Unconditional Quantile Partial Effects via Conditional Quantile Regression*

Javier Alejo, Antonio F. Galvao, Julian Martinez-Iriarte, and Gabriel Montes-Rojas

November 2, 2022

Abstract

This paper develops a semi-parametric procedure for estimation of unconditional quantile partial effects using quantile regression coefficients. The main result is based on the fact that, for continuous covariates, unconditional quantile effects are a weighted average of conditional ones at particular quantile levels that depend on the covariates. We propose a two-step estimator for the unconditional effects where in the first step one estimates a structural quantile regression model, and in the second step a non-parametric regression is applied to the first step coefficients. We establish the asymptotic properties of the estimator, say consistency and asymptotic normality. Monte Carlo simulations show numerical evidence that the estimator has very good finite sample performance and is robust to the selection of bandwidth and kernel. To illustrate the proposed method, we study the canonical application of the Engel's curve, i.e. food expenditures as a share of income.

Keywords: Quantile regression, unconditional quantile regression, nonparametric regression.

JEL: C14, C21.

*Alejo: IECON-Universidad de la Republica, Montevideo, Uruguay. E-mail: javier.alejo@fcea.edu.uy; Galvao: Department of Economics, Michigan State University, E-mail: agalvao@msu.edu; Martinez-Iriarte: Department of Economics, UC Santa Cruz. E-mail: jmart425@ucsc.edu; Montes-Rojas: CONICET and Universidad de Buenos Aires. E-mail: gabriel.montes@fce.uba.ar. We thank Yixiao Sun, Takuya Ura and seminar participants at MSU for helpful comments. All errors remain our own.

1 Introduction

Conditional quantile regression (CQR) (see, e.g., Koenker and Bassett, 1978; Koenker and Hallock, 2001; Koenker, 2005; Koenker, Chernozhukov, He, and Peng, 2017, for comprehensive analyses of CQR) is a general approach to estimate conditional quantile partial effects (CQPE), i.e., the effect of a covariate variable of interest (*ceteris paribus*) on the conditional quantile distribution of the outcome. CQR is a useful way to represent heterogeneity using a set of parameters to characterize the entire conditional distribution of an outcome variable given a list of observable covariates.

More recently, unconditional quantile regression (UQR), proposed in Firpo, Fortin, and Lemieux (2009), has attracted interest in both applied and theoretical literatures. UQR is an important tool for practitioners since it provides a method to evaluate the impact of changes in the distribution of the explanatory variables on the quantiles of the unconditional (marginal) distribution of the outcome variable. This method allows researchers to investigate important heterogeneity in the variable of interest. Naturally, UQR leads to the unconditional quantile partial effect (UQPE), which refers to the effect of a covariate (*ceteris paribus*) on the unconditional quantile distribution of the outcome variable.[1]

The combination of standard CQR with simulation exercises is usually implemented to evaluate distributional effects, such as UQPE. While it is feasible to calculate the unconditional distribution of an outcome variable using CQR (see, e.g., Autor, Katz, and Kearney (2005), Machado and Mata (2005), Melly (2005), and Chernozhukov, Fernández-Val, and Melly (2013)), this task is not obvious, at least compared to the ordinary least-squares (OLS) for the conditional mean. Since an analogue of the law of iterated expectations does not hold in the case of quantiles, the CQR analysis cannot be directly employed to analyze unconditional quantiles (see the discussion in Fortin, Lemieux, and Firpo, 2011).

Firpo, Fortin, and Lemieux (2009) seminal paper proposes several ways to estimate the UQPE. The most popular approach is the recentered influence function regression method, commonly referred to as RIF regression. It is a two-step procedure, where in the first stage one estimates the RIF, and in the second step, a standard OLS regression of the RIF on covariates estimates the UQPE. While the method is appealing due to its simplicity, it relies on the ablity of the ability of the researcher to specify a regression equation for the influence function, a relatively abstract object. Our methodology, on the contrary, starts from an specification of the conditional quantiles, which is a common practice when researches want to explore heterogeneity in conditional effects of a certain covariate.

An interesting theoretical derivation connecting CQPE and UQPE is that, when considering a continuous covariate, the UQPE can be expressed as a weighted average of the CQPE, a result

[1]Conditional here means that one is conditioning on a set of observable variables, while partial means that one is looking at the effect of one particular covariate controlling for the rest of the covariates. On the other hand, Conditional Quantile Regression (CQR) and Unconditional Quantile Regression (UQR) refer to two regression methodologies to estimate CQPE and UQPE, respectively. Sometimes the acronym of the method is informally interchanged with the parameter of interest, and, therefore, can be somewhat confusing if read lightly.

derived in Firpo, Fortin, and Lemieux (2009, p.959). Based on this result, first we show that one is able to express the UQPE as a function of CQR coefficients. In particular, we start by considering the CQR as a process indexed by quantiles $\eta \in (0,1)$. Thus, a useful by-product of the CQR analysis is the ability to express UQPE, for a given quantile $\tau \in (0,1)$, as a function of CQR.[2] Indeed, we show that if we start with the common assumption of linear conditional quantiles, then a simple reweighting of the CQR coefficients using density functions delivers the UQPE.

We propose a new two-step semi-parametric estimator that employs CQR coefficients to estimate the UQPE. The practical implementation is simple and makes use of the usual practice of estimating the CQR process, that is, for many conditional quantiles. In the first step one uses standard QR methods to estimate CQPE from the conditional model of interest over a grid of quantiles $\{\eta_1, \ldots, \eta_m\}$, and also estimates the unconditional τ-quantile of the outcome of interest. Then, one applies a matching function to select the CQR coefficient of that correspond to each particular value of the covariates. In the second step, one employs a non-parametric regression of the matched CQR coefficients on the outcome, and evaluates this at the unconditional τ-quantile. This is a one-dimensional (reverse) regression: the regressor is the outcome. Mild sufficient conditions are provided for the two-step estimator to have desired asymptotic properties, namely, consistency and asymptotic normality. We derive the convergence rate of the estimator and show that, as expected, it converges at a standard non-parametric rate. In addition, we suggest statistical inference procedures.

The proposed method offers at least two important advantages over available techniques to compute UQPE. First, the effect of covariates on the unconditional model is usually difficult to conceptualize. As an example, it is not obvious how to handle covariates in the data generating process representation for the RIF regression. There is, thus, a risk of misspecifying the unconditional model. However, in the proposed method covariates enter in structural standard way in the CQR, i.e. the first step in our case. This approach is simpler for the researcher. The conditional QR model allows for a simple and intuitive modelling framework, and it is more familiar to model main output variable as a function of the covariates using CQR. Second, there has been considerable improvements in estimating CQR specifications. It is a very well known semi-parametric model that has been applied in many contexts, such as panel data or models with endogeneity. Moreover, the second non-parametric step does not suffer from the curse of dimensionality, since it is a reverse regression where the one-dimensional outcome is the regressor. Finally, the main motivation of this paper is that we can use all the accumulated knowledge related to CQR to apply to UQPE estimation. The proposed method also has pedagogical merits as it clearly illustrates the link between CQR and UQR.

Although the literature on applications of UQR methods is extensive, the literature on theoretical developments is relatively small. Rothe (2012) generalizes the method of Firpo, Fortin,

[2]Other procedures where statistics of interest are based on a combination of CQR coefficients are the following: Bera, Galvao, Montes-Rojas, and Park (2016) propose to estimate a unique representative CQPE based on an asymmetric Laplace framework, and Lee (2021) considers a general weighted average quantile derivative using the CQR coefficients for a *fixed* quantile level.

and Lemieux (2009), and for other recent developments, see, e.g., Sasaki, Ura, and Zhang (2022), Martinez-Iriarte, Montes-Rojas, and Sun (2022), Inoue, Li, and Xu (2021), and Martinez-Iriarte (2021). For a comprehensive survey on counterfactual distributions and decomposition methods, see Fortin, Lemieux, and Firpo (2011). Moreover, the theoretical derivations of the statistical properties of the two-step estimator are related to a small literature on non-parametric regression with estimated variables, see. e.g., Andrews (1995), Song (2008), Sperlich (2009), Mammen, Rothe, and Schienle (2012), and Escanciano, Jacho-Chávez, and Lewbel (2014).

The remaining of the paper is organized as follows. Section 2 presents the main result that motivates the UQPE estimator based on CQR. Section 3 proposes an estimator for the UQPE and Section 4 derives its asymptotic properties. Section 5 studies the estimator's finite sample performance using Monte Carlo experiments. Section 6 provides an empirical application. Section 7 concludes.

2 Quantile Partial Effects

In this section we define the unconditional quantile partial effect (UQPE) and the conditional quantile partial effect (CQPE), and discuss the relationship between them using a matching function. This relationship is the foundation for the UQPE estimator we discuss in the next section.

2.1 UQPE in terms of CQPE

Consider a general model $Y = r(X, U)$, where $X = (X_1, X_2')'$. Here, Y is the dependent variable, X_1 is the target variable of interest and is a scalar, X_2 is a $(d-1) \times 1$ vector consisting of other observable covariates, and U consists of unobservables. A leading example is the simple linear model $Y = \beta_0 + \beta_1 X_1 + X_2'\beta_2 + U$, such that the conditional τ-quantile of Y given (X_1, X_2) is

$$Q_Y[\tau|X_1, X_2] = \beta_0(\tau) + \beta_1(\tau)X_1 + X_2'\beta_2(\tau). \tag{1}$$

The typical object of study of the standard conditional quantile regression (CQR) is the conditional quantile partial effect (CQPE) defined as

$$CQPE_{X_1}(\tau, x) := \left.\frac{\partial Q_Y[\tau|X_1 = z, X_2 = x_2]}{\partial z}\right|_{z=x_1}, \tag{2}$$

and corresponds to the marginal effect of X_1 on the conditional quantiles of the outcome when $X_1 = x_1$ and $X_2 = x_2$. In the case of model (1), $CQPE_{X_1}(\tau, x) = \beta_1(\tau)$ and estimation of this parameter follows from standard quantile regression methods.

To define an unconditional counterpart to $CQPE_{X_1}(\tau, x)$, we follow Firpo, Fortin, and Lemieux (2009). To that end, consider the counterfactual outcome

$$Y_{\delta, X_1} = r(X_1 + \delta, X_2, U),$$

where δ captures a small location change in the variable X_1. Let $Q_Z[\tau]$ be the unconditional τ-quantile of the random variable Z. Then the unconditional quantile partial effect (UQPE) is defined as[3]

$$UQPE_{X_1}(\tau) := \lim_{\delta \to 0} \frac{Q_{Y_{\delta,X_1}}[\tau] - Q_Y[\tau]}{\delta}. \tag{3}$$

The $UQPE_{X_1}(\tau)$ is the marginal effect of a location shift in X_1 on the unconditional τ-quantile of the outcome. The $CQPE_{X_1}(\tau, x)$ amounts to manipulating X_1 locally at x and evaluating a local impact on Y: the effect on the τ-conditional quantile of Y. The $UQPE_{X_1}(\tau)$ looks at a global change in X_1 and its associated global impact in the τ-unconditional quantile of Y.

Firpo, Fortin, and Lemieux (2009) show that under some mild conditions the following identification result holds:

$$UQPE_{X_1}(\tau) = -\frac{1}{f_Y(Q_Y[\tau])} \int \frac{\partial F_{Y|X}(Q_Y[\tau]|z, x_2)}{\partial z}\bigg|_{z=x_1} dF_X(x), \tag{4}$$

where $F_X(x)$ is short for the joint distribution, $F_{X_1,X_2}(x_1, x_2)$. In a similar manner, assuming differentiability of $F_{Y|X}(y|\cdot)$, from (2) we have that

$$CQPE_{X_1}(\tau, x) = -\frac{1}{f_{Y|X}(Q_Y[\tau|X = x]|x)} \frac{\partial F_{Y|X}(Q_Y[\tau|X = x]|z, x_2)}{\partial z}\bigg|_{z=x_1}. \tag{5}$$

It is interesting to see that $CQPE_{X_1}(\tau, x)$ in equation (5) has a similar structure to $UQPE_{X_1}(\tau)$ in (4). Comparing the formulas in (4) and (5), one is able to see that even if the conditional quantile is equal to the corresponding unconditional, that is, $Q_Y[\tau|X = x] = Q_Y[\tau]$, one is *not* able to recover $UQPE_{X_1}(\tau)$ from $CQPE_{X_1}(\tau, \cdot)$ by simply integrating the latter over X. Moreover, it is usually the case that $Q_Y[\tau|X = x] \neq Q_Y[\tau]$. Thus, first we need to match conditional and unconditional quantiles and then re weight them appropriately to recover $UQPE_{X_1}(\tau)$ from $CQPE_{X_1}(\cdot, \cdot)$.

The following matching map, introduced by Firpo, Fortin, and Lemieux (2009, p.959), is an important tool to relate the CQPE and UQPE:

$$\xi_\tau(x) = \{\eta : Q_Y[\eta|X = x] = Q_Y[\tau]\}. \tag{6}$$

The map $\xi_\tau(x) : (0, 1) \times \mathbb{R}^d \mapsto (0, 1)$ corresponds to the quantile index(es) in the conditional model, η, that produces the closest match with the unconditional quantiles τ for different values of x. In Section 2.3 we analyze this map in detail. For now, we assume that $\xi_\tau(x)$ is a singleton. Therefore, we have that, for every x, $Q_Y[\xi_\tau(x)|X = x] = Q_Y[\tau]$. Under this condition, it is simple to formalize the relationship between CQPE and UQPE. Note that the $CQPE_{X_1}$ in equation (5)

[3]The UQPE can be defined for a vector of covariates as in Firpo, Fortin, and Lemieux (2009) resulting in a vector of UQPEs. Martinez-Iriarte, Montes-Rojas, and Sun (2022) provide an interpretation of the linear combination of UQPEs as a compensated counterfactual change.

evaluated at the $\xi_\tau(x)$ quantile for $X = x$ can be written as

$$CQPE_{X_1}(\xi_\tau(x), x) = -\frac{1}{f_{Y|X}(Q_Y[\tau]|x)} \frac{\partial F_{Y|X}(Q_Y[\tau]|z, x_2)}{\partial z}\Bigg|_{z=x_1}.$$

Now rearranging $CQPE_{X_1}(\xi_\tau(x), x)$ above and substituting into equation (4) yields

$$UQPE_{X_1}(\tau) = \int CQPE_{X_1}(\xi_\tau(x), x) \frac{f_{Y|X}(Q_Y[\tau]|x)}{f_Y(Q_Y[\tau])} dF_X(x). \tag{7}$$

The weights in (7) can be rearranged as

$$\frac{f_{Y|X}(Q_Y[\tau]|x)}{f_Y(Q_Y[\tau])} f_X(x) = \frac{f_{Y,X}(Q_Y[\tau], x)}{f_Y(Q_Y[\tau]) f_X(x)} f_X(x) = f_{X|Y}(x|Q_Y[\tau]).$$

Finally, equation (7) becomes a reverse projection as

$$UQPE_{X_1}(\tau) = E\big[CQPE_{X_1}(\xi_\tau(X), X)|Y = Q_Y[\tau]\big]. \tag{8}$$

The preceding informal discussion is summarized in the lemma below.

Lemma 1. *Let the following assumptions hold: (i) $UQPE_{X_1}(\tau)$ is identified by (4); (ii) the matching function defined in (6) is a singleton; (iii) $F_{Y|X}(y|x)$ and $Q_Y[\tau|X = x]$ are differentiable with respect to x_1; (iv) $f_{Y|X}$, f_Y and f_X are strictly positive. Then (8) holds.*

As mentioned above, sufficient conditions for (i) are laid out in Firpo, Fortin, and Lemieux (2009). Regarding (ii), see Assumption 1 stated below for sufficient conditions for $\xi_\tau(x)$ to be singleton. The rest of the assumptions are customary regularity conditions.

2.2 Intuition of the procedure

Equation (8) shows that the UQPE is in fact a *local* weighted average of CQPE effects "near" the unconditional τ-quantile of Y. As noted above, the τ^{th} unconditional quantile of interest may be different from the (random) $\xi_\tau(X)^{\text{th}}$ conditional quantiles used inside the integral.

When conditional quantiles are linear as in (1), then $CQPE_{X_1}(\xi_\tau(x), x) = \beta_1(\xi_\tau(x))$. This implies that (8) becomes a weighted average of matched slopes. Figures 1 and 2 illustrate how the procedure works in two different linear cases. The figures plot both the unconditional quantile, $Q_Y[\tau]$ (red line) and conditional quantiles, $Q_{Y|X}[\eta]$ (blue lines), as well as the conditional density $f_{X|Y}$ (green curve). The intuition is as following:

1. Identify the unconditional τ quantile, $Q_Y[\tau]$, say $Q_Y[0.50]$ as illustrated in the figures for the unconditional median, and drawn in a horizontal (red) line.

