economically large figure, mostly from the perspective that the swings in inequality in the period under analysis range around 4 points.

The year 2008 presents a completely different picture. The levels of the effects are now similar to those of 1992, but the heterogeneity reduces drastically, as can be seen in the third graph of the first row of Figure 3.2. Now CQR effects stay rather stable around the mean effect (0.08), while UQR effects now range from 0.063 to 0.11. The effect of education on the Gini index is still unequalizing, but considerably smaller (0.49 in 2008).

Next we compute results measuring education by maximum levels of achievement, in order to allow for 'sheepskin effects' (Hungerford and Solon (1987)). To this purpose we run the same regressions as before but now replacing years of education with binary variables indicating the highest level of education reached by the individual. The results (OLS and QR) for the conditional distribution are reported in Table 3.6. Table 3.7 shows the results for the unconditional distribution (RIF-regressions). Given that we are dealing with binary regressors, it is common practice to interpret the estimations as the effect of finishing a certain level of education. However, the interpretation of the

[3] We are not simulating a policy exercise here, since that is not possible with these tools. The nature of changes in education distribution analyzed in this paper (small location shifts) can be considered as a simple exercise compared to a real educational policy seeking to improve human capital. However, our exercise can contribute to shed light about the effects on inequality of a certain policy that changes slightly the level of education from its status quo in order to improve human capital.

TABLE 3.5

MARGINAL EFFECTS ON MARGINAL WAGE DISTRIBUTION
RIF REGRESSION - MEN BETWEEN 16 AND 64 YEARS OLD

Argentina 1992

	q(0.10)	q(0.20)	q(0.30)	q(0.40)	q(0.50)	q(0.60)	q(0.70)	q(0.80)	q(0.90)	Gini
Indicator	4.64	5.81	6.83	7.89	9.32	10.92	13.26	17.06	24.52	40.5
Marginal Effects										
Age	0.005	0.006	0.008	0.009	0.010	0.012	0.013	0.014	0.016	0.14
(1)	(3.16)**	(5.69)**	(7.54)**	(8.77)**	(8.37)**	(10.12)**	(9.59)**	(8.43)**	(6.89)**	(3.86)**
(2)	(7.58)**	(10.40)**	(14.25)**	(17.08)**	(18.88)**	(18.93)**	(18.03)**	(16.23)**	(14.47)**	(9.47)**
Years of education	0.046	0.047	0.054	0.065	0.075	0.087	0.107	0.119	0.140	1.83
(1)	(13.23)**	(17.39)**	(20.81)**	(23.46)**	(24.44)**	(25.81)**	(27.15)**	(23.58)**	(15.89)**	(17.07)**
(2)	(18.14)**	(24.67)**	(27.98)**	(29.76)**	(31.16)**	(31.13)**	(28.95)**	(27.60)**	(21.11)**	(19.30)**
Married	0.149	0.168	0.142	0.134	0.170	0.172	0.213	0.205	0.197	2.05
(1)	(4.02)**	(5.85)**	(5.18)**	(4.58)**	(5.44)**	(5.14)**	(5.68)**	(4.70)**	(3.22)**	(2.03)*
(2)	(9.81)**	(11.91)**	(13.11)**	(13.47)**	(12.27)**	(12.94)**	(12.69)**	(10.86)**	(7.96)**	(12.76)**
Region 2 (Pampa)	–0.244	–0.213	–0.216	–0.236	–0.253	–0.256	–0.255	–0.260	–0.268	–0.76
(1)	(8.99)**	(10.28)**	(11.03)**	(11.55)**	(11.53)**	(10.94)**	(9.64)**	(8.55)**	(6.16)**	(0.68)
(2)	(7.47)**	(10.56)**	(11.99)**	(11.23)**	(10.56)**	(9.72)**	(9.42)**	(8.72)**	(7.13)**	(9.62)**
Region 3 (Cuyo)	–0.667	–0.550	–0.469	–0.453	–0.428	–0.414	–0.394	–0.386	–0.341	5.85
(1)	(14.98)**	(19.06)**	(18.89)**	(18.70)**	(16.98)**	(16.08)**	(13.90)**	(12.64)**	(8.11)**	(2.38)*
(2)	(10.98)**	(14.14)**	(15.92)**	(16.59)**	(15.98)**	(14.52)**	(14.09)**	(11.84)**	(9.37)**	(14.18)**
Region 4 (NOA)	–0.614	–0.467	–0.423	–0.398	–0.398	–0.368	–0.366	–0.344	–0.337	4.11
(1)	(18.43)**	(20.27)**	(20.44)**	(18.95)**	(18.25)**	(16.05)**	(14.62)**	(12.26)**	(8.78)**	(2.14)*
(2)	(13.95)**	(17.06)**	(18.96)**	(18.42)**	(17.98)**	(16.29)**	(14.18)**	(13.21)**	(9.86)**	(17.85)**
Region 5 (Patagonia)	–0.045	0.022	0.067	0.113	0.131	0.163	0.191	0.219	0.125	–0.18
(1)	(1.65)	(1.08)	(3.37)**	(5.32)**	(5.64)**	(6.53)**	(6.77)**	(6.72)**	(2.78)**	(0.08)
(2)	(0.18)	(2.17)*	(3.74)**	(7.13)**	(7.66)**	(9.41)**	(9.56)**	(8.38)**	(5.84)**	(1.51)
Constant	–0.407	–0.272	–0.214	–0.245	–0.221	–0.293	–0.343	–0.202	–0.165	0.15
(1)	(5.74)**	(5.04)**	(4.27)**	(4.78)**	(4.12)**	(5.30)**	(5.96)**	(2.94)**	(1.49)	(8.66)**
(2)	(13.36)**	(11.75)**	(10.00)**	(8.38)**	(7.81)**	(7.21)**	(5.82)**	(1.55)	(0.19)	(35.45)**
Sample size	12,196	12,196	12,196	12,196	12,196	12,196	12,196	12,196	12,196	12,196

