*Econometrica*, Vol. 77, No. 3 (May, 2009), 953–973

# UNCONDITIONAL QUANTILE REGRESSIONS

BY SERGIO FIRPO, NICOLE M. FORTIN, AND THOMAS LEMIEUX[1]

We propose a new regression method to evaluate the impact of changes in the distribution of the explanatory variables on quantiles of the unconditional (marginal) distribution of an outcome variable. The proposed method consists of running a regression of the (recentered) influence function (RIF) of the unconditional quantile on the explanatory variables. The influence function, a widely used tool in robust estimation, is easily computed for quantiles, as well as for other distributional statistics. Our approach, thus, can be readily generalized to other distributional statistics.

KEYWORDS: Influence functions, unconditional quantile, RIF regressions, quantile regressions.

## 1. INTRODUCTION

IN THIS PAPER, we propose a new computationally simple regression method to estimate the impact of changing the distribution of explanatory variables, $X$, on the marginal quantiles of the outcome variable, $Y$, or other functional of the marginal distribution of $Y$. The method consists of running a regression of a transformation—the (recentered) influence function defined below—of the outcome variable on the explanatory variables. To distinguish our approach from commonly used *conditional* quantile regressions (Koenker and Bassett (1978), Koenker (2005)), we call our regression method an *unconditional quantile regression*.[2]

Empirical researchers are often interested in changes in the quantiles, denoted $q_\tau$, of the marginal (unconditional) distribution, $F_Y(y)$. For example, we may want to estimate the direct effect $dq_\tau(p)/dp$ of increasing the proportion of unionized workers, $p = \Pr[X = 1]$, on the $\tau$th quantile of the distribution of wages, where $X = 1$ if the workers is unionized and $X = 0$ otherwise. In the case of the mean $\mu$, the coefficient $\beta$ of a standard regression of $Y$ on $X$ is a measure of the impact of increasing the proportion of unionized

[1]We thank the co-editor and three referees for helpful suggestions. We are also indebted to Joe Altonji, Richard Blundell, David Card, Vinicius Carrasco, Marcelo Fernandes, Chuan Goh, Jinyong Hahn, Joel Horowitz, Guido Imbens, Shakeeb Khan, Roger Koenker, Thierry Magnac, Ulrich Müller, Geert Ridder, Jean-Marc Robin, Hal White, and seminar participants at CESG2005, UCL, CAEN–UFC, UFMG, Econometrics in Rio 2006, PUC-Rio, IPEA-RJ, SBE Meetings 2006, Tilburg University, Tinbergen Institute, KU Leuven, ESTE-2007, Harvard–MIT Econometrics Seminar, Yale, Princeton, Vanderbilt, and Boston University for useful comments on earlier versions of the manuscript. Fortin and Lemieux thank SSHRC for financial support. Firpo thanks CNPq for financial support. Usual disclaimers apply.

[2]The "unconditional quantiles" are the quantiles of the marginal distribution of the outcome variable $Y$. Using "marginal" instead of "unconditional" would be confusing, however, since we also use the word "marginal" to refer to the impact of small changes in covariates (marginal effects).

 DOI: 10.3982/ECTA6822

workers on the mean wage, $d\mu(p)/dp$. As is well known, the same coefficient $\beta$ can also be interpreted as an impact on the conditional mean.[3] Unfortunately, the coefficient $\beta_\tau$ from a single conditional quantile regression, $\beta_\tau = F_Y^{-1}(\tau|X=1) - F_Y^{-1}(\tau|X=0)$, is generally different from $dq_\tau(p)/dp = (\Pr[Y > q_\tau|X=1] - \Pr[Y > q_\tau|X=0])/f_Y(q_\tau)$, the effect of increasing the proportion of unionized workers on the $\tau$th quantile of the unconditional distribution of $Y$.[4] A new approach is therefore needed to provide practitioners with an easy way to compute $dq_\tau(p)/dp$, especially when $X$ is not univariate and binary as in the above example.

Our approach builds upon the concept of the influence function (IF), a widely used tool in the robust estimation of statistical or econometric models. As its name suggests, the influence function $\mathrm{IF}(Y;\nu,F_Y)$ of a distributional statistic $\nu(F_Y)$ represents the influence of an individual observation on that distributional statistic. Adding back the statistic $\nu(F_Y)$ to the influence function yields what we call the *recentered influence function* (RIF). One convenient feature of the RIF is that its expectation is equal to $\nu(F_Y)$.[5] Because influence functions can be computed for most distributional statistics, our method easily extends to other choices of $\nu$ beyond quantiles, such as the variance, the Gini coefficient, and other commonly used inequality measures.[6]

For the $\tau$th quantile, the influence function $\mathrm{IF}(Y;q_\tau,F_Y)$ is known to be equal to $(\tau - \mathbb{1}\{Y \le q_\tau\})/f_Y(q_\tau)$. As a result, $\mathrm{RIF}(Y;q_\tau,F_Y)$ is simply equal to $q_\tau + \mathrm{IF}(Y;q_\tau,F_Y)$. We call the conditional expectation of the $\mathrm{RIF}(Y;\nu,F_Y)$ modeled as a function of the explanatory variables, $E[\mathrm{RIF}(Y;\nu,F_Y)|X] = m_\nu(X)$, the RIF *regression model*.[7] In the case of quantiles, $E[\mathrm{RIF}(Y;q_\tau, F_Y)|X] = m_\tau(X)$ can be viewed as an *unconditional quantile regression*. We show that the average derivative of the unconditional quantile regression, $E[m'_\tau(X)]$, corresponds to the marginal effect on the unconditional quantile of a small location shift in the distribution of covariates, holding everything else constant.

Our proposed approach can be easily implemented as an ordinary least squares (OLS) regression. In the case of quantiles, the dependent variable in the regression is $\mathrm{RIF}(Y;q_\tau,F_Y) = q_\tau + (\tau - \mathbb{1}\{Y \le q_\tau\})/f_Y(q_\tau)$. It is easily

[3]The conditional mean interpretation is the wage change that a worker would expect when her union status changes from non-unionized to unionized, or $\beta = E(Y|X=1) - E(Y|X=0)$. Since the unconditional mean is $\mu(p) = pE(Y|X=1) + (1-p)E(Y|X=0)$, it follows that $d\mu(p)/dp = E(Y|X=1) - E(Y|X=0) = \beta$.

[4]The expression for $dq_\tau(p)/dp$ is obtained by implicit differentiation applied to $F_Y(q_\tau) = p \cdot (\Pr[Y \le q_\tau|X=1] - \Pr[Y \le q_\tau|X=0]) + \Pr[Y \le q_\tau|X=0]$.

[5]Such property is important in some situations, although for the marginal effects in which we are interested in this paper the recentering is not fundamental. In Firpo, Fortin, and Lemieux (2007b), the recentering is useful because it allows us to identify the intercept and perform Oaxaca-type decompositions at various quantiles.

[6]See Firpo, Fortin, and Lemieux (2007b) for such regressions on the variance and Gini.

[7]In the case of the mean, since the RIF is simply the outcome variable $Y$, a regression of $\mathrm{RIF}(Y;\mu)$ on $X$ is the same as an OLS regression of $Y$ on $X$.

computed by estimating the sample quantile $q_\tau$, estimating the density $f_Y(q_\tau)$ at that point $q_\tau$ using kernel (or other) methods, and forming a dummy variable $\mathbb{1}\{Y \leq q_\tau\}$, indicating whether the value of the outcome variable is below $q_\tau$. Then we can simply run an OLS regression of this new dependent variable on the covariates, although we suggest more sophisticated estimation methods in Section 3.

We view our approach as an important complement to the literature concerned with the estimation of quantile functions. However, unlike Imbens and Newey (2009), Chesher (2003), and Florens, Heckman, Meghir, and Vytlacil (2008), who considered the identification of structural functions defined from conditional quantile restrictions in the presence of endogenous regressors, our approach is concerned solely with parameters that capture changes in unconditional quantiles in the presence of exogenous regressors.

