

Alejo, Javier; Gabrielli, María Florencia; Escudero, Walter Sosa

Working Paper

The distributive Effects of Education: An Unconditional Quantile Regression Approach

Documento de Trabajo, No. 125

Provided in Cooperation with:

Centro de Estudios Distributivos, Laborales y Sociales (CEDLAS), Universidad Nacional de La Plata

Suggested Citation: Alejo, Javier; Gabrielli, María Florencia; Escudero, Walter Sosa (2011) : The distributive Effects of Education: An Unconditional Quantile Regression Approach, Documento de Trabajo, No. 125, Universidad Nacional de La Plata, Centro de Estudios Distributivos, Laborales y Sociales (CEDLAS), La Plata

This Version is available at:

<https://hdl.handle.net/10419/127636>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

C | E | D | L | A | S

Centro de Estudios
Distributivos, Laborales y Sociales

Maestría en Economía
Facultad de Ciencias Económicas
Universidad Nacional de La Plata



**The Distributive Effects of Education: An
Unconditional Quantile Regression Approach**

Javier Alejo, Maria Florencia Gabrielli y Walter Sosa-
Escudero

Documento de Trabajo Nro. 125
Noviembre, 2011

ISSN 1853-0168

The Distributive Effects of Education: An Unconditional Quantile Regression Approach

Javier Alejo

Universidad Nacional de La Plata and CONICET

Maria Florencia Gabrielli

Universidad Nacional de Cuyo and CONICET

Walter Sosa-Escudero*

Universidad de San Andrés and CONICET

August 30, 2011

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 counterfactual 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.

JEL Classification: C21, I24, I31, D3

Keywords: unconditional quantile regression, income inequality, education, Argentina.

*Corresponding author: Walter Sosa Escudero, Departamento de Economía, Universidad de San Andrés. Vito Dumas 284, Victoria, Argentina, wsosa@udesa.edu.ar. 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

1 Introduction

Considerable attention has been awarded to the effects of education on economic outcomes. The massive literature on returns to education focuses on the causal effect of increasing education (usually measured in years) on expected earnings. In such context, a major concern of this body of research relates to the likely endogenous nature of education, which biases standard OLS estimates, and usually calls for instrumental variables strategies (See Card (2001) for a review). On the other hand, the literature on poverty and inequality focuses on aspects of the distribution of earnings other than the mean, such as its left tail, as the case of poverty, or its dispersion. This literature has moved gradually from unconditional analysis (i.e., measuring income based poverty or inequality) to conditional models that help explain the sources and causes of deprivation and/or inequality. From this perspective, standard returns to education analysis is seen as one particular step (focused on the mean) towards the final goal of quantifying the effect of the determinants of income (including education) on the whole income distribution, and eventually on functionals other than the mean, like poverty ratios or inequality indexes.

Regarding the goal of measuring distributional outcomes, a major step forward relates to the increasing popularity of *quantile regression* (QR) methods (Koenker (2005)), that help researchers focus on the effects of education (and other determinants of income) on the whole *conditional distribution*, beyond the conditional mean as in standard regression analysis. An important and recurrent result of this literature, triggered mostly by Buchinsky (1994)'s seminal article, is that education has a markedly heterogeneous effect on the conditional distribution of income. More concretely, several studies (see Martins and Pereira (2004)) suggest that an additional year of education has a monotonically increasing effect along the quantiles of the conditional distribution of income. Intuitively, this implies that for higher levels of education, the distribution of incomes is, both, shifted to the right and more disperse; that is, education has the double impact of increasing the center and the dispersion of the distribution of income. This previous result has lead some researchers to worry about an undesirable, unequalizing effect of education, as long as this effect translates from the conditional to the unconditional distribution of income.

Quantile regression results focus on the effects of measured determinants on the

conditional distribution of incomes, even though the ultimate interest lies on the *unconditional*, marginal, distribution. The transition from conditional to unconditional effects is not trivial. There are several approaches available. For example, Mata and Machado (2005) propose integrating observed determinants (like education) in conditional QR models, so an unconditional empirical distribution is obtained. In this strategy, the effects of education can be quantified by integrating with respect to alternative distributions of education and comparing resulting distributions. This approach is similar, in spirit, to decomposition approaches (see Firpo, Fortin, and Lemieux (2011) and Bourguignon, Lustig, and Ferreira (2004)), who try to decompose observed changes in inequality (or any other distributive measure) in changes in the observed determinants, changes in the way these determinants affect income, and usually a residual term. The nature of these decompositions depend on how much structure is given to the model that links the distribution of incomes to that of its observed and unobserved determinants.

In this paper we use recent advances in *unconditional quantile regressions* (UQR) (Firpo, Fortin, and Lemieux (2009)) to measure the effect of education (or any other observed determinant) on the whole (*unconditional*) distribution of incomes, with minimal assumptions like those in standard Mincerian mean income analysis, and without the need to rely on the passage of time to address the issue. That is, unlike decomposition analysis, UQR's allow researchers to measure the effect of a small change in education on the Gini index (or any functional of the unconditional income distribution), in a similar sense the coefficients of a linear model capture marginal effects on the mean in standard regression. The key analytical tool is the *recentered influence function* (RIF) regression, explained in detail in section 2.

We implement this method on data for Argentina. As well documented (Gasparini and Cruces (2009) and Sosa Escudero and Petralia (2011, forthcoming)), in the last 30 years Argentina went through several institutional and social episodes that altered the distribution of incomes dramatically, which led to unusually wide movements in poverty and inequality measures. Also, in the last 20 years the country has experienced dramatic improvements in educational attainments, see Gasparini and Cruces (2009) for a detailed description of such changes. This scenario provides relevant variability to explore distributive effects. To our knowledge, ours is the first study applying an UQR strategy for this case, and, in general, this relevant approach is still incipient, being a handful the only relevant studies.