2. Notice that for each η, the conditional quantiles $Q_{Y|X}[\eta]$ (blue lines) intersect the unconditional quantile $Q_Y[\tau]$ (horizontal red line); in the figures, we illustrate this for $\{x_1, x_2, x_3, x_4\}$ values of X that correspond to $\eta = \xi_\tau(x)$ values $\{0.80, 0.60, 0.40, 0.20\}$ respectively.

3. The UQPE is the weighted average – with weights given by density $f_{X|Y}(x|Q_Y[\tau])$ (green curve) – of the intersected slopes on the conditional quantile models.

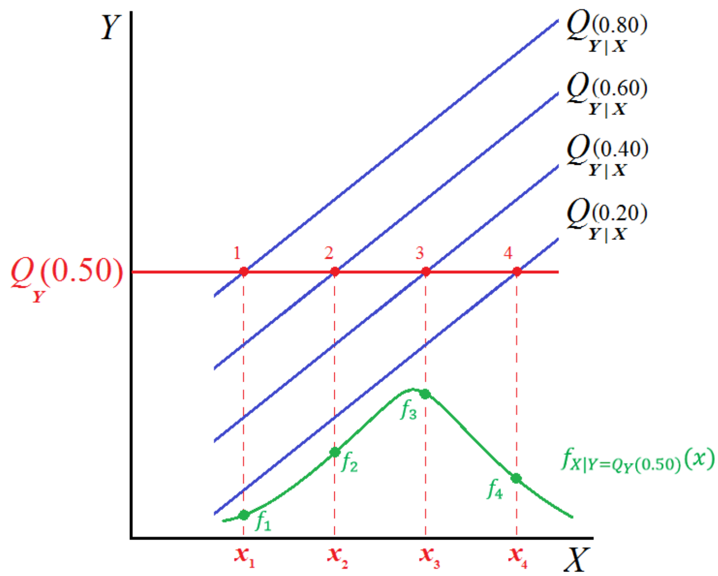


Figure 1: Constant CQPE.

The figures are useful for analyzing the source of the variation in the UQPE across different unconditional quantiles. For example, in Figure 1, the CQPE slopes are the same across conditional quantiles η, which implies that the CQPE is constant across quantiles. Even if the weights change with τ, this is irrelevant, because the slopes are constant.

On the other hand, in Figure 2, the CQPE slopes exhibit some variation across conditional quantiles η. This heterogeneity can be present even if the weights are not a function of τ. The UQPE is then constructed as a weighted average of those. An additional source of variation is given by the potential different conditional densities of X given $Y = Q_Y[\tau]$. The UQPE will then be the based on the different CQPE and the corresponding density weights.

2.3 The matching map ξ_τ

In equation (6), we defined the matching map as

$$\xi_\tau(x) = \{\eta : Q_Y[\eta|X = x] = Q_Y[\tau]\}.$$

For a fixed covariate value $X = x$, the map $\tau \mapsto \xi_\tau(x)$ describes how the unconditional distribution maps on the conditional one. In general, $\xi_\tau(x)$ may vary across the value of covariates as well. Note that it is entirely possible that $\tau \neq \xi_\tau(x)$.

For the purpose of this paper it is important that $\xi_\tau(x)$ is unique. But, generally, three situations may occur. First, $\xi_\tau(x)$ is unique when $F_{Y|X}(y|x)$ is strictly increasing. In this case,

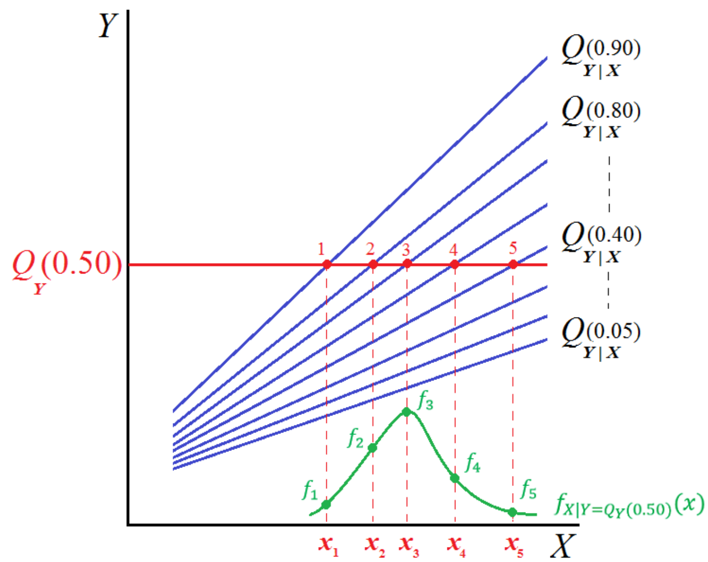


Figure 2: Heterogeneous CQPE.

there can be at most one η that satisfies equation (6). To see this, note that $Q_Y[\eta|X=x]=Q_Y[\tau]$ is identical to $\eta = F_{Y|X}(Q_Y[\tau]|x)$, so that $\xi_\tau(x)$ is unique. Second, $\xi_\tau(x)$ might be an interval. For example if $F_{Y|X}(y|x)$ has a jump discontinuity at $y = Q_Y[\tau]$, but it is otherwise continuous and strictly increasing, then $\xi_\tau(x) = \left[\lim_{y\uparrow Q_Y[\tau]} F_{Y|X}(y|x), F_{Y|X}(Q_Y[\tau]|x)\right]$. See Figure 3a below. Third, $\xi_\tau(x)$ might be emptyfor some x. For example, suppose that $F_{Y|X}(y|x)$ is continuous, and aside for a flat interval, it is strictly increasing. If $Q_Y[\tau]$ is in the interior of the interval mapping to the flat interval, then we cannot have $Q_Y[\eta|X=x]=Q_Y[\tau]$. This is illustrated in Figure 3b below.

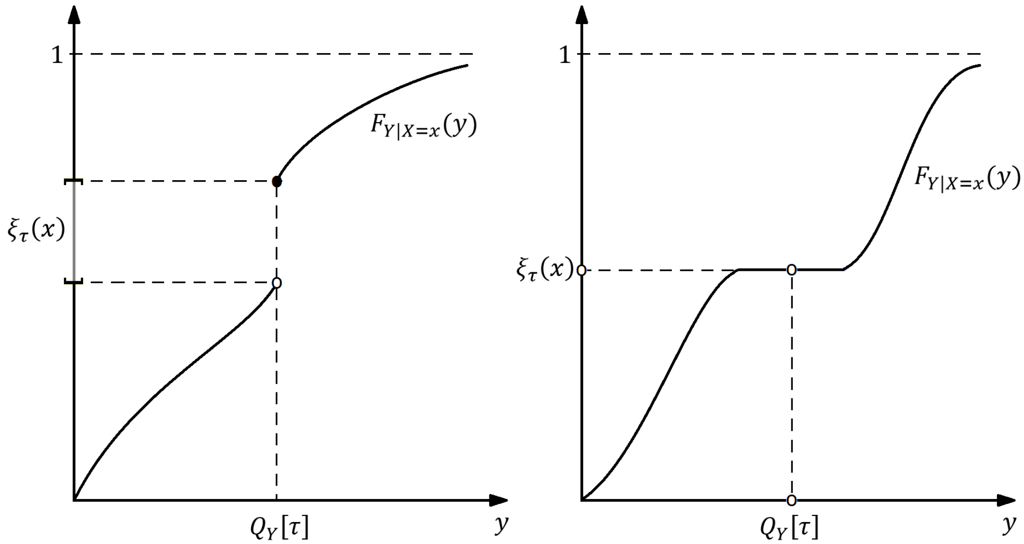
Example 1. *Consider the model $Y = \alpha_0 + \alpha_1 X_1 + (1+\theta X_1)U$ with $X \perp U$. By standard computations, if $1+\theta x_1 > 0$, then $Q_Y[\eta|X_1 = x_1] = \alpha_0 + \alpha_1 x_1 + (1+\theta x_1)Q_U[\eta]$. To find $\xi_\tau(x_1)$ we need the level η such that $Q_Y[\eta|X_1=x_1]=Q_Y[\tau]$. Thus,*

$$\xi_\tau(x_1) = F_U\left(\frac{Q_Y[\tau]-\alpha_0-\alpha_1 x_1}{1+\theta x_1}\right). \tag{9}$$

Remark 1. *If the matching function is the identity function: $\xi_\tau(X) = \{\eta : Q_Y[\eta|X=x] = Q_Y[\tau]\} = \tau$, Then, by equation (8), $UQPE_{X_1}(\tau)$ can be written as*

$$UQPE_{X_1}(\tau) = \int CQPE_{X_1}(\tau,x)\frac{f_{X|Y}(x|Q_Y[\tau])}{f_X(x)} f_X(x)dx,$$

which is the parameter of interest of Lee (2021): a weighted average quantile derivative. Here the weight is $\frac{f_{X|Y}(x|Q_Y[\tau])}{f_X(x)}$. If the matching function is not the identity, then our parameter is not covered by the methods of Lee (2021).



(a) Interval valued $\xi_\tau(x)$.

(b) Empty $\xi_\tau(x)$.

Figure 3: Non-uniqueness of $\xi_\tau(x)$.

3 Estimator

In this section we describe a two-step estimator of $UQPE_{X_1}(\tau)$, which is based on the reverse projection in equation (8). The asymptotic properties are discussed in later sections. In the following, the *unconditional* quantiles of Y are indexed by $\tau \in (0,1)$, while the *conditional* quantiles Y given X are indexed by $\eta \in (0,1)$.

Assume first that

$$Q_Y[\eta|X_1 = x_1, X_2 = x_2] = x_1\beta_1(\eta) + x_2'\beta_2(\eta) = x'\beta(\eta), \tag{10}$$

where $\beta = (\beta_1, \beta_2')'$. Note that x_2 has to include a constant for correct specification. In this paper, we use the conditional quantile function in (10) to estimate the UQPE. Using this quantile regression model has advantages. First, it allows the researcher to directly model the outcome variable Y as a function of observable covariates X, instead of modeling the recentered influence function. This is important because it may be simpler to relate the variable of interest directly from the economic theory or existing literature, than modeling the influence function. Second, practical estimation of (10) is simple, as we discuss below.

Under (10), $CQPE_{X_1}(\xi_\tau(x), x) = \beta_1(\xi_\tau(x))$. Equation (8) then has the convenient form

$$UQPE_{X_1}(\tau) = E\left[\beta_1(\xi_\tau(X))|Y = Q_Y[\tau]\right]. \tag{11}$$

Our proposed estimator is a non-parametric regression of $\{\beta_1(\xi_\tau(x_i))\}_{i=1}^n$ on $\{y_i\}_{i=1}^n$ evaluated at $Q_Y[\tau]$. To implement this method in practice we are required to estimate $\beta_1(\xi_\tau(x))$ and $Q_Y[\tau]$.

To estimate $\beta_1(\xi_\tau(x_i))$ we first use CQR methods, and estimate $\beta(\eta)$ for a grid of m values of η's given by $\mathcal{H}_m = \{\epsilon < \eta_1 < \cdots < \eta_j < \cdots < \eta_m < 1-\epsilon\}$, $\epsilon \in (0, \frac{1}{2})$. In the standard linear case we have that for a given value of η_j, and a sample $\{y_i, x_i\}_{i=1}^n$, we simply apply standard quantile regression methods as

$$(\hat{\beta}_1(\eta_j), \hat{\beta}_2(\eta_j)')' = \hat{\beta}(\eta_j) = \arg\min_b \frac{1}{n} \sum_{i=1}^n \rho_{\eta_j}(y_i - x_i' b),$$

where $\rho_\tau(u) = u(\tau - 1[u < 0])$ is the Koenker and Bassett (1978) check function. We also estimate the unconditional quantile $Q_Y[\tau]$ by

$$\hat{Q}_Y[\tau] = \arg\min_q \frac{1}{n} \sum_{i=1}^n \rho_\tau(y_i - q).$$

To find the matched coefficient $\hat{\beta}_1(\hat{\xi}_\tau(x_i))$, we employ the two previous estimates as following

$$\hat{\xi}_\tau(x_i) = \begin{cases} \eta_1 \in \mathcal{H}_m & \text{if} \quad \hat{Q}_Y[\tau] < x_i'\hat{\beta}(\eta_1) \\ \eta_j \in \mathcal{H}_m & \text{if} \quad \{x_i'\hat{\beta}(\eta_j) \leq \hat{Q}_Y[\tau] < x_i'\hat{\beta}(\eta_{j+1})\} \\ \eta_m \in \mathcal{H}_m & \text{if} \quad \hat{Q}_Y[\tau] \geq x_i'\hat{\beta}(\eta_m) \end{cases}, \tag{12}$$

for $i = 1, ..., n$. The above estimator relies on monotonicity of CQR such that there is only one match. In practice, this needs to be checked in small samples as multiple crossings may occur if x_i is very different from $\bar{x}$. Then an algorithm could be implemented such as taking the average of the selected β_1 or a rearrangement of estimated quantiles (see, for instance, Chernozhukov, Fernández-Val, and Galichon (2010) for a discussion about quantile crossings). Furthermore, if the unconditional quantile does not lie inside the intervals generated by the grid of conditional quantile values, then we impute either the minimum if it is below, or the maximum if it is above.

Finally, to estimate the $UQPE_{X_1}(\tau)$, we can use a Nadaraya-Watson type-estimator, using the preliminary estimators:

$$\hat{E}\left[\hat{\beta}_1(\hat{\xi}_\tau(X)) | Y = \hat{Q}_Y[\tau]\right] = \frac{\sum_{i=1}^n K_h(y_i - \hat{Q}_Y[\tau]) \cdot \hat{\beta}_1(\hat{\xi}_\tau(x_i))}{\sum_{i=1}^n K_h(y_i - \hat{Q}_Y[\tau])}, \tag{13}$$

where K_h is the rescaled kernel $K_h(u) := \frac{1}{h} K\left(\frac{u}{h}\right)$. The estimator in (13) avoids the curse of dimensionality because it is a regression on just one regressor: Y. Indeed, the dimension of X enters in the CQR estimation and in the matching function.

Equation (13) highlights the main benefit of our proposed approach: obtaining the unconditional effect is an easy follow-up from the conditional effects. If the researcher, as is usually the case, has estimated a grid of CQR coefficients, then, after they are "matched" according to (12), they can be averaged following (13) to yield the unconditional effect for the desired quantile level.

Remark 2. *An alternative approach to estimating $UQPE_{X_1}(\tau)$ based on* (11) *is a linear regression of*

$\hat{\beta}_1(\hat{\xi}_\tau(X))$ on a constant and Y. The predicted fit at $Y = \hat{Q}_Y[\tau]$ is an easy-to-compute approximation to $UQPE_{X_1}(\tau)$.

Another option is to do a local linear regression. This estimator may help reduce the bias in lower or higher quantiles. The estimator is $\hat{a}_{\tau,0} + \hat{a}_{\tau,1}\hat{Q}_Y[\tau]$, where $(\hat{a}_{\tau,0}, \hat{a}_{\tau,1})'$ solve

$$(\hat{a}_{\tau,0}, \hat{a}_{\tau,1})' = \arg\min_{a_{\tau,0}, a_{\tau,1}} \sum_{i=1}^{n} K_h(y_i - \hat{Q}_Y[\tau]) \left[\beta_1(\hat{\xi}_\tau(x_i)) - a_{\tau,0} - a_{\tau,1}\left(\frac{y_i - \hat{Q}_Y[\tau]}{h}\right)\right]^2.$$

A study of the properties of this estimator in this particular setting is left for future research.

4 Asymptotic Theory

This section derives the asymptotic properties of the two-step estimator. First, we study the first step, and establish an asymptotic linear representation and rate of convergence for the conditional quantile regression coefficients as a function of the matched quantiles. Second, we study the asymptotic properties of the non-parametric regression in the second step.

4.1 Structural QR and Matched Quantiles

The following assumptions are needed to establish that $\hat{\beta}_1(\hat{\xi}_\tau(x)) - \beta_1(\xi_\tau(x)) = O_p(n^{-1/2})$, where $\hat{\beta}_1(\hat{\xi}_\tau(x))$ is computed according to (12).