The structure of the paper is as follows. In the next section, we define the key object of interest, the "unconditional quantile partial effect" (UQPE) and show how RIF regressions for the quantile can be used to estimate the UQPE. We also link this parameter to the structural parameters of a general model and the conditional quantile partial effects (CQPE). The estimation issues are addressed in Section 3. Section 4 presents an empirical application of our method that illustrates well the difference between our method and conditional quantiles regressions. We conclude in Section 5.

## 2. UNCONDITIONAL PARTIAL EFFECTS

### 2.1. *General Concepts*

We assume that $Y$ is observed in the presence of covariates $X$, so that $Y$ and $X$ have a joint distribution, $F_{Y,X}(\cdot,\cdot): \mathbb{R} \times \mathcal{X} \to [0,1]$, and $\mathcal{X} \subset \mathbb{R}^k$ is the support of $X$. By analogy with a standard regression coefficient, our object of interest is the effect of a small increase in the location of the distribution of the explanatory variable $X$ on the $\tau$th quantile of the unconditional distribution of $Y$. We represent this small location shift in the distribution of $X$ in terms of the counterfactual distribution $G_X(x)$. By definition, the unconditional (marginal) distribution function of $Y$ can be written as

$$F_Y(y) = \int F_{Y|X}(y|X=x) \cdot dF_X(x). \tag{1}$$

Under the assumption that the conditional distribution $F_{Y|X}(\cdot)$ is unaffected by this small manipulation of the distribution of $X$, a counterfactual distribution of $Y$, $G_Y^*$, can be obtained by replacing $F_X(x)$ with $G_X(x)$[8]:

$$G_Y^*(y) \equiv \int F_{Y|X}(y|X=x) \cdot dG_X(x). \tag{2}$$

Our regression method builds on some elementary properties of the influence function, a measure introduced by Hampel (1968, 1974) to study the infinitesimal behavior of real-valued functionals $\nu(F_Y)$, where $\nu : \mathcal{F}_\nu \to \mathbb{R}$, and where $\mathcal{F}_\nu$ is a class of distribution functions such that $F_Y \in \mathcal{F}_\nu$ if $|\nu(F)| < +\infty$. Let $G_Y$ be another distribution in the same class. Let $F_{Y,t\cdot G_Y} \in \mathcal{F}_\nu$ represent the mixing distribution, which is $t$ away from $F_Y$ in the direction of the probability distribution $G_Y$: $F_{Y,t\cdot G_Y} = (1-t)\cdot F_Y + t\cdot G_Y = t\cdot(G_Y - F_Y) + F_Y$, where $0 \le t \le 1$. The directional derivative of $\nu$ in the direction of the distribution $G_Y$ can be written as

$$\begin{aligned}\lim_{t\downarrow 0} \frac{\nu(F_{Y,t\cdot G_Y}) - \nu(F_Y)}{t} &= \left.\frac{\partial \nu(F_{Y,t\cdot G_Y})}{\partial t}\right|_{t=0} \\ &= \int \mathrm{IF}(y;\nu,F_Y)\cdot d(G_Y - F_Y)(y),\end{aligned} \tag{3}$$

where $\mathrm{IF}(y;\nu,F_Y) = \partial\nu(F_{Y,t\cdot\Delta_y})/\partial t|_{t=0}$, with $\Delta_y$ denoting the probability measure that puts mass 1 at the value $y$. The von Mises (1947) linear approximation of the functional $\nu(F_{Y,t\cdot G_Y})$ is

$$\begin{aligned}\nu(F_{Y,t\cdot G_Y}) = \nu(F_Y) &+ t\cdot \int \mathrm{IF}(y;\nu,F_Y)\cdot d(G_Y - F_Y)(y) \\ &+ r(t;\nu;G_Y,F_Y),\end{aligned}$$

where $r(t;\nu;G_Y,F_Y)$ is a remainder term. We define the *recentered influence function* (RIF) more formally as the leading terms of the above expansion for the particular case where $G_Y = \Delta_y$ and $t=1$. Since $\int \mathrm{IF}(y;\nu,F_Y)\cdot dF_Y(y) = 0$ by definition, it follows that

$$\begin{aligned}\mathrm{RIF}(y;\nu,F_Y) &= \nu(F_Y) + \int \mathrm{IF}(s;\nu,F_Y)\cdot d\Delta_y(s) \\ &= \nu(F_Y) + \mathrm{IF}(y;\nu,F_Y).\end{aligned}$$

Finally, note that the last equality in equation (3) also holds for $\mathrm{RIF}(y;\nu,F_Y)$.

In the presence of covariates $X$, we can use the law of iterated expectations to express $\nu(F_Y)$ in terms of the conditional expectation of $\mathrm{RIF}(y;\nu,F_Y)$

[8]Instead of assuming a constant conditional distribution $F_{Y|X}(\cdot|\cdot)$, we could allow the conditional distributions to vary as long as they converge as the marginal distributions of $X$ converge to one another.

given $X$:

$$\begin{aligned}(4)\qquad \nu(F_Y) &= \int \mathrm{RIF}(y;\nu,F_Y)\cdot dF_Y(y)\\ &= \int\int \mathrm{RIF}(y;\nu,F_Y)\cdot dF_{Y|X}(y|X=x)\cdot dF_X(x)\\ &= \int E[\mathrm{RIF}(Y;\nu,F_Y)|X=x]\cdot dF_X(x),\end{aligned}$$

where the first equality follows from the fact that the influence function integrates to zero, and the second equality comes from substituting in equation (1).

Equation (4) shows that when we are interested in the impact of covariates on a specific distributional statistic $\nu(F_Y)$ such as a quantile, we simply need to integrate over $E[\mathrm{RIF}(Y;\nu,F_Y)|X]$, which is easily done using regression methods. By contrast, in equation (1) we need to integrate over the whole conditional distribution $F_{Y|X}(y|X=x)$, which is, in general, more difficult to estimate.[9]

We now state our main result on how the impact of a marginal change in the distribution of $X$ on $\nu(F_Y)$ can be obtained using the conditional expectation of the $\mathrm{RIF}(Y;\nu,F_Y)$. Note that all proofs are provided in the Appendix.

THEOREM 1—Marginal Effect of a Change in the Distribution of $X$: *Suppose we can induce a small perturbation in the distribution of covariates, from $F_X$ in the direction of $G_X$, maintaining the conditional distribution of $Y$ given $X$ unaffected. The marginal effect of this distributional change on the functional $\nu(F_Y)$ is given by integrating up the conditional expectation of the (recentered) influence function with respect to the changes in distribution of the covariates $d(G_X - F_X)$:*

$$\left.\frac{\partial \nu(F_{Y,t\cdot G_Y^*})}{\partial t}\right|_{t=0} = \int E[\mathrm{RIF}(Y;\nu,F_Y)|X=x]\cdot d(G_X-F_X)(x),$$

*where $F_{Y,t\cdot G_Y^*} = (1-t)\cdot F_Y + t\cdot G_Y^*$.*

We next consider a particular change, a small location shift $t$, in the distribution of covariates $X$. Let $X_j$ be a continuous covariate in the vector $X$, where

[9]Most other approaches, such as the conditional quantile regression method of Machado and Mata (2005), have essentially proposed to estimate and integrate the whole conditional distribution, $F_{Y|X}(y|X=x)$ over a new distribution $G_X$ of $X$ to obtain the counterfactual unconditional distribution of $Y$. See also Albrecht, Björklund, and Vroman (2003) and Melly (2005). By contrast, we show in Section 3 that our approach requires estimating the conditional distribution $F_{Y|X}(y|X=x)=\Pr[Y>y|X=x]$ only at one point of the distribution. Note that these approaches do not generate a marginal effect parameter, but instead a total effect of changes in the distribution of $X$ on selected features (e.g., quantiles) of the unconditional distribution of $Y$.