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 expectation of the variable of interest. Only under very stringent assumptions such model can be used to extrapolate the effect of altering a covariate on other moments 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, like any quantile, the variance, or its Gini coefficient, that is, we will be interested in a particular feature of the distribution of income. In this sense, standard regression models focus 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 recover such effects 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 functionals that can be expressed as

$$\nu(F_Y) = \int \psi(y) dF_Y(y),$$

for some function $\psi(y)$. For example, the mean, μ_Y , corresponds to $\psi(y) = y$. In this context, the *influence function* of ν at F_Y is given by

$$IF(y, F_Y) \equiv \psi(y) - \int \psi(y) dF_Y(y)$$

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

$$E[IF(Y, F_Y)] = 0.$$

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

$$RIF(y, F_Y) \equiv IF(y, F_Y) + \nu(F_Y) = \psi(y),$$

and, trivially,

$$E[RIF(Y, F_Y)] = \nu(F_Y).$$

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,

$$\nu(F_Y) = \int RIF(y; \nu) dF_Y(y) = \int E[RIF(Y; \nu) | X = x] dF_X(x),$$

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:

$$\alpha(\nu) = \int \frac{dE[RIF(y; \nu) | X = x]}{dx} dF(x).$$

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 a 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 (Firpo, Fortin, and Lemieux (2009)) to be given by

$$\begin{aligned} RIF(y; q_\tau) = q_\tau + IF(y; q_\tau) &= q_\tau + \frac{\tau - \mathbb{I}\{y \leq q_\tau\}}{f_Y(q_\tau)} \\ &= \frac{\mathbb{I}\{y > q_\tau\}}{f_Y(q_\tau)} + q_\tau - \frac{1 - \tau}{f_Y(q_\tau)} \\ &= c_{1,\tau} \cdot \mathbb{I}\{y > q_\tau\} + c_{2,\tau}. \end{aligned}$$

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

$$\begin{aligned} E[RIF(Y; q_\tau)|X = x] &= c_{1,\tau} \cdot E[\mathbb{I}\{Y > q_\tau\}|X = x] + c_{2,\tau} \\ &= c_{1,\tau} Pr[Y > q_\tau|X = x] + c_{2,\tau}. \end{aligned}$$

This last expression is the *unconditional quantile regression*, that is, a regression model that links the expected value of quantiles (as measured by the RIF) to covariates. Particular specifications on $Pr[Y > q_\tau|X = x]$ lead to alternative regressions.

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

$$\beta = \frac{dPr[Y > q_\tau|X = x]}{dx}.$$

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

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

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

$$\mathbb{I}[y > q_\tau] = x'\beta + u.$$

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

$$\begin{aligned} \mathbb{I}[y > q_\tau]c_{1,\tau} + c_{2,\tau} &= c_{2,\tau} + c_{1,\tau} x'\beta + u \\ &= c_{2,\tau} + x'\beta^* + u, \end{aligned}$$

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

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 q_τ and $f_Y(q_\tau)$, which are estimated by the sample τ -th quantile of Y , and a standard non-parametric 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). 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 functionals 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), to which we refer for details.

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

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

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

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

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 results (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. The ultimate effect on the unconditional distribution (the subject of interest of distributive analysis) requires to ‘average’ 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 example, and as seen in the empirical part of this paper, the observed unequalizing effect in the conditional quantile regression might

be enhanced if takes place over an already unequal distribution of education, or damped if increases in education result in a more equal distribution of education.

3 Exploring the distributive effects of education: Argentina 1992-2009

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. It is worth mentioning that this survey has gone through some methodological changes. In 1998 thirteen cities were added to the sample and beginning in May 2003 data collection started to be done continuously instead of twice a year. In this respect, our results must be interpreted carefully. On the other hand, this aspect would not be too much of a problem since our method considers only the cross-sectional dimension of the data. However, for better comparison of our results across time we only consider the sample that includes the cities present in the EPH between 1992 and 1998. 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, Río Gallegos, Salta, San Luis - El Chorrillo, Greater San Juan, Santa Rosa - Toay, Ushuaia - Río 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 (2011, forthcoming) 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 more clear 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%. Similarly, the same proportion for complete high school raised from 16.4% to 22.4%. Educational levels increased monotonically, with most of the action taking place in the center of the distribution (around complete high school). These changes can be more drastically appreciated in Figure 3.1, which shows the estimated densities of education for the three periods.

In light of these results, it is natural to explore the interaction between changes in the distribution of education along those in the distribution of income. Gasparini, Marchionni, and Sosa Escudero (2001) is the first application for Argentina that explores this link using a microeconomic decomposition framework, 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 for higher levels of education. Alejo (2006) explores the statistical significance of all these results.

In this section we use RIF–regressions as introduced by Firpo, Fortin, and Lemieux (2009) and discussed in the previous section. This approach has some advantages, namely, (i) less data requirement since only one cross section sample is needed, in contrast with previous work that requires repeated cross section data with at least two periods, and the construction (by simulation) of counterfactual distributions; (ii) RIF–regressions are easier to compute, given that in order to recover the marginal distribution of income it is not necessary to use a large number of simulations as in Mata and Machado (2005) or other numerical solutions (e.g. Melly (2005)); finally, (iii) the marginal effects can be directly interpreted from the estimation results.

As discussed in the previous section, we estimate RIF regressions for several unconditional quantiles, using a linear probability specification. We also estimate a 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 3.2 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 3.3 presents results based on unconditional quantile regression, and the last columns shows 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 Table 3.2, 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 suggest that after controlling for all covariates, all quantiles of the conditional distribution increase when education is enhanced, but at an increasing rate the higher the quantiles. A common difficulty associated with interpreting 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, within the CQR it is difficult to see if this unequalizing effect translates to the unconditional distribution of incomes, hence the usefulness of the UQR approach, that studies effect directly on the distribution of income. Remarkably, UQR results show an even more pronounced heterogeneous behavior, with effect ranging from 0.046 to 0.140. UQR results are more directly interpretable since, now they suggest that the effects of education are 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 3.3) 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 that shifting the distribution of education marginally, leads to an *unequalizing* increase of 1.83 points in the Gini coefficient. At this point 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 (2001) and Bustelo (2004)). A major advantage is that the UQR requires cross sectional information only, like in standard Mincer equations, and unlike previous results who require either two points in time or the construction of counterfactual distributions by simulation.