Assumption 1. *Let $\{y_i, x_i\}_{i=1}^n$ be a random sample of independent and identically distributed (iid) observations with y_i a scalar and $x_i \in \mathbb{R}^d$ that satisfy the following properties:*

1. *The conditional quantiles are linear: $Q_Y[\eta|X = x] = x'\beta(\eta)$, $\eta \in [\epsilon, 1-\epsilon]$, $\epsilon \in (0, \frac{1}{2})$, with $X \in \mathbb{R}^d$ and $E|X| < \infty$.*

2. *For every x in the support of X, $f_{Y|X}(y|x)$ is bounded away from zero.*

3. *The conditional quantile regression estimators satisfy*

$$\begin{aligned}\hat{\beta}(\eta) - \beta(\eta) &= E\left[f_{Y|X}(X'\beta(\eta)|X)XX'\right]^{-1} \frac{1}{n}\sum_{i=1}^{n}\left(\eta - \mathbb{1}\left\{y_i \le x_i'\beta(\eta)\right\}\right)x_i + o_p(n^{-1/2}) \\ &= \frac{1}{n}\sum_{i=1}^{n}\Psi_i(\eta) + o_p(n^{-1/2}),\end{aligned}$$

   uniformly in $\eta \in [\epsilon, 1-\epsilon]$, $\epsilon \in (0, \frac{1}{2})$, and $\eta \mapsto E\left[f_{Y|X}(X'\beta(\eta)|X)XX'\right]$ has uniformly bounded derivatives.

4. *The unconditional quantile estimator satisfies*

$$\hat{Q}_Y[\tau] - Q_Y[\tau] = f_Y(Q_Y[\tau])^{-1}\frac{1}{n}\sum_{i=1}^{n}(\tau - \mathbb{1}\{y_i \le Q_Y[\tau]\})$$
$$= \frac{1}{n}\sum_{i=1}^{n}\psi_i(\tau) + o_p(n^{-1/2}).$$

5. *The grid of quantiles* $\{\epsilon < \eta_1 < \ldots < \eta_j < \ldots < \eta_m < 1-\epsilon\}$, $\epsilon \in (0, \frac{1}{2})$, *satisfies* $\Delta\eta = o(n^{-1/2})$ *as* $n \to \infty$ *for* $\Delta\eta := \eta_j - \eta_{j-1}$, $j = 2, ..., m$, *and* $\eta_1 = \epsilon$ *and* $\eta_m = 1-\epsilon$ *for a small* $\epsilon > 0$.

The conditions in Assumption 1 are very mild. Assumption 1.1 imposes linearity of the model, and condition 1.2 is very standard in the QR literature, see, e.g., Koenker (2005). Assumptions 1.1 and 1.2, allow us to write $F_{Y|X}(x'\beta(\eta)|x) = \eta$, so that $x'\dot{\beta}(\eta) = f_{Y|X}(x'\beta(\eta)|x)^{-1} > 0$, where $\dot{\beta}(\eta)$ is the Jacobian vector: the derivative of the map $\eta \mapsto \beta(\eta)$. This quantity appears in the denominator of the influence function of $\hat{\xi}_\tau(x)$. Assumption 1.3 is a uniform Bahadur representation for the QR estimator. It is established in Lemma 3 in Ota, Kato, and Hara (2019). See also Theorem 3 in Angrist, Chernozhukov, and Fernández-Val (2006). It implies $\sup_{\eta\in[\epsilon,1-\epsilon]}|\hat{\beta}(\eta) - \beta(\eta)| = O_p(n^{-1/2})$ and the stochastic equicontinuity of the process $\tau \mapsto \sqrt{n}(\hat{\beta}(\eta) - \beta(\eta))$ on $[\epsilon, 1-\epsilon]$. Condition 1.4 is a simple linear representation for the unconditional quantile. Sufficient conditions for Assumption 1.4 are given in Serfling (1980). Finally, Assumption 1.5 requires that the grid for the matching function becomes denser as the sample size increases. This condition has appeared in the QR literature. Chernozhukov, Fernández-Val, and Melly (2013, Remark 3.1 p.2220) provide a similar condition when computing counterfactual distributions.

The next result provides a rate of convergence and a linear representation for $\hat{\beta}_1(\hat{\xi}_\tau(x)) - \beta_1(\xi_\tau(x)) = O_p(n^{-1/2})$.

Theorem 1. *Under Assumption 1, the CQR coefficient of* X_1 *evaluated at the random quantile* $\hat{\xi}_\tau(x)$ *satisfies* $\hat{\beta}_1(\hat{\xi}_\tau(x)) - \beta_1(\xi_\tau(x)) = O_p(n^{-1/2})$ *and can be represented as*

$$\hat{\beta}_1(\hat{\xi}_\tau(x)) - \beta_1(\xi_\tau(x)) = \hat{\beta}_1(\xi_\tau(x)) - \beta_1(\xi_\tau(x)) + \dot{\beta}_1(\xi_\tau(x))(\hat{\xi}_\tau(x) - \xi_\tau(x)) + o_p(n^{-1/2}),$$

where

$$\hat{\xi}_\tau(x) - \xi_\tau(x) = -\frac{1}{x'\dot{\beta}(\xi_\tau(x))}\frac{1}{n}\sum_{i=1}^{n}x'\Psi_i(\xi_\tau(x)) + \frac{1}{x'\dot{\beta}(\xi_\tau(x))}\frac{1}{n}\sum_{i=1}^{n}\psi_i(\tau) + o_p(n^{-1/2}).$$

Here, $\dot{\beta}_1(\xi_\tau(x))$ *is the* β_1 *component of the Jacobian vector* $\dot{\beta}(\xi_\tau(x))$.

4.2 Nadaraya-Watson Estimator

Our parameter of interest given in (11) is

$$UQPE_{X_1}(\tau) = E\left[\beta_1(\xi_\tau(X))|Y = Q_Y[\tau]\right],$$

and we propose the following non-parametric regression Nadaraya-Watson-type estimator:

$$\widehat{UQPE}_{X_1}(\tau) = \hat{E}\left[\hat{\beta}_1(\hat{\xi}_\tau(X))|Y = \hat{Q}_Y[\tau]\right] = \frac{\sum_{i=1}^n K_h(y_i - \hat{Q}_Y[\tau]) \cdot \hat{\beta}_1(\hat{\xi}_\tau(x_i))}{\sum_{i=1}^n K_h(y_i - \hat{Q}_Y[\tau])}.$$

The unfeasible (oracle) version is denoted by

$$\widetilde{UQPE}_{X_1}(\tau) = \hat{E}\left[\beta_1(\xi_\tau(X))|Y = Q_Y[\tau]\right] = \frac{\sum_{i=1}^n K_h(y_i - Q_Y[\tau]) \cdot \beta_1(\xi_\tau(x_i))}{\sum_{i=1}^n K_h(y_i - Q_Y[\tau])}.$$

We maintain the following assumptions.

Assumption 2. *$K(u)$ is a symmetric function around 0 that satisfies: (i) $\int K(u)du = 1$; (ii) For $r > 2$, $\int u^j K(u)du = 0$ when $j = 1, ..., r-1$, and $\int |u|^r K(u)du < \infty$; (iii) $\int K'(u)du = 0$; (iv) $u^j K(u) \to 0$ as $u \to \pm\infty$ for j=1,...,r+1; (v) $\sup_{u\in\mathbb{R}} |K''(u)| < \infty$; (vi) $\int K'(u)^2 du < \infty$ and $\int uK'(u)^2 du < \infty$.*

This assumption requires we use a r^{th} order kernel with $r > 2$. This is to remove the bias introduced by evaluating the kernel estimator at the estimated quantile. A popular one is the 4^{th} order Gaussian kernel (see Section 2.7.2 in Pagan and Ullah (1999)): $K(u) = \frac{3-u^2}{2}\phi(u)$ where $\phi(u)$ is the pdf of a standard normal. This kernel satisfies Assumption 2.[4]

Assumption 3. *(i) The density of Y is $r+1$ times continuously differentiable, with uniformly bounded derivatives; (ii) The joint density $f_{Y,X}(y,x)$ is $r+1$ times continuously differentiable, with uniformly bounded derivatives for every x in the support of X.*

Assumption 4. *As $n \to \infty$, the bandwidth satisfies: (i) $h \to 0$; (ii) $(nh)^{1/2}h^r \to 0$, (iii) $nh^5 \to \infty$.*

In order for a Assumption 4 to hold, we need that $1 + 2r > 5$, which implies $r > 2$. This is in line with the requirement of $r > 2$ in Assumption 2. For example, for $r = 4$, if $h \propto n^{-1/6}$, then Assumption 4 holds.

Assumption 5. *The following approximation rate holds for $\hat{\xi}_\tau$: $E\left[\left(n^{1/4}\left[\beta_1(e(X)) - \beta_1(\xi_\tau(X))\right]\right)^2\right]|_{e=\hat{\xi}_\tau} = o_p(1)$.*

[4] Recall that the odd moments of a standard normal random variable are 0, and the first three even moments are 1, 3, and 15. We have $\int_{\mathbb{R}} K(u)du = 1$, $\int_{\mathbb{R}} uK(u)du = 0$, $\int_{\mathbb{R}} u^2K(u)du = 0$, $\int_{\mathbb{R}} u^3K(u)du = 0$, and $\int_{\mathbb{R}} u^4K(u)du = -3$. The first derivative of the kernel, using the recursive fact that $\phi'(u) = -u\phi(u)$, is $K'(u) = -\frac{5}{2}u\phi(u) + \frac{1}{2}u^3\phi(u)$, so that $\int_{\mathbb{R}} K'(u)du = 0$. Moreover, $u^jK(u) \to 0$ as $u \to \pm\infty$ for all j since the exponential rate of $\phi(u)$ decreases faster than the polynomial rate. By integration by parts, and using $u^jK(u) \to 0$ as $u \to \pm\infty$, we have that $\int u^jK'(u)du = -\int ju^{j-1}K(u)du$ which is 0 for $j = 2, 3, 4$ and $\int u^5K'(u)du = -\int 5u^4K(u)du = 15$. Finally, the second derivative is given by $K''(u) = -\frac{5}{2}\phi(u) + 4u^2\phi(u) - \frac{1}{2}u^4\phi(u)$ which is bounded since $u^jK(u) \to 0$ as $u \to \pm\infty$. Finally, using $K'(u) = -\frac{5}{2}u\phi(u) + \frac{1}{2}u^3\phi(u)$ it is easily verified that $\int K'(u)^2du < \infty$ and $\int K'(u)^2udu < \infty$.

Theorem 2. *Let Assumptions 1, 2, 3, 4, and 5 hold. Then, as $n \to \infty$,*

$$\widehat{UQPE}_{X_1}(\tau) = \widetilde{UQPE}_{X_1}(\tau) + o_p(n^{-1/2}h^{-1/2}).$$

This theorem states that the preliminary estimators of the CQR slopes, the matched quantiles and the unconditional quantile of Y vanish asymptotically because they converge at a faster rate: $n^{-1/2}$ as opposed to $n^{-1/2}h^{-1/2}$. Moreover, the asymptotic distribution of the unfeasible estimator $\widetilde{UQPE}_{X_1}(\tau)$ is well-known and can be readily established.

The following assumption is customary in order to apply the Lindeberg-Feller Central Limit Theorem.

Assumption 6. *(i) For $U_\tau := \beta_1(\xi_\tau(X)) - E[\beta_1(\xi_\tau(X))|Y]$, and $\delta > 0$, $E[|U_\tau|^{2+\delta}|Y] < C < \infty$ a.s. for some C; (ii) $\int |K(u)|^{2+\delta}du < \infty$; (iii) The map $y \mapsto E[\beta_1(\xi_\tau(X))|Y = y]$ is $r+1$ times continuously differentiable, with uniformly bounded derivatives; (iv) The map $y \mapsto \sigma_\tau^2(y) := E[U_\tau^2|Y = y]$ is continuous.*

Corollary 1. *Let Assumptions 1, 2, 3, 4, 5 and 6 hold. Then, as $n \to \infty$,*

$$\sqrt{nh}\left(\widehat{UQPE}_{X_1}(\tau) - UQPE_{X_1}(\tau)\right) \xrightarrow{d} N\left(0, \sigma_\tau^2(Q_Y[\tau])f_Y(Q_Y[\tau])^{-1}\int K(u)^2du\right).$$

Remark 3. *The practical computation of the asymptotic variance-covariance matrix in Corollary 1 is difficult due to the presence of generated regressors in the nonparametric regression. Thus, in practice, we employ resampling approach for inference. There is an extensive literature on constructing nonparametric confidence bands for functions, we refer the reader to Härdle and Bowman (1988) and Hall and Horowitz (2013) and references therein for resampling methods. Moreover, we refer the reader to Mammen, Rothe, and Schienle (2016) for results establishing the validity of the bootstrap for a general class of standard semiparametric estimators when the nuisance parameter is estimated using generated covariates. Relatedly, Mammen, Rothe, and Schienle (2012) derive a formula for the asymptotic variance in a nonparametric regression with nonparametrically generated covariates.*

We describe now the implementation of the bootstrap procedure.

1. Estimate $\{\hat{\beta}(\eta)\}$ for a given grid $\mathcal{H}_m = \{\eta_1, ..., \eta_m\}$ and $\hat{Q}_Y[\tau]$, then compute $\widehat{UQPE}_{X_1}(\tau)$ using the sample $\{y_i, x_i\}_{i=1}^n$.
2. Compute samples with replacement $\{y_i^{*b}, x_i^{*b}\}_{i=1}^n$, for $b = 1, ..., B$, and estimators $\{\hat{\beta}^{*b}(\eta)\}$ for $\mathcal{H}_m$, $\hat{Q}_Y^{*b}[\tau]$ and $\widehat{UQPE}_{X_1}^{*b}(\tau)$.
3. Compute the standard deviation from the bootstrap sample,

$$\hat{\sigma}_{UQPE}^* = \sqrt{\frac{1}{B}\sum_{b=1}^{B}\left(\widehat{UQPE}_{X_1}^{*b}(\tau) - \overline{\widehat{UQPE}}_{X_1}^{*}(\tau)\right)^2}$$

where $\overline{\widehat{UQPE}^{*}_{X_1}(\tau)} = \frac{1}{B}\sum_{b=1}^{B} \widehat{UQPE}^{*b}_{X_1}(\tau)$.

4. Compute the $1-\alpha$ confidence interval $[\widehat{UQPE}^{*[\alpha/2]}_{X_1}(\tau), \widehat{UQPE}^{*[1-\alpha/2]}_{X_1}(\tau)]$ using the ordered statistics of the bootstrap sample.

5 Monte Carlo experiments

This section presents several simulation exercises to study the finite sample performance of the proposed estimator. First, we assess the matching function estimator. Second, we evaluate the unconditional quantile partial effect (UQPE) estimation.

The first data generating process (DGP) we consider is as following:

$$y_i = 1 + x_i + (1 + \theta x_i)u_i, \tag{14}$$

where $x_i \sim N(10, 1)$ and u_i is a random variable with $E(u_i) = 0$, $V(u_i) = 1$ and independent of x_i. The distribution of u_i is specified below as either standard Gaussian or (standardized) Chi-squared with 1 degree of freedom. The parameter θ controls the type of effect of the covariate x on the distribution of $y|x$: when $\theta = 0$ the effect is a *location shift*, and if $\theta \neq 0$ is a *location-scale shift*. In the former case the conditional quantile regression (CQR) effects are constant across quantiles, while in the latter case they vary.

Second, we use a DGP with an additional covariate

$$y_i = 1 + w_i + x_i + (1 + \theta x_i)u_i, \tag{15}$$

where we consider two cases: (i) $w_i \sim N(10, 1)$ (independent of x_i); (ii) $w_i = 10 + (x_i + N(10, 1) - 20)/\sqrt{2}$, where we make w_i correlated with x_i.

5.1 Matching function estimator

The proposed UQPE estimator relies on the estimator of the matching function for the quantiles, $\hat{\xi}_\tau(x)$. This subsection presents simulations exercises for assessing the accuracy of the matching estimator as given in equation (12). Recall from Example 1, equation (9), that in the simple linear case we have an explicit formula for the population matching function, $\xi_\tau(x)$. Thus, we are able to use simulations to assess the finite sample performance of the estimator.