$1 \leq j \leq k$. The new distribution $G_X$ will be the distribution of a random $k \times 1$ vector $Z$, where $Z_l = X_l$ for $l \neq j$ and $l = 1, \ldots, k$, and $Z_j = X_j + t$. In this special case, let $\alpha_j(\nu)$ denote the partial effect of a small change in the distribution of covariates from $F_X$ to $G_X$ on the functional $\nu(F_Y)$. Collecting all $j$ entries, we construct the $k \times 1$ vector $\alpha(\nu) = [\alpha_j(\nu)]_{j=1}^k$. We can write the unconditional partial effect $\alpha(\nu)$ as an average derivative.

COROLLARY 1—Unconditional Partial Effect: *Assume that $d\mathcal{X}$, the boundary of the support $\mathcal{X}$ of $X$, is such that if $x \in d\mathcal{X}$, then $f_X(x) = 0$. Then the vector $\alpha(\nu)$ of partial effects of small location shifts in the distribution of a continuous covariate $X$ on $\nu(F_Y)$ can be written using the vector of average derivatives*[10]

$$\alpha(\nu) = \int \frac{dE[\mathrm{RIF}(Y;\nu)|X=x]}{dx} \cdot dF(x). \tag{5}$$

### 2.2. *The Case of Quantiles*

Turning to the specific case of quantiles, consider the $\tau$th quantile $q_\tau = \nu_\tau(F_Y) = \inf_q\{q : F_Y(q) \geq \tau\}$. It follows from the definition of the influence function that

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

where $c_{1,\tau} = 1/f_Y(q_\tau)$, $c_{2,\tau} = q_\tau - c_{1,\tau} \cdot (1-\tau)$, and $f_Y(q_\tau)$ is the density of $Y$ evaluated at $q_\tau$. Thus

$$E[\mathrm{RIF}(Y;q_\tau)|X=x] = c_{1,\tau} \cdot \Pr[Y > q_\tau|X=x] + c_{2,\tau}.$$

From equation (5), the unconditional partial effect, that we denote $\alpha(\tau)$ in the case of the $\tau$th quantile, simplifies to

$$\alpha(\tau) = \left.\frac{\partial \nu_\tau(F_{Y,t\cdot G_Y^*})}{\partial t}\right|_{t=0} = c_{1,\tau} \cdot \int \frac{d\Pr[Y > q_\tau|X=x]}{dx} \cdot dF_X(x), \tag{6}$$

where the last term is the average marginal effect from the probability response model $\Pr[Y > q_\tau|X]$. We call the parameter $\alpha(\tau) = E[dE[\mathrm{RIF}(Y, q_\tau)|X]/dx]$ the *unconditional quantile partial effect* (UQPE), by analogy with the Wooldridge (2004) unconditional average partial effect (UAPE), which is defined as $E[dE[Y|X]/dx]$.[11]

[10]The expression $dE[\mathrm{RIF}(Y;\nu)|X=x]/dx$ is the $k$ vector of partial derivatives $[\partial E[\mathrm{RIF}(Y;\nu)|X=x]/\partial x_j]_{j=1}^k$.

[11]The UAPE is a special case of Corollary 1 for the mean ($\nu = \mu$), where $\alpha(\mu) = E[dE[Y|X]/dx]$ since $\mathrm{RIF}(Y, \mu) = y$.

Our next result provides an interpretation of the UQPE in terms of a general structural model, $Y = h(X, \varepsilon)$, where the unknown mapping $h(\cdot, \cdot)$ is invertible on the second argument, and $\varepsilon$ is an unobservable determinant of the outcome variable $Y$. We also show that the UQPE can be written as a weighted average of a family of conditional quantile partial effects (CQPE), which is the effect of a small change of $X$ on the conditional quantile of $Y$:

$$\mathrm{CQPE}(\tau, x) = \frac{\partial Q_\tau[h(X, \varepsilon)|X = x]}{\partial x} = \frac{\partial h(x, Q_\tau[\varepsilon])}{\partial x},$$

where $Q_\tau[Y|X = x] \equiv \inf_q\{q : F_{Y|X}(q|x) \geq \tau\}$ is the conditional quantile operator. For the sake of simplicity and comparability between the CQPE and the UQPE, we consider the case where $\varepsilon$ and $X$ are independent. Thus, we can use the unconditional form for $Q_\tau[\varepsilon]$ in the last term of the above equation.[12]

In a linear model $Y = h(X, \varepsilon) = X^\intercal\beta + \varepsilon$, both the UQPE and the CQPE are trivially equal to the parameter $\beta_j$ of the structural form for any quantile. While this specific result does not generalize beyond the linear model, useful connections can still be drawn between the UQPE and the underlying structural form, and between the UQPE and the CQPE. To establish these connections, we define three auxiliary functions. The first function, $\omega_\tau : \mathcal{X} \to \mathbb{R}^+$, is a weighting function defined as the ratio between the conditional density given $X = x$ and the unconditional density: $\omega_\tau(x) \equiv f_{Y|X}(q_\tau|x)/f_Y(q_\tau)$. The second function, $\varepsilon_\tau : \mathcal{X} \to \mathbb{R}$, is the inverse function $h^{-1}(\cdot, q_\tau)$, which shall exist under the assumption that $h$ is strictly monotonic in $\varepsilon$. The third function, $\zeta_\tau : \mathcal{X} \to (0, 1)$, is a "matching" function indicating where the unconditional quantile $q_\tau$ falls in the conditional distribution of $Y$:

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

PROPOSITION 1—UQPE and the Structural Form:

(i) *Assuming that the structural form* $Y = h(X, \varepsilon)$ *is strictly monotonic in* $\varepsilon$ *and that* $X$ *and* $\varepsilon$ *are independent*, *the parameter* UQPE$(\tau)$ *will be*

$$\mathrm{UQPE}(\tau) = E\left[\omega_\tau(X) \cdot \frac{\partial h(X, \varepsilon_\tau(X))}{\partial x}\right]$$

(ii) *We can also represent* UQPE$(\tau)$ *as a weighted average of* CQPE$(\zeta_\tau(x), x)$:

$$\mathrm{UQPE}(\tau) = E\big[\omega_\tau(X) \cdot \mathrm{CQPE}(\zeta_\tau(X), X)\big].$$

[12]In this setting, the identification of the UQPE requires $F_{\varepsilon|X}$ to be unaffected by changes in the distribution of covariates. The identification of the CQPE requires quantile independence between $\varepsilon$ and $X$, that is, the $\tau$-conditional quantile of $\varepsilon$ given $X$ equals the $\tau$-unconditional quantile of $\varepsilon$. Independence between $\varepsilon$ and $X$ guarantees, therefore, that both the UQPE and the CQPE parameters are identified.

Result (i) of the Proposition 1 shows formally that $\mathrm{UQPE}(\tau)$ is equal to a weighted average (over the distribution of $X$) of the partial derivatives of the structural function. In the simple case of the linear model mentioned above, it follows that $\partial h(X, \varepsilon_\tau(X))/\partial x = \beta$ and $\mathrm{UQPE}(\tau) = \beta$ for all $\tau$. More generally, the UQPE will typically depend on $\tau$ in nonlinear settings. For example, when $h(X, \varepsilon) = \tilde{h}(X^\intercal \beta + \varepsilon)$, where $\tilde{h}$ is differentiable and strictly monotonic, simple algebra yields $\mathrm{UQPE}(\tau) = \beta \cdot \tilde{h}'(\tilde{h}^{-1}(q_\tau))$, which depends on $\tau$. Finally, note that independence plays a crucial role here. If, instead, we had dropped the independence assumption between $\varepsilon$ and $X$, we would not be able, even in a linear model, to express $\mathrm{UQPE}(\tau)$ as a simple function of the structural parameter $\beta$.[13]

Result (ii) shows that $\mathrm{UQPE}(\tau)$ is a weighted average (over the distribution of $X$) of a family of $\mathrm{CQPE}(\zeta_\tau(X), X)$ at $\zeta_\tau(X)$, the conditional quantile corresponding to the $\tau$th unconditional quantile of the distribution of $Y$, $q_\tau$. But while result (ii) of Proposition 1 provides a more structural interpretation of the UQPE, it is not practical from an estimation point of view as it would require estimating $h$ and $F_\varepsilon$, the distribution of $\varepsilon$, using nonparametric methods. As shown below, we propose a simpler way to estimate the UQPE based on the estimation of average marginal effects.