We then explore these effects for the remaining two periods (1998 and 2002). First, in 1998 all effects move to the right, for example, the mean effect moves from 0.084 in 1992 to 0.104. Interestingly, the CQR results, though still positive and increasing along the quantiles of the conditional distribution, are less dispersed, with a difference now 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 a 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 reduced 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 effect 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).

Even though it seems reasonable to measure the amount of human capital by years of formal education, this information is sometimes not considered as a satisfactory or adequate measure of qualification in the labor market and as a result the reference taken is if the worker has finished certain level of education. This is known in the literature as “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 3.4. Table 3.5 shows the results for the unconditional distribution (RIF-regressions). In both cases the base category is unfinished primary school. Results are shown graphically as before, in the 2nd to 4th row of graphs in Figure 3.2.

Finishing 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. Moving to other levels, the heterogeneity starts to increase and now follows a pattern closer to that found when measuring education in years: heterogeneity in effects 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.

4 Concluding remarks

Even though abundant literature exists on the effect of education on expected earnings, distributive effects are more difficult to assess and quantify. This paper shows that unconditional quantile regression analysis is a powerful and simple tool to characterize changes in inequality and, in general, in other aspects of the distribution of income, like its quantiles. RIF regressions exploit cross-sectional variability and can be easily reproduced over time to measure the evolution of these effects.

The case of Argentina is a very relevant one, in light of the drastic movements in its income distribution and 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, indeed, 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. 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 functionals alike. For example, when the interest lies in inequality, a biased counterfactual distribution that arises by ignoring endogeneities does not necessarily bias the functionals 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, CEDLAS, UNLP, Argentina.*, (36).
- ANGRIST, J., AND J.-S. PISCHKE (2008): *Mostly Harmless Econometrics: An Empiricist’s Companion*. Princeton University Press, 1 edition edn.
- BOURGUIGNON, F., N. LUSTIG, AND F. FERREIRA (2004): *The Microeconomics of Income Distribution Dynamics*. Oxford University Press, Washington.
- BUCHINSKY, M. (1994): “Changes in the U.S. Wage Structure 1963-1987: Application of Quantile Regression,” *Econometrica*, 62(2), 405–458.
- BUSTELO, M. (2004): “Caracterización de los Cambios en la Desigualdad y la Pobreza en Argentina Haciendo Uso de Técnicas de Descomposiciones Microeconómicas,” *CEDLAS, Documento de Trabajo*, (13).
- CARD, D. (2001): “Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems,” *Econometrica*, 69, 1127–1160.
- FIRPO, S., N. FORTIN, AND T. LEMIEUX (2009): “Unconditional Quantile Regressions,” *Econometrica*, 77(3), 953–973.
- (2011): *Handbook of Labor Economics* chap. Decomposition Method in Economics. Elsevier, in press.
- GASPARINI, L., AND G. CRUCES (2009): “Desigualdad En Argentina: Una Revisión De La Evidencia Empírica,” *Desarrollo Economico*, 1.
- GASPARINI, L., M. MARCHIONNI, AND W. SOSA ESCUDERO (2001): *Distribución del Ingreso en la Argentina: Perspectivas y Efectos sobre el Bienestar*. Premio Fulvio S. Pagani, Fundacin Arcor.
- HUNGERFORD, T., AND G. SOLON (1987): “Sheepskin effects in the returns to education,” *The Review of Economics and Statistics*, 69(1), 175–177.
- KOENKER, R. (2005): *Quantile Regression*. Cambridge University Press, Cambridge.

- KOENKER, R., AND B. G. (1978): “Regression Quantiles,” *Econometrica*, 46(1), 33–50.
- MARTINS, P., AND P. PEREIRA (2004): “Does Education Reduce Wage Inequality? Quantile Regression Evidence from 16 Countries,” *Labour Economics*, 11(3), 355–371.
- MATA, J., AND J. MACHADO (2005): “Counterfactual decomposition of changes in wage distributions using quantile regression,” *Journal of Applied Econometrics*, 20(445-465).
- MELLY, B. (2005): “Decomposition of Differences in Distribution Using Quantile Regressions,” *Labour Economics*, 12, 577–90.
- POWELL, D. (2011): “Unconditional Quantile Regression for Exogenous or Endogenous Treatment Variables,” *Labor and Population working paper series*, (WR-824).
- SOSA ESCUDERO, W., AND S. PETRALIA (2011, forthcoming): *Comparative Growth and Development: Brazil and Argentina* chap. I Can Hear the Grass Grow: The Anatomy of Distributive Changes in Argentina. Edward Elgar Publishers.

Tables and Figures

Table 3.1: Summary statistics of survey data. Argentina 1992 - 2008 Sample:
Men between 16 and 64 years old

Year	Variable					
	Hourly wage					
	Mean	Std. Dev.	Quantile 0.10	Median	Quantile 0.90	Range 90-10
1992	11.5	12.0	4.1	8.2	21.6	17.5
1998	12.6	14.3	3.8	8.6	25.7	21.9
2008	11.4	15.9	3.6	8.7	21.1	17.5
	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 of education					
	Mean	Std. Dev.	Quantile 0.10	Median	Quantile 0.90	Range 90-10
1992	9.9	3.8	7	10	15	8
1998	10.0	3.8	7	10	16	9
2008	10.8	3.7	7	12	16	9
	Educational Level					
	Prim. incom.	Prim. compl.	Dropouts	Highschool	Coll. incom.	Coll. compl.
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%
	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 SEDLAC (CEDLAS and World Bank).