Argentina 1998

	q(0.10)	q(0.20)	q(0.30)	q(0.40)	q(0.50)	q(0.60)	q(0.70)	q(0.80)	q(0.90)	Gini
Indicator	4.29	5.66	6.9	8.08	9.73	11.66	14.14	18.28	29.16	44.0
Marginal Effects										
Age	0.005	0.007	0.010	0.011	0.012	0.014	0.016	0.019	0.024	0.35
(1)	(2.74)**	(5.87)**	(8.42)**	(9.38)**	(10.21)**	(11.61)**	(12.60)**	(12.19)**	(9.51)**	(9.23)**
(2)	(6.98)**	(11.46)**	(15.58)**	(16.82)**	(18.28)**	(20.04)**	(19.59)**	(19.11)**	(14.73)**	(7.86)**
Years of education	0.066	0.059	0.070	0.079	0.090	0.104	0.124	0.148	0.190	1.99
(1)	(14.50)**	(19.87)**	(24.91)**	(29.23)**	(31.53)**	(33.87)**	(35.15)**	(30.45)**	(20.08)**	(18.02)**
(2)	(18.33)**	(21.89)**	(28.12)**	(31.35)**	(32.97)**	(31.18)**	(31.31)**	(27.74)**	(19.75)**	(15.56)**
Married	0.238	0.144	0.112	0.120	0.117	0.104	0.072	0.108	0.063	–4.00
(1)	(5.34)**	(4.78)**	(3.86)**	(4.22)**	(3.93)**	(3.38)**	(2.12)*	(2.65)**	(0.97)	(3.99)**
(2)	(7.04)**	(8.86)**	(9.49)**	(9.69)**	(9.29)**	(9.05)**	(8.30)**	(6.29)**	(2.92)**	(10.01)**
Region 2 (Pampa)	–0.143	–0.155	–0.192	–0.219	–0.228	–0.235	–0.243	–0.261	–0.421	–4.25
(1)	(3.99)**	(6.16)**	(7.94)**	(9.46)**	(9.50)**	(9.49)**	(8.81)**	(7.79)**	(8.54)**	(3.55)**
(2)	(3.96)**	(6.60)**	(8.62)**	(10.11)**	(10.34)**	(9.45)**	(9.30)**	(8.48)**	(8.36)**	(4.21)**
Region 3 (Cuyo)	–0.501	–0.365	–0.391	–0.379	–0.371	–0.331	–0.304	–0.292	–0.355	2.08
(1)	(10.26)**	(12.17)**	(14.53)**	(15.25)**	(14.75)**	(12.98)**	(11.00)**	(8.87)**	(7.22)**	(0.88)
(2)	(8.68)**	(11.17)**	(13.20)**	(15.27)**	(13.62)**	(13.77)**	(13.39)**	(10.44)**	(7.48)**	(10.41)**
Region 4 (NOA)	–0.794	–0.545	–0.520	–0.436	–0.432	–0.411	–0.372	–0.343	–0.453	3.42
(1)	(15.86)**	(18.51)**	(20.11)**	(18.23)**	(17.92)**	(17.06)**	(14.19)**	(10.96)**	(10.02)**	(1.82)
(2)	(12.29)**	(16.83)**	(17.56)**	(18.50)**	(15.47)**	(15.63)**	(13.87)**	(11.60)**	(9.76)**	(14.35)**
Region 5 (Patagonia)	0.007	0.018	0.028	0.064	0.082	0.097	0.128	0.137	0.059	0.10
(1)	(0.19)	(0.73)	(1.14)	(2.71)**	(3.27)**	(3.72)**	(4.44)**	(4.05)**	(1.18)	(0.04)
(2)	(2.25)*	(3.35)**	(4.57)**	(5.52)**	(7.19)**	(8.10)**	(9.15)**	(8.40)**	(5.20)**	(2.20)*
Constant	–0.597	–0.278	–0.265	–0.237	–0.200	–0.239	–0.339	–0.435	–0.527	0.14
(1)	(6.48)**	(4.62)**	(4.85)**	(4.74)**	(3.98)**	(4.89)**	(6.45)**	(6.62)**	(4.30)**	(7.73)**
(2)	(10.91)**	(7.40)**	(8.46)**	(7.26)**	(7.30)**	(5.90)**	(4.53)**	(4.54)**	–1.48	(32.37)**
Sample size	11,228	11,228	11,228	11,228	11,228	11,228	11,228	11,228	11,228	11,228