## 3. ESTIMATION

In this section, we discuss the estimation of $\mathrm{UQPE}(\tau)$ using RIF regressions. Equation (6) shows that three components are involved in the estimation of $\mathrm{UQPE}(\tau)$: the quantile $q_\tau$, the density of the unconditional distribution of $Y$ that appears in the constant $c_{1,\tau} = 1/f_Y(q_\tau)$, and the average marginal effect $E[d\Pr[Y > q_\tau|X]/dX]$. We discuss the estimation of each component in turn and then briefly address the asymptotic properties of related estimators.

The estimator of the $\tau$th population quantile of the marginal distribution of $Y$ is $\widehat{q}_\tau$, the usual $\tau$th sample quantile, which can be represented, as in Koenker and Bassett (1978), as

$$\widehat{q}_\tau = \arg\min_q \sum_{i=1}^{N} (\tau - \mathbb{1}\{Y_i - q \leq 0\}) \cdot (Y_i - q).$$

[13] These examples are worked in detail in the working paper version of this article. See Firpo, Fortin, and Lemieux (2007a).

We estimate the density of $Y$, $\widehat{f}_Y(\cdot)$, using the kernel density estimator[14]

$$\widehat{f}_Y(\widehat{q}_\tau) = \frac{1}{N \cdot b} \cdot \sum_{i=1}^{N} \mathcal{K}_Y\left(\frac{Y_i - \widehat{q}_\tau}{b}\right),$$

where $\mathcal{K}_Y(z)$ is a kernel function and $b$ a positive scalar bandwidth.

We suggest three estimation methods for the UQPE based on three ways, among many, to estimate the average marginal effect $E[d\Pr[Y > q_\tau|X]/dX]$. As discussed in Firpo, Fortin, and Lemieux (2009), the first two estimators will be consistent if we correctly impose functional form restrictions. The third estimator involves a fully nonparametric first stage and, therefore, will be consistent quite generally for the average derivative parameter.

The first method estimates the average marginal effect $E[d\Pr[Y > q_\tau| X]/dX]$ with an OLS regression, which provides consistent estimates if $\Pr[Y > q_\tau|X = x]$ is linear in $x$. This method, that we call RIF-OLS, consists of regressing $\widehat{\mathrm{RIF}}(Y; \widehat{q}_\tau) = \widehat{c}_{1,\tau} \cdot \mathbb{1}\{Y > \widehat{q}_\tau\} + \widehat{c}_{2,\tau}$ on $X$. The second method uses a logistic regression of $\mathbb{1}\{Y > \widehat{q}_\tau\}$ on $X$ to estimate the average marginal effect, which is then multiplied by $\widehat{c}_{1,\tau}$. Again, the average marginal effect from this logit model will be consistent if $\Pr[Y > q_\tau|X = x] = \Lambda(x^\top \theta_\tau)$, where $\Lambda(\cdot)$ is the cumulative distribution function (c.d.f.) of a logistic distribution and $\theta_\tau$ is a vector of coefficients. We call this method RIF-Logit. In the empirical section, we use these two estimators and find that, in our application, they yield estimates very close to the fully nonparametric estimator.

The last estimation method, called RIF-NP, is based on a nonparametric estimator that does not require any functional form assumption on $\Pr[Y > q_\tau|X = x]$ to be consistent. We use the method discussed by Newey (1994) and estimate $\Pr[Y > q_\tau|X = x]$ by polynomial series. As the object of interest is the average of $d\Pr[Y > q_\tau|X = x]/dx$, once we have a polynomial function that approximates the conditional probability, we can easily take derivatives of polynomials and average them. As shown by Stoker (1991) for the average derivative case and later formalized in a more general setting by Newey (1994), the choice of the nonparametric estimator for the derivative is not crucial in large samples. Averaging any regular nonparametric estimator with respect to $X$ yields an estimator that converges at the usual parametric rate and has the same limiting distribution as other estimators based on different nonparametric methods.[15]

[14]In the empirical section we propose using the Gaussian kernel. The requirements for the kernel and the bandwidth are described in Firpo, Fortin, and Lemieux (2009). We propose using the kernel density estimator, but other consistent estimators of the density could be used as well.

[15]Nonparametric estimation of $\Pr[Y > q_\tau|X = x]$ could also be performed by series approximation of the log-odds ratio, which would keep predictions between 0 and 1 (Hirano, Imbens, and Ridder (2003)). Note, however, that we are mainly interested in another object, the derivative $d\Pr[Y > q_\tau|X = x]/dx$, and imposing that the conditional probability lies in the unit interval does not necessarily add much structure to its derivative.

We study the asymptotic properties of our estimators in detail in Firpo, Fortin, and Lemieux (2009), where we establish the limiting distributions of these estimators, discuss how to estimate their asymptotic variances, and show how to construct test statistics. Three important results from Firpo, Fortin, and Lemieux (2009) are summarized here. The first result is that the asymptotic linear expression of each one of the three estimators consists of three components. The first component is associated with uncertainty regarding the density; the second component is associated with the uncertainty regarding the population quantile; and the third component is associated with the average derivative term $E[d\Pr[Y > q_\tau|X]/dX]$. The second result states that because the density is nonparametrically estimated by kernel methods, the rate of convergence of the three estimators will be dominated by this slower term. In Firpo, Fortin, and Lemieux (2009), we use a higher order expansion type of argument to allow for the quantile and the average derivative components to be explicitly included. By doing so, we can introduce a refinement in the expression of the asymptotic variance. Finally, the third result is that to test the null hypothesis that UQPE $= 0$, we do not need to estimate the density, as $E[d\Pr[Y > q_\tau|X]/dX] = 0 \Leftrightarrow$ UQPE $= 0$. Thus, we can use test statistics that converge at the parametric rate. In this case, the only components that contribute to the asymptotic variance are the quantile and the average derivative.

As with standard average marginal effects, we can also estimate the UQPE for a dummy covariate by estimating $E[\Pr[Y > q_\tau|X = 1]] - E[\Pr[Y > q_\tau|X = 0]]$ instead of $E[d\Pr[Y > q_\tau|X]/dX]$ using any of the three methods discussed above. Like in the example of union status mentioned in the Introduction, the UQPE in such cases represents the impact of a small change in the probability $p = \Pr[X = 1]$, instead of the small location shift for a continuous covariate considered in Section 2.[16]

## 4. EMPIRICAL APPLICATION

In this section, we present an empirical application to illustrate how the unconditional quantile regressions work in practice using the three estimators discussed above.[17] We also show how the results compare to standard (conditional) quantile regressions. Our application considers the direct effect of union status on male log wages, which is well known to be different at different points of the wage distribution.[18] We use a large sample of 266,956 observa-

[16]See Firpo, Fortin, and Lemieux (2007a) for more detail.

[17]A Stata ado file that implements the RIF-OLS estimator is available on the author's website, http://www.econ.ubc.ca/nfortin/.

[18]See, for example, Chamberlain (1994) and Card (1996). For simplicity, we maintain the assumption that union coverage status is exogenous. Studies that have used selection models or longitudinal methods to allow the union status to be endogenously determined (e.g., Lemieux (1998)) suggest that the exogeneity assumption only introduces small biases in the estimation.

tions on U.S. males from the 1983–1985 Outgoing Rotation Group (ORG) supplement of the Current Population Survey.[19]

Looking at the impact of union status on log wages illustrates well the difference between conditional and unconditional quantiles regressions. Consider, for example, the effect of union status estimated at the 90th and 10th quantiles. Finding that the effect of unions (for short) estimated using conditional quantile regressions is smaller at the 90th than at the 10th quantile simply means that unions reduce within-group dispersion, where the "group" consists of workers who share the same values of the covariates $X$ (other than union status). This does not mean, however, that increasing the rate of unionization would reduce overall wage dispersion as measured by the difference between the 90th and the 10th quantiles of the unconditional wage dispersion. To answer this question we have to turn to unconditional quantile regressions.