We consider experiments using DGP model in (14) for a pure location model, $\theta = 0$, as well as a location-scale model, $\theta = 1$. We use $x_i \sim N(10, 1)$ and $u_i \sim N(10, 1)$. Each experiment has 100 simulations of the DGP with sample sizes $n = \{250, 500, 2500, 5000\}$, and quantile grid sizes $m = \{9, 24, 99, 199\}$, respectively. We consider three quantiles $\tau \in \{0.25, 0.50, 0.75\}$. Figure 4 reports results for the location case, and Figure 5 displays results for the location-scale case. In each figure, we plot the parameter of interest (the true value of the matching function), the estimates

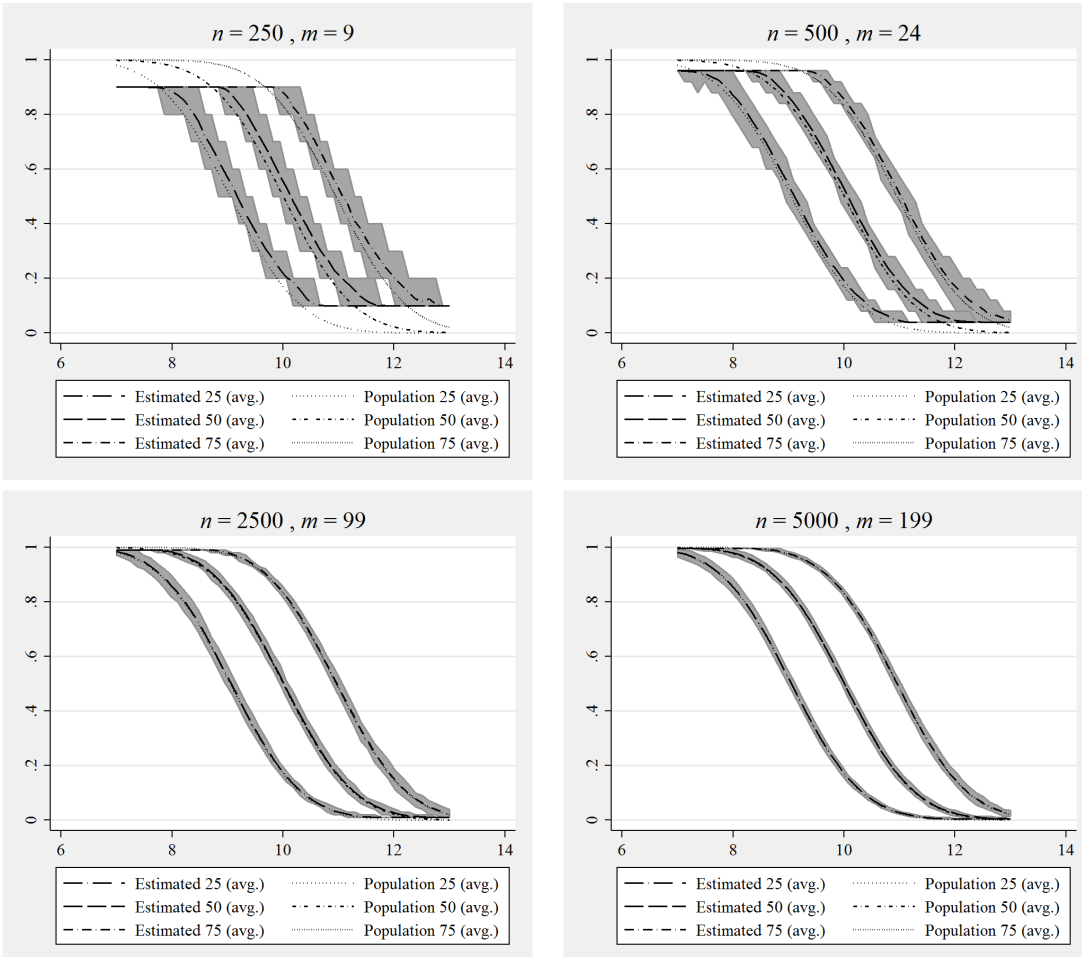


Figure 4: Estimation of matching functions, $u \sim N(0,1)$ and $\theta = 0$ (pure location).

(average estimates over the number of simulations), as well as the 95% empirical confidence interval.[5]

Simulation results show evidence that the matching function estimator provides an approximately asymptotically unbiased estimator for both the pure location and location-scale models with a better performance of sample sizes of $n \geq 500$. Point estimates are close to the populations counterparts even for small samples and grids. As sample size and grid increase together, point estimates become very close to the population and confidence intervals shrink. Section A.5 of the appendix contains a proposal for conducting inference on the matching function should this be of interest.

5.2 UQPE estimation

Now we investigate the finite sample performance of the proposed UQPE estimator as in equation (13). In what follows, We label this estimator as Nadaraya-Watson (NW). For comparison,

[5] These are computed as the 2.5^{th} and 97.5^{th} empirical percentiles of estimates across simulations.

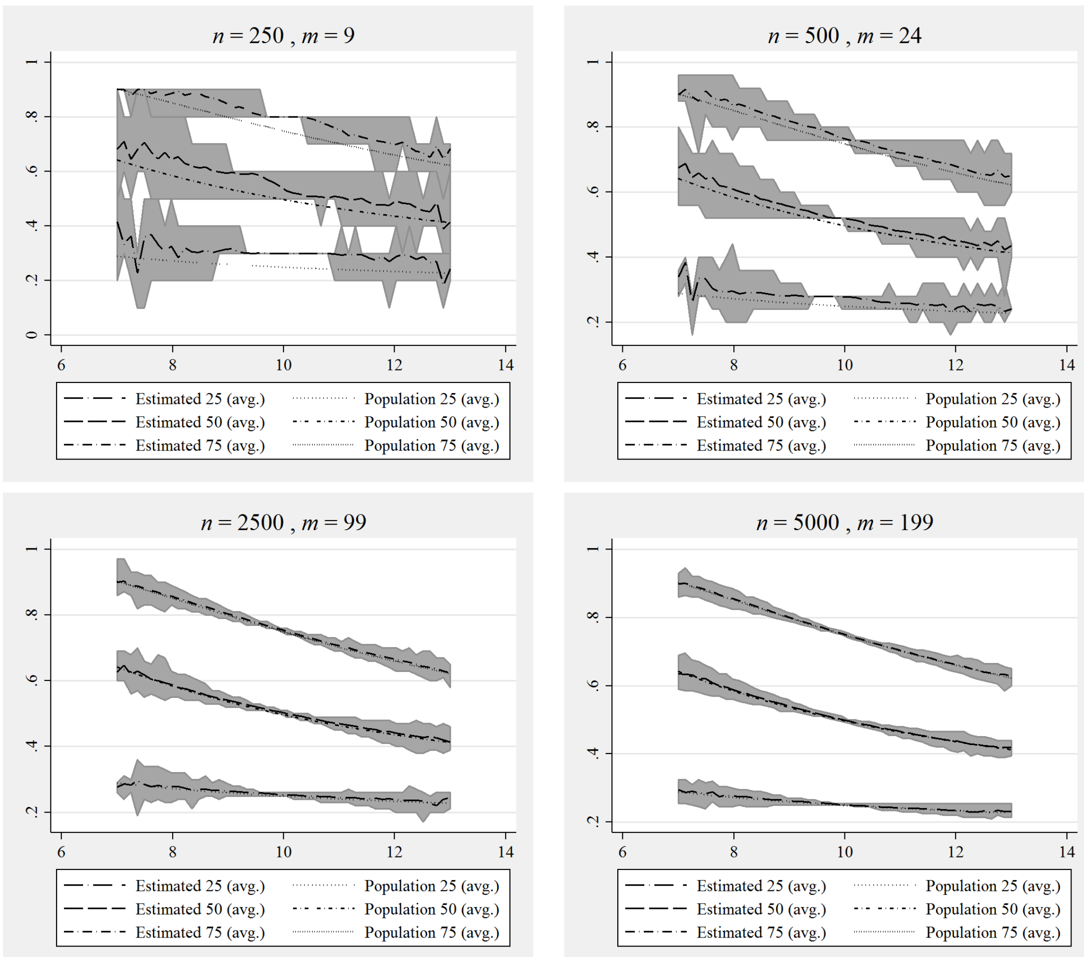


Figure 5: Estimation of matching functions, $u \sim N(0,1)$ and $\theta = 1$ (location-scale).

we also implement the RIF regression model for UQR for each unconditional quantile using the `rifvar` STATA command (Rios-Avila, 2020). In this case, the UQPE is estimated by OLS with a linear and a cubic polynomial model of the RIF for each quantile as a function of x_i (or as a function of x_i and w_i for the model with additional covariates).

Each experiment has 1,000 simulations of the DGP with sample sizes $n = \{250, 500, 2500, 5000\}$, and quantiles grid sizes $m = \{9, 24, 99, 199\}$, respectively. We consider three quantiles $\tau \in \{0.25, 0.50, 0.75\}$. Moreover, we use the Naradaya-Watson (NW) estimator described in equation (13) with the bandwidth $h_n = 0.9\hat{\sigma}_y n^{-1/6}$ and the 4th order Gaussian kernel function, $K(u) = \frac{3-u^2}{2}\phi(u)$, where $\phi(u)$ is the pdf of a standard normal, as discussed above. To evaluate the procedures we report the sample average of the point estimates, bias, variance, and mean-squared error (MSE).

Table 1 presents results for the baseline model for the simple location-shift model (i.e. $\theta = 0$) and Gaussian covariate and innovation. Both RIF and NW estimators have a good performance in terms of bias, variance and MSE. These three statistics decrease for both estimators as sample size increases, for all three quantiles.

Tables 2 and 3 present simulations results for Gaussian and Chi-squared innovations, respectively, for the location-scale shift model (i.e. $\theta = 1$) with a Gaussian covariate. For all cases we observe that for the proposed NW estimator the bias and variance reduces as n increases. The relative performance to the RIF-regression model varies depending on the simulation exercises, but in most cases the NW estimator outperforms the RIF one.

Tables 4 and 5 collect simulation results for cases where there is an additional covariate, w_i. The former case uses an independent additional covariate and in the latter case w_i is correlated with x_i. In both cases we use the model with $\theta = 1$ and $x_i \sim N(10, 1)$. The results are also in line with previous ones, highlighting a good performance of the NW estimator in terms of bias, variance, and MSE.

Overall, these simulation results indicate that our proposed method produces a consistent estimator, where both bias and variance reduce as n increases. In some cases, however, the bias improvement applies only for $n \geq 500$.

Finally, Table 6 presents simulation exercises where we consider different bandwidth and kernels choices. In particular, we use $h_n = 0.9\hat{\sigma}_y n^{-1/4}$, $h_n = 0.9\hat{\sigma}_y n^{-1/5}$ and $h_n = 0.9\hat{\sigma}_y n^{-1/6}$. In order to explore our proposed optimal choice $h_n \propto n^{-1/6}$ we compare with with $h_n \propto n^{-1/5}$, the standard bandwidth choice in non-parametric kernel estimators and also with $h_n \propto n^{-1/4}$. Next we also evaluate the proposed 4th order Gaussian kernel with the standard Gaussian one. We consider the location-scale model with $\theta = 1$ and Gaussian errors. The results show evidence that there are only small differences across bandwidths and kernels which suggest that the estimator is robust to these choices. For the empirical researcher this suggests that our proposed estimator can be combined with the standard non-parametric implementation.[6]

[6] Other model specification results are available from the authors upon request. In all cases, they gave similar results.

Table 1: Location model: $\theta = 0$, and $x_i \sim N(10, 1)$, $u_i \sim N(0, 1)$.

Estimator	τ	n	Parameter	Average	Bias	Variance	MSE
RIF	25	250	1.000	1.0251	0.02518	0.01559	0.01623
		500	1.000	1.0249	0.02496	0.00876	0.00938
		2500	1.000	1.0135	0.01354	0.00202	0.00221
		5000	1.000	1.0091	0.00915	0.00122	0.00130
	50	250	1.000	1.0454	0.04503	0.01215	0.01417
		500	1.000	1.0391	0.03865	0.00735	0.00884
		2500	1.000	1.0232	0.02276	0.00166	0.00218
		5000	1.000	1.0176	0.01716	0.00090	0.00119
	75	250	1.000	1.0375	0.03755	0.01592	0.01733
		500	1.000	1.0251	0.02517	0.00846	0.00909
		2500	1.000	1.0116	0.01167	0.00219	0.00233
		5000	1.000	1.0107	0.01082	0.00120	0.00132
Nadaraya-Watson	25	250	1.000	0.9963	-0.00367	0.00455	0.00457
		500	1.000	1.0005	0.00060	0.00240	0.00240
		2500	1.000	0.9993	-0.00061	0.00046	0.00046
		5000	1.000	0.9999	-0.00006	0.00022	0.00022
	50	250	1.000	0.9996	-0.00077	0.00421	0.00421
		500	1.000	1.0021	0.00166	0.00219	0.00219
		2500	1.000	0.9994	-0.00100	0.00043	0.00043
		5000	1.000	0.9999	-0.00055	0.00020	0.00020
	75	250	1.000	1.0018	0.00187	0.00495	0.00496
		500	1.000	1.0034	0.00350	0.00257	0.00258
		2500	1.000	0.9996	-0.00036	0.00047	0.00047
		5000	1.000	0.9999	-0.00003	0.00023	0.00023

Notes: Monte Carlo experiments based on 1000 simulations.

Table 2: Location-shift model: $\theta = 1$, and $x_i \sim N(10,1)$, $u_i \sim N(0,1)$.

Estimator	τ	n	Parameter	Expectation	Bias	Variance	MSE
RIF	25	250	0.330	0.3210	-0.00912	1.01506	1.01514
		500	0.330	0.3543	0.02423	0.47570	0.47629
		2500	0.330	0.3234	-0.00672	0.09271	0.09276
		5000	0.330	0.3288	-0.00125	0.04431	0.04431
	50	250	1.020	1.0693	0.04934	0.86653	0.86897
		500	1.020	1.0844	0.06442	0.43005	0.43420
		2500	1.020	1.0253	0.00535	0.08020	0.08023
		5000	1.020	1.0239	0.00396	0.03970	0.03971
	75	250	1.682	1.7460	0.06373	1.06619	1.07026
		500	1.682	1.7099	0.02761	0.50398	0.50474
		2500	1.682	1.7014	0.01909	0.10109	0.10145
		5000	1.682	1.6902	0.00787	0.04821	0.04827
Nadaraya-Watson	25	250	0.330	0.4745	0.14443	0.80540	0.82626
		500	0.330	0.4176	0.08749	0.42508	0.43273
		2500	0.330	0.3306	0.00046	0.08557	0.08557
		5000	0.330	0.3316	0.00146	0.04142	0.04143
	50	250	1.020	1.1269	0.10694	0.68828	0.69972
		500	1.020	1.0844	0.06441	0.35801	0.36216
		2500	1.020	1.0161	-0.00386	0.07001	0.07003
		5000	1.020	1.0126	-0.00733	0.03350	0.03355
	75	250	1.682	1.8656	0.18328	0.82577	0.85936
		500	1.682	1.7607	0.07842	0.42324	0.42939
		2500	1.682	1.6887	0.00642	0.08076	0.08080
		5000	1.682	1.6823	-0.00001	0.03743	0.03743

Notes: Monte Carlo experiments based on 1000 simulations.

Table 3: Location-scale shift model: $\theta = 1$, and $x_i \sim N(10, 1)$, $u_i \sim (\chi^2_1 - 1)/\sqrt{2}$.

Estimator	τ	n	Parameter	Expectation	Bias	Variance	MSE
RIF	25	250	0.367	0.7425	0.37510	0.12921	0.26991
		500	0.367	0.6731	0.30575	0.05070	0.14418
		2500	0.367	0.5425	0.17506	0.00637	0.03702
		5000	0.367	0.5018	0.13445	0.00289	0.02096
	50	250	0.622	0.5380	-0.08427	0.22020	0.22730
		500	0.622	0.5261	-0.09615	0.10207	0.11132
		2500	0.622	0.5343	-0.08798	0.02002	0.02776
		5000	0.622	0.5453	-0.07698	0.01075	0.01667
	75	250	1.204	1.1749	-0.02864	1.39024	1.39106
		500	1.204	1.1446	-0.05896	0.70950	0.71297
		2500	1.204	1.2052	0.00166	0.14136	0.14136
		5000	1.204	1.1986	-0.00497	0.06550	0.06552
Nadaraya-Watson	25	250	0.367	0.4103	0.04288	0.05469	0.05653
		500	0.367	0.3820	0.01455	0.02024	0.02045
		2500	0.367	0.3740	0.00661	0.00316	0.00321
		5000	0.367	0.3705	0.00307	0.00157	0.00158
	50	250	0.622	0.7019	0.07965	0.34568	0.35202
		500	0.622	0.6460	0.02373	0.14878	0.14934
		2500	0.622	0.6201	-0.00214	0.02376	0.02377
		5000	0.622	0.6123	-0.00995	0.01236	0.01246
	75	250	1.204	1.4710	0.26743	1.94604	2.01756
		500	1.204	1.2765	0.07291	0.80686	0.81218
		2500	1.204	1.2486	0.04506	0.13421	0.13624
		5000	1.204	1.2260	0.02250	0.05981	0.06032

Notes: Monte Carlo experiments based on 1000 simulations.