Table 3.5 (Continuation)

Argentina 2008

	q(0.10)	q(0.20)	q(0.30)	q(0.40)	q(0.50)	q(0.60)	q(0.70)	q(0.80)	q(0.90)	Gini
Indicator	3.62	4.97	6.21	7.37	8.69	10.14	12.28	15.35	21.12	39.8
Marginal Effects										
Age	0.005	0.006	0.008	0.009	0.010	0.011	0.012	0.013	0.015	0.18
(1)	(2.87)**	(5.68)**	(8.50)**	(10.76)**	(12.44)**	(13.23)**	(13.24)**	(12.65)**	(10.18)**	(3.39)**
(2)	(5.93)**	(8.96)**	(13.48)**	(15.58)**	(17.21)**	(19.10)**	(18.45)**	(15.21)**	(13.86)**	(2.13)*
Years of education	0.063	0.071	0.072	0.076	0.077	0.082	0.087	0.097	0.110	0.49
(1)	(13.46)**	(21.14)**	(26.65)**	(31.17)**	(33.46)**	(33.71)**	(31.84)**	(26.63)**	(19.25)**	(3.04)**
(2)	(17.69)**	(24.69)**	(28.25)**	(32.46)**	(33.09)**	(34.03)**	(31.90)**	(27.95)**	(23.13)**	(14.00)**
Married	0.276	0.136	0.071	0.070	0.066	0.062	0.054	0.010	0.046	–0.56
(1)	(6.52)**	(4.69)**	(2.97)**	(3.14)**	(3.06)**	(2.76)**	(2.24)*	(0.36)	(1.19)	(0.40)
(2)	(8.26)**	(8.42)**	(7.26)**	(6.75)**	(5.98)**	(6.02)**	(5.60)**	(4.61)**	(2.79)**	(7.28)**
Region 2 (Pampa)	–0.114	–0.087	–0.118	–0.112	–0.096	–0.094	–0.091	–0.090	–0.162	–2.50
(1)	(3.21)**	(3.47)**	(5.51)**	(5.65)**	(4.94)**	(4.65)**	(4.19)**	(3.44)**	(4.61)**	(1.49)
(2)	(2.67)**	(3.25)**	(4.53)**	(4.94)**	(4.95)**	(4.66)**	(3.33)**	(3.77)**	(3.53)**	(1.78)
Region 3 (Cuyo)	–0.430	–0.420	–0.372	–0.305	–0.289	–0.284	–0.312	–0.292	–0.311	1.14
(1)	(7.94)**	(11.27)**	(12.64)**	(11.63)**	(11.86)**	(11.64)**	(12.74)**	(10.67)**	(9.28)**	(0.33)
(2)	(8.11)**	(12.50)**	(13.86)**	(14.20)**	(13.30)**	(13.18)**	(11.24)**	(10.85)**	(8.59)**	(9.00)**
Region 4 (NOA)	–1.000	–0.714	–0.587	–0.484	–0.431	–0.385	–0.371	–0.354	–0.336	7.66
(1)	(18.56)**	(22.25)**	(23.99)**	(22.78)**	(21.89)**	(19.55)**	(18.39)**	(15.44)**	(11.21)**	(2.87)**
(2)	(16.12)**	(20.64)**	(21.37)**	(21.79)**	(22.18)**	(20.10)**	(17.06)**	(15.74)**	(11.32)**	(18.43)**
Region 5 (Patagonia)	0.173	0.241	0.260	0.311	0.336	0.405	0.428	0.510	0.554	3.70
(1)	(5.36)**	(10.23)**	(12.48)**	(15.39)**	(16.28)**	(18.33)**	(17.30)**	(16.19)**	(12.39)**	(1.10)
(2)	(7.44)**	(12.33)**	(15.90)**	(19.56)**	(20.79)**	(21.28)**	(20.27)**	(17.99)**	(14.82)**	(0.36)
Constant	0.330	0.558	0.753	0.823	0.924	0.999	1.131	1.182	1.291	0.28
(1)	(3.43)**	(8.68)**	(14.75)**	(18.43)**	(22.81)**	(24.40)**	(25.84)**	(21.30)**	(16.51)**	(9.91)**
(2)	(3.84)**	(10.60)**	(14.86)**	(19.83)**	(24.76)**	(26.70)**	(27.16)**	(27.60)**	(26.65)**	(33.23)**
Sample size	14,580	14,580	14,580	14,580	14,580	14,580	14,580	14,580	14,580	14,580

Source: Own calculations based on EPH (INDEC).
Notes: Absolute value of t statistics in parentheses, * significant at 5%; ** significant at 1%. Indicator refers to the functional (quantiles, Gini, etc.). (1) Standard Errors computed by rifreg.ado. (2) Standard Errors computed by a 1000 resampling bootstrap.

marginal effect is slightly different when it comes to RIF-regressions. As an example, if we pick the variable that indicates if the individual finished primary school (i.e. takes on the value of 1 for such individuals), the estimated coefficient of a RIF-regression measures the distributive effect of a small increase in the proportion of people in that educational level (see Firpo, Fortin, and Lemieux (2007b)). Results are shown graphically as before, in the 2nd to 4th row of graphs in Figure 3.2.

