

Unconditional Quantile Regressions

Sergio Firpo

Nicole M. Fortin

Thomas Lemieux

July 2007

Presented by Castiel Chen Zhuang on December 12, 2013

Authors



Firpo

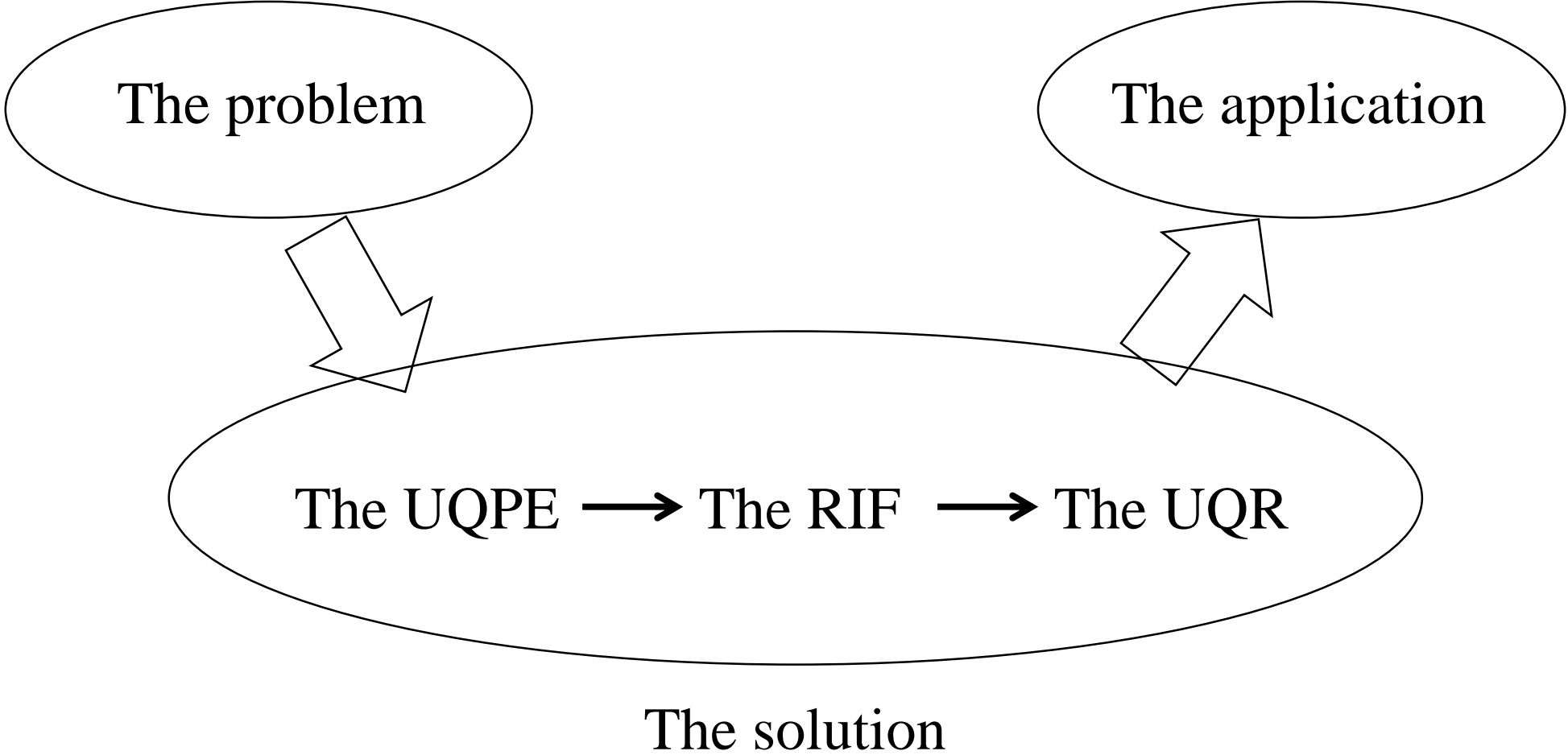


Fortin



Lemieux

Roadmap



1. The problem (Introduction)

- OLS Regression
 - impact of X on $E[Y]$ (unconditional mean)
 - the law of iterated expectations: $E[E[Y|X]] = E[Y]$
 - $E[Y] = E[E[Y|X]] = E[X\beta] = E[X]\beta$
- Conditional Quantile Regression
 - impact of X on $Q_\tau[Y|X]$ (conditional quantile)
 - $E[Q_\tau[Y|X]] \neq Q_\tau[Y]$
 - β_τ in $Q_\tau[Y|X] = X\beta_\tau$ and $Q_\tau[Y] = X\beta_\tau$ are quite different

2. The UQPE and the “policy effect”

- The UQPE
- The CQPE
- The quantile “policy effect”
- The mean “policy effect”

The UQPE

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

observed unobserved

- The UQPE (unconditional quantile partial effect) is the effect of a small increase t in X_j on an unconditional quantile $Q_\tau[Y]$.

$$\alpha_j(\tau) = \frac{\partial Q_\tau[Y]}{\partial x_j} = \lim_{t \downarrow 0} \frac{Q_\tau[h([X_j + t_j; X_{-j}], \varepsilon)] - Q_\tau[Y]}{t_j}$$

- Compare to the UAPE: $\partial E[Y]/\partial x_j = E[\partial E[Y|X]/\partial x_j]$.

A: average

The CQPE

- The CQPE (conditional quantile partial effect) is the effect of a small increase t in X_j on a conditional quantile $Q_\tau[Y|X]$.

$$\frac{\partial Q_\tau[Y|X]}{\partial x_j} = \lim_{t \downarrow 0} \frac{Q_\tau[h([X_j + t_j; X_{-j}], \varepsilon)|X] - Q_\tau[Y|X]}{t_j}$$

- Compare to the CAPE: $\partial E[Y|X]/\partial x_j$.
- Note that while $UAPE = E[CAPE]$, $UQPE \neq E[CQPE]$. By OLS regressions we can only derive the CAPE.

The quantile “policy effect”

- We might consider the effect of a more general change in X instead of a simple increase t .
- A policy change might cause changes not only in the means of X_j but also in the “pattern” of X . We can define a function $\ell(X), \ell: \mathcal{X} \rightarrow \mathcal{X}$ to illustrate this policy change.
- Since the coming up of a policy change is binary, the effect of the policy on the τ th quantile of Y is

$$\delta_\ell(\tau) = Q_\tau[h(\ell(X), \varepsilon)] - Q_\tau[h(X, \varepsilon)]$$

The mean “policy effect”

$$\begin{aligned}\delta_{\ell}(\mu) &= E[h(\ell(X), \varepsilon)] - E[h(X, \varepsilon)] \\ &= E[E[h(\ell(X), \varepsilon)] - E[h(X, \varepsilon)]|X]\end{aligned}$$

- This mean effect can be easily derived from OLS regressions under the law of iterated expectations.
- However, the problem is how to derive the UQPE and the quantile “policy effect”. (Note that the policy effects we are talking about here are both unconditional.)