Table 4: Location-scale shift model with independent covariate: $\theta = 1$, $w_i \sim N(10, 1)$, and $x_i \sim N(10, 1)$, $u_i \sim N(0, 1)$.

Estimator	τ	n	**Parameter**	**Expectation**	**Bias**	**Variance**	**MSE**
RIF	25	250	0.3247	0.3113	-0.01345	0.97543	0.97561
		500	0.3247	0.3515	0.02682	0.52336	0.52408
		2500	0.3247	0.3359	0.01113	0.09618	0.09630
		5000	0.3247	0.3298	0.00504	0.04465	0.04468
	50	250	0.9975	1.0265	0.02893	0.92922	0.93006
		500	0.9975	1.0734	0.07590	0.41277	0.41853
		2500	0.9975	1.0341	0.03654	0.07985	0.08119
		5000	0.9975	1.0237	0.02615	0.03849	0.03918
	75	250	1.6642	1.6825	0.01830	1.01380	1.01414
		500	1.6642	1.7358	0.07167	0.45838	0.46351
		2500	1.6642	1.7010	0.03687	0.10087	0.10223
		5000	1.6642	1.6868	0.02263	0.04800	0.04851
Nadaraya-Watson	25	250	0.3247	0.4466	0.12185	0.76641	0.78126
		500	0.3247	0.4191	0.09434	0.40590	0.41480
		2500	0.3247	0.3488	0.02407	0.08349	0.08407
		5000	0.3247	0.3343	0.00961	0.03924	0.03934
	50	250	0.9975	1.0955	0.09798	0.68349	0.69309
		500	0.9975	1.0866	0.08902	0.32540	0.33332
		2500	0.9975	1.0201	0.02254	0.07001	0.07052
		5000	0.9975	1.0133	0.01578	0.03170	0.03195
	75	250	1.6642	1.7889	0.12473	0.85360	0.86916
		500	1.6642	1.7475	0.08335	0.37254	0.37949
		2500	1.6642	1.6960	0.03188	0.07782	0.07883
		5000	1.6642	1.6777	0.01349	0.03863	0.03882

Notes: Monte Carlo experiments based on 1000 simulations.

Table 5: Location-shift model with correlated covariate: $\theta = 1$, $w_i = 10 + (x_i + N(10,1) - 20)/\sqrt{2}$, and $x_i \sim N(10,1)$, $u_i \sim N(0,1)$.

Estimator	τ	n	Parameter	Expectation	Bias	Variance	MSE
RIF	25	250	0.3247	0.2623	-0.06247	2.05660	2.06050
		500	0.3247	0.3357	0.01093	0.98728	0.98740
		2500	0.3247	0.3272	0.00252	0.16915	0.16916
		5000	0.3247	0.3368	0.01203	0.08660	0.08674
	50	250	1.0176	1.0264	0.00885	1.90909	1.90917
		500	1.0176	1.0553	0.03769	0.82247	0.82389
		2500	1.0176	1.0055	-0.01204	0.15643	0.15657
		5000	1.0176	1.0205	0.00294	0.07593	0.07594
	75	250	1.6966	1.6676	-0.02897	2.03511	2.03595
		500	1.6966	1.7081	0.01155	1.03865	1.03879
		2500	1.6966	1.6819	-0.01464	0.19409	0.19430
		5000	1.6966	1.6666	-0.02997	0.09898	0.09988
Nadaraya-Watson	25	250	0.3247	0.4033	0.07859	1.61797	1.62415
		500	0.3247	0.3987	0.07394	0.78275	0.78822
		2500	0.3247	0.3391	0.01433	0.14653	0.14673
		5000	0.3247	0.3401	0.01539	0.07626	0.07650
	50	250	1.0176	1.0923	0.07473	1.42727	1.43285
		500	1.0176	1.0541	0.03653	0.65508	0.65641
		2500	1.0176	1.0033	-0.01424	0.13169	0.13190
		5000	1.0176	1.0095	-0.00809	0.06319	0.06325
	75	250	1.6966	1.7819	0.08536	1.60994	1.61723
		500	1.6966	1.7356	0.03897	0.78620	0.78772
		2500	1.6966	1.6860	-0.01054	0.14993	0.15004
		5000	1.6966	1.6602	-0.03635	0.07755	0.07887

Notes: Monte Carlo experiments based on 1000 simulations.

Table 6: Kernel and bandwidth choice ($\theta = 1$ and $u_i \sim N(0,1)$).

Kernel	η	n	Bias			Variance			MSE		
			$n^{-1/4}$	$n^{-1/5}$	$n^{-1/6}$	$n^{-1/4}$	$n^{-1/5}$	$n^{-1/4}$	$n^{-1/4}$	$n^{-1/5}$	$n^{-1/6}$
Gauss	25	250	0.14441	0.14415	0.14396	0.80470	0.80311	0.80181	0.82556	0.82389	0.82253
		500	0.08744	0.08701	0.08667	0.42508	0.42354	0.42244	0.43273	0.43111	0.42995
		2500	0.00062	0.00029	0.00009	0.08563	0.08548	0.08537	0.08563	0.08548	0.08537
		5000	0.00160	0.00133	0.00116	0.04145	0.04139	0.04135	0.04145	0.04139	0.04135
	50	250	0.10681	0.10492	0.10360	0.68763	0.68532	0.68358	0.69904	0.69633	0.69431
		500	0.06464	0.06303	0.06194	0.35777	0.35710	0.35651	0.36195	0.36108	0.36035
		2500	-0.00376	-0.00435	-0.00483	0.06997	0.06996	0.06994	0.06998	0.06998	0.06996
		5000	-0.00732	-0.00769	-0.00805	0.03348	0.03349	0.03348	0.03353	0.03355	0.03355
	75	250	0.18308	0.18106	0.17962	0.82528	0.82485	0.82442	0.85880	0.85763	0.85668
		500	0.07868	0.07782	0.07728	0.42280	0.42367	0.42411	0.42899	0.42973	0.43008
		2500	0.00671	0.00659	0.00663	0.08074	0.08089	0.08103	0.08078	0.08094	0.08107
		5000	0.00006	0.00022	0.00039	0.03745	0.03748	0.03754	0.03745	0.03748	0.03754
Gauss Mod	25	250	0.14477	0.14463	0.14443	0.80707	0.80633	0.80540	0.82803	0.82724	0.82626
		500	0.08804	0.08775	0.08749	0.42747	0.42610	0.42508	0.43522	0.43380	0.43273
		2500	0.00104	0.00068	0.00046	0.08582	0.08568	0.08557	0.08582	0.08568	0.08557
		5000	0.00192	0.00163	0.00146	0.04152	0.04146	0.04142	0.04152	0.04146	0.04143
	50	250	0.11000	0.10825	0.10694	0.69141	0.68978	0.68828	0.70351	0.70150	0.69972
		500	0.06709	0.06556	0.06441	0.35864	0.35841	0.35801	0.36314	0.36271	0.36216
		2500	-0.00311	-0.00352	-0.00386	0.06996	0.07000	0.07001	0.06997	0.07001	0.07003
		5000	-0.00699	-0.00715	-0.00733	0.03346	0.03349	0.03350	0.03351	0.03354	0.03355
	75	250	0.18633	0.18485	0.18328	0.82606	0.82631	0.82577	0.86079	0.86048	0.85936
		500	0.07997	0.07916	0.07842	0.42129	0.42263	0.42324	0.42768	0.42889	0.42939
		2500	0.00691	0.00664	0.00642	0.08060	0.08066	0.08076	0.08065	0.08071	0.08080
		5000	-0.00009	-0.00002	-0.00001	0.03745	0.03741	0.03743	0.03745	0.03741	0.03743

Notes: Monte Carlo experiments based on 1000 simulations.

6 Empirical Application

This section illustrates the UQPE estimator with an analysis of Engel's curve. The original concept corresponds to Ernst Engel (1857, cited in Koenker (2005), pp. 78-79) who studied the European working class households consumption in the 19th century. Engel curves describe how household expenditures on particular goods and services depend on household income. The analysis of Engel curves has a long history of estimating the expenditure-income relationship. They are regression functions where the dependent variable is the level or the budget share of total expenses used to purchase a commodity of goods or services, and the explanatory variable, total expenditure, is usually used as a proxy for income.[7] An empirical result commonly referred to as "Engel's law" states that the poorer a family is, the larger the budget share it spends on food. Other categories of expenditure present a less robust pattern. Hence, we investigate the hypothesis that food expenditure constitutes a declining share of household income.

We apply this framework to household expenditures in Argentina using the national survey of expenditures (Encuesta Nacional de Gasto de los Hogares, known as ENGHO 2017-2018), implemented by the Instituto Nacional de Estadística y Censos (INDEC). The survey was carried out between November 2017 and November 2018. The ENGHO 2017-2018 surveys the households' living conditions in terms of their access to goods and services, as well as their income. The data contains information about household expenditures on different goods and services. About 21,547 households were randomly selected and participated on the survey. We consider both food household expenditures and total non-durable consumption for comparison.[8] We also use the following covariates set: age, education of the household head, the number of children under 14 years of age in the household, the number of exclusive rooms in the household, a dummy that indicates whether the head is the owner of the dwelling, and regional dummies.

We estimate UQPE and CQPE. The former analysis corresponds to evaluating effect of an increase in income for every household in a uniform pattern (using the same additional controls, if any) on the unconditional quantile of food expenditure while focusing on the entire distribution of expenditure. The latter effect corresponds to the study of how expenditure changes when marginally increasing income conditional on income and other controls. For comparison, we also provide estimate results for the RIF regression of Firpo, Fortin, and Lemieux (2009) using the `rifvar` STATA command (Rios-Avila, 2020). We use both income and expenditures in logarithm, so that the coefficient estimates can be interpreted as an elasticity. Confidence intervals are computed using 200 wild bootstrap replications.

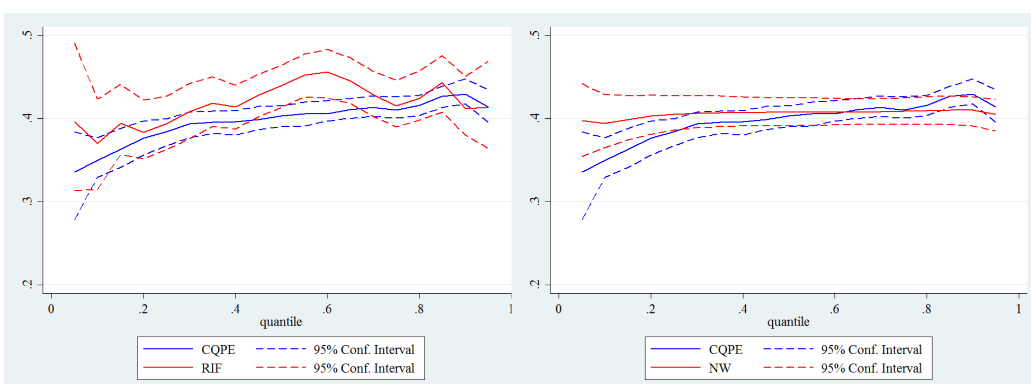
[7] There is a large literature on Engel's curve, see, e.g., for example, among many others, Lewbel (1997, 2008), Blundell, Chen, and Kristensen (2007), Chai and Moneta (2010), Chernozhukov, Fernández-Val, and Kowalski (2015).

[8] In particular, the sample has information on: (*i*) food and non-alcoholic beverages, (*ii*) alcoholic beverages and tobacco, (*iii*) clothing and footwear, (*iv*) housing, water, electricity, gas and other fuels, (*v*) home equipment and maintenance, (*vi*) health, (*vii*) transportation, (*viii*) communications, (*ix*) recreation and culture, (*x*) education, (*xi*) restaurants and hotels, and (*xii*) miscellaneous goods and services. Both expenses and income are transformed to represent monthly values. Since the monetary values of each household are expressed in current currency at the time of the survey, an inflation adjustment was made to transform them into constant currency for December of the fourth quarter of 2018 using the Consumer Price Index (CPI) computed by national statistical office, INDEC.

Table 7: Engel's curve for food expenditures (no additional covariates).

	Quantile Partial Effect				
	10	**25**	**50**	**75**	**90**
Conditional distribution					
CQR	0.383***	0.407***	0.408***	0.408***	0.425***
	(0.000571)	(0.000422)	(0.000246)	(0.000278)	(0.000336)
Unconditional distribution					
RIF (linear model)	0.367***	0.388***	0.427***	0.396***	0.393***
	(0.0285)	(0.0170)	(0.0139)	(0.0130)	(0.0181)
RIF (quadratic model)	0.360***	0.383***	0.427***	0.403***	0.406***
	(0.0275)	(0.0166)	(0.0140)	(0.0129)	(0.0182)
RIF (cubic model)	0.370***	0.394***	0.440***	0.415***	0.412***
	(0.0279)	(0.0169)	(0.0143)	(0.0137)	(0.0183)
NW	0.395***	0.405***	0.408***	0.409***	0.410***
	(0.0166)	(0.0111)	(0.00851)	(0.00809)	(0.00870)
Observations	21,017	21,017	21,017	21,012	21,017

Notes: Standard errors in parentheses (analytical for CQR, bootstrap with 200 replications for RIF and NW). * indicates significance at 10 %, ** at 5 % and *** at 1 %.



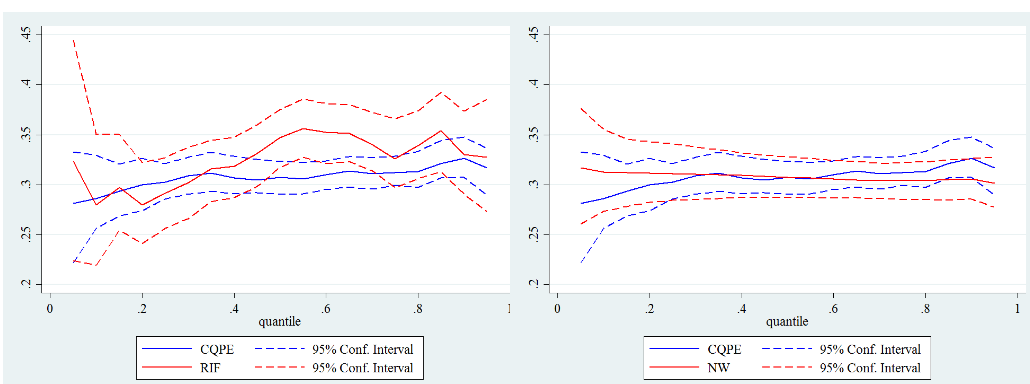
(a) RIF and CQR. (b) NW and CQR.

Figure 6: Engel's curves for food expenditures (no covariates).

Table 8: Engel's curve for food expenditures (with additional covariates).

	Quantile Partial Effect				
	10	**25**	**50**	**75**	**90**
Conditional distribution					
CQR	0.308***	0.303***	0.308***	0.291***	0.324***
	(0.000733)	(0.000353)	(0.000416)	(0.000366)	(0.000507)
Unconditional distribution					
RIF (linear model)	0.272***	0.282***	0.330***	0.302***	0.309***
	(0.0315)	(0.0186)	(0.0148)	(0.0149)	(0.0213)
RIF (quadratic model)	0.266***	0.278***	0.330***	0.310***	0.323***
	(0.0307)	(0.0183)	(0.0149)	(0.0149)	(0.0210)
RIF (cubic model)	0.280***	0.292***	0.347***	0.326***	0.330***
	(0.0318)	(0.0188)	(0.0157)	(0.0162)	(0.0215)
NW	0.313***	0.311***	0.307***	0.304***	0.306***
	(0.0210)	(0.0142)	(0.0107)	(0.00948)	(0.0109)
Observations	21,017	21,017	21,017	21,012	21,017

Notes: Other variables included are age, education and ownership situation of the household head, number of children in the household, number of exclusive rooms and regional dummies. Standard errors in parentheses (analytical for CQR, bootstrap with 200 replications for RIF and NW). * indicates significance at 10 %, ** at 5 % and *** at 1 %.



(a) RIF and CQR.