Table 3.2 marginal effects on conditional wage distributions: 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 (3.29)**	0.007 (6.67)**	0.008 (7.07)**	0.009 (11.59)**	0.010 (10.65)**	0.011 (11.79)**	0.012 (11.19)**	0.015 (12.10)**	0.015 (9.00)**	0.010 (20.33)**
Years of education	0.063 (16.46)**	0.065 (24.55)**	0.071 (22.87)**	0.076 (33.80)**	0.079 (28.00)**	0.086 (28.84)**	0.089 (25.06)**	0.093 (22.21)**	0.095 (16.09)**	0.084 (59.12)**
Married	0.154 (3.88)**	0.173 (6.62)**	0.175 (5.84)**	0.157 (7.38)**	0.181 (6.96)**	0.168 (6.24)**	0.196 (6.23)**	0.159 (4.47)**	0.153 (3.19)**	0.180 (13.52)**
Region 2 (Pampa)	-0.208 (7.62)**	-0.214 (11.52)**	-0.208 (9.60)**	-0.227 (14.86)**	-0.239 (12.75)**	-0.253 (13.07)**	-0.265 (11.71)**	-0.274 (10.82)**	-0.292 (8.51)**	-0.241 (16.53)**
Region 3 (Cuyo)	-0.472 (13.62)**	-0.427 (18.43)**	-0.424 (15.88)**	-0.446 (23.85)**	-0.470 (20.59)**	-0.464 (19.83)**	-0.468 (17.29)**	-0.472 (15.64)**	-0.529 (13.11)**	-0.461 (14.27)**
Region 4 (NOA)	-0.476 (16.95)**	-0.409 (21.60)**	-0.397 (18.03)**	-0.402 (25.91)**	-0.423 (22.21)**	-0.419 (21.33)**	-0.421 (18.32)**	-0.429 (16.63)**	-0.462 (13.27)**	-0.430 (17.02)**
Region 5 (Patagonia)	-0.031 (1.08)	0.055 (2.84)**	0.093 (4.13)**	0.112 (7.12)**	0.126 (6.54)**	0.138 (7.00)**	0.146 (6.38)**	0.133 (5.26)**	0.109 (3.21)**	0.088 (2.86)**
Constant	-0.525 (7.86)**	-0.411 (9.22)**	-0.372 (7.05)**	-0.314 (8.30)**	-0.258 (5.51)**	-0.220 (4.52)**	-0.160 (2.78)**	-0.065 (2.78)**	0.200 (2.11)**	-0.256 (10.99)**
Sample size	12196	12196	12196	12196	12196	12196	12196	12196	12196	12196

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 (2.71)**	0.007 (5.75)**	0.011 (9.76)**	0.014 (14.20)**	0.015 (14.69)**	0.016 (16.40)**	0.018 (19.68)**	0.020 (16.16)**	0.021 (12.90)**	0.013 (24.84)**
Years of education	0.091 (23.77)**	0.090 (29.63)**	0.096 (32.78)**	0.100 (37.58)**	0.100 (34.87)**	0.102 (34.93)**	0.107 (35.89)**	0.111 (26.70)**	0.113 (19.05)**	0.104 (66.76)**
Married	0.204 (5.37)**	0.162 (5.44)**	0.124 (4.41)**	0.106 (4.23)**	0.098 (3.78)**	0.122 (4.78)**	0.128 (5.16)**	0.121 (3.60)**	0.066 (1.41)	0.116 (8.21)**
Region 2 (Pampa)	-0.226 (7.11)**	-0.208 (8.22)**	-0.221 (9.40)**	-0.217 (10.48)**	-0.228 (10.65)**	-0.239 (11.41)**	-0.235 (11.63)**	-0.243 (9.07)**	-0.252 (6.69)**	-0.230 (13.61)**
Region 3 (Cuyo)	-0.370 (10.66)**	-0.325 (11.95)**	-0.339 (13.24)**	-0.337 (14.96)**	-0.359 (15.39)**	-0.378 (16.64)**	-0.395 (18.10)**	-0.424 (14.67)**	-0.382 (9.32)**	-0.361 (10.73)**
Region 4 (NOA)	-0.529 (15.64)**	-0.471 (17.80)**	-0.458 (18.52)**	-0.456 (20.93)**	-0.462 (20.53)**	-0.482 (21.92)**	-0.474 (22.38)**	-0.475 (16.97)**	-0.472 (12.11)**	-0.475 (17.88)**
Region 5 (Patagonia)	0.013 (0.39)	0.026 (1.01)	0.025 (1.05)	0.052 (2.47)**	0.062 (2.89)**	0.062 (4.38)**	0.091 (4.62)**	0.091 (4.31)**	0.111 (4.16)**	0.148 (1.99)**
Constant	-0.693 (9.42)**	-0.495 (8.87)**	-0.485 (9.40)**	-0.468 (10.29)**	-0.347 (7.32)**	-0.286 (6.11)**	-0.277 (5.96)**	-0.179 (2.84)**	0.036 (0.40)	-0.359 (14.09)**
Sample size	11228	11228	11228	11228	11228	11228	11228	11228	11228	11228