Primary school has a positive but homogeneous effect on both the conditional and unconditional distribution of income, along the whole period. Moreover, as measured by the RIF/Gini regression, this step induces an overall equalizing effect of education on the distribution of income. When moving to other levels of educational achievement, the heterogeneity starts to increase and now follows a pattern closer to that found when measuring education in years: heterogeneity is important but dampens in 2008. Also, it is interesting to see that higher education (as compared to the base category), shows a highly heterogeneous performance that peaked in 1998, coinciding with the period where inequality peaked in Argentina, suggesting that education had a markedly different effect which fueled inequality up.

TABLE 3.6

MARGINAL EFFECTS OF EDUCATION ON CONDITIONAL WAGE DISTRIBUTION QUANTILE REGRESSION - MEN BETWEEN 16 AND 64 YEARS OLD

Argentina 1992

	q(0.10)	q(0.20)	q(0.30)	q(0.40)	q(0.50)	q(0.60)	q(0.70)	q(0.80)	q(0.90)	Media
Primary complete	0.266	0.223	0.190	0.188	0.179	0.152	0.116	0.143	0.164	0.21
	(5.67)**	(5.79)**	(6.28)**	(5.58)**	(4.31)**	(3.44)**	(2.18)*	(3.17)**	(2.63)**	(9.83)**
Secondary incomplete	0.353	0.319	0.304	0.338	0.310	0.312	0.308	0.362	0.401	0.36
	(7.09)**	(7.90)**	(9.55)**	(9.57)**	(7.09)**	(6.69)**	(5.48)**	(7.48)**	(5.98)**	(16.31)**
Secondary complete	0.499	0.492	0.496	0.539	0.551	0.589	0.610	0.633	0.635	0.59
	(9.69)**	(11.76)**	(15.02)**	(14.67)**	(12.08)**	(12.13)**	(10.41)**	(12.60)**	(9.05)**	(25.77)**
College incomplete	0.640	0.669	0.694	0.757	0.758	0.798	0.806	0.892	0.956	0.81
	(10.68)**	(13.97)**	(18.35)**	(17.90)**	(14.38)**	(14.15)**	(11.81)**	(15.31)**	(11.99)**	(31.00)**
College complete	0.968	1.019	1.052	1.098	1.079	1.130	1.224	1.282	1.465	1.18
	(16.10)**	(20.77)**	(26.78)**	(25.15)**	(19.96)**	(19.70)**	(17.71)**	(21.73)**	(17.73)**	(44.48)**
Sample size	10,618	10,618	10,618	10,618	10,618	10,618	10,618	10,618	10,618	10,618

Argentina 1998

	q(0.10)	q(0.20)	q(0.30)	q(0.40)	q(0.50)	q(0.60)	q(0.70)	q(0.80)	q(0.90)	Media
Primary complete	0.212	0.137	0.185	0.148	0.142	0.166	0.215	0.230	0.213	0.19
	(3.40)**	(2.93)**	(3.99)**	(3.68)**	(3.89)**	(3.95)**	(5.42)**	(5.00)**	(2.96)**	(8.02)**
Secondary incomplete	0.371	0.270	0.298	0.291	0.307	0.354	0.389	0.406	0.384	0.36
	(5.90)**	(5.65)**	(6.26)**	(7.06)**	(8.24)**	(8.18)**	(9.51)**	(8.49)**	(5.10)**	(14.94)**
Secondary complete	0.611	0.475	0.538	0.520	0.535	0.591	0.630	0.676	0.753	0.61
	(9.40)**	(9.63)**	(11.00)**	(12.26)**	(13.90)**	(13.22)**	(14.94)**	(13.74)**	(9.72)**	(24.42)**
College incomplete	0.936	0.780	0.888	0.886	0.880	0.902	0.942	0.994	1.016	0.93
	(13.20)**	(14.39)**	(16.53)**	(18.94)**	(20.68)**	(18.19)**	(20.13)**	(18.14)**	(11.69)**	(33.66)**
College complete	1.300	1.217	1.301	1.286	1.350	1.468	1.533	1.542	1.535	1.42
	(18.38)**	(22.69)**	(24.26)**	(27.51)**	(31.68)**	(29.44)**	(32.33)**	(27.92)**	(17.62)**	(51.67)**
Sample size	11,231	11,231	11,231	11,231	11,231	11,231	11,231	11,231	11,231	11,231