One remark

- Both changes in X can be modelled as changes in the distribution of X , $F_X(x)$. We can also denote its counterfactual distribution as $G_X(x)$.
- The unconditional distribution function of Y can be written as:

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

- If we assume that the conditional distribution $F_{Y|X}(\cdot)$ is unaffected by manipulations of the distribution X , then we can write the counterfactual distribution of Y , $G_Y(y)$, as follow:

$$G_Y(y) = \int F_{Y|X}(y|X = x) dG_X(x)$$

3. The RIF

- The IF (Influence Function)
- The RIF (*Re-centered IF*)
- Properties of the RIF
- The RIF-regression
- The approximation of the UPE
- The approximation of the “policy effect”

The IF

- Hampel (1974) introduces the **influence function (IF)** as a measure to study the infinitesimal behavior of **real-valued functionals (泛函) $v(F)$, $v: \mathcal{F}_v \rightarrow \mathbb{R}$** .
- In our setting, F is the cumulative distribution function (CDF) of Y , and $v(F)$ is a distributional statistic such as a quantile.
- Denote Δ_y as the **probability measure (概率测度)** that gives mass 1 to $\{y\}$.

The IF

- The **Gâteaux derivative** (加托导数) of v at F in the direction G is defined by

$$L_F(G) = \lim_{t \downarrow 0} \frac{v((1-t)F + tG) - v(F)}{t}$$

- If v is **Gâteaux differentiable** at F , then there exists a real kernel function $a(\cdot)$ such that for all G in F_v :

$$L_F(G) = \int a(y) dG(y) = \int a(y) d(G - F)(y)$$

The IF

- Denote $(1 - t)F + tG$ as $F_{t,G}$. The concept of the IF arises from the special case where G is replaced by Δ_y . Then we have

$$IF(y; v, F) = \lim_{t \downarrow 0} \frac{v(F_{t, \Delta_y}) - v(F)}{t} = \int a(y) d\underline{\Delta_y(y)} = a(y)$$

mass 1

- By a normalization argument, $IF(y; v, F)$ will be written as $IF(y; v)$.

The IF

- The functional $v(F_{t,G})$ per se can be represented as a **von Mises linear approximation** (冯米塞斯线性近似):

$$v(F_{t,G}) = v(F) + t \int IF(y; v) d(G - F)(y) + r(t; v; G, F)$$

- where $r(t; v; G, F)$ is a remainder term that converges to zero as t goes to 0 at the general rate $o(t)$.
- For the mean μ , $r(t; \mu; G, F) = 0$, while for the quantile, $r(t; q_\tau; G, F) = o(t)$.

The RIF

- Then we have the approximation for $v(G)$:

$$v(G) = v(F_{1,G}) \approx v(F) + \int IF(y; v) dG(y)$$

- By replacing G with Δ_y again we can get the **re-centered influence function (RIF)**:

$$RIF(y; v, F) = v(F) + \int IF(y; v) d\overset{\text{mass 1}}{\Delta_y}(y) = v(F) + IF(y; v)$$

- Again, by a normalization argument, we write $RIF(y; v, F)$ as $RIF(y; v)$.

Properties of the RIF

- **Mean** of the RIF:

$$\int RIF(y; v) dF(y) = v(F)$$

- **Variance** of the RIF:

$$\int (RIF(y; v) - v(F))^2 dF(y) = AV(v, F)$$

- where $AV(v, F)$ is the **asymptotic variance** of functional v under the probability distribution F .

Properties of the RIF

- The **Gâteaux derivative** of v at F in the direction G can be obtained by integrating up the RIF at F over the distributional differences between G and F :

$$L_F(G) = \int RIF(y; v) d(G - F)(y)$$

- Recall that

$$L_F(G) = \int a(y) d(G - F)(y)$$

The RIF-regression

- From the **mean of the RIF** and the **unconditional distribution function of Y** mentioned before, we can simply get that:

$$\begin{aligned}v(F_Y) &= \int RIF(y; v) dF_Y(y) \\ &= \int E[RIF(Y; v) | X = x] dF_X(x)\end{aligned}$$

- where $E[RIF(Y; v) | X = x]$ is the mean of the RIF-regression.

The RIF-regression

- Recall that

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

- [Machado and Mata \(2005\)](#) just directly estimate the whole conditional distribution by running (conditional) quantile regressions for each and every possible quantile.
- However, we might be more interested in $v(F_Y)$ instead of F_Y itself.

The RIF-regression

- The RIF-regression provides a much more convenient way to get a specific distribution statistic $v(F_Y)$: we simply need to integrate over $E[RIF(Y; v)|X]$, which is easily estimated using regression methods.
- Consider a general change in the distribution of covariates X from $F_X(x)$ to the counterfactual $G_X(x)$, the distribution of dependent variable Y thus changing from $F_Y(y)$ to $G_Y(y)$.

The RIF-regression

- We then can consider the **marginal (unconditional) effect** of a change in the distribution of X as the **directional derivative**:

$$\begin{aligned}\pi_G(v) &= \lim_{t \downarrow 0} \frac{v((1-t)F_Y + tG_Y) - v(F_Y)}{t} && L_{F_Y}(G_Y) \\ &= \int RIF(y; v) d(G_Y - F_Y)(y) \\ &= \int E[RIF(Y; v) | X = x] d(G_X - F_X)(x)\end{aligned}$$

The approximation of the UPE

- Consider increasing a **continuous** covariate X by t , from X to $X + t$. This change results in the counterfactual distribution $F_{Y,t}^*(y) = \int F_{Y|X}(y|x) dF_X(x - t)$. The effect of X on the distribution statistic v , $\alpha(v)$, is

$$\begin{aligned}\alpha(v) &= \lim_{t \downarrow 0} \frac{v(F_{Y,t}^*) - v(F_Y)}{t} \\ &= \int \frac{dE[RIF(Y; v)|X = x]}{dx} dF(x)\end{aligned}$$

The approximation of the UPE

- Consider the case where X is a **discrete (dummy)** variable, $X \in \{0,1\}$. Define $P_X = \Pr[X = 1]$.
- Consider an increase from P_X to $P_X + t$. This results in the counterfactual distribution $F_{Y,t}^*(y) = F_{Y|X}(y|1)(P_X + t) + F_{Y|X}(y|0)(1 - P_X - t)$.
- The effect of a small increase in the probability that $X = 1$ is given by $\alpha_D(v) = \lim_{t \downarrow 0} \frac{v(F_{Y,t}^*) - v(F_Y)}{t} = E[RIF(Y; v, F)|X = 1] - E[RIF(Y; v, F)|X = 0]$