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 (3.84)**	0.006 (6.49)**	0.008 (8.39)**	0.008 (12.78)**	0.010 (13.93)**	0.011 (14.27)**	0.013 (21.44)**	0.014 (13.80)**	0.014 (14.48)**	0.009 (20.31)**
Years of education	0.082 (22.06)**	0.081 (28.21)**	0.080 (29.39)**	0.079 (38.77)**	0.078 (34.52)**	0.080 (30.96)**	0.082 (39.96)**	0.086 (23.50)**	0.084 (21.50)**	0.080 (57.67)**
Married	0.183 (5.19)**	0.123 (4.76)**	0.093 (3.89)**	0.096 (5.50)**	0.080 (4.19)**	0.060 (2.88)**	0.056 (3.47)**	0.055 (1.96)**	0.027 (0.92)	0.102 (8.40)**
Region 2 (Pampa)	-0.163 (5.23)**	-0.102 (4.44)**	-0.084 (3.94)**	-0.085 (5.42)**	-0.101 (5.95)**	-0.103 (5.53)**	-0.089 (6.19)**	-0.125 (5.16)**	-0.118 (4.64)**	-0.109 (7.51)**
Region 3 (Cuyo)	-0.386 (9.99)**	-0.331 (11.28)**	-0.294 (10.77)**	-0.312 (15.35)**	-0.308 (13.92)**	-0.324 (13.35)**	-0.311 (16.81)**	-0.306 (9.65)**	-0.338 (10.65)**	-0.336 (11.18)**
Region 4 (NOA)	-0.723 (20.96)**	-0.619 (24.31)**	-0.554 (23.35)**	-0.558 (32.35)**	-0.536 (28.66)**	-0.500 (24.42)**	-0.475 (30.44)**	-0.465 (17.43)**	-0.452 (16.36)**	-0.543 (23.46)**
Region 5 (Patagonia)	0.252 (7.41)**	0.318 (12.60)**	0.348 (14.98)**	0.334 (19.92)**	0.361 (20.03)**	0.365 (18.70)**	0.383 (26.09)**	0.407 (16.46)**	0.434 (16.95)**	0.351 (12.07)**
Constant	0.270 (3.71)**	0.537 (10.07)**	0.693 (14.17)**	0.835 (23.31)**	0.930 (23.76)**	1.028 (23.82)**	1.085 (32.41)**	1.182 (20.51)**	1.431 (23.59)**	0.896 (37.09)**
Sample size	14580	14580	14580	14580	14580	14580	14580	14580	14580	14580

Source: own calculations based on SEDLAC (CEDLAS and World Bank).

Table 3.3: Marginal effects on marginal wage distribution RIF Regression - Men between 16 and 64 years old

Argentina 1992										
<i>Indicator</i>	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
<i>Indicator</i>	4.64	5.81	6.83	7.89	9.32	10.92	13.26	17.06	24.52	40.5
<i>Marginal Effects</i>										
Age	0.005 (3.16)**	0.006 (5.69)**	0.008 (7.54)**	0.009 (8.77)**	0.010 (8.37)**	0.012 (10.12)**	0.013 (9.59)**	0.014 (8.43)**	0.016 (6.89)**	0.14 (3.86)**
Years of education	0.046 (13.23)**	0.047 (17.39)**	0.054 (20.81)**	0.065 (23.46)**	0.075 (24.44)**	0.087 (25.81)**	0.107 (27.15)**	0.119 (23.58)**	0.140 (15.89)**	1.83 (17.07)**
Married	0.149 (4.02)**	0.168 (5.85)**	0.142 (5.18)**	0.134 (4.58)**	0.170 (5.44)**	0.172 (5.14)**	0.213 (5.68)**	0.205 (4.70)**	0.197 (3.22)**	2.05 (2.03)**
Region 2 (Pampa)	-0.244 (8.99)**	-0.213 (10.28)**	-0.216 (11.03)**	-0.236 (11.55)**	-0.253 (11.53)**	-0.256 (10.94)**	-0.255 (9.64)**	-0.260 (8.55)**	-0.268 (6.16)**	-0.76 (0.68)
Region 3 (Cuyo)	-0.667 (14.98)**	-0.550 (19.06)**	-0.469 (18.89)**	-0.453 (18.70)**	-0.428 (16.98)**	-0.414 (16.08)**	-0.394 (13.90)**	-0.386 (12.64)**	-0.341 (8.11)**	5.85 (2.38)**
Region 4 (NOA)	-0.614 (18.43)**	-0.467 (20.27)**	-0.423 (20.44)**	-0.398 (18.95)**	-0.398 (18.25)**	-0.368 (16.05)**	-0.366 (14.62)**	-0.344 (12.26)**	-0.337 (8.78)**	4.11 (2.14)**
Region 5 (Patagonia)	-0.045 (1.65)	0.022 (1.08)	0.067 (3.37)**	0.113 (5.32)**	0.131 (5.64)**	0.163 (6.53)**	0.191 (6.77)**	0.219 (6.72)**	0.125 (2.78)**	-0.18 (0.08)
Constant	-0.407 (5.74)**	-0.272 (5.04)**	-0.214 (4.27)**	-0.245 (4.78)**	-0.221 (4.12)**	-0.293 (5.30)**	-0.343 (5.96)**	-0.202 (2.94)**	-0.165 (1.49)	0.15 (8.66)**
Sample size	12196	12196	12196	12196	12196	12196	12196	12196	12196	12196