Argentina 2008

	q(0.10)	q(0.20)	q(0.30)	q(0.40)	q(0.50)	q(0.60)	q(0.70)	q(0.80)	q(0.90)	Media
Primary complete	0.212	0.156	0.134	0.122	0.137	0.186	0.190	0.220	0.195	0.19
	(3.11)**	(4.06)**	(2.93)**	(2.87)**	(3.27)**	(3.94)**	(4.63)**	(5.18)**	(3.68)**	(8.14)**
Secondary incomplete	0.238	0.215	0.227	0.245	0.262	0.345	0.348	0.366	0.350	0.30
	(3.41)**	(5.55)**	(4.94)**	(5.71)**	(6.21)**	(7.22)**	(8.40)**	(8.48)**	(6.46)**	(12.38)**
Secondary complete	0.527	0.511	0.469	0.437	0.450	0.505	0.523	0.567	0.536	0.52
	(7.80)**	(13.64)**	(10.51)**	(10.49)**	(10.96)**	(10.92)**	(13.03)**	(13.57)**	(10.12)**	(22.07)**
College incomplete	0.825	0.787	0.743	0.702	0.706	0.759	0.790	0.860	0.848	0.79
	(11.28)**	(19.51)**	(15.50)**	(15.62)**	(15.94)**	(15.15)**	(18.12)**	(18.81)**	(14.73)**	(30.69)**
College complete	0.955	0.941	0.940	0.926	0.949	1.029	1.062	1.125	1.086	1.00
	(13.48)**	(23.94)**	(20.13)**	(21.20)**	(22.03)**	(21.13)**	(24.97)**	(25.31)**	(19.37)**	(40.24)**
Sample size	14,608	14,608	14,608	14,608	14,608	14,608	14,608	14,608	14,608	14,608

Source: Own calculations based on EPH (INDEC).
Note: Years old, marital status and regional *dummies* also was included in regression.

TABLE 3.7

MARGINAL EFFECTS OF EDUCATION ON UNCONDITIONAL WAGE DISTRIBUTION RIF REGRESSION - MEN BETWEEN 16 AND 64 YEARS OLD

Argentina 1992

	q(0.10)	q(0.20)	q(0.30)	q(0.40)	q(0.50)	q(0.60)	q(0.70)	q(0.80)	q(0.90)	Gini
Indicator	4.09	5.12	6.02	6.96	8.21	9.63	11.69	15.04	21.61	40.5
Marginal Effects										
Primary complete	0.280	0.217	0.165	0.149	0.169	0.168	0.174	0.188	0.145	–3.26
(1)	(4.10)**	(4.29)**	(3.52)**	(3.13)**	(3.49)**	(3.47)**	(3.72)**	(4.55)**	(2.95)**	(2.06)*
(2)	(6.35)**	(7.48)**	(8.00)**	(8.34)**	(8.28)**	(7.69)**	(7.39)**	(5.27)**	(5.44)**	(4.69)**
Secondary incomplete	0.339	0.320	0.302	0.299	0.344	0.354	0.380	0.432	0.307	–2.10
(1)	(4.83)**	(6.15)**	(6.28)**	(5.99)**	(6.73)**	(6.80)**	(7.28)**	(8.47)**	(5.02)**	(1.26)
(2)	(7.95)**	(10.04)**	(11.54)**	(12.17)**	(12.26)**	(13.01)**	(13.43)**	(12.91)**	(9.72)**	(1.76)
Secondary complete	0.497	0.462	0.466	0.532	0.580	0.626	0.739	0.752	0.601	–1.42
(1)	(7.27)**	(9.09)**	(9.81)**	(10.78)**	(11.09)**	(11.47)**	(12.79)**	(12.36)**	(7.30)**	(0.82)
(2)	(11.95)**	(15.82)**	(16.99)**	(17.91)**	(19.93)**	(19.96)**	(20.16)**	(17.98)**	(13.77)**	(0.13)
College incomplete	0.600	0.577	0.613	0.685	0.786	0.875	0.981	1.078	1.108	2.071
(1)	(8.54)**	(10.84)**	(12.42)**	(12.79)**	(13.63)**	(13.92)**	(14.01)**	(13.14)**	(8.79)**	(1.06)
(2)	(11.86)**	(16.93)**	(19.64)**	(20.21)**	(21.66)**	(20.24)**	(20.33)**	(18.13)**	(13.05)**	(2.44)*
College complete	0.625	0.611	0.652	0.806	0.956	1.151	1.437	1.632	2.133	37.08
(1)	(9.85)**	(12.56)**	(14.14)**	(16.94)**	(18.51)**	(20.73)**	(22.61)**	(19.49)**	(13.56)**	(18.64)**
(2)	(14.46)**	(19.02)**	(23.11)**	(25.35)**	(27.15)**	(25.74)**	(26.17)**	(22.86)**	(15.50)**	(10.82)**
Sample size	10,618	10,618	10,618	10,618	10,618	10,618	10,618	10,618	10,618	10,618