The approximation of the “policy effect”

- If a policy change from X to $\ell(X)$ can be described as a change in the distribution of covariates, which is, $\ell(X) \sim G_X$, where $G_X(x) = F_X(\ell^{-1}(X))$, then the policy effect on the functional v consists of the **marginal effect of the policy** $\pi_\ell(v)$ and a **remainder term** $r(v, G_Y, F_Y)$.

$$\delta_\ell(v) = v(G_Y) - v(F_Y) = \pi_\ell(v) + r(v, G_Y, F_Y)$$

The approximation of the “policy effect”

- where

$$\pi_\ell(\nu) = \int E[RIF(Y; \nu) | X = x] \left(dF_X(\ell^{-1}(X)) - dF_X(x) \right)$$

- and

$$\begin{aligned} & r(\nu, G_Y, F_Y) \\ &= \int (E[RIF(Y; \nu, G_Y) | X] - E[RIF(Y; \nu) | X]) dF_X(\ell^{-1}(X)) \end{aligned}$$

4. The UQR using the RIF

- The RIF for quantiles
- The approximated UQPE
- The UQPE and the structural form
 - Case 1: Linear, additively separable model
 - Case 2: Non-linear, additively separable model
 - Case 3: Linear, separable, but heteroskedastic model
 - General case

The RIF for quantiles

- Consider $v(F) = q_\tau$, then we have

$$IF(y; q_\tau) = \frac{\tau - \mathbb{I}\{y \leq q_\tau\}}{f_Y(q_\tau)}$$

- The RIF can thus be written as

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

The RIF for quantiles

- We can simply get that

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

- Since the conditional expectation $E[RIF(Y; q_\tau) | X = x]$ is a linear function of $\Pr[Y > q_\tau | X = x]$, it can be estimated using **Probit** or **Logit** regressions, or a simple **OLS** regression (linear probability model).

The approximated UQPE

- Based on the above analysis, we can derive that

$$\begin{aligned} UQPE(\tau) &= \alpha(\tau) = \int \frac{dE[RIF(Y; q_\tau) | X = x]}{dx} dF_X(x) \\ &= \frac{1}{f_Y(q_\tau)} \cdot \int \frac{d\Pr[Y > q_\tau | X = x]}{dx} dF_X(x) \end{aligned}$$

- This is the case where X is **continuous**.

The approximated UQPE

- For the case where X is **discrete**, we have

$$\begin{aligned} UQPE(\tau) &= \alpha_D(\tau) \\ &= \frac{1}{f_Y(q_\tau)} \cdot (\Pr[Y > q_\tau | X = 1] - \Pr[Y > q_\tau | X = 0]) \end{aligned}$$

- Here we can see that estimating the **RIF-regression for quantiles** is closely linked to the estimation of a **probability response model**.

The UQPE and the structural form

- Recall that we defined the structural form $Y = h(X, \varepsilon)$.
- Now we consider three specific cases and the general case:
 - Case 1: $Y = h(X, \varepsilon) = X^T \beta + \varepsilon$
 - Case 2: $Y = h(X, \varepsilon) = \tilde{h}(X^T \beta + \varepsilon)$
 - Case 3: $Y = h(X, \varepsilon) = X^T \beta + \sigma(X)\varepsilon$
 - General case

Case 1: Linear, additively separable model

- The simplest linear model: $Y = h(X, \varepsilon) = X^T \beta + \varepsilon$, where X and ε are independent.
- The UQPE for any quantile is equal to β_j :

$$UQPE(\tau) = \frac{1}{f_Y(q_\tau)} \cdot \beta_j \cdot f_Y(q_\tau) = \beta_j$$

Case 2: Non-linear, additively separable model

- A simple extension of the linear model is the **index model**:
 $Y = h(X, \varepsilon) = \tilde{h}(X^T \beta + \varepsilon)$, where a small change t in a covariate X_j does not correspond to a simple location shift of the distribution of Y , thus the UQPE is no longer equal to β :

$$\begin{aligned} UQPE(\tau) &= \frac{1}{f_Y(q_\tau)} \cdot \beta_j \cdot E[f_\varepsilon(\tilde{h}^{-1}(q_\tau) - X^T \beta)] \\ &= \beta_j \cdot \tilde{h}'(\tilde{h}^{-1}(q_\tau)) \end{aligned}$$

- The UQPE is proportional, but not equal, to β .

Case 3: Linear, separable, heteroskedastic model

- A more standard model used in economics is the linear, but **heteroskedastic (异方差)** model: $Y = h(X, \varepsilon) = X^T \beta + \sigma(X)\varepsilon$, where $Var(Y|X) = \sigma^2(X)$.
- It is no longer possible to express the marginal effects as simple functions of the structural parameter β .
- One could estimate the implied non-linear Probit model using maximum likelihood (ML) and then compute the Probit marginal effects to get the UQPE or estimate a more standard flexible probability response model and compute the average marginal effects.

General case

- In order to draw a connection between UQPE and the underlying structural form, we can try to establish the link between the UQPE and the CQPE.
- We first define three auxiliary functions.
- $w_\tau(x) = \frac{f_{Y|X}(q_\tau|x)}{f_Y(q_\tau)}$ (weighting function)
- $\varepsilon_\tau(x) = h^{-1}(x, q_\tau)$ (inverse h function)
- $\zeta_\tau(x) = F_{Y|X}(q_\tau|X = x)$ (matching function)

General case

- Then we can derive that

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

- or

$$UQPE(\tau) = E[w_\tau(X) \cdot CQPE(\zeta_\tau(X), X)]$$

- We can see that $UQPE(\tau) \neq E[CQPE(\tau, X)]$, but $UQPE(\tau)$ is equal to a weighted average of the $CQPE(\zeta_\tau(X), X)$.

General case

- If most of the variation in Y is in the residual, then the matching function $\zeta_\tau(X)$ will not tend to vary very much and is more or less equal to τ for all values of X .
- Then we will have

$$UQPE(\tau) \approx E[w_\tau(X) \cdot CQPE(\tau, X)]$$

5. The UQR Estimation

- The RIF and its components
- Three estimation methods
 - RIF-OLS
 - RIF-Logit
 - RIF-NP
 - (I won't discuss the quantile “policy effect” hereinafter.)

The RIF and its components

- Since $\mathbb{I}\{y > q_\tau\}$ is non-observable random variable that depends on the true unconditional quantile q_τ , we use a **feasible** version of that variable:

$$\hat{T}_\tau = \mathbb{I}\{y > \hat{q}_\tau\}$$

- The corresponding **feasible** version of the RIF is

$$\widehat{RIF}(Y; \hat{q}_\tau) = \hat{c}_{1,\tau} \cdot \hat{T}_\tau + \hat{c}_{2,\tau}$$

- which involves two unknown quantiles to be estimated, \hat{q}_τ and $\hat{f}_Y(\hat{q}_\tau)$.