In addition to the within-group wage compression effect captured by conditional quantile regressions, unconditional quantile regressions also capture an inequality-enhancing between-group effect linked to the fact that unions increase the conditional mean of wages of union workers. This creates a wedge between otherwise comparable union and non-union workers.[20] As a result, unions tend to increase wages for low wage quantiles where both the between- and within-group effects go in the same direction, but can decrease wages for high wage quantiles where the between- and within-group effects go in opposite directions.

As a benchmark, Table I reports the RIF-OLS estimated coefficients of the log wages model for the 10th, 50th, and 90th quantiles. The results (labeled UQR for unconditional quantile regressions) are compared with standard OLS (conditional mean) estimates and with standard (conditional) quantile regressions (CQR) at the corresponding quantiles. For the sake of comparability, we use simple linear specifications for all estimated models. We also show in Figure 1 how the estimated UQPE of unions changes when we use the RIF-Logit and RIF-NP methods instead.

Interestingly, the UQPE of unions first increases from 0.198 at the 10th quantile to 0.349 at the median, before turning negative ($-0.137$) at the 90th quantile. These findings strongly confirm the point discussed above that unions

[19]We start with 1983 because it is the first year in which the ORG supplement asked about union status. The dependent variable is the real log hourly wage for all wage and salary workers, and the explanatory variables include six education classes, married, non-white, and nine experience classes. The hourly wage is measured directly for workers paid by the hour and is obtained by dividing usual earnings by usual hours of work for other workers. Other data processing details can be found in Lemieux (2006).

[20]In the case of the variance, it is easy to write down an analytical expression for the between- and within-group effects (see, for example, Card, Lemieux, and Riddell (2004)) and find the conditions under which one effect dominates the other. It is much harder to ascertain, however, whether the between- or the within-group effect tends to dominate at different points of the wage distribution.

TABLE I

COMPARING OLS, UNCONDITIONAL QUANTILE REGRESSIONS (UQR), AND CONDITIONAL QUANTILE REGRESSIONS (CQR); 1983–1985 CPS DATA FOR MEN[a]

| | | 10th Centile | | 50th Centile | | 90th Centile | |
|---|---|---|---|---|---|---|---|
| | OLS | UQR | CQR | UQR | CQR | UQR | CQR |
| Union status | 0.179 | 0.195 | 0.288 | 0.337 | 0.195 | −0.135 | 0.088 |
| | (0.002) | (0.002) | (0.003) | (0.004) | (0.002) | (0.004) | (0.003) |
| Non-white | −0.134 | −0.116 | −0.139 | −0.163 | −0.134 | −0.099 | −0.120 |
| | (0.003) | (0.005) | (0.004) | (0.004) | (0.003) | (0.005) | (0.005) |
| Married | 0.140 | 0.195 | 0.166 | 0.156 | 0.146 | 0.043 | 0.089 |
| | (0.002) | (0.004) | (0.004) | (0.003) | (0.002) | (0.004) | (0.003) |
| Education | | | | | | | |
| Elementary | −0.351 | −0.307 | −0.279 | −0.452 | −0.374 | −0.240 | −0.357 |
| | (0.004) | (0.009) | (0.006) | (0.006) | (0.005) | (0.005) | (0.007) |
| HS dropout | −0.190 | −0.344 | −0.127 | −0.195 | −0.205 | −0.068 | −0.227 |
| | (0.003) | (0.007) | (0.004) | (0.004) | (0.003) | (0.003) | (0.005) |
| Some college | 0.133 | 0.058 | 0.058 | 0.179 | 0.133 | 0.154 | 0.172 |
| | (0.002) | (0.004) | (0.003) | (0.004) | (0.003) | (0.005) | (0.004) |
| College | 0.406 | 0.196 | 0.252 | 0.464 | 0.414 | 0.582 | 0.548 |
| | (0.003) | (0.004) | (0.005) | (0.005) | (0.004) | (0.008) | (0.006) |
| Post-graduate | 0.478 | 0.138 | 0.287 | 0.522 | 0.482 | 0.844 | 0.668 |
| | (0.004) | (0.004) | (0.007) | (0.005) | (0.004) | (0.012) | (0.006) |
| Constant | 1.742 | 0.970 | 1.145 | 1.735 | 1.744 | 2.511 | 2.332 |
| | (0.004) | (0.005) | (0.006) | (0.006) | (0.004) | (0.008) | (0.005) |

[a]Robust standard errors (OLS) and bootstrapped standard errors (200 replications) for UQR and CQR are given in parentheses. All regressions also include a set of dummies for labor market experience categories.

have different effects at different points of the wage distribution.[21] The conditional quantile regression estimates reported in the corresponding columns show, as in Chamberlain (1994), that unions shift the location of the conditional wage distribution (i.e., positive effect on the median) but also reduce conditional wage dispersion.

The difference between the estimated effect of unions for conditional and unconditional quantile regression estimates is illustrated in more detail in panel A of Figure 1, which plots both conditional and unconditional quantile regression estimates of union status at 19 different quantiles (from the 5th to the 95th).[22] As indicated in Table I, the unconditional union effect is highly nonmonotonic, while the conditional effect declines monotonically. More precisely, the unconditional effect first increases from about 0.1 at the 5th quantile to about 0.4 at the 35th quantile, before declining and eventually reaching

[21]Note that the effects are very precisely estimated for all specifications and the $R$-squared (close to 0.40) are sizeable for cross-sectional data.

[22]Bootstrapped standard errors are provided for both estimates. Analytical standard errors for the UQPE are nontrivial and derived in Firpo, Fortin, and Lemieux (2009).

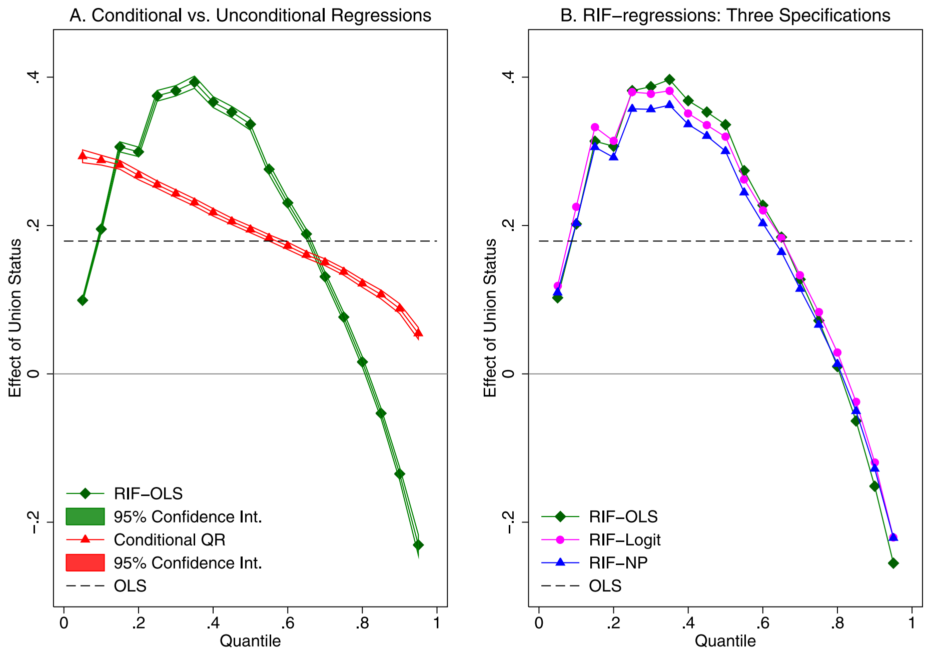


FIGURE 1.—Unconditional and conditional quantile regression estimates of the effect of union status on log wages.

a large negative effect of over −0.2 at the 95th quantile. By contrast, standard (conditional) quantile regression estimates decline almost linearly from about 0.3 at the 5th quantile to barely more than 0 at the 95th quantile.