Argentina 1998

	q(0.10)	q(0.20)	q(0.30)	q(0.40)	q(0.50)	q(0.60)	q(0.70)	q(0.80)	q(0.90)	Gini
Indicator	3.78	4.99	6.08	7.12	8.57	10.27	12.46	16.11	25.7	44.0
Marginal Effects										
Primary complete	0.267	0.181	0.136	0.164	0.178	0.158	0.180	0.171	0.143	–1.30
(1)	(2.76)**	(2.97)**	(2.43)*	(3.19)**	(3.49)**	(3.28)**	(3.87)**	(3.87)**	(2.87)**	(0.78)
(2)	(6.10)**	(6.76)**	(6.16)**	(5.25)**	(7.28)**	(7.03)**	(7.58)**	(7.16)**	(5.23)**	(2.76)**
Secondary incomplete	0.437	0.302	0.308	0.322	0.322	0.342	0.350	0.377	0.332	–1.44
(1)	(4.55)**	(4.93)**	(5.46)**	(6.16)**	(6.14)**	(6.82)**	(7.01)**	(7.44)**	(5.34)**	(0.85)
(2)	(7.76)**	(9.36)**	(9.57)**	(10.37)**	(12.50)**	(11.85)**	(12.61)**	(12.05)**	(7.77)**	(3.95)**
Secondary complete	0.696	0.512	0.523	0.571	0.599	0.630	0.670	0.659	0.720	–3.15
(1)	(7.57)**	(8.56)**	(9.38)**	(10.86)**	(11.12)**	(11.79)**	(11.98)**	(10.86)**	(8.40)**	(1.79)
(2)	(10.42)**	(13.63)**	(15.26)**	(15.50)**	(17.92)**	(17.07)**	(17.88)**	(17.27)**	(11.78)**	(3.51)**
College incomplete	0.852	0.716	0.799	0.871	0.949	1.017	1.100	1.155	0.948	–2.47
(1)	(9.02)**	(11.84)**	(14.26)**	(16.15)**	(16.78)**	(17.44)**	(16.89)**	(14.71)**	(8.40)**	(1.29)
(2)	(12.87)**	(16.79)**	(19.09)**	(20.29)**	(22.46)**	(21.05)**	(21.65)**	(19.90)**	(13.75)**	(0.26)
College complete	0.819	0.714	0.825	0.959	1.127	1.332	1.653	2.054	2.870	37.62
(1)	(9.06)**	(12.59)**	(15.79)**	(19.76)**	(22.73)**	(26.86)**	(29.97)**	(27.95)**	(19.15)**	(19.62)**
(2)	(14.16)**	(18.13)**	(22.63)**	(25.28)**	(29.14)**	(27.13)**	(28.91)**	(26.31)**	(18.71)**	(11.07)**
Sample size	11,231	11,231	11,231	11,231	11,231	11,231	11,231	11,231	11,231	11,231

Table 3.7 (Continuation)

Argentina 2008

	q(0.10)	q(0.20)	q(0.30)	q(0.40)	q(0.50)	q(0.60)	q(0.70)	q(0.80)	q(0.90)	Gini
Indicator	3.62	4.97	6.21	7.37	8.69	10.14	12.28	15.35	21.12	39.8
Marginal Effects										
Primary complete	0.247	0.213	0.195	0.212	0.187	0.157	0.138	0.097	0.083	–1.88
(1)	(2.46)*	(2.80)**	(3.26)**	(4.26)**	(4.24)**	(3.64)**	(3.18)**	(1.92)	(2.15)*	(0.69)
(2)	(5.48)**	(5.95)**	(6.46)**	(7.31)**	(7.31)**	(7.13)**	(6.06)**	(4.66)**	(2.43)*	(3.78)**
Secondary incomplete	0.257	0.330	0.305	0.350	0.320	0.308	0.288	0.233	0.224	–1.95
(1)	(2.52)*	(4.36)**	(5.09)**	(7.01)**	(7.17)**	(6.95)**	(6.39)**	(4.41)**	(5.09)**	(0.69)
(2)	(6.38)**	(7.64)**	(8.91)**	(10.78)**	(10.48)**	(10.58)**	(9.81)**	(8.39)**	(5.74)**	(3.72)**
Secondary complete	0.620	0.648	0.590	0.567	0.524	0.497	0.467	0.392	0.313	–3.04
(1)	(6.48)**	(9.00)**	(10.30)**	(11.81)**	(12.10)**	(11.52)**	(10.54)**	(7.44)**	(6.90)**	(1.12)
(2)	(10.46)**	(13.58)**	(16.36)**	(18.40)**	(18.09)**	(17.73)**	(17.15)**	(13.97)**	(8.56)**	(4.65)**
College incomplete	0.821	0.873	0.849	0.864	0.822	0.793	0.724	0.715	0.750	–5.70
(1)	(8.44)**	(12.13)**	(14.77)**	(17.53)**	(17.93)**	(16.71)**	(14.32)**	(11.43)**	(10.40)**	(1.92)
(2)	(10.46)**	(13.58)**	(16.36)**	(18.40)**	(18.09)**	(17.73)**	(17.15)**	(13.97)**	(8.56)**	(4.65)**
College complete	0.747	0.848	0.857	0.925	0.945	1.020	1.096	1.192	1.361	7.06
(1)	(7.79)**	(11.97)**	(15.33)**	(19.72)**	(21.95)**	(23.10)**	(23.01)**	(19.08)**	(17.06)**	(2.46)*
(2)	(12.18)**	(16.00)**	(18.62)**	(21.07)**	(20.54)**	(19.87)**	(18.90)**	(17.00)**	(11.54)**	(6.19)**
Sample size	14,608	14,608	14,608	14,608	14,608	14,608	14,608	14,608	14,608	14,608

Source: Own calculations based on EPH (INDEC).
Note: Years old, marital status and regional dummies also was included in regression. Note: (1) Standard Errors computed by rifreg.ado, (2) Standard Errors computed by a 1000 resampling bootstrap.