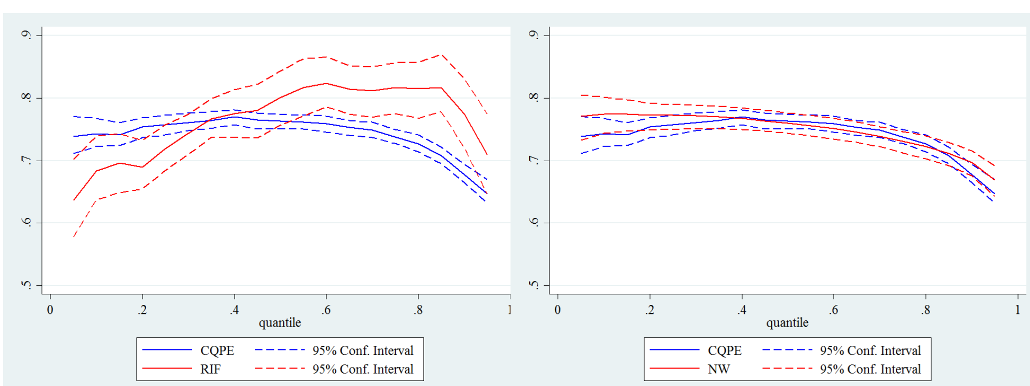
(b) NW and CQR.

Figure 7: Engel's curves for food expenditures (with additional covariates).

Table 9: Engel's curve for total non-durables (no covariates).

	Quantile Partial Effect				
	10	**25**	**50**	**75**	**90**
Conditional distribution					
CQR	0.778***	0.784***	0.776***	0.738***	0.662***
	(0.000359)	(0.000265)	(0.000211)	(0.000283)	(0.000295)
Unconditional distribution					
RIF (linear model)	0.700***	0.706***	0.761***	0.768***	0.731***
	(0.0295)	(0.0194)	(0.0216)	(0.0229)	(0.0292)
RIF (quadratic model)	0.674***	0.693***	0.765***	0.792***	0.769***
	(0.0278)	(0.0183)	(0.0217)	(0.0251)	(0.0338)
RIF (cubic model)	0.683***	0.719***	0.801***	0.817***	0.774***
	(0.0267)	(0.0183)	(0.0217)	(0.0229)	(0.0275)
NW	0.774***	0.772***	0.760***	0.731***	0.697***
	(0.0142)	(0.0102)	(0.00837)	(0.00855)	(0.0103)
Observations	21,461	21,461	21,461	21,461	21,461

Notes: Standard errors in parentheses (analytical for CQR, bootstrap with 200 replications for RIF and NW). * indicates significance at 10 %, ** at 5 % and *** at 1 %.



(a) RIF and CQR.

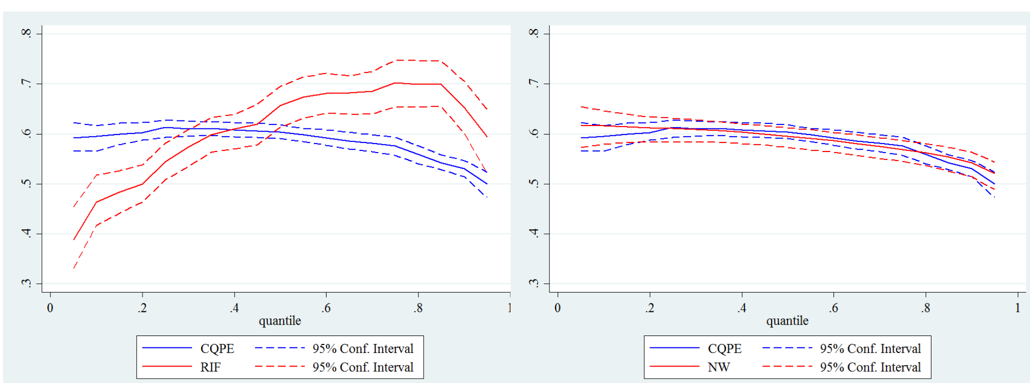
(b) NW and CQR.

Figure 8: Engel's curves for non-durable expenditures (no covariates).

Table 10: Engel's curve for total non-durables (with additional covariates).

	Quantile Partial Effect				
	10	**25**	**50**	**75**	**90**
Conditional distribution					
CQR	0.633***	0.639***	0.614***	0.568***	0.517***
	(0.000461)	(0.000359)	(0.000259)	(0.000297)	(0.000383)
Unconditional distribution					
RIF (linear model)	0.485***	0.524***	0.602***	0.644***	0.615***
	(0.0284)	(0.0197)	(0.0208)	(0.0249)	(0.0291)
RIF (quadratic model)	0.460***	0.510***	0.605***	0.667***	0.653***
	(0.0279)	(0.0189)	(0.0209)	(0.0275)	(0.0375)
RIF (cubic model)	0.464***	0.545***	0.657***	0.703***	0.653***
	(0.0252)	(0.0186)	(0.0211)	(0.0253)	(0.0269)
NW	0.616***	0.611***	0.596***	0.569***	0.543***
	(0.0161)	(0.0123)	(0.0107)	(0.0106)	(0.0117)
Observations	21,461	21,461	21,461	21,461	21,461

Notes: Other variables included are age, education and ownership situation of the household head, number of children in the household, number of exclusive rooms and regional dummies. Standard errors in parentheses (analytical for CQR, bootstrap with 200 replications for RIF and NW). * indicates significance at 10 %, ** at 5 % and *** at 1 %.



(a) RIF and CQR.

(b) NW and CQR.

Figure 9: Engel's curves for non-durable expenditures (with additional covariates).

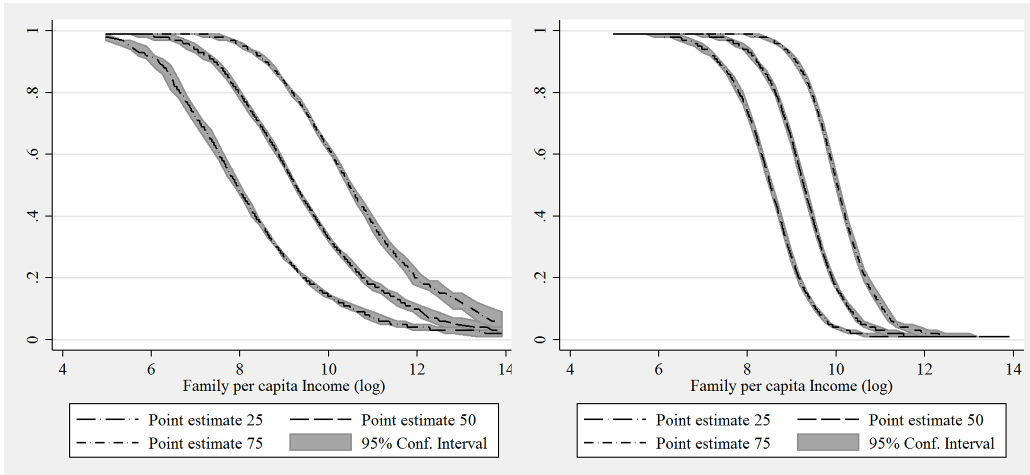
We estimate these models for different quantiles. The results are collected in Figures 6 and 7 for food expenditures for the model without and with additional covariates, respectively, for CQR, RIF (cubic model in the second step), and NW estimators. Figures 8 and 9 repeat the same empirical estimates for the case of non-durable expenditures. In each case, the left (a) panels display RIF estimates together with the CQR estimators, while the right (b) panels plot the UQPE NW estimator also together with the same CQR. The horizontal axis corresponds to the quantile index, which should be τ for RIF and NW unconditional estimators and η for CQR. In Tables 7-10 we provide additional results for RIF using linear and quadratic models in the second step.

The results in Figures 6 and 7 show evidence that CQR coefficients are roughly constant across η, although mildly increasing. The UQPE NW is roughly constant across τ. The RIF estimate, however, shows an increasing pattern. In all cases, the estimated effects can be interpreted as elasticities, implying that a 1% increase in income increase food consumption in less than 1%. For the case of non-durables, Figures 8 and 9, the RIF estimates are increasing along τ, while the UQPE NW estimator is decreasing. The fact that RIF estimates have a larger range of variation than CQR and that it gives the counter-intuitive increasing pattern suggest that it might be misspecified. Nevertheless, our UQPE NW estimates clearly remain within the CQR variation. Moreover, since the CQR coefficients are mildly increasing, the variation in the UQPE has to be coming from the variation in the density of X given $Y = Q_\tau[Y]$. As τ increases, for the UQPE to increase, higher CQR coefficients must be getting higher weight. This happens if the density of "income | food= $Q_\tau[food]$" is moving to the right. For this to happen, this shift has to happen and happen relatively quickly: given food expenditure is getting higher, we expect income to become higher, but at a faster rate than the increase in food expenditure, so that the proportion is falling.

In order to explore the results in more detail, we plot the by-product of this analysis that is the matching quantile function. This was implicitly used for the estimation of the UQPE NW estimator. Figure 10 plots the estimated match for $\tau = \{0.25, 0.50, 0.75\}$ for different values of log income (this corresponds to the case with no additional covariates). The figures illustrate that for each τ there is a full range of variation in the corresponding CQPE model indexed by η.

7 Conclusion

This paper considers the use of conditional quantile regression analysis to estimate unconditional quantile partial effects. The proposed methodology is based on a matching and reweighting result to link the unconditional effects to the conditional ones. This method thus benefits from the usual conditional quantile regression estimation techniques, and suggests a two-step estimator for the unconditional effects. In the first step one estimates a structural quantile regression model, and in the second stage a non-parametric regression is applied. We establish the asymptotic properties of the estimator. Monte Carlo simulations show evidence that the estimator has good finite sample performance and is robust to the selection of bandwidth and kernel. To illustrate



(a) Food (b) Non-durables

Figure 10: Estimated matching function

the proposed methods, we study Engel's curves in Argentina.

The current paper can be extended in several directions. First, the proposed model uses a simple linear quantile regression framework and is based on its coefficient estimators. The current framework can be applied to any other $\sqrt{n}$ consistent estimation procedure. In particular, as an example, instrumental variables quantile regression and/or panel data models deliver consistent estimators for the conditional effects in several related statistical models. The current methodology could be extended to evaluate unconditional effects, starting from any initial consistent conditional estimation procedure. Second, the current proposed framework can be used to evaluate any other functional analysis related to the unconditional quantile regression one. In other words, to recover general distributional effects. Third, the Nadaraya-Watson estimator is the first approximation to a larger family of estimators that can be used to estimate the unconditional effects. Local linear regression models is a proposed refinement to obtain possibly better asymptotic properties.

Appendix A

A.1 Proof of Lemma 1

Recall that $X = (X_1, X_2')'$, $x = (x_1, x_2')'$, and X_1 and x_1 are one-dimensional. First we show that $CQPE_{X_1}(\tau, x)$, defined in (2) as

$$CQPE_{X_1}(\tau, x) := \frac{\partial Q_Y[\tau|X_1 = z, X_2 = x_2]}{\partial z}\bigg|_{z=x_1}$$

can be written in the way of (5) as

$$CQPE_{X_1}(\tau, x) = -\frac{1}{f_{Y|X}(Q_Y[\tau|X = x]|x)} \frac{\partial F_{Y|X}(Q_Y[\tau|X = x]|z, x_2)}{\partial z}\bigg|_{z=x_1}.$$

By definition of quantiles, we have that this identity holds for all $x \in \mathcal{X}$ for a given fixed τ:

$$F_{Y|X}(Q_Y[\tau|X = x]|x) = \tau.$$

Differentiating both sides with respect to x_1, we obtain

$$f_{Y|X}(Q_Y[\tau|X = x]|x) CQPE_{X_1}(\tau, x) + \frac{\partial F_{Y|X}(Q_Y[\tau|X = x]|x)}{\partial z}\bigg|_{z=x_1} = 0.$$

Since $f_{Y|X}(Q_Y[\tau|X = x]|x) \neq 0$, then the result follows by solving for $CQPE_{X_1}(\tau, x)$. Since the matching is a singleton, then for every x, and any τ, we have $Q_Y[\xi_\tau(x)|X = x] = Q_Y[\tau]$. Thus, we evaluate $CQPE_{X_1}(\tau, x)$ at $\tau = \xi_\tau(x)$ to yield

$$CQPE_{X_1}(\xi_\tau(x), x) = -\frac{1}{f_{Y|X}(Q_Y[\tau]|x)} \frac{\partial F_{Y|X}(Q_Y[\tau]|z, x_2)}{\partial z}\bigg|_{z=x_1}.$$

Given the identification result for $UQPE_{X_1}(\tau)$ in equation (4), we have that

$$UQPE_{X_1}(\tau) = \int CQPE_{X_1}(\xi_\tau(x), x) \frac{f_{Y|X}(Q_Y[\tau]|x)}{f_Y(Q_Y[\tau])} dF_X(x),$$

which is the result in (7). Since f_Y and f_X are non-zero, then

$$\frac{f_{Y|X}(Q_Y[\tau]|x)}{f_Y(Q_Y[\tau])} f_X(x) = \frac{f_{Y,X}(Q_Y[\tau], x)}{f_Y(Q_Y[\tau]) f_X(x)} f_X(x) = f_{X|Y}(x|Q_Y[\tau]).$$

Therefore, we obtain (8):

$$UQPE_{X_1}(\tau) = E\big[CQPE_{X_1}(\xi_\tau(X), X)|Y = Q_Y[\tau]\big].$$

A.2 Proof of Theorem 1

Let

$$\Psi_\tau(\eta|x) = Q_Y[\eta|X = x] - Q_Y[\tau].$$

Here, τ and x are fixed, and the criterion function is the map $[\epsilon, 1-\epsilon] \ni \eta \mapsto \Psi_\tau(\eta|x)$ for $0 < \epsilon < 1/2$. Under Assumption 1.2, $y \mapsto F_{Y|X}(y|x)$ is strictly increasing, and hence $\Psi_\tau(\eta|x)$ has a unique zero given by $\xi_\tau(x) = F_{Y|X}(Q_\tau[Y]|x)$. This shows that $Q_\tau[Y]$ is $\xi_\tau(x)$-conditional quantile of $Y|X = x$. By Assumption 1.1, this can be written as $x'\beta(\xi_\tau(x)) = Q_\tau[Y]$.

Now we will show consistency: $\hat{\xi}_\tau(x) \xrightarrow{p} \xi_\tau(x)$. The matching function is defined to be the (approximate) zero of the random criterion function $\Psi_{\tau,n}(\eta|x)$:

$$\Psi_{\tau,n}(\hat{\xi}_\tau(x)|x) = x'\hat{\beta}(\hat{\xi}_\tau(x)) - \hat{Q}_Y[\tau].$$

Indeed, the computational procedure outlined in equation (12) implicitly defines $\hat{\xi}_\tau(x)$ as η_j for some j in $\{1, 2, \ldots, m\}$ such that

$$x'\hat{\beta}(\eta_j) \leq \hat{Q}_Y[\tau] < x'\hat{\beta}(\eta_{j+1}).$$

We want to show that this, together with Assumption 1.5 that ensures $\Delta\eta = \eta_{j+1} - \eta_j = o(n^{-1/2})$, imply that $\Psi_{\tau,n}(\hat{\xi}_\tau(x)|x) = o_p(n^{-1/2})$. For a given n, let $\eta_j = \hat{\xi}_\tau(x)$, that is

$$\begin{aligned}
x'\hat{\beta}(\eta_j) \leq \hat{Q}_Y[\tau] &< x'\hat{\beta}(\eta_{j+1}) \\
0 \leq \hat{Q}_Y[\tau] - x'\hat{\beta}(\eta_j) &< x'\hat{\beta}(\eta_{j+1}) - x'\hat{\beta}(\eta_j) \\
0 \leq -\Psi_{\tau,n}(\hat{\xi}_\tau(x)|x) &< x'\left(\hat{\beta}(\eta_{j+1}) - \hat{\beta}(\eta_j)\right).
\end{aligned}$$

We focus on the difference $\hat{\beta}(\eta_{j+1}) - \hat{\beta}(\eta_j)$. We write

$$\begin{aligned}
\hat{\beta}(\eta_{j+1}) - \hat{\beta}(\eta_j) &= \hat{\beta}(\eta_{j+1}) - \beta(\eta_{j+1}) - \left(\hat{\beta}(\eta_j) - \beta(\eta_j)\right) \\
&\quad + \beta(\eta_{j+1}) - \beta(\eta_j).
\end{aligned}$$