Three estimation methods: RIF-OLS

- Recall that $UQPE(\tau) = \int \frac{dE[RIF(Y; q_\tau) | X=x]}{dx} dF_X(x)$, so we need to calculate the **derivative** in order to estimate the UQPE.
- The RIF-OLS regression is based on this perspective:

$$\widehat{UQPE}_{RIF-OLS}(\tau) = (XX^T)^{-1} X \widehat{RIF}(Y; \hat{q}_\tau)$$

- Recall that for $Y = X^T \beta + \varepsilon$, we have

$$\hat{\beta} = (XX^T)^{-1} XY$$

Three estimation methods: RIF-Logit

- The second estimator exploit the fact that the regression model is closely connected to a **probability response model** since $E[RIF(Y; q_\tau) | X = x] = c_{1,\tau} \cdot \Pr[Y > q_\tau | X = x] + c_{2,\tau}$.
- Consider a logistic model:

$$\Pr[Y > q_\tau | X = x] = \Lambda(x^T \theta_\tau)$$

- where $\Lambda(\cdot)$ is the logistic CDF. we can then estimate θ_τ by ML by using the feasible version.

Three estimation methods: RIF-Logit

- The main advantage of the Logit model over the linear specification for $E[RIF(Y; q_\tau) | X = x]$ is that it **allows heterogenous marginal effects** (i.e. depending on x).

- We know that

$$\begin{aligned} \frac{dE[RIF(Y; q_\tau) | X = x]}{dx} &= c_{1,\tau} \frac{d\Pr[Y > q_\tau | X = x]}{dx} \\ &= c_{1,\tau} \cdot \theta_\tau \cdot \Lambda(x^T \theta_\tau) (1 - \Lambda(x^T \theta_\tau)) \end{aligned}$$

- Then we can derive the UQPE. (see next page)

Three estimation methods: RIF-Logit

- We can derive the UQPE as:

$$\widehat{UQPE}_{RIF-Logit}(\tau) = \hat{c}_{1,\tau} \cdot \hat{\theta}_\tau \cdot \frac{1}{N} \cdot \underbrace{\Lambda(X^T \hat{\theta}_\tau) (1 - \Lambda(X^T \hat{\theta}_\tau))}_{\text{in matrix form}}$$

- Recall that here we make the functional form assumption about $\Pr[Y > q_\tau | X = x]$ as the logistic CDF. If we make no assumption about the conditional probability, then we come to the third method.

Three estimation methods: RIF-NP

- The third method is called RIF-NP (RIF-nonparametric). Based on Hirano, Imbens and Ridder (2003), we can estimate a probability response model nonparametrically by means of **polynomial approximation (多项式近似)** of the log-odds ratio of $\Pr[Y > q_\tau | X = x]$.
- The log-odds ratio is:

$$\log \left(\frac{\Pr[Y > q_\tau | X = x]}{1 - \Pr[Y > q_\tau | X = x]} \right)$$

Three estimation methods: RIF-NP

- Then we can derive that

$$\begin{aligned} & \widehat{UQPE}_{RIF-NP}(\tau) \\ &= \hat{c}_{1,\tau} \cdot \hat{\rho}_K(\tau) \cdot \frac{1}{N} \cdot \sum_{i=1}^N \frac{dH_{K(\tau)}(X_i)^T}{dx} \hat{\rho}_{K,\tau}(X_i) \left(1 - \hat{\rho}_{K,\tau}(X_i)\right) \end{aligned}$$

- If $H_{K(\tau)}(x) = x$ for all x , then RIF-Logit and RIF-NP will coincide.

6. Empirical Applications

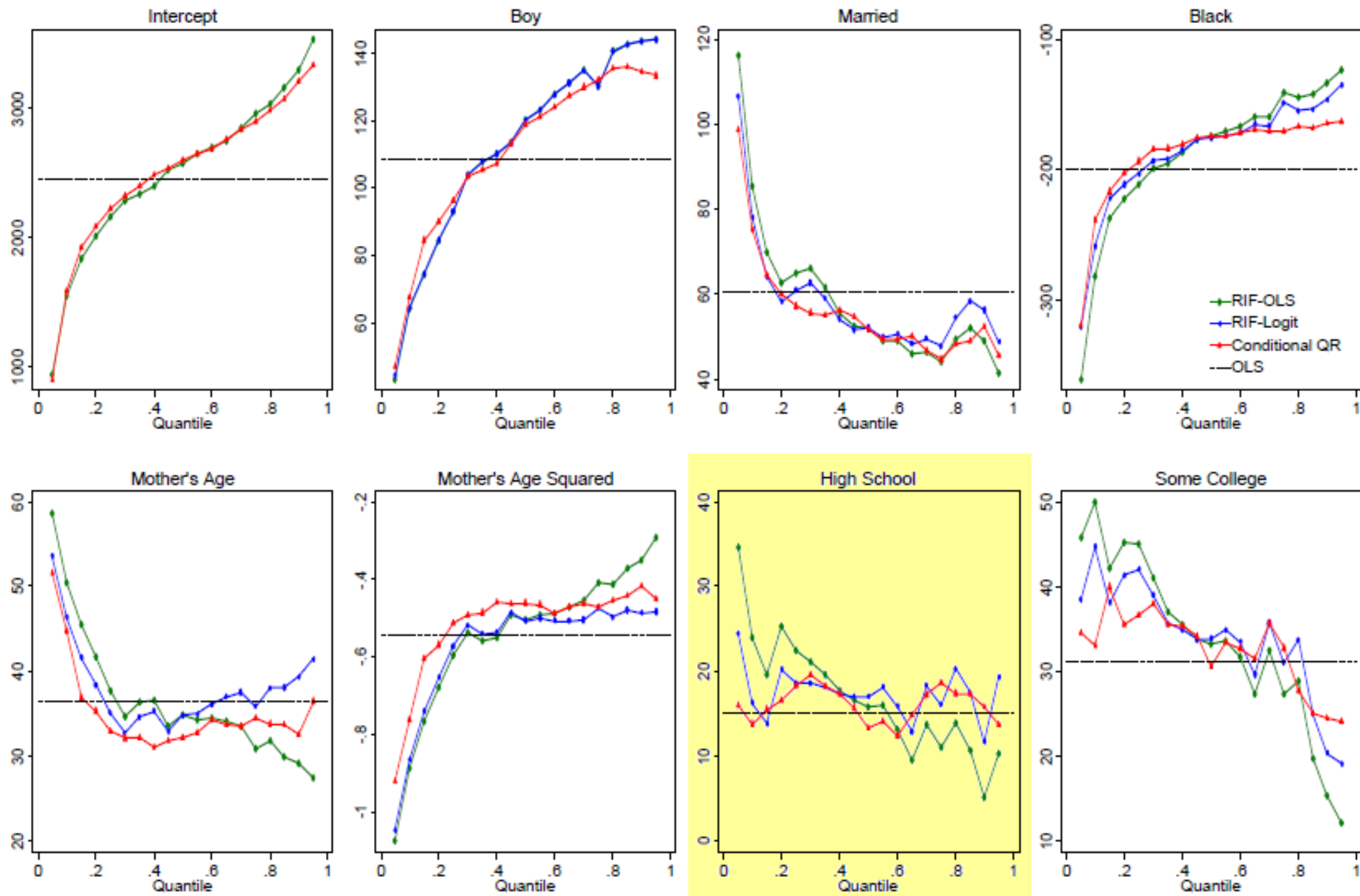
- Revisit Koenker and Hallock (2001)
- Unions and wage inequality
- (Here I only concentrate on the partial effect.)