At first glance, the fact that the effect of unions is uniformly positive for conditional quantile regressions, but negative above the 80th quantile for unconditional quantile regressions may seem puzzling. Since Proposition 1 states that the UQPE is a weighted average of the CQPEs, for the UQPE to be negative it must be that some of the CQPEs are negative too. Unlike the UQPE, however, the CQPE generally depends on $X$. For the sake of clarity, in Figure 1 we report the conditional quantile regressions using a highly restricted specification where the effect of unions is not allowed to depend on a rich set of other covariates (no interaction terms). When we relax this assumption, we find that conditional quantile regressions estimates are often negative for more "skilled" workers (in high education/high labor market experience cells). However, these negative effects are averaged away by positive effects in the more parsimonious conditional quantile regressions. On the other hand, because the matching function $\zeta_\tau(x)$ from Proposition 1 reassigns some of the negative union effects from the $s$-conditional quantiles to the $\tau$-unconditional quantiles at the top of the wage distribution and because the weighting function $\omega_\tau(x)$ puts more weight on these workers, the UQPE becomes negative for workers at the top end of the wage distribution.

Panel B shows that the RIF-OLS and RIF-Logit estimates of the UQPE are very similar, which confirms the "folk wisdom" in empirical work that, in many instances, using a linear probability model or a logit gives very similar average marginal effects. More importantly, Figure 1 shows that the RIF-NP estimates are also very similar to the estimates obtained using these two simpler methods.[23] This suggests that, at least for this particular application, using a simple linear specification for the unconditional quantile regressions provides fairly accurate estimates of the UQPE. The small difference between RIF-OLS and RIF-NP estimates stands in sharp contrast to the large differences between the RIF-OLS estimates and the conditional quantile regression estimates in panel A.

The large differences between the conditional and unconditional quantile regressions results have important implications for understanding recent changes in wage inequality. There is a long tradition in labor economics of attempting to estimate the effect of unionization on the (unconditional) distribution of wages.[24] The unconditional quantile regressions provide a simple and direct way to estimate this effect at all points of the distribution. The estimates reported in Figure 1 show that unionization progressively increases wages in the three lower quintiles of the distribution, peaking around the 35th quantile, and actually reduces wages in the top quintile of the distribution. As a result, the decline in unionization over the last three decades should have contributed to a reduction in wage inequality at the bottom end of the distribution and to an increase in wage inequality at the top end. This precisely mirrors the actual U-shaped changes observed in the data.[25] By contrast, conditional quantile regressions results describe a positive but monotonically declining effect of unionization on wages, which fails to account for the observed pattern of changes in the wage distribution.

## 5. CONCLUSION

In this paper, we propose a new regression method to estimate the effect of explanatory variables on the unconditional quantiles of an outcome variable. The proposed unconditional quantile regression method consists of running a regression of the (recentered) influence function of the unconditional quantile of the outcome variable on the explanatory variables. The influence

[23]The RIF-NP is estimated using a model fully saturated with all possible interactions (up to 432 parameters) of our categorical variables, omitting for each estimated quantile the interactions that would result in perfect predictions. For the RIF-OLS, the figure graphs the estimated coefficients, while for the RIF-Logit and RIF-NP, the average unconditional partial effects are displayed.

[24]See, for example, Card (1996) and DiNardo, Fortin, and Lemieux (1996).

[25]See, for example, Autor, Katz, and Kearney (2008) and Lemieux (2008).

function is a widely used tool in robust estimation that can easily be computed for each quantile of interest. We show how standard partial effects, that we call unconditional quantile partial effects (UQPE), can be estimated using our regression approach.

Another important advantage of the proposed method is that it can be easily generalized to other distributional statistics such as the Gini, the log variance, or the Theil coefficient. Once the recentered influence function for these statistics is computed, all that is required is running a regression of the resulting RIF on the covariates. We discuss in a companion paper (Firpo, Fortin, and Lemieux (2007b)) how our regression method can be used to generalize traditional Oaxaca–Blinder decompositions, devised for means, to other distributional statistics.

Finally, our method can be useful even when the independence assumption is relaxed. However, the interpretation of the identified parameter in terms of its relation to the structural function linking observed and unobserved factors to the dependent variable would change. Yet, the UQPE parameter would still be defined by holding unobserved variables and other components of $X$ fixed when evaluating the marginal effect of changes in the distribution of $X_j$ on a given quantile of the unconditional distribution of $Y$. Such structural averaged marginal effects can be useful in practice. We plan to show in future work how our approach can be used when instrumental variables are available for the endogenous covariates and how consistent estimates of marginal effects can be obtained by adding a control function in the unconditional quantile regressions.

## APPENDIX

PROOF OF THEOREM 1: The effect on the functional $\nu$ of the distribution of $Y$ of an infinitesimal change in the distribution of $X$ from $F_X$ toward $G_X$ is defined as $\partial\nu(F_{Y,t\cdot G_Y^*})/\partial t|_{t=0}$. Given that equation (3) also applies to RIF$(y;\nu)$, it follows that

$$\left.\frac{\partial\nu(F_{Y,t\cdot G_Y^*})}{\partial t}\right|_{t=0} = \int \mathrm{RIF}(y;\nu)\cdot d(G_Y^* - F_Y)(y).$$

Substituting in equations (1) and (2), and applying the fact that $E[\mathrm{RIF}(Y;\nu)|X=x] = \int_y \mathrm{RIF}(y;\nu)\cdot dF_{Y|X}(y|X=x)$ yields

$$\begin{aligned}\left.\frac{\partial\nu(F_{Y,t\cdot G_Y^*})}{\partial t}\right|_{t=0} &= \int\left(\int \mathrm{RIF}(y;\nu)\cdot dF_{Y|X}(y|X=x)\right)\cdot d(G_X - F_X)(x)\\ &= \int E[\mathrm{RIF}(Y;\nu)|X=x]\cdot d(G_X - F_X)(x).\end{aligned}$$

*Q.E.D.*

PROOF OF COROLLARY 1: Consider the distribution $\widetilde{G}_X(\cdot; t)$ of the random vector $Z = X + t_j$, where $t_j = t \cdot \mathbf{e}_j$ and $\mathbf{e}_j = [0, \dots, 0, 1, 0, \dots, 0]^\intercal$, which is a $k$ vector of zeros except at the $j$th entry, which equals 1. The density of $Z$ is $\widetilde{g}_X(x; t) = f_X(x - t_j)$.[26] The counterfactual distribution $\widetilde{G}^*_Y(\cdot; t)$ of $Y$ using $F_{Y|X}$ and $\widetilde{G}_X(\cdot; t)$ will be

$$\begin{aligned}\widetilde{G}^*_Y(y; t) &= \int F_{Y|X}(y|x) \cdot f_X(x - t_j) \cdot dx \\ &= \int F_{Y|X}(y|x) \cdot f_X(x) \cdot dx \\ &\quad - t \cdot \int F_{Y|X}(y|x) \cdot \frac{\partial f_X(x)/\partial x_j}{f_X(x)} \cdot f_X(x) \cdot dx + \chi_t \\ &= F_Y(y) + t \cdot \int F_{Y|X}(y|x) \cdot \mathbf{e}_j^\intercal \cdot l_X(x) \cdot f_X(x) \cdot dx + \chi_t,\end{aligned}$$

where the second line is obtained using a first-order expansion, where $l_X(x) = -d\ln(f_X(x))/dx = -f'_X(x)/f_X(x)$, and $f'_X(x) = [\partial f_X(x)/\partial x_l]_{l=1}^k$ is the $k$ vector of partial derivatives of $f_X(x)$. Therefore, $\chi_t = O(t^2)$. Now, define

$$g_X(x) = f_X(x) \cdot (1 + \mathbf{e}_j^\intercal \cdot l_X(x)) \quad \text{and} \quad G_X(x) = \int^x g_X(\xi) \cdot d\xi.$$

By the usual definition of the counterfactual distribution $G^*_Y$ of $Y$ using $F_{Y|X}$ and $G_X$, we have

$$\begin{aligned}G^*_Y(y) &= \int F_{Y|X}(y|x) \cdot g_X(x) \cdot dx \\ &= F_Y(y) + \int F_{Y|X}(y|x) \cdot \mathbf{e}_j^\intercal \cdot l_X(x) \cdot f_X(x) \cdot dx.\end{aligned}$$