For the last term, we can write $\beta(\eta_{j+1}) - \beta(\eta_j) = \beta'(\tilde{\eta})(\eta_{j+1} - \eta_j) = o(n^{-1/2})$ because the derivative is bounded by Assumption 1.3. To alleviate notation, define:

$$J(\eta) = E\left[f_{Y|X}(X'\beta(\eta)|X)XX'\right],$$

which is differentiable with bounded derivative by 1.3. Using Assumption 1.3 we have that

$$\hat{\beta}(\eta) - \beta(\eta) = J(\eta)^{-1}\frac{1}{n}\sum_{i=1}^{n}\left(\eta - \mathbb{1}\left\{y_i \leq x_i'\beta(\eta)\right\}\right)x_i + o_p(n^{-1/2}),$$

so that

$$
\begin{aligned}
\hat{\beta}(\eta_{j+1}) - \beta(\eta_{j+1}) - \left(\hat{\beta}(\eta_j) - \beta(\eta_j)\right) &= J(\eta_{j+1})^{-1}\frac{1}{n}\sum_{i=1}^{n}\left(\eta_{j+1} - \mathbb{1}\left\{y_i \leq x_i'\beta(\eta_{j+1})\right\}\right)x_i \\
&- J(\eta_j)^{-1}\frac{1}{n}\sum_{i=1}^{n}\left(\eta_j - \mathbb{1}\left\{y_i \leq x_i'\beta(\eta_j)\right\}\right)x_i + o_p(n^{-1/2}) \\
&= J(\eta_{j+1})^{-1}\frac{1}{n}\sum_{i=1}^{n}\Bigg[\left(\eta_{j+1} - \mathbb{1}\left\{y_i \leq x_i'\beta(\eta_{j+1})\right\}\right)x_i \\
&- \left(\eta_j - \mathbb{1}\left\{y_i \leq x_i'\beta(\eta_j)\right\}\right)x_i\Bigg] \\
&+ \Big(J(\eta_{j+1})^{-1} - J(\eta_j)^{-1}\Big)\frac{1}{n}\sum_{i=1}^{n}\left(\eta_j - \mathbb{1}\left\{y_i \leq x_i'\beta(\eta_j)\right\}\right)x_i + o_p(n^{-1/2}).
\end{aligned}
$$

By Assumption 1.2, $J(\eta)$ is bounded away from zero, so that $J(\eta)^{-1}$ is bounded. We focus first on the difference in the sums.

$$
\frac{1}{n}\sum_{i=1}^{n}\left(\eta_{j+1} - \mathbb{1}\left\{y_i \leq x_i'\beta(\eta_{j+1})\right\} - \eta_j + \mathbb{1}\left\{y_i \leq x_i'\beta(\eta_j)\right\}\right)x_i.
$$

We note first that since $\eta_j < \eta_{j+1}$, by definition of quantiles, $x_i'\beta(\eta_j) < x_i'\beta(\eta_{j+1})$. This means that if $\mathbb{1}\left\{y_i \leq x_i'\beta(\eta_{j+1})\right\} = 0$, then $\mathbb{1}\left\{y_i \leq x_i'\beta(\eta_j)\right\} = 0$, and if $\mathbb{1}\left\{y_i \leq x_i'\beta(\eta_{j+1})\right\} = 1$, then either $\mathbb{1}\left\{y_i \leq x_i'\beta(\eta_j)\right\} = 0$, or $\mathbb{1}\left\{y_i \leq x_i'\beta(\eta_j)\right\} = 1$. Thus, the difference

$$
\mathbb{1}\left\{y_i \leq x_i'\beta(\eta_{j+1})\right\} - \mathbb{1}\left\{y_i \leq x_i'\beta(\eta_j)\right\}
$$

is either 1 or 0. Using this, we have

$$
\begin{aligned}
&\left|\frac{1}{n}\sum_{i=1}^{n}\left(\eta_{j+1} - \mathbb{1}\left\{y_i \leq x_i'\beta(\eta_{j+1})\right\} - \eta_j + \mathbb{1}\left\{y_i \leq x_i'\beta(\eta_j)\right\}\right)x_i\right| \\
&\leq \frac{1}{n}\sum_{i=1}^{n}\left|\eta_{j+1} - \mathbb{1}\left\{y_i \leq x_i'\beta(\eta_{j+1})\right\} - \eta_j + \mathbb{1}\left\{y_i \leq x_i'\beta(\eta_j)\right\}\right||x_i| \\
&\leq \left|\eta_{j+1} - \eta_j\right|\frac{1}{n}\sum_{i=1}^{n}|x_i| \\
&= o(n^{-1/2})O_p(1) = o_p(n^{-1/2}).
\end{aligned}
$$

Now, for the second term, we note that

$$\begin{aligned}
\left| \left(J(\eta_{j+1})^{-1} - J(\eta_j)^{-1} \right) \frac{1}{n} \sum_{i=1}^{n} \left(\eta_j - \mathbb{1}\left\{ y_i \leq x_i' \beta(\eta_j) \right\} \right) x_i \right| &\leq \left| \frac{J(\eta_j) - J(\eta_{j+1})}{J(\eta_j) J(\eta_{j+1})} \right| \frac{1}{n} \sum_{i=1}^{n} |x_i| \\
&= \left| \frac{J'(\tilde{\eta})(\eta_{j+1} - \eta_j)}{J(\eta_j) J(\eta_{j+1})} \right| \frac{1}{n} \sum_{i=1}^{n} |x_i| \\
&= o(n^{-1/2}) O_p(1) = o_p(n^{-1/2}).
\end{aligned}$$

where $\tilde{\eta}$ is between η_j and η_{j+1}. Therefore, we have that

$$\Psi_{\tau,n}(\hat{\xi}_\tau(x)|x) = o_p(n^{-1/2}).$$

Since $\hat{\xi}_\tau(x)$ is a Z-estimator, we follow Theorem 5.9 in van der Vaart (1998). We need to show (i) that the criterion function converges uniformly in probability:

$$\sup_{\eta \in [\epsilon, 1-\epsilon]} |\Psi_{\tau,n}(\eta|x) - \Psi_\tau(\eta|x)| \xrightarrow{p} 0, \tag{A.1}$$

and (ii) that the zero is well-separated: for any $\Delta > 0$

$$\inf_{\eta : |\eta - \xi_\tau(x)| > \Delta} |\Psi_\tau(\eta|X)| > 0.$$

To show (A.1), we note that by Assumption 1.1

$$\begin{aligned}
\sup_{\eta \in [\epsilon, 1-\epsilon]} |\Psi_{\tau,n}(\eta|x) - \Psi_\tau(\eta|x)| &= \sup_{\eta \in [\epsilon, 1-\epsilon]} |x'\hat{\beta}(\eta) - \hat{Q}_Y[\tau] - x'\beta(\eta) + Q_Y[\tau]| \\
&\leq ||x|| \sup_{\eta \in [\epsilon, 1-\epsilon]} |\hat{\beta}(\eta) - \beta(\eta)| + |\hat{Q}_Y[\tau] - Q_Y[\tau]| \\
&= ||x|| O_p(n^{-1/2}) + O_p(n^{-1/2}) \\
&= o_p(1).
\end{aligned}$$

where $\| \cdot \|$ is the Euclidean norm, and the bounds follow from Assumptions 1.3 and 1.4. To show that the zero is well-separated, we note that by Assumption 1.2, $y \mapsto F_{Y|X}(y|x)$ is strictly increasing, so that for any $\Delta > 0$, if $|\eta - \xi_\tau(x)| > \Delta$, then by the mean-value theorem

$$|\eta - F_{Y|X}(Q_\tau[Y]|x)| = |F_{Y|X}(Q_\eta[Y|X=x]|x) - F_{Y|X}(Q_\tau[Y]|x)| = f_{Y|X}(\tilde{Q}|x)|Q_\eta[Y|X=x] - Q_\tau[Y]|,$$

where $\tilde{Q}$ is between $Q_\eta[Y|X=x]$ and $Q_\tau[Y]$. Now, $f_{Y|X}(\tilde{Q}|x) > 0$. Moreover, if $\eta > \xi_\tau(x)$, then $\eta > \xi_\tau(x) + \Delta$, so that $Q_\eta[Y|X=x] > Q_{\xi_\tau(x)+\Delta}[Y|X=x] > Q_{\xi_\tau(x)}[Y|X=x] = Q_\tau[Y]$, where we take Δ small enough such that $\eta < 1$. The same analysis can be carried out for $\eta < \xi_\tau(x)$, in which case: $Q_\eta[Y|X=x] < Q_{\xi_\tau(x)-\Delta}[Y|X=x] < Q_{\xi_\tau(x)}[Y|X=x] = Q_\tau[Y]$. Therefore,

$$
\begin{aligned}
\inf_{\eta:|\eta-\xi_\tau(x)|>\Delta} |\Psi_\tau(\eta|X)| &= \inf_{\eta:|\eta-\xi_\tau(x)|>\Delta} f_{Y|X}(\tilde{Q}|x)|Q_\eta[Y|X=x] - Q_\tau[Y]| \\
&\geq \inf_{\eta:|\eta-\xi_\tau(x)|>\Delta} f_{Y|X}(\tilde{Q}|x) \times \inf_{\eta:|\eta-\xi_\tau(x)|>\Delta} |Q_\eta[Y|X=x] - Q_\tau[Y]| \\
&\geq \inf_{y\in\mathbb{R}} f_{Y|X}(y|x) \times \min\left\{|Q_{\xi_\tau(x)+\Delta}[Y|X=x] - Q_\tau[Y]|, |Q_{\xi_\tau(x)-\Delta}[Y|X=x] - Q_\tau[Y]|\right\} \\
&> 0,
\end{aligned}
$$

where we have used that $f_{Y|X}(y|x)$ is bounded away from zero. Finally, we can invoke Theorem 5.9 in van der Vaart (1998), since we also showed that $\Psi_{\tau,n}(\hat{\xi}_\tau(x)|x) = o_p(n^{-1/2})$, therefore, $\hat{\xi}_\tau(x) \xrightarrow{p} \xi_\tau(x)$.

Having shown consistency, we now prove it is actually $\sqrt{n}$-consistent. To that end, we use Assumption 1.5:

$$
\begin{aligned}
o_p(n^{-1/2}) &= x'\hat{\beta}(\hat{\xi}_\tau(x)) - \hat{Q}_Y[\tau] \\
&= x'\left(\hat{\beta}(\hat{\xi}_\tau(x)) - \beta(\hat{\xi}_\tau(x))\right) + x'\left(\beta(\hat{\xi}_\tau(x)) - \beta(\xi_\tau(x))\right) \\
&+ \underbrace{x'\beta(\xi_\tau(x)) - Q_Y[\tau]}_{=\Psi_\tau(\xi_\tau(x)|x)=0} - \left(\hat{Q}_Y[\tau] - Q_Y[\tau]\right).
\end{aligned}
$$

By Assumptions 1.1, 1.2, we have that $F_{Y|X}(x'\beta(\eta)|x) = \eta$, so that $x'\beta'(\eta) = f_{Y|X}(x'\beta(\eta)|x)^{-1} > 0$, so that we can do a first order term-by-term Taylor expansion to obtain

$$
\beta(\hat{\xi}_\tau(x)) - \beta(\xi_\tau(x)) = \dot{\beta}(\xi_\tau(x))\left(\hat{\xi}_\tau(x) - \xi_\tau(x)\right) + o_p(|\hat{\xi}_\tau(x) - \xi_\tau(x)|).
$$

Here, $\dot{\beta}(\xi_\tau(X))$ is the Jacobian vector: the derivative of the map $\tau \mapsto \beta(\tau)$ and $o_p(|\hat{\xi}_\tau(x) - \xi_\tau(x)|)$ is a vector of residuals of the expansion. Plugging this into the previous display, we obtain

$$
\begin{aligned}
o_p(n^{-1/2}) &= x'\left(\hat{\beta}(\hat{\xi}_\tau(x)) - \beta(\hat{\xi}_\tau(x))\right) + x'\dot{\beta}(\xi_\tau(x))\left(\hat{\xi}_\tau(x) - \xi_\tau(x)\right) \\
&+ o_p(|\hat{\xi}_\tau(x) - \xi_\tau(x)|) - \left(\hat{Q}_Y[\tau] - Q_Y[\tau]\right).
\end{aligned}
\tag{A.2}
$$

Here the term $o_p(|\hat{\xi}_\tau(x) - \xi_\tau(x)|)$ is scalar-valued and collects all the terms from $x'o_p(|\hat{\xi}_\tau(x) - \xi_\tau(x)|)$. Now, $\hat{Q}_Y[\tau] - Q_Y[\tau] = O_p(n^{-1/2})$ by Assumption 1.4. Also, by Assumption 1.3 $\hat{\beta}(\hat{\xi}_\tau(x)) - \beta(\hat{\xi}_\tau(x)) \leq \sup_{\eta\in[\epsilon,1-\epsilon]} |\hat{\beta}(\eta) - \beta(\eta)| = O_p(n^{-1/2})$. Therefore, since $x'\dot{\beta}(\xi_\tau(x)) \neq 0$, we have that $|\hat{\xi}_\tau(x) - \xi_\tau(x)| = O_p(n^{-1/2})$.

Finally to obtain the asymptotic distribution, we go back to (A.2), and using the stochastic equicontinuity guaranteed by Assumption 1.3, we replace $\hat{\beta}(\hat{\xi}_\tau(x) - \beta(\hat{\xi}_\tau(x)$ by $\hat{\beta}(\xi_\tau(x) -$

$\beta(\xi_\tau(x)) + o_p(n^{-1/2})$. Therefore, we obtain

$$\begin{aligned}
\hat{\xi}_\tau(x) - \xi_\tau(x) &= -\frac{1}{x'\dot{\beta}(\xi_\tau(X))} x' \left(\hat{\beta}(\xi_\tau(x)) - \beta(\xi_\tau(x))\right) + \frac{1}{x'\dot{\beta}(\xi_\tau(x))} \left(\hat{Q}_Y[\tau] - Q_Y[\tau]\right) + o_p(n^{-1/2}) \\
&= -\frac{1}{x'\dot{\beta}(\xi_\tau(x))} \frac{1}{n} \sum_{i=1}^{n} x'\Psi_i(\xi_\tau(x)) + \frac{1}{x'\dot{\beta}(\xi_\tau(x))} \frac{1}{n} \sum_{i=1}^{n} \psi_i(\tau) + o_p(n^{-1/2}).
\end{aligned}$$

To obtain the main statement of the theorem we write

$$\begin{aligned}
\hat{\beta}_1(\hat{\xi}_\tau(x)) - \beta_1(\xi_\tau(x)) &= \hat{\beta}_1(\hat{\xi}_\tau(x)) - \beta_1(\hat{\xi}_\tau(x)) + \beta_1(\hat{\xi}_\tau(x)) - \beta_1(\xi_\tau(x)) \\
&= \hat{\beta}_1(\xi_\tau(x)) - \beta_1(\xi_\tau(x)) + \dot{\beta}_1(\xi_\tau(x))(\hat{\xi}_\tau(x) - \xi_\tau(x)) + o_p(n^{-1/2}).
\end{aligned}$$

A.3 Proof of Theorem 2

To alleviate notation, we write:

$$\begin{aligned}
\hat{m}_1(q, b, e) &:= \frac{1}{n} \sum_{i=1}^{n} K_h(y_i - q) \cdot b(e(x_i)) \\
\hat{m}_2(q) &:= \frac{1}{n} \sum_{i=1}^{n} K_h(y_i - q).
\end{aligned}$$

Thus, our estimator of $UQPE_{X_1}(\tau)$ can be written as

$$\widehat{UQPE}_{X_1}(\tau) = \frac{\hat{m}_1(\hat{Q}_Y[\tau], \hat{\beta}_1, \hat{\xi}_\tau)}{\hat{m}_2(\hat{Q}_Y[\tau])}.$$

The unfeasible version is then

$$\widetilde{UQPE}_{X_1}(\tau) = \frac{\hat{m}_1(Q_Y[\tau], \beta_1, \xi_\tau)}{\hat{m}_2(Q_Y[\tau])}.$$

Consider the difference

$$\begin{aligned}
\widehat{UQPE}_{X_1}(\tau) - \widetilde{UQPE}_{X_1}(\tau) &= \frac{\hat{m}_1(\hat{Q}_Y[\tau], \hat{\beta}_1, \hat{\xi}_\tau)}{\hat{m}_2(\hat{Q}_Y[\tau])} - \frac{\hat{m}_1(Q_Y[\tau], \beta_1, \xi_\tau)}{\hat{m}_2(Q_Y[\tau])} \\
&= \frac{\hat{m}_2(Q_Y[\tau])\hat{m}_1(\hat{Q}_Y[\tau], \hat{\beta}_1, \hat{\xi}_\tau) - \hat{m}_2(\hat{Q}_Y[\tau])\hat{m}_1(Q_Y[\tau], \beta_1, \xi_\tau)}{\hat{m}_2(\hat{Q}_Y[\tau])\hat{m}_2(Q_Y[\tau])} \\
&= \frac{\hat{m}_1(\hat{Q}_Y[\tau], \hat{\beta}_1, \hat{\xi}_\tau) - \hat{m}_1(Q_Y[\tau], \beta_1, \xi_\tau)}{\hat{m}_2(\hat{Q}_Y[\tau])} \\
&\quad - \frac{\hat{m}_1(Q_Y[\tau], \beta_1, \xi_\tau)}{\hat{m}_2(\hat{Q}_Y[\tau])\hat{m}_2(Q_Y[\tau])} \left(\hat{m}_2(\hat{Q}_Y[\tau]) - \hat{m}_2(Q_Y[\tau])\right). \qquad \text{(A.3)}
\end{aligned}$$

First we focus on the second term of (A.3). We note that

$$\hat{m}_2(\hat{Q}_Y[\tau]) := \frac{1}{n}\sum_{i=1}^{n} K_h(y_i - \hat{Q}_Y[\tau]) = \hat{f}_Y(\hat{Q}_Y[\tau])$$

is an estimator of the density of Y evaluated at $\hat{Q}_Y[\tau]$, the estimator of $Q_Y[\tau]$; while

$$\hat{m}_2(Q_Y[\tau]) := \frac{1}{n}\sum_{i=1}^{n} K_h(y_i - Q_Y[\tau]) = \hat{f}_Y(Q_Y[\tau])$$

is an estimator of the density of Y evaluated at $Q_Y[\tau]$. We now show that

$$\begin{aligned}\hat{m}_2(\hat{Q}_Y[\tau]) - \hat{m}_2(Q_Y[\tau]) &= \hat{f}_Y(\hat{Q}_Y[\tau]) - \hat{f}_Y(Q_Y[\tau]) \\ &= f'_Y(Q_Y[\tau])(\hat{Q}_Y[\tau] - Q_Y[\tau]) + o_p(n^{-1/2}h^{-1/2}),\end{aligned} \tag{A.4}$$

which implies that $\hat{m}_2(\hat{Q}_Y[\tau]) - \hat{m}_2(Q_Y[\tau]) = o_p(n^{-1/2}h^{-1/2})$.