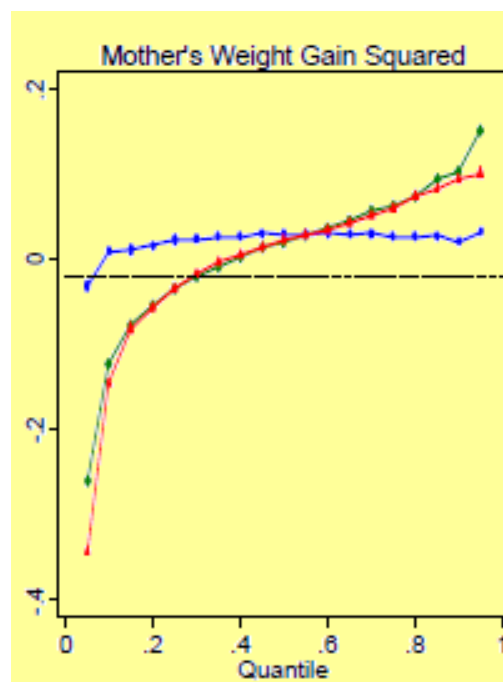
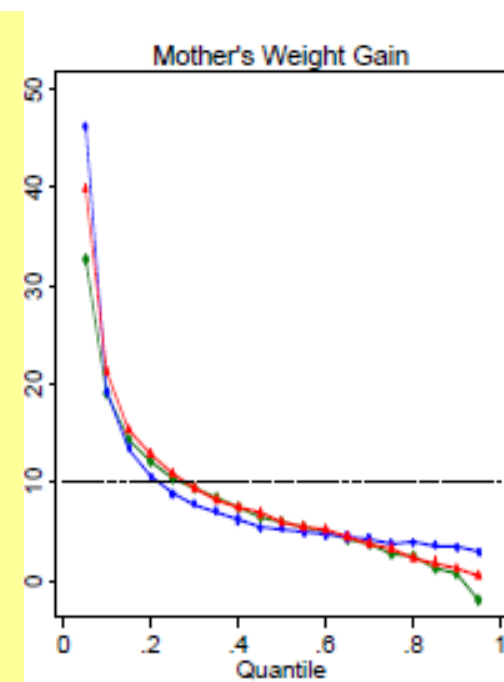
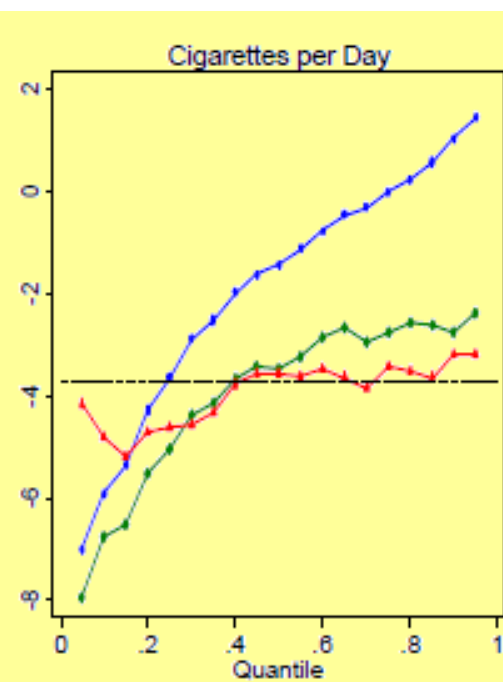
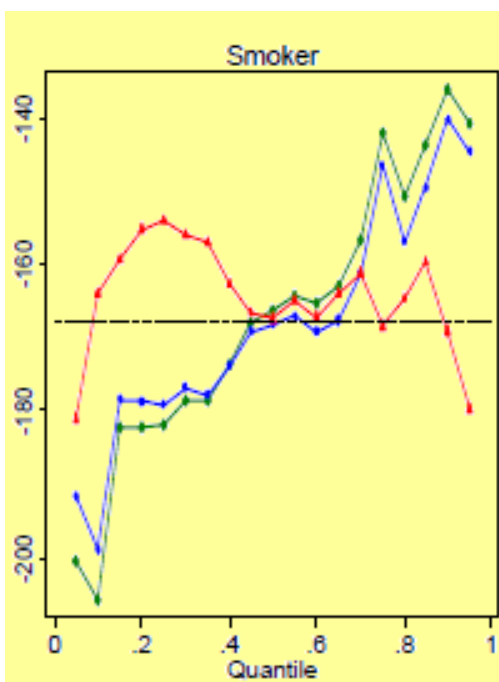
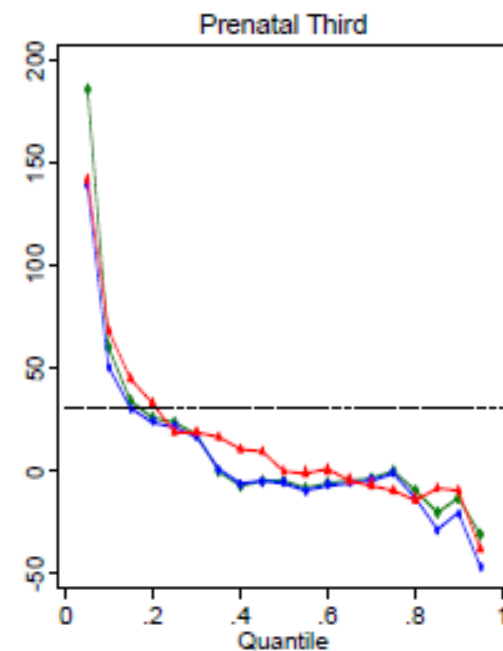
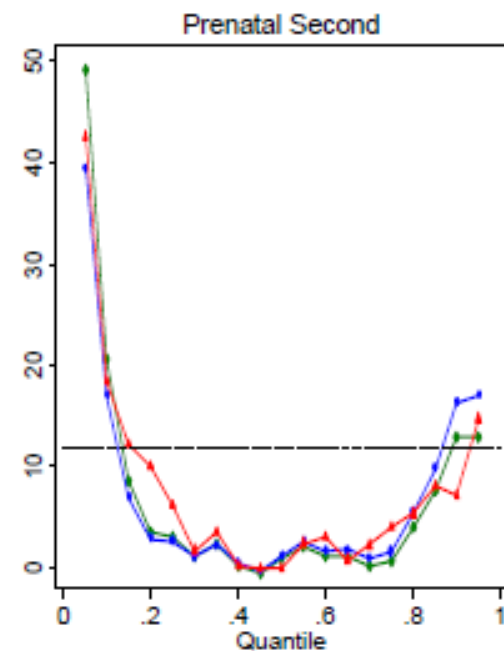
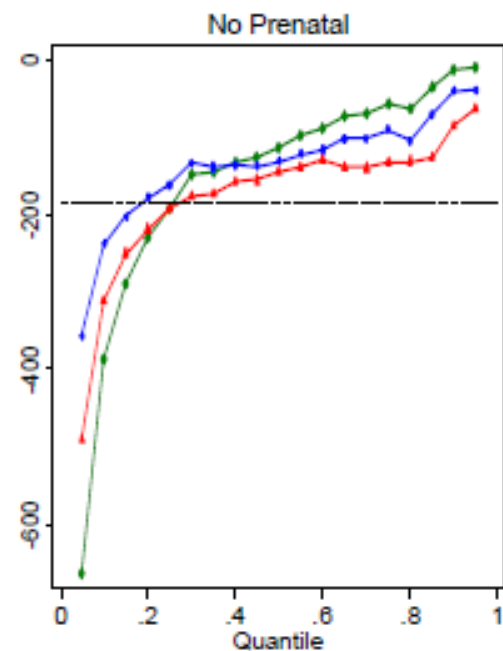
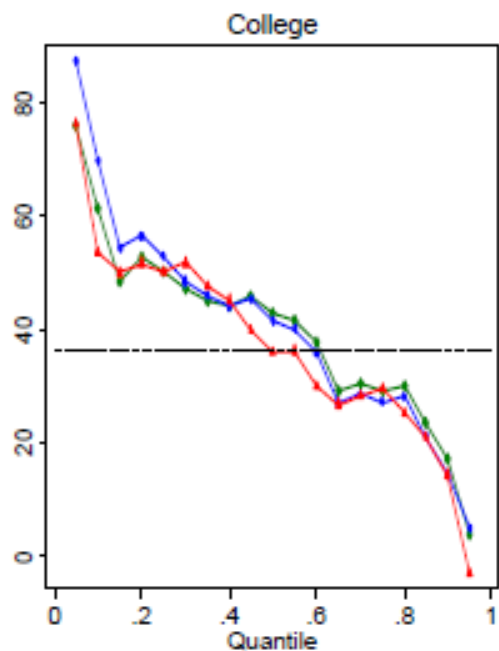
Revisit Koenker and Hallock (2001)