Argentina 1998										
<i>Indicator</i>	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
<i>Indicator</i>	4.29	5.66	6.9	8.08	9.73	11.66	14.14	18.28	29.16	44.0
<i>Marginal Effects</i>										
Age	0.005 (2.74)**	0.007 (5.87)**	0.010 (8.42)**	0.011 (9.38)**	0.012 (10.21)**	0.014 (11.61)**	0.016 (12.60)**	0.019 (12.19)**	0.024 (9.51)**	0.35 (9.23)**
Years of education	0.066 (14.50)**	0.059 (19.87)**	0.070 (24.91)**	0.079 (29.23)**	0.090 (31.53)**	0.104 (33.87)**	0.124 (35.15)**	0.148 (30.45)**	0.190 (20.08)**	1.99 (18.02)**
Married	0.238 (5.34)**	0.144 (4.78)**	0.112 (3.86)**	0.120 (4.22)**	0.117 (3.93)**	0.104 (3.38)**	0.072 (2.12)**	0.108 (2.65)**	0.063 (0.97)	-4.00 (3.99)**
Region 2 (Pampa)	-0.143 (3.99)**	-0.155 (6.16)**	-0.192 (7.94)**	-0.219 (9.46)**	-0.228 (9.50)**	-0.235 (9.49)**	-0.243 (8.81)**	-0.261 (7.79)**	-0.421 (8.54)**	-4.25 (3.55)**
Region 3 (Cuyo)	-0.501 (10.26)**	-0.365 (12.17)**	-0.391 (14.53)**	-0.379 (15.25)**	-0.371 (14.75)**	-0.331 (12.98)**	-0.304 (11.00)**	-0.292 (8.87)**	-0.355 (7.22)**	2.08 (0.88)
Region 4 (NOA)	-0.794 (15.86)**	-0.545 (18.51)**	-0.520 (20.11)**	-0.436 (18.23)**	-0.432 (17.92)**	-0.411 (17.06)**	-0.372 (14.19)**	-0.343 (10.96)**	-0.453 (10.02)**	3.42 (1.82)
Region 5 (Patagonia)	0.007 (0.19)	0.018 (0.73)	0.028 (1.14)	0.064 (2.71)**	0.082 (3.27)**	0.097 (3.72)**	0.128 (4.44)**	0.137 (4.05)**	0.059 (1.18)	0.10 (0.04)
Constant	-0.597 (6.48)**	-0.278 (4.62)**	-0.265 (4.85)**	-0.237 (4.74)**	-0.200 (3.98)**	-0.239 (4.89)**	-0.339 (6.45)**	-0.435 (6.62)**	-0.527 (4.30)**	0.14 (7.73)**
Sample size	11228	11228	11228	11228	11228	11228	11228	11228	11228	11228
Argentina 2008										
<i>Indicator</i>	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
<i>Indicator</i>	3.62	4.97	6.21	7.37	8.69	10.14	12.28	15.35	21.12	39.8
<i>Marginal Effects</i>										
Age	0.005 (2.87)**	0.006 (5.68)**	0.008 (8.50)**	0.009 (10.76)**	0.010 (12.44)**	0.011 (13.23)**	0.012 (13.24)**	0.013 (12.65)**	0.015 (10.18)**	0.18 (3.39)**
Years of education	0.063 (13.46)**	0.071 (21.14)**	0.072 (26.65)**	0.076 (31.17)**	0.077 (33.46)**	0.082 (33.71)**	0.087 (31.84)**	0.097 (26.63)**	0.110 (19.25)**	0.49 (3.04)**
Married	0.276 (6.52)**	0.136 (4.69)**	0.071 (2.97)**	0.070 (3.14)**	0.066 (3.06)**	0.062 (2.76)**	0.054 (2.24)**	0.010 (0.36)	0.046 (1.19)	-0.56 (0.40)
Region 2 (Pampa)	-0.114 (3.21)**	-0.087 (3.47)**	-0.118 (5.51)**	-0.112 (5.65)**	-0.096 (4.94)**	-0.094 (4.65)**	-0.091 (4.19)**	-0.090 (3.44)**	-0.162 (4.61)**	-2.50 (1.49)
Region 3 (Cuyo)	-0.430 (7.94)**	-0.420 (11.27)**	-0.372 (12.64)**	-0.305 (11.63)**	-0.289 (11.86)**	-0.284 (11.64)**	-0.312 (12.74)**	-0.292 (10.67)**	-0.311 (9.28)**	1.14 (0.33)
Region 4 (NOA)	-1.000 (18.56)**	-0.714 (22.25)**	-0.587 (23.99)**	-0.484 (22.78)**	-0.431 (21.89)**	-0.385 (19.55)**	-0.371 (18.39)**	-0.354 (15.44)**	-0.336 (11.21)**	7.66 (2.87)**
Region 5 (Patagonia)	0.173 (5.36)**	0.241 (10.23)**	0.260 (12.48)**	0.311 (15.39)**	0.336 (16.28)**	0.405 (18.33)**	0.428 (17.30)**	0.510 (16.19)**	0.554 (12.39)**	3.70 (1.10)
Constant	0.330 (3.43)**	0.558 (8.68)**	0.753 (14.75)**	0.823 (18.43)**	0.924 (22.81)**	0.999 (24.40)**	1.131 (25.84)**	1.182 (21.30)**	1.291 (16.51)**	0.28 (9.91)**
Sample size	14580	14580	14580	14580	14580	14580	14580	14580	14580	14580