4. CONCLUDING REMARKS

Standard conditional quantile regression suggest an unequalizing effect of education on the within dispersion of education for cells of individuals with similar education and other observed covariates, as documented by Martins and Pereira (2004). This does not necessarily mean that education has an unequalizing effect on the unconditional distribution of earnings. To this purpose, this paper uses unconditional quantile regressions as a powerful tool to explore the effects of education on the distribution of earnings directly.

The case of Argentina is a very relevant one, in light of the drastic movements in its income distribution and of the improvements in terms of educational achievement. In line with existing results for several countries and periods (including Argentina), the conditional quantile regression results in this paper suggest the presence of an unequalizing effect of education through positive and heterogeneous returns, increasing along the quantiles of the conditional distribution, a particularly strong effect for the nineties. Our unconditional quantile regression results suggest that in the nineties these heterogeneous returns were further enhanced and co-moved positively with the observed increases in inequality, as measured by the Gini index. Interestingly, results for the year 2008 suggest that these unequalizing effects reduced dramatically, revitalizing the role of education as a powerful policy variable to improve welfare. We think welfare as something affected by growth and inequality. That is, since education affects both the wage level (growth) and its distribution (inequality) in opposite directions, we find that the unequalizing effect becomes less strong. To summarize, the rapid increase in inequality in the nineties coincided with a period of increased education, particularly successful in moving individuals into high school, and a markedly heterogeneous performance in terms of how discrepancies in education were remunerated in the market. That is, in this period, the strong unequalizing effect of education is not due to increased education per-se, but on discrepancies in either quality of education, the way the market remunerates these discrepancies, and the interaction with abilities and their own remunerations. The results for the end of the nineties, suggest that this strong unequalizing effect has disappeared, reaffirming the relevant role fostering education has on improving welfare. Another relevant result, that reinforces the previous result, is that the channel that increases inequality through heterogeneity is almost absent when education increases at the lowest levels.

Finally, this paper refrains from exploring the effect of treating education as an endogenous variable. Unlike mean results, methods for handling such problem when the interest lies in distributive effects are still in their infancy (Powell (2011)), and, surely, are a top priority for further work. Nevertheless, it is relevant to remark that it is not clear ex-ante that the concerns that affect mean estimates translate into other functional alike. For example, when the interest lies in inequality, a biased counterfactual distribution that arises by ignoring endogeneities does not necessarily biased the functional of interest for distributive purposes. For example, if neglected endogeneities bias the whole conditional distribution up (or down), this affects negatively the estimation of the mean effect, but not necessarily that of distributive effects, which depend on distances between quantiles and not on their levels. A detailed exploration of these effects is a relevant route for further exploration, once reliable and ready to implement models and techniques become available.

REFERENCES

ALEJO, J. (2006). Desigualdad salarial en el Gran Buenos Aires: una aplicación de regresión por cuantiles en microdescomposiciones, Documento de Trabajo 36, CEDLAS, UNLP, Argentina.

ANGRIST, J.D. and J.S. PISCHKE (2008). *Mostly Harmless Econometrics: An Empiri- cist's Companion*. Princeton University Press, 1 edition.

BEHRMAN, J.R. (2011). "How much might human capital policies affect earnings inequalities and poverty?", *Estudios de Economía* 38 (1), pp. 9-41.

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

BUSTELO, M. (2004). Caracterización de los Cambios en la Desigualdad y la Pobreza en Argentina Haciendo uso de Técnicas de Descomposiciones Microeconométricas, Documento de Trabajo 13, CEDLAS.

CARD, D. (2001). "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems", *Econometrica* 69, pp. 1127-1160.

FIRPO, S., N. FORTIN and T. LEMIEUX (2007a). Decomposing Wage Distributions using Recentered Influence Functions Regressions, Mimeo, University of British Columbia.

FIRPO, S., N. FORTIN and T. LEMIEUX (2007b). Unconditional Quantile Regressions, Technical Working Paper 339, National Bureau of Economic Research, Cambridge, MA.

FIRPO, S., N. FORTIN, and T. LEMIEUX (2009). "Unconditional Quantile Regressions", *Econometrica* 77 (3), pp. 953-973.

GASPARINI, L. and G. CRUCES (2009). "Desigualdad En Argentina: Una Revisión de la Evidencia Empírica", *Desarrollo Económico* 1.