- The limitation: low birth-weight threshold may well fall at different quantiles depending on mother's characteristics.
- *For example:* the 10th quantile for infants of **black high school dropout mothers who also smoke** is **2183 grams**, which is well below the low birth-weight threshold **2500 grams**; but the 10th quantile for infants of **white college educated mothers who do not smoke** is **2880 grams**, which is well above the low birth-weight threshold. The CQR estimate at the 10th conditional quantile will mix the impact for some infants.

Revisit Koenker and Hallock (2001)

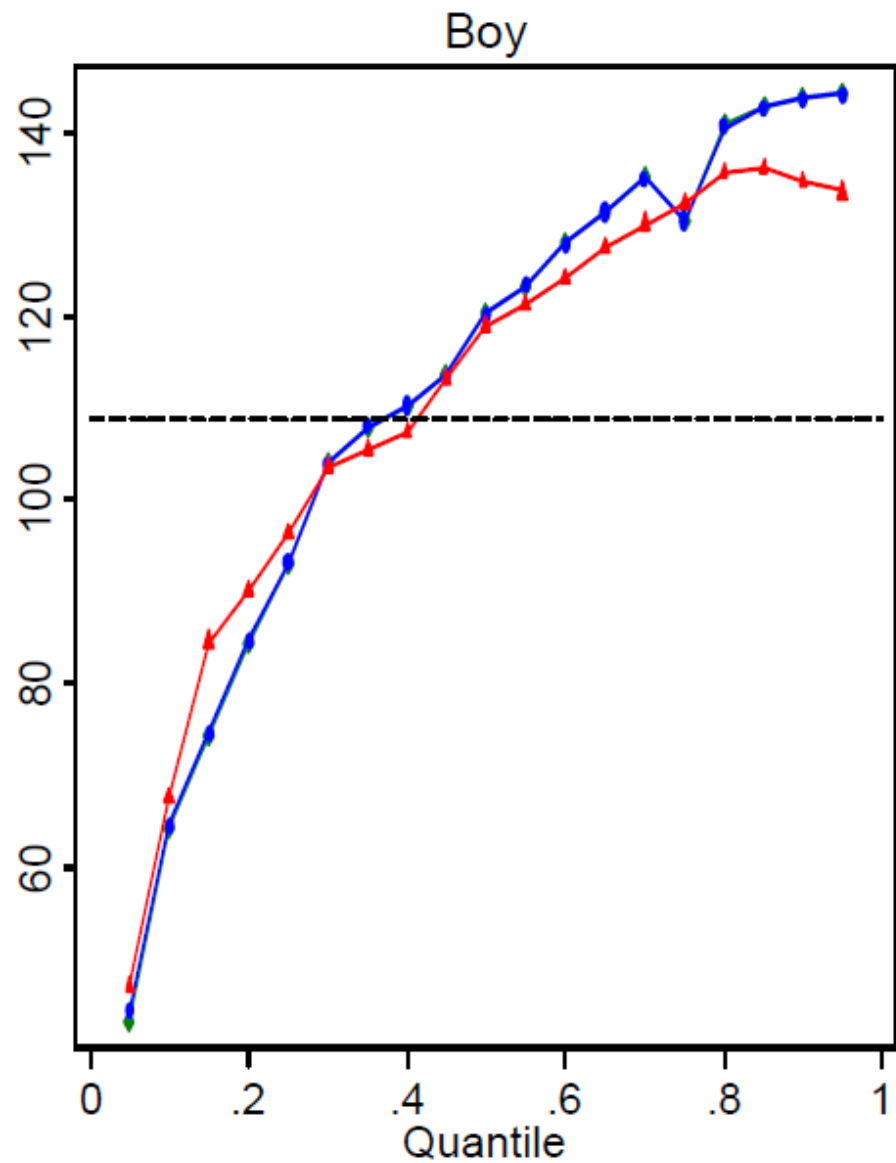
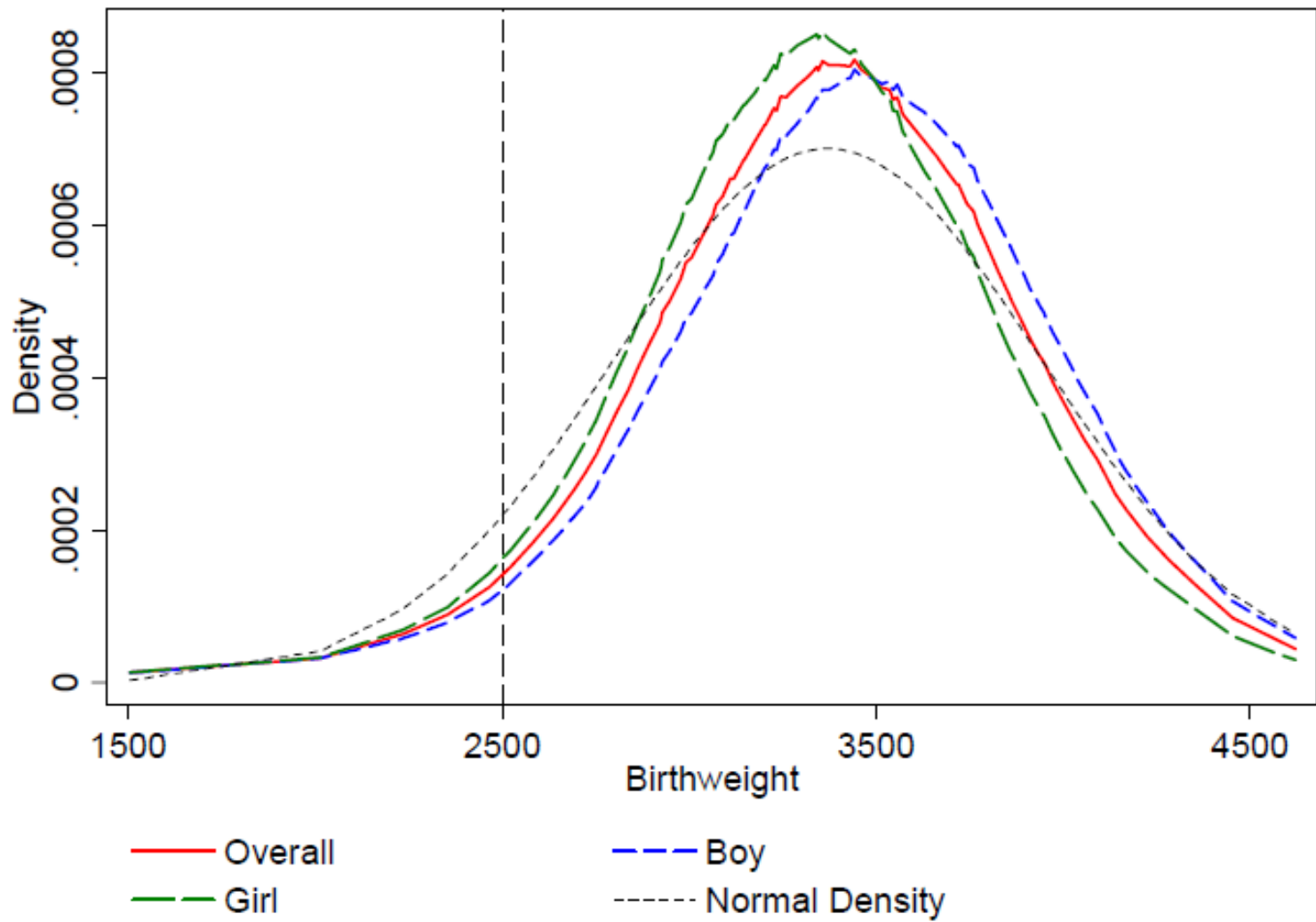
- We can see from the result that the CQPE and the UQPE turn out to be fairly similar in the case of birth weight.
- However, there still exist some differences.
- Moreover, there even exists some differences between the RIF-OLS and RIF-Logit (e.g. cigarettes), which may derive from **the difficulty of defining marginal effects** for variables whose distributions are actually **mixtures of categorical and continuous variables**.