Source: own calculations based on SEDLAC (CEDLAS and World Bank).

Table 3.4: 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 (5.67)**	0.223 (5.79)**	0.190 (6.28)**	0.188 (5.58)**	0.179 (4.31)**	0.152 (3.44)**	0.116 (2.18)*	0.143 (3.17)**	0.164 (2.63)**	0.21 (9.83)**
Secondary incomplete	0.353 (7.09)**	0.319 (7.90)**	0.304 (9.55)**	0.338 (9.57)**	0.310 (7.09)**	0.312 (6.69)**	0.308 (5.48)**	0.362 (7.48)**	0.401 (5.98)**	0.36 (16.31)**
Secondary complete	0.499 (9.69)**	0.492 (11.76)**	0.496 (15.02)**	0.539 (14.67)**	0.551 (12.08)**	0.589 (12.13)**	0.610 (10.41)**	0.633 (12.60)**	0.635 (9.05)**	0.59 (25.77)**
College incomplete	0.640 (10.68)**	0.669 (13.97)**	0.694 (18.35)**	0.757 (17.90)**	0.758 (14.38)**	0.798 (14.15)**	0.806 (11.81)**	0.892 (15.31)**	0.956 (11.99)**	0.81 (31.00)**
College complete	0.968 (16.10)**	1.019 (20.77)**	1.052 (26.78)**	1.098 (25.15)**	1.079 (19.96)**	1.130 (19.70)**	1.224 (17.71)**	1.282 (21.73)**	1.465 (17.73)**	1.18 (44.48)**
Sample size	10618	10618	10618	10618	10618	10618	10618	10618	10618	10618
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 (3.40)**	0.137 (2.93)**	0.185 (3.99)**	0.148 (3.68)**	0.142 (3.89)**	0.166 (3.95)**	0.215 (5.42)**	0.230 (5.00)**	0.213 (2.96)**	0.19 (8.02)**
Secondary incomplete	0.371 (5.90)**	0.270 (5.65)**	0.298 (6.26)**	0.291 (7.06)**	0.307 (8.24)**	0.354 (8.18)**	0.389 (9.51)**	0.406 (8.49)**	0.384 (5.10)**	0.36 (14.94)**
Secondary complete	0.611 (9.40)**	0.475 (9.63)**	0.538 (11.00)**	0.520 (12.26)**	0.535 (13.90)**	0.591 (13.22)**	0.630 (14.94)**	0.676 (13.74)**	0.753 (9.72)**	0.61 (24.42)**
College incomplete	0.936 (13.20)**	0.780 (14.39)**	0.888 (16.53)**	0.886 (18.94)**	0.880 (20.68)**	0.902 (18.19)**	0.942 (20.13)**	0.994 (18.14)**	1.016 (11.69)**	0.93 (33.66)**
College complete	1.300 (18.38)**	1.217 (22.69)**	1.301 (24.26)**	1.286 (27.51)**	1.350 (31.68)**	1.468 (29.44)**	1.533 (32.33)**	1.542 (27.92)**	1.535 (17.62)**	1.42 (51.67)**
Sample size	11231	11231	11231	11231	11231	11231	11231	11231	11231	11231
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 (3.11)**	0.156 (4.06)**	0.134 (2.93)**	0.122 (2.87)**	0.137 (3.27)**	0.186 (3.94)**	0.190 (4.63)**	0.220 (5.18)**	0.195 (3.68)**	0.19 (8.14)**
Secondary incomplete	0.238 (3.41)**	0.215 (5.55)**	0.227 (4.94)**	0.245 (5.71)**	0.262 (6.21)**	0.345 (7.22)**	0.348 (8.40)**	0.366 (8.48)**	0.350 (6.46)**	0.30 (12.38)**
Secondary complete	0.527 (7.80)**	0.511 (13.64)**	0.469 (10.51)**	0.437 (10.49)**	0.450 (10.96)**	0.505 (10.92)**	0.523 (13.03)**	0.567 (13.57)**	0.536 (10.12)**	0.52 (22.07)**
College incomplete	0.825 (11.28)**	0.787 (19.51)**	0.743 (15.50)**	0.702 (15.62)**	0.706 (15.94)**	0.759 (15.15)**	0.790 (18.12)**	0.860 (18.81)**	0.848 (14.73)**	0.79 (30.69)**
College complete	0.955 (13.48)**	0.941 (23.94)**	0.940 (20.13)**	0.926 (21.20)**	0.949 (22.03)**	1.029 (21.13)**	1.062 (24.97)**	1.125 (25.31)**	1.086 (19.37)**	1.00 (40.24)**
Sample size	14608	14608	14608	14608	14608	14608	14608	14608	14608	14608

Source: own calculations based on SEDLAC (CEDLAS and World Bank).
 Note: years old, marital status and regional dummies also was included in regression

Table 3.5: 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
<i>Indicator</i>	4.09	5.12	6.02	6.96	8.21	9.63	11.69	15.04	21.61	40.5
<i>Marginal Effects</i>										
Primary complete	0.280 (4.10)**	0.217 (4.29)**	0.165 (3.52)**	0.149 (3.13)**	0.169 (3.49)**	0.168 (3.47)**	0.174 (3.72)**	0.188 (4.55)**	0.145 (2.95)**	-3.26 (2.06)*
Secondary incomplete	0.339 (4.83)**	0.320 (6.15)**	0.302 (6.28)**	0.299 (5.99)**	0.344 (6.73)**	0.354 (6.80)**	0.380 (7.28)**	0.432 (8.47)**	0.307 (5.02)**	-2.10 (1.26)
Secondary complete	0.497 (7.27)**	0.462 (9.09)**	0.466 (9.81)**	0.532 (10.78)**	0.580 (11.09)**	0.626 (11.47)**	0.739 (12.79)**	0.752 (12.36)**	0.601 (7.30)**	-1.42 (0.82)
College incomplete	0.600 (8.54)**	0.577 (10.84)**	0.613 (12.42)**	0.685 (12.79)**	0.786 (13.63)**	0.875 (13.92)**	0.981 (14.01)**	1.078 (13.14)**	1.108 (8.79)**	2.07 (1.06)
College complete	0.625 (9.85)**	0.611 (12.56)**	0.652 (14.14)**	0.806 (16.94)**	0.956 (18.51)**	1.151 (20.73)**	1.437 (22.61)**	1.632 (19.49)**	2.133 (13.56)**	37.08 (18.64)**
Sample size	10618	10618	10618	10618	10618	10618	10618	10618	10618	10618

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
<i>Indicator</i>	3.78	4.99	6.08	7.12	8.57	10.27	12.46	16.11	25.7	44.0
<i>Marginal Effects</i>										
Primary complete	0.267 (2.76)**	0.181 (2.97)**	0.136 (2.43)*	0.164 (3.19)**	0.178 (3.49)**	0.158 (3.28)**	0.180 (3.87)**	0.171 (3.87)**	0.143 (2.87)**	-1.30 (0.78)
Secondary incomplete	0.437 (4.55)**	0.302 (4.93)**	0.308 (5.46)**	0.322 (6.16)**	0.322 (6.14)**	0.342 (6.82)**	0.350 (7.01)**	0.377 (7.44)**	0.332 (5.34)**	-1.44 (0.85)
Secondary complete	0.696 (7.57)**	0.512 (8.56)**	0.523 (9.38)**	0.571 (10.86)**	0.599 (11.12)**	0.630 (11.79)**	0.670 (11.98)**	0.659 (10.86)**	0.720 (8.40)**	-3.15 (1.79)
College incomplete	0.852 (9.02)**	0.716 (11.84)**	0.799 (14.26)**	0.871 (16.15)**	0.949 (16.78)**	1.017 (17.44)**	1.100 (16.89)**	1.155 (14.71)**	0.948 (8.40)**	-2.47 (1.29)
College complete	0.819 (9.06)**	0.714 (12.59)**	0.825 (15.79)**	0.959 (19.76)**	1.127 (22.73)**	1.332 (26.86)**	1.653 (29.97)**	2.054 (27.95)**	2.870 (19.15)**	37.62 (19.62)**
Sample size	11231	11231	11231	11231	11231	11231	11231	11231	11231	11231