Thus we can write

$$\widetilde{G}^*_Y(y; t) = F_Y(y) + t \cdot (G^*_Y(y) - F_Y(y)) + \chi_t = F_{Y,t\cdot G^*_Y} + \chi_t.$$

Hence,

$$\begin{aligned}\alpha_j(\nu) &\equiv \lim_{t\downarrow 0} \frac{\nu(\widetilde{G}^*_Y(\cdot; t)) - \nu(F_Y)}{t} \\ &= \lim_{t\downarrow 0}\left(\frac{\nu(F_{Y,t\cdot G^*_Y}) - \nu(F_Y)}{t}\right) + \lim_{t\downarrow 0}\left(\frac{\nu(\widetilde{G}^*_Y(\cdot; t)) - \nu(F_{Y,t\cdot G^*_Y})}{t}\right) \\ &= \left.\frac{\partial \nu(F_{Y,t\cdot G^*_Y})}{\partial t}\right|_{t=0} + \lim_{t\downarrow 0}\left(\frac{\nu(F_{Y,t\cdot G^*_Y} + \chi_t) - \nu(F_{Y,t\cdot G^*_Y})}{t}\right),\end{aligned}$$

[26]The density of $X$ is $f_X(\cdot)$ and, by definition of densities, $\int^x f_X(\xi) \cdot d\xi = F_X(x)$.

where the last term vanishes:

$$\lim_{t\downarrow 0}\left(\frac{\nu(F_{Y,t\cdot G_Y^*}+\chi_t)-\nu(F_{Y,t\cdot G_Y^*})}{t}\right)=\lim_{t\downarrow 0}\left(\frac{O(\chi_t)}{t}\right)$$
$$=\lim_{t\downarrow 0}\big(O(|t|)\big)=O(1)\cdot\lim_{t\downarrow 0}t.$$

Using Theorem 1, it follows that

$$\alpha_j(\nu)=\frac{\partial\nu(F_{Y,t\cdot G_Y^*})}{\partial t}\bigg|_{t=0}=\int E[\mathrm{RIF}(Y;\nu)|X=x]\cdot d(G_X-F_X)(x)$$
$$=\int E[\mathrm{RIF}(Y;\nu)|X=x]\cdot \mathbf{e}_j^{\intercal}\cdot l_X(x)\cdot f_X(x)\cdot dx.$$

Applying partial integration and using the condition that $f_X(x)$ is zero at the boundary of the support yields

$$\mathbf{e}_j^{\intercal}\cdot\int E[\mathrm{RIF}(Y;\nu)|X=x]\cdot l_X(x)\cdot f_X(x)\cdot dx$$
$$=\int \mathbf{e}_j^{\intercal}\cdot\frac{dE[\mathrm{RIF}(Y;\nu)|X=x]}{dx}\cdot f_X(x)\cdot dx$$
$$=\int\frac{\partial E[\mathrm{RIF}(Y;\nu)|X=x]}{\partial x_j}\cdot f_X(x)\cdot dx.$$

Hence

$$\alpha_j(\nu)=\int\frac{\partial E[\mathrm{RIF}(Y;\nu,F)|X=x]}{\partial x_j}\cdot f_X(x)\cdot dx. \qquad \textit{Q.E.D.}$$

PROOF OF PROPOSITION 1:

(i) Starting from equation (6),

$$\mathrm{UQPE}(\tau)=-\frac{1}{f_Y(q_\tau)}\cdot\int\frac{d\Pr[Y\le q_\tau|X=x]}{dx}\cdot dF_X(x),$$

and assuming that the structural form $Y=h(X,\varepsilon)$ is monotonic in $\varepsilon$, so that $\varepsilon_\tau(x)=h^{-1}(x,q_\tau)$, we can write

$$\Pr[Y\le q_\tau|X=x]=\Pr[\varepsilon\le\varepsilon_\tau(X)|X=x]$$
$$=F_{\varepsilon|X}(\varepsilon_\tau(x)|x)=F_\varepsilon(\varepsilon_\tau(x)).$$

Taking the derivative with respect to $x$, we get

$$d\frac{\Pr[Y\le q_\tau|X=x]}{dx}=f_\varepsilon(\varepsilon_\tau(x))\cdot\frac{\partial h^{-1}(x,q_\tau)}{\partial x}.$$

Defining $H(x, \varepsilon_\tau(x), q_\tau) = h(x, \varepsilon_\tau(x)) - q_\tau$, it follows that

$$\frac{\partial h^{-1}(x, q_\tau)}{\partial x} = \frac{\partial \varepsilon_\tau(x)}{\partial x} = -\frac{\partial H(x, \varepsilon_\tau, q_\tau)/\partial x}{\partial H(x, \varepsilon_\tau, q_\tau)/\partial \varepsilon_\tau}$$

$$= -\frac{\partial h(x, \varepsilon_\tau)/\partial x}{\partial h(x, \varepsilon_\tau)/\partial \varepsilon_\tau} = -\frac{\partial h(x, \varepsilon_\tau)}{\partial x} \cdot \left( \left. \frac{\partial h(x, \varepsilon)}{\partial \varepsilon} \right|_{\varepsilon = \varepsilon_\tau} \right)^{-1}.$$

Similarly,

$$\frac{\partial h^{-1}(x, q_\tau)}{\partial q_\tau} = -\frac{\partial H(x, \varepsilon_\tau, q_\tau)/\partial q_\tau}{\partial H(x, \varepsilon_\tau, q_\tau)/\partial \varepsilon_\tau} = \left( \left. \frac{\partial h(x, \varepsilon)}{\partial \varepsilon} \right|_{\varepsilon = \varepsilon_\tau} \right)^{-1}.$$

Hence,

$$f_{Y|X}(q_\tau; x) = d\frac{\Pr[Y \leq q_\tau | X = x]}{dq_\tau} = d\frac{F_\varepsilon(h^{-1}(x, q_\tau))}{dq_\tau}$$

$$= f_\varepsilon(\varepsilon_\tau(x)) \cdot \frac{\partial h^{-1}(x, q_\tau)}{\partial q_\tau}$$

$$= \left( \left. \frac{\partial h(x, \varepsilon)}{\partial \varepsilon} \right|_{\varepsilon = \varepsilon_\tau} \right)^{-1} \cdot f_\varepsilon(\varepsilon_\tau(x)).$$

Substituting in these expressions yields

$$\begin{aligned}
&\mathrm{UQPE}(\tau) \\
&= -(f_Y(q_\tau))^{-1} \cdot \int d\frac{\Pr[Y \leq q_\tau | X = x]}{dx} \cdot dF_X(x) \\
&= (f_Y(q_\tau))^{-1} \\
&\quad \cdot \int \left( f_\varepsilon(\varepsilon_\tau(x)) \cdot \frac{\partial h(x, \varepsilon_\tau)}{\partial x} \cdot \left( \left. \frac{\partial h(x, \varepsilon)}{\partial \varepsilon} \right|_{\varepsilon = \varepsilon_\tau} \right)^{-1} \right) \cdot dF_X(x) \\
&= (f_Y(q_\tau))^{-1} \cdot E\left[ f_{Y|X}(q_\tau | X) \cdot \frac{\partial h(X, \varepsilon_\tau(X))}{\partial x} \right] \\
&= E\left[ \frac{f_{Y|X}(q_\tau | X)}{f_Y(q_\tau)} \cdot \frac{\partial h(X, \varepsilon_\tau(X))}{\partial x} \right] \\
&= E\left[ \omega_\tau(X) \cdot \frac{\partial h(X, \varepsilon_\tau(X))}{\partial x} \right].
\end{aligned}$$

(ii) Let the CQPE be defined as

$$\mathrm{CQPE}(\tau, x) = d\frac{Q_\tau[Y | X = x]}{dx},$$

where $\tau$ denote the quantile of the conditional distribution: $\tau = \Pr[Y \leq Q_\tau[Y|X = x]|X = x]$. Since $Y = h(X, \varepsilon)$ is monotonic in $\varepsilon$,

$$\begin{aligned}\tau &= \Pr[Y \leq Q_\tau[Y|X = x]|X = x] \\ &= \Pr\big[\varepsilon \leq h^{-1}(X, Q_\tau[Y|X = x])|X = x\big] \\ &= F_\varepsilon\big(h^{-1}(x, Q_\tau[Y|X = x])\big).\end{aligned}$$