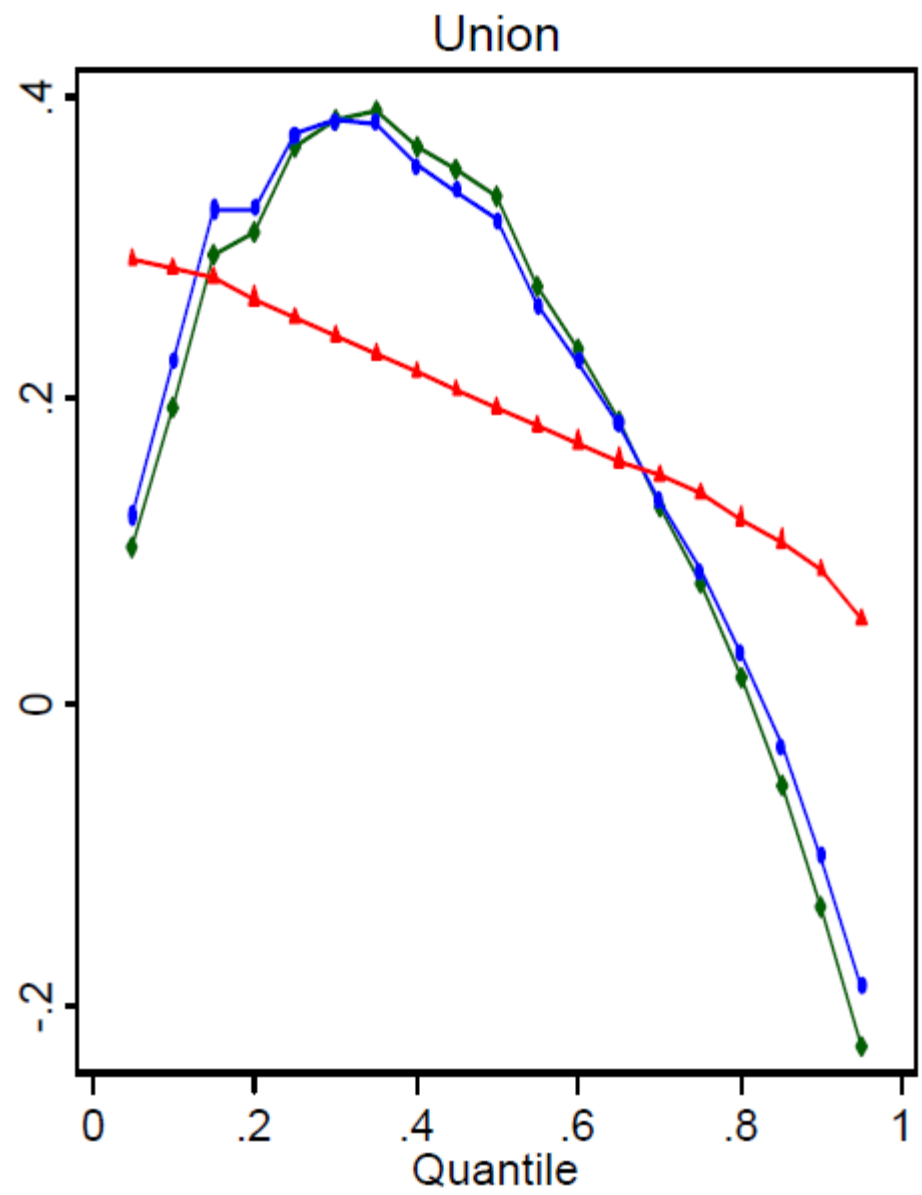
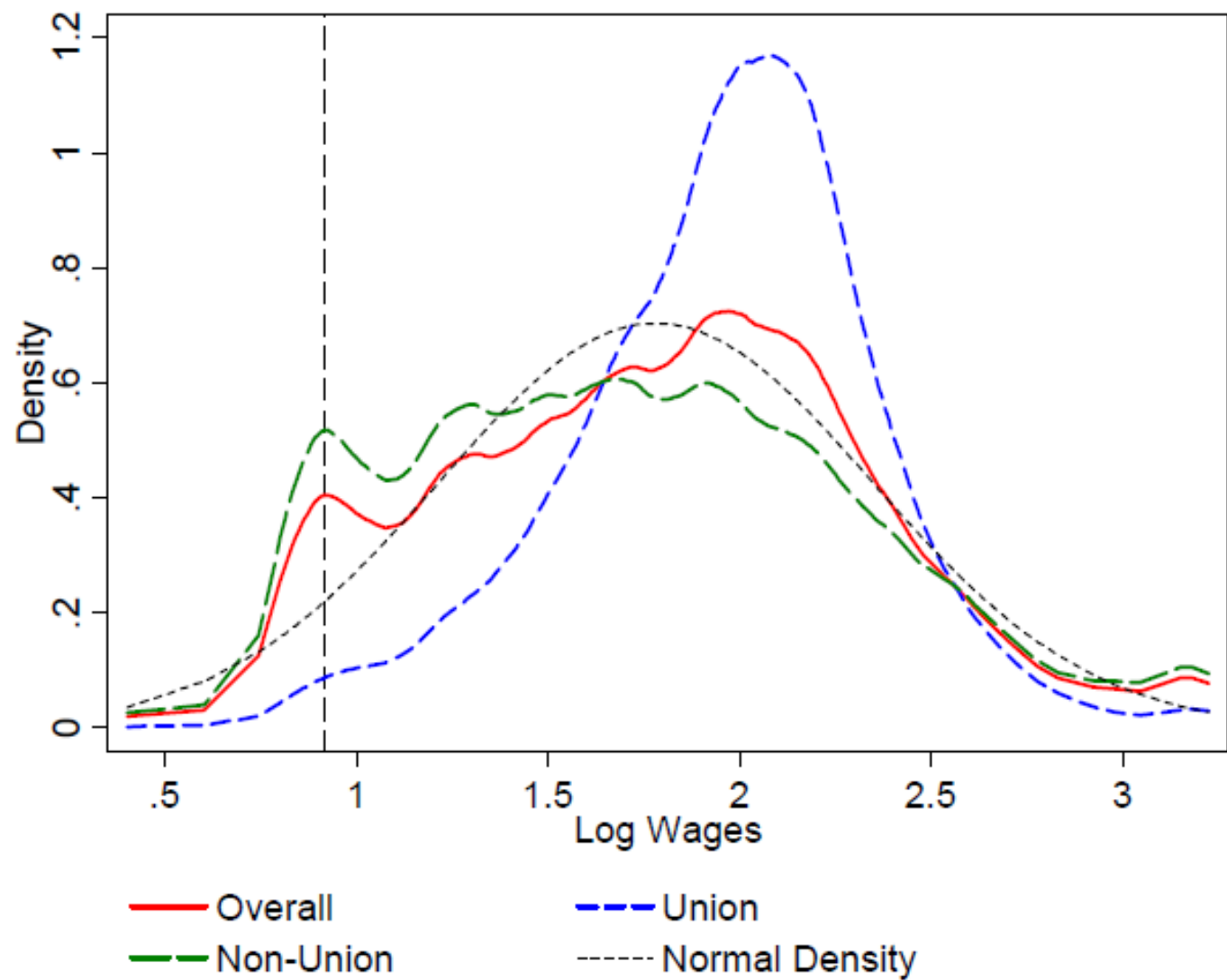
Revisit Koenker and Hallock (2001)

- One explanation for the similarity of the CQPE and the UQPE in this case is, **the covariates do not seem to be explaining much of the overall variation in birth weight.**
- Recall that: if most of the variation in Y is in the residual, $UQPE(\tau) \approx E[w_\tau(X) \cdot CQPE(\tau, X)]$. When the conditional and unconditional distributions are similar, then the CQPE and the UQPE will be similar.



Unions and wage inequality

- There are two effects from unions:
 - The “between” effect: increase the conditional mean of wages;
 - The “within” effect: decrease the conditional dispersion of wages.
- Unions tend to increase wages in low wage quantiles where both effects go in the same direction.
- Can the two effects go in opposite directions in high wage quantiles?



7. Conclusion

- The UQR consists of running a regression of the RIF of the unconditional quantile of Y on the explanatory variables X .
- There are three methods to estimate the UQPE: RIF-OLS, RIF-Logit, and RIF-NP, and they all yield similar estimates.
- The estimates of UQR and CQR could be very different.
- Moreover, the RIF-regression can be easily generalized to other distributional statistics such as the Gini or the Theil coefficient.

- However, the regression method also has its limitations.
- One of the limitations is, the method is under the assumption that the covariates X are independent of the unobservables ε , which is restrictive in many problems of economic interest, e.g., the effect of schooling on the distribution of wages.
- (Since the variable schooling X_S is generally regarded as an endogenous variable, which is, $\text{cov}(X_S, \varepsilon) \neq 0$.)
- Further researches are needed to show how the assumption can be relaxed when IVs are available for the endogenous covariates and how consistent estimates of the UQPE can be obtained by adding a control function in the unconditional quantile regressions.

Contact the Presenter

<http://cczhuang.weebly.com/>



Castiel Chen Zhuang

B.A. Student in Economics

B.S. Student in Statistics

Department of Economics, School of Economics, Peking University

No.5 Yiheyuan Road, Haidian District, Beijing 100871, China