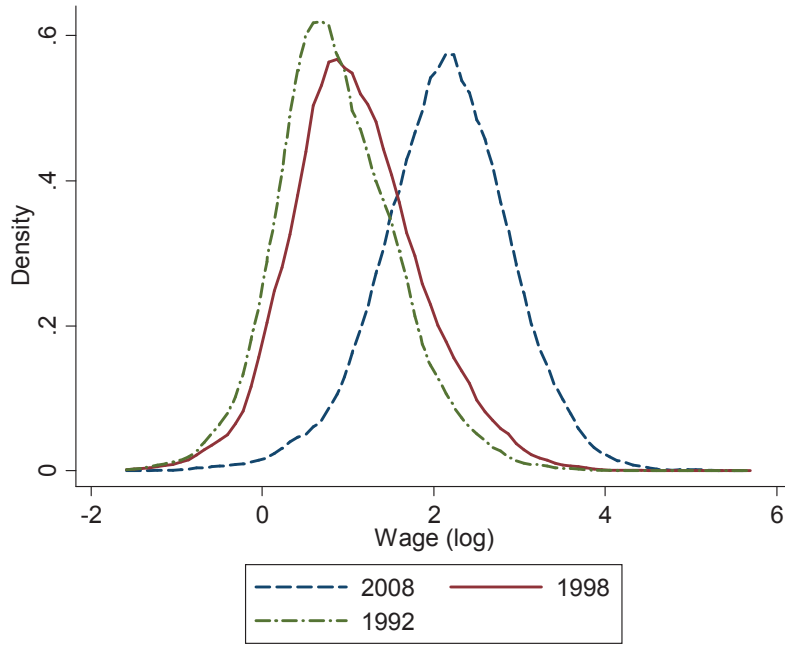
Argentina 2008

	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
<i>Indicator</i>	3.62	4.97	6.21	7.37	8.69	10.14	12.28	15.35	21.12	39.8
<i>Marginal Effects</i>										
Primary complete	0.247 (2.46)*	0.213 (2.80)**	0.195 (3.26)**	0.212 (4.26)**	0.187 (4.24)**	0.157 (3.64)**	0.138 (3.18)**	0.097 (1.92)	0.083 (2.15)*	-1.88 (0.69)
Secondary incomplete	0.257 (2.52)*	0.330 (4.36)**	0.305 (5.09)**	0.350 (7.01)**	0.320 (7.17)**	0.308 (6.95)**	0.288 (6.39)**	0.233 (4.41)**	0.224 (5.09)**	-1.95 (0.69)
Secondary complete	0.620 (6.48)**	0.648 (9.00)**	0.590 (10.30)**	0.567 (11.81)**	0.524 (12.10)**	0.497 (11.52)**	0.467 (10.54)**	0.392 (7.44)**	0.313 (6.90)**	-3.04 (1.12)
College incomplete	0.821 (8.44)**	0.873 (12.13)**	0.849 (14.77)**	0.864 (17.53)**	0.822 (17.93)**	0.793 (16.71)**	0.724 (14.32)**	0.715 (11.43)**	0.750 (10.40)**	-5.70 (1.92)
College complete	0.747 (7.79)**	0.848 (11.97)**	0.857 (15.33)**	0.925 (19.72)**	0.945 (21.95)**	1.020 (23.10)**	1.096 (23.01)**	1.192 (19.08)**	1.361 (17.06)**	7.06 (2.46)*
Sample size	14608	14608	14608	14608	14608	14608	14608	14608	14608	14608

Source: own calculations based on SEDLAC (CEDLAS and World Bank).

Note: years old, marital status and regional dummies also was included in regression.

Figure 3.1: Wage marginal distribution. Argentina 1992 - 2008 Sample: Men between 16 and 64 years old



Source: own calculations based on SEDLAC (CEDLAS and World Bank).

Figure 3.2: Marginal effects of education on unconditional wage distribution - Men between 16 and 64 years old

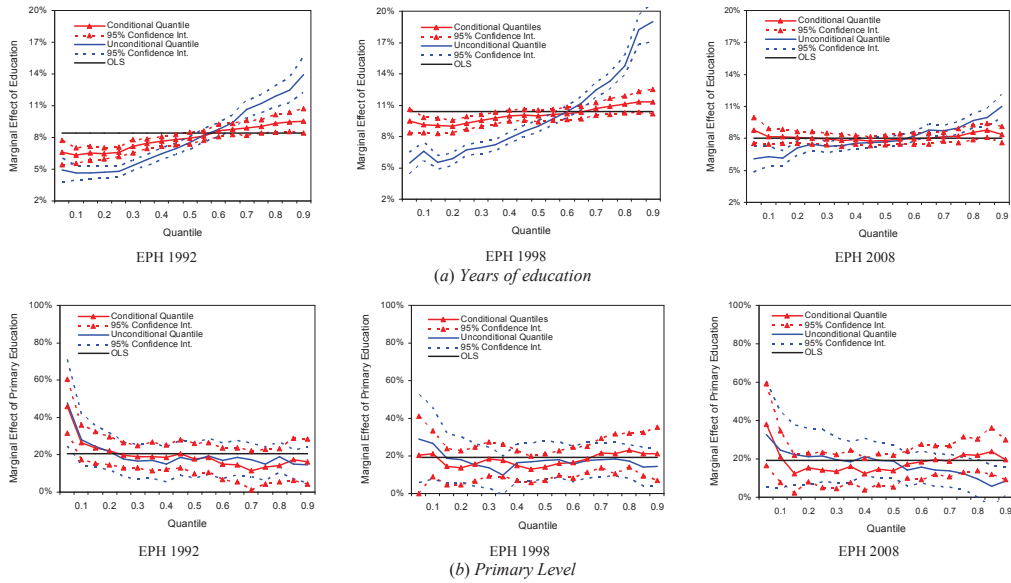


Figure 3.2 (cont.): Marginal effects of education on unconditional wage distribution - Men between 16 and 64 years old (cont.)