Thus, by the implicit function theorem,

$$\begin{aligned}&\mathrm{CQPE}(\tau, x) \\ &\quad = -\frac{f_\varepsilon(h^{-1}(x, Q_\tau[Y|X = x])) \cdot \partial h^{-1}(x, Q_\tau[Y|X = x])/\partial x}{f_\varepsilon(h^{-1}(x, Q_\tau[Y|X = x])) \cdot \partial h^{-1}(x, q)/\partial q|_{q=Q_\tau[Y|X=x]}} \\ &\quad = -\big(-\partial h\big(x, h^{-1}(x, Q_\tau[Y|X = x])\big)/\partial x\big) \\ &\qquad \cdot \left(\frac{\partial h(x, \varepsilon)}{\partial \varepsilon}\bigg|_{\varepsilon = h^{-1}(x, Q_\tau[Y|X=x])}\right)^{-1} \Bigg/ \left(\frac{\partial h(x, \varepsilon)}{\partial \varepsilon}\bigg|_{\varepsilon = h^{-1}(x, Q_\tau[Y|X=x])}\right)^{-1} \\ &\quad = \frac{\partial h(x, h^{-1}(x, Q_\tau[Y|X = x]))}{\partial x}.\end{aligned}$$

Using the matching function $\zeta_\tau(x) \equiv \{s : Q_s[Y|X = x] = q_\tau\}$, we can write $\mathrm{CQPE}(s, x)$ for the $\tau$th conditional quantile at a fixed $x$ ($Q_s[Y|X = x]$) that equals (matches) the $\tau$th unconditional quantile ($q_\tau$) as

$$\begin{aligned}\mathrm{CQPE}(s, x) &= \mathrm{CQPE}(\zeta_\tau(x), x) \\ &= \frac{\partial h(x, h^{-1}(x, Q_s[Y|X = x]))}{\partial x} \\ &= \frac{\partial h(x, h^{-1}(x, q_\tau))}{\partial x} = \frac{\partial h(X, \varepsilon_\tau(X))}{\partial x}.\end{aligned}$$

Therefore,

$$\begin{aligned}\mathrm{UQPE}(\tau) &= E\left[\omega_\tau(X) \cdot \frac{\partial h(X, \varepsilon_\tau(X))}{\partial x}\right] \\ &= E[\omega_\tau(X) \cdot \mathrm{CQPE}(\zeta_\tau(X), X)]. \qquad \textit{Q.E.D.}\end{aligned}$$

## REFERENCES

ALBRECHT, J., A. BJÖRKLUND, AND S. VROMAN (2003): "Is There a Glass Ceiling in Sweden?" *Journal of Labor Economics*, 21, 145–178. [957]

AUTOR, D. H., L. F. KATZ, AND M. S. KEARNEY (2008): "Trends in U.S. Wage Inequality: Revising the Revisionists," *Review of Economics and Statistics*, 90, 300–323. [966]

CARD, D. (1996): "The Effect of Unions on the Structure of Wages: A Longitudinal Analysis," *Econometrica*, 64, 957–979. [962,966]

CARD, D., T. LEMIEUX, AND W. C. RIDDELL (2004): "Unions and Wage Inequality," *Journal of Labor Research*, 25, 519–562. [963]

CHAMBERLAIN, G. (1994): "Quantile Regression Censoring and the Structure of Wages," in *Advances in Econometrics*, ed. by C. Sims. New York: Elsevier. [962,964]

CHESHER, A. (2003): "Identification in Nonseparable Models," *Econometrica*, 71, 1401–1444. [955]

DINARDO, J., N. M. FORTIN, AND T. LEMIEUX (1996): "Labor Market Institutions and the Distribution of Wages: A Semi-Parametric Approach, 1973–1992," *Econometrica*, 64, 1001–1044. [966]

FIRPO, S., N. M. FORTIN, AND T. LEMIEUX (2007a): "Unconditional Quantile Regressions," Technical Working Paper 339, National Bureau of Economic Research, Cambridge, MA. [960,962]

——— (2007b): "Decomposing Wage Distributions Using Recentered Influence Function Regressions," Unpublished Manuscript, University of British Columbia. [954,967]

——— (2009): "Supplement to 'Unconditional Quantile Regressions'," *Econometrica Supplemental Material*, 77, http://www.econometricsociety.org/ecta/Supmat/6822_extensions.pdf; http://www.econometricsociety.org/ecta/Supmat/6822_data and programs.zip. [961,962,964]

FLORENS, J. P., J. J. HECKMAN, C. MEGHIR, AND E. VYTLACIL (2008): "Identification of Treatment Effects Using Control Functions in Models With Continuous, Endogenous Treatment and Heterogeneous Effect," *Econometrica*, 76, 1191–1207. [955]

HAMPEL, F. R. (1968): "Contribution to the Theory of Robust Estimation," Ph.D. Thesis, University of California at Berkeley. [956]

——— (1974): "The Influence Curve and Its Role in Robust Estimation," *Journal of the American Statistical Association*, 60, 383–393. [956]

HIRANO, K., G. W. IMBENS, AND G. RIDDER (2003): "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score," *Econometrica*, 71, 1161–1189. [961]

IMBENS, G. W., AND W. K. NEWEY (2009): "Identification and Estimation of Triangular Simultaneous Equations Models Without Additivity," *Econometrica* (forthcoming). [955]

KOENKER, R. (2005): *Quantile Regression*, New York: Cambridge University Press. [953]

KOENKER, R., AND G. BASSETT (1978): "Regression Quantiles," *Econometrica*, 46, 33–50. [953, 960]

LEMIEUX, T. (1998): "Estimating the Effects of Unions on Wage Inequality in a Panel Data Model With Comparative Advantage and Non-Random Selection," *Journal of Labor Economics*, 16, 261–291. [962]

——— (2006): "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?" *American Economic Review*, 96, 461–498. [963]

——— (2008): "The Changing Nature of Wage Inequality," *Journal of Population Economics*, 21, 21–48. [966]

MACHADO, A. F., AND J. MATA (2005): "Counterfactual Decomposition of Changes in Wage Distributions Using Quantile Regression," *Journal of Applied Econometrics*, 20, 445–465. [957]

MELLY, B. (2005): "Decomposition of Differences in Distribution Using Quantile Regression," *Labour Economics*, 12, 577–590. [957]

NEWEY, W. K. (1994): "The Asymptotic Variance of Semiparametric Estimators," *Econometrica*, 62, 1349–1382. [961]

STOKER, T. M. (1991): "Equivalence of Direct, Indirect and Slope Estimators of Average Derivatives," in *Nonparametric and Semiparametric Methods*, ed. by W. A. Barnett, J. L. Powell, and G. Tauchen. Cambridge, U.K.: Cambridge University Press. [961]

VON MISES, R. (1947): "On the Asymptotic Distribution of Differentiable Statistical Functions," *The Annals of Mathematical Statistics*, 18, 309–348. [956]

WOOLDRIDGE, J. M. (2004): "Estimating Average Partial Effects Under Conditional Moment Independence Assumptions," Unpublished Manuscript, Michigan State University. [958]

*Escola de Economia de São Paulo, Fundação Getúlio Vargas, Rua Itapeva 474, São Paulo, SP 01332-000, Brazil; sergio.firpo@fgv.br,*

*Dept. of Economics, University of British Columbia, 997-1873 East Mall, Vancouver, BC V6T 1Z1, Canada and Canadian Institute for Advanced Research, Toronto, Canada nifortin@interchange.ubc.ca,*

*and*

*Dept. of Economics, University of British Columbia, 997-1873 East Mall, Vancouver, BC V6T 1Z1, Canada; tlemieux@interchange.ubc.ca.*

*Manuscript received November, 2006; final revision received December, 2008.*