GASPARINI, L., M. MARCHIONNI and W. SOSA ESCUDERO (2004). *A characterization of inequality changes in Argentina through microeconometric decompositions*. Bourguignon, F., N. Lustig and F. Ferreira (eds.), The Microeconomics of Income Distribution Dynamics. Oxford.

HUNGERFORD, T. and G. SOLON (1987). "Sheepskin effects in the returns to education", *The Review of Economics and Statistics* 69 (1), pp. 175-177.

KOENKER, R. and G. BASSETT (1978). "Regression Quantiles", *Econometrica* 46 (1), pp. 33-50.

MARTINS, P. and P. PEREIRA (2004). "Does Education Reduce Wage Inequality? Quantile Regression Evidence from 16 Countries", *Labour Economics* 11 (3), pp. 355-371.

MATA, J. and J. MACHADO (2005). "Counterfactual decomposition of changes in wage distributions using quantile regression", *Journal of Applied Econometrics* 20, pp. 445-465.

MONTI, A.C. (1991). "The Study of the Gini Concentration Ratio by Means of the Influence Function", *Statistica* 51, pp. 561-577.

POWELL, D. (2011). Unconditional Quantile Regression for Exogenous or Endogenous Treatment Variables, Labor and Population working paper series WR 824.

RUIZ-TAGLE, J. (2007). "Forecasting wage inequality", *Estudios de Economía* 34 (1), pp. 141-162.

SOSA ESCUDERO, W. and S. PETRALIA (2012). *Anatomy of Distributive Changes in Argentina.* Baer, W. and D. Fleischer (eds.), The Economies of Argentina and Brazil: A Comparative Perspective. Edward Elgar Publishers, Northampton.

APPENDIX A: THE INFLUENCE FUNCTION

The *influence function* of ν at F_Y is given by

$$IF(y,v) \equiv \psi(y) - \int \psi(y)\, dF_Y\ (y) \tag{1}$$

Intuitively, it measures the influence a single point y has on a particular functional. For example, for the mean, the influence function is given by $y - \mu$.

It is important, to observe that

$$E[IF(y,v)] = 0 \tag{2}$$

APPENDIX B: LINEAR PROBABILITY MODEL AND RIF

If we assume the linear probability model $Pr[Y>q_\tau \mid X = x] = x'\beta$, trivially

$$\beta = \frac{d\,\Pr[I(Y > q_\tau) \mid X = x]}{dx} \tag{3}$$

Then, replacing in the result for the unconditional partial effect, for the case of quantiles we get

$$\alpha(\nu) = c_{1,\tau}\beta. \tag{4}$$

This leads to a very simple way to estimate these partial effects. Consider the regression model

$$I(y > q_\tau) = x'\beta + u \tag{5}$$

Note that under the linear probability assumption, $E(u|x) = 0$. Now

$$\begin{aligned} I(y > q_\tau)c_{1,\tau} + c_{2,\tau} &= c_{2,\tau} + c_{1,\tau}x'\beta + u \\ & c_{2,\tau} + c_{1,\tau}x'\beta^* + u \end{aligned} \tag{6}$$

with $\beta^* \equiv c_{1,\tau}\,\beta = \alpha(\nu)$. Then, if $RIF(y, q_\tau) = \mathrm{I}(y > q\tau)\, c_{1,\tau} + c_{2,\tau}$ were observable, a regression of $RIF(y,q_\tau)$ on x would provide a consistent estimate of $\beta^* = = \alpha(\nu)$.

APPENDIX C:
RECENTERED INFLUENCE FUNCTION FOR GINI INDEX

In this Appendix we follow Firpo, Fortin, and Lemieux (2007a). Let Y be a random variable distributed according to F_Y, then the Generalized Lorenz Curve is

$$GL(p,F_Y) = \int_{-\infty}^{F_{Y(p)}^{-1}} z dF_Y(z) \tag{7}$$

Define,

$$R(F_Y) = \int_0^1 GL(p,F_Y)\,dp \tag{8}$$

Then, the RIF for the Gini Index evaluated at y is given by

$$RIF(y,v) = 1 + B(F_Y) + C(F_Y), \tag{9}$$

where

$$B(F_Y) = 2\mu^{-2}R(F_Y) \tag{10}$$

$$C(y,F_Y) = 2\mu^{-1}\{y[1-p(y)] + GL[p(y)(F_Y)]\} \tag{11}$$

with $\mu = E(Y)$ and $p(y) = FY(y)$, (see Monti (1991)).

The RIF estimation uses the sample analogs in the previous formula. Let $y_1 \leq y_2 \leq \ldots \leq y_n$ be the ordered sample observations, then we have the following

$$\hat{p}(y) = \frac{1}{n}\sum_{j=1}^{n} I(y_j \leq y_i) \tag{12}$$

$$\widehat{GL}[\hat{p}(y_i)] = \frac{1}{n}\sum_{j=1}^{n} I(y_j \leq y_i)y_j \tag{13}$$

$R(F_Y)$ is obtained by numeric integration of $GL[p(y_i)]$. The next step is to replace each estimation in the previous formulas.