By Taylor's theorem, we have that for some $\tilde{Q}_Y[\tau]$ between $\hat{Q}_Y[\tau]$ and $Q_Y[\tau]$

$$\begin{aligned}\hat{f}_Y(\hat{Q}_Y[\tau]) - \hat{f}_Y(Q_Y[\tau]) &= \hat{f}'_Y(Q_Y[\tau])(\hat{Q}_Y[\tau] - Q_Y[\tau]) + \frac{1}{2}\hat{f}''_Y(\tilde{Q}_Y[\tau])(\hat{Q}_Y[\tau] - Q_Y[\tau])^2 \\ &= f'_Y(Q_Y[\tau])(\hat{Q}_Y[\tau] - Q_Y[\tau]) \\ &+ (\hat{f}'_Y(Q_Y[\tau]) - f'_Y(Q_Y[\tau]))(\hat{Q}_Y[\tau] - Q_Y[\tau]) \\ &+ \frac{1}{2}\hat{f}''_Y(\tilde{Q}_Y[\tau])(\hat{Q}_Y[\tau] - Q_Y[\tau])^2.\end{aligned}$$

We need to show that the second and third terms are $o_p(n^{-1/2}h^{-1/2})$. We start with the third term. The second derivative of the kernel estimator is

$$\hat{f}''(q) = \frac{1}{nh^3}\sum_{i=1}^{n} K''\left(\frac{y_i - q}{h}\right).$$

Since the second derivative K'' is bounded by Assumption 2: $\sup_u |K''(u)| < \infty$, then

$$\hat{f}''(\tilde{Q}_Y[\tau]) \leq \frac{1}{nh^3}\sum_{i=1}^{n}\sup_u |K''(u)| = \frac{1}{h^3}\sup_u |K''(u)|,$$

therefore

$$\hat{f}''(\tilde{Q}_Y[\tau]) = O_p(h^{-3}).$$

This means that

$$\frac{1}{2}\hat{f}''(\tilde{Q}_Y[\tau])(\hat{Q}_Y[\tau] - Q_Y[\tau])^2 = O_p(n^{-1}h^{-3}).$$

Now,

$$n^{1/2}h^{1/2}\frac{1}{2}\hat{f}''(\tilde{Q}_Y[\tau])(\hat{Q}_Y[\tau]-Q_Y[\tau])^2 = n^{1/2}h^{1/2}O_p(n^{-1}h^{-3}) = O_p(n^{-1/2}h^{-5/2}) = o_p(1),$$

because $nh^5 \to \infty$ by Assumption 4. Therefore,

$$\frac{1}{2}\hat{f}''(\tilde{Q}_Y[\tau])(\hat{Q}_Y[\tau]-Q_Y[\tau])^2 = o_p(n^{-1/2}h^{-1/2}).$$

Now, for the other term, $(\hat{f}'_Y(Q_Y[\tau]) - f'_Y(Q_Y[\tau]))(\hat{Q}_Y[\tau]-Q_Y[\tau])$, we focus on the derivative. The first derivative of the kernel estimator is

$$\hat{f}'(q) = -\frac{1}{nh^2}\sum_{i=1}^{n} K'\left(\frac{y_i - q}{h}\right).$$

We approximate its mean and variance under Assumption 2. We will use the fact that, by Assumption 2, $\int K'(u)du = 0$, $\int K'(u)udu = -1$, $\int K'(u)u^j du = 0$ for $j = 2, ..., r$, and $\int K'(u)u^{r+1}du < \infty$. In the following, $f_Y^{(j)}$ denotes the j^{th} derivative.

$$\begin{aligned}
E[\hat{f}'(q)] &= -\frac{1}{h^2}E\left[K'\left(\frac{Y-q}{h}\right)\right] \\
&= -\frac{1}{h^2}\int_{\mathbb{R}} K'\left(\frac{y-q}{h}\right) f_Y(y)dy \\
&= -\frac{1}{h}\int_{\mathbb{R}} K'(u) f_Y(q+hu)du \\
&= -\frac{1}{h}\int_{\mathbb{R}} K'(u)\left[f_Y(q) + \sum_{j=1}^{r}\frac{h^j u^j f_Y^{(j)}(q)}{j!} + \frac{h^{r+1}u^{r+1}f_Y^{(r+1)}(\tilde{q})}{(r+1)!}\right]du \\
&= f'_Y(q) - \frac{h^r}{(r+1)!}\int_{\mathbb{R}} K'(u)u^{r+1}f_Y^{(r+1)}(\tilde{q})du.
\end{aligned}$$

Since, by Assumption 3, the derivatives are bounded, the bias is

$$E[\hat{f}'(q)] - f'_Y(q) = O(h^r).$$

For the variance term, we have

$$
\begin{aligned}
Var[\hat{f}'(q)] &= \frac{1}{nh^4} Var\left[K'\left(\frac{Y-q}{h}\right)\right] \\
&= \frac{1}{nh^4} E\left[K'\left(\frac{Y-q}{h}\right)^2\right] - \frac{1}{nh^4} E\left[K'\left(\frac{Y-q}{h}\right)\right]^2 \\
&= \frac{1}{nh^4} \int_{\mathbb{R}} K'\left(\frac{y-q}{h}\right)^2 f_Y(y)dy - \frac{1}{nh^4}\left[\int_{\mathbb{R}} K'\left(\frac{y-q}{h}\right) f_Y(y)dy\right]^2 \\
&= \frac{1}{nh^3} \int_{\mathbb{R}} K'(u)^2 f_Y(q+hu)du - \frac{1}{nh^2}\left[\int_{\mathbb{R}} K'(u) f_Y(q+hu)du\right]^2 \\
&= \frac{1}{nh^3} \int_{\mathbb{R}} K'(u)^2 \left[f_Y(q) + huf_Y'(\tilde{q})\right] du - O(n^{-1}h^{-2}) \\
&= \frac{f_Y(q)}{nh^3} \int_{\mathbb{R}} K'(u)^2 du + o(n^{-1}h^{-2}) - O(n^{-1}h^{-2}).
\end{aligned}
$$

Therefore,

$$
\hat{f}'(q) - f'(q) = O_p(n^{-1/2}h^{-3/2} + h^r).
$$

Finally, we have

$$
\begin{aligned}
n^{1/2}h^{1/2}(\hat{f}_Y'(Q_Y[\tau]) - f_Y'(Q_Y[\tau]))(\hat{Q}_Y[\tau] - Q_Y[\tau]) &= n^{1/2}h^{1/2}O_p(n^{-1/2}h^{-3/2} + h^r)O_p(n^{-1/2}) \\
&= O_p(n^{-1/2}h^{-1} + h^{r+1/2})O_p(1) \\
&= o_p(1),
\end{aligned}
$$

since $n^{1/2}h = (nh^2)^{1/2} \to \infty$ since $nh^2 \to \infty$ because by Assumption 4 $nh^5 \to \infty$. Putting all the results together, we obtain

$$
\hat{f}_Y(\hat{Q}_Y[\tau]) - \hat{f}_Y(Q_Y[\tau]) = f_Y'(Q_Y[\tau])(\hat{Q}_Y[\tau] - Q_Y[\tau]) + o_p(n^{-1/2}h^{-1/2}).
$$

Thus, we obtain that

$$
\hat{m}_2(\hat{Q}_Y[\tau]) - \hat{m}_2(Q_Y[\tau]) = o_p(n^{-1/2}h^{-1/2}). \tag{A.5}
$$

Now we focus on the first term of (A.3). For the numerator, consider the following decomposition

$$\begin{aligned}
\hat{m}_1(\hat{Q}_Y[\tau],\hat{\beta}_1,\hat{\xi}_\tau) - \hat{m}_1(Q_Y[\tau],\beta_1,\xi_\tau) &= \frac{1}{n}\sum_{i=1}^{n} K_h(y_i - \hat{Q}_Y[\tau]) \cdot \hat{\beta}_1(\hat{\xi}_\tau(x_i)) - \frac{1}{n}\sum_{i=1}^{n} K_h(y_i - Q_Y[\tau]) \cdot \beta_1(\xi_\tau(x_i)) \\
&= \underbrace{\frac{1}{n}\sum_{i=1}^{n} \left[K_h(y_i - \hat{Q}_Y[\tau]) - K_h(y_i - Q_Y[\tau])\right] \cdot \beta_1(\xi_\tau(x_i))}_{:=T_1} \\
&+ \underbrace{\frac{1}{n}\sum_{i=1}^{n} K_h(y_i - Q_Y[\tau]) \cdot \left[\hat{\beta}_1(\hat{\xi}_\tau(x_i)) - \beta_1(\xi_\tau(x_i))\right]}_{:=T_2} \\
&+ \underbrace{\frac{1}{n}\sum_{i=1}^{n} \left[K_h(y_i - \hat{Q}_Y[\tau]) - K_h(y_i - Q_Y[\tau])\right] \cdot \left[\hat{\beta}_1(\hat{\xi}_\tau(x_i)) - \beta_1(\xi_\tau(x_i))\right]}_{:=T_3} \\
&= T_1 + T_2 + T_3. \qquad \text{(A.6)}
\end{aligned}$$

We start with T_1 and we do a second order Taylor expansion:

$$\begin{aligned}
T_1 &:= \frac{1}{n}\sum_{i=1}^{n} \left[K_h(y_i - \hat{Q}_Y[\tau]) - K_h(y_i - Q_Y[\tau])\right] \cdot \beta_1(\xi_\tau(x_i)) \\
&= \left(\hat{Q}_Y[\tau] - Q_Y[\tau]\right) \frac{1}{n}\sum_{i=1}^{n} \left.\frac{\partial K_h(y_i - q)}{\partial q}\right|_{q=Q_Y[\tau]} \cdot \beta_1(\xi_\tau(x_i)) \\
&+ \frac{1}{2}\left(\hat{Q}_Y[\tau] - Q_Y[\tau]\right)^2 \frac{1}{n}\sum_{i=1}^{n} \left.\frac{\partial^2 K_h(y_i - q)}{\partial^2 q}\right|_{q=\tilde{q}} \cdot \beta_1(\xi_\tau(x_i)). \qquad \text{(A.7)}
\end{aligned}$$

Consider the first term. Let $f^{(j)}_{Y,X}(y,x)$ denote the j^{th} partial derivative of $f_{Y,X}(y,x)$ with respect to y. The expected value is

$$\begin{aligned}
E\left[\frac{1}{n}\sum_{i=1}^{n} \left.\frac{\partial K_h(Y_i - q)}{\partial q}\right|_{q=Q_Y[\tau]} \cdot \beta_1(\xi_\tau(X_i))\right] &= E\left[\left.\frac{\partial K_h(Y - q)}{\partial q}\right|_{q=Q_Y[\tau]} \cdot \beta_1(\xi_\tau(X))\right] \\
&= -\frac{1}{h^2} E\left[K'\left(\frac{Y - Q_Y[\tau]}{h}\right) \cdot \beta_1(\xi_\tau(X))\right] \\
&= -\frac{1}{h^2}\int\int K'\left(\frac{y - Q_Y[\tau]}{h}\right)\beta_1(\xi_\tau(x)) f_{Y,X}(y,x)dydx \\
&= -\frac{1}{h}\int\int K'(u)\beta_1(\xi_\tau(x)) f_{Y,X}(Q_Y[\tau] + hu, x)dudx \\
&= -\frac{1}{h}\int\int K'(u)\beta_1(\xi_\tau(x))\Bigg[f_{Y,X}(Q_Y[\tau], x) \\
&+ \sum_{j=1}^{r} \frac{h^j u^j f^{(j)}_{Y,X}(Q_Y[\tau], x)}{j!} + \frac{h^{r+1}u^{r+1} f^{(r+1)}_{Y,X}(\tilde{Q}_Y[\tau], x)}{(r+1)!}\Bigg]dudx,
\end{aligned}$$

where we used Assumption 3 to expand the joint density. The properties of the kernel of Assumption 2 yield

$$
\begin{aligned}
E\left[\frac{1}{n}\sum_{i=1}^{n}\frac{\partial K_h(Y_i-q)}{\partial q}\bigg|_{q=Q_Y[\tau]}\cdot\beta_1(\xi_\tau(X_i))\right] &= -\frac{1}{h}\int\int K'(u)\beta_1(\xi_\tau(x))\bigg[f_{Y,X}(Q_Y[\tau],x)\\
&\quad+\sum_{j=1}^{r}\frac{h^j u^j f_{Y,X}^{(j)}(Q_Y[\tau],x)}{j!}+\frac{h^{r+1}u^{r+1}f_{Y,X}^{(r+1)}(\tilde{Q}_Y[\tau],x)}{(r+1)!}\bigg]dudx\\
&= \int\beta_1(\xi_\tau(x))f_{Y,X}^{(1)}(Q_Y[\tau],x)dx\\
&\quad-h^r\int\int K'(u)\beta_1(\xi_\tau(x))u^{r+1}f_{Y,X}^{(r+1)}(\tilde{Q}_Y[\tau],x)\beta_1(\xi_\tau(x))dxdu.
\end{aligned}
$$

Therefore, the bias is of order $O(h^r)$:

$$
E\left[\frac{1}{n}\sum_{i=1}^{n}\frac{\partial K_h(Y_i-q)}{\partial q}\bigg|_{q=Q_Y[\tau]}\cdot\beta_1(\xi_\tau(X_i))\right] = \int\beta_1(\xi_\tau(x))f_{Y,X}^{(1)}(Q_Y[\tau],x)dx+O(h^r).
$$

For the variance, we have

$$
\begin{aligned}
Var\left[\frac{1}{n}\sum_{i=1}^{n}\frac{\partial K_h(Y_i-q)}{\partial q}\bigg|_{q=Q_Y[\tau]}\cdot\beta_1(\xi_\tau(X_i))\right] &= \frac{1}{nh^4}Var\left[K'\left(\frac{Y-Q_Y[\tau]}{h}\right)\beta_1(\xi_\tau(X))\right]\\
&= \frac{1}{nh^4}E\left[K'\left(\frac{Y-Q_Y[\tau]}{h}\right)^2\beta_1(\xi_\tau(X))^2\right]\\
&\quad-\frac{1}{nh^4}E\left[K'\left(\frac{Y-Q_Y[\tau]}{h}\right)\beta_1(\xi_\tau(X))\right]^2.
\end{aligned}
$$

We take care of each term at a time.

$$
\begin{aligned}
\frac{1}{nh^4}E\left[K'\left(\frac{Y-Q_Y[\tau]}{h}\right)^2\beta_1(\xi_\tau(X))^2\right] &= \frac{1}{nh^4}\int\int K'\left(\frac{y-q}{h}\right)^2\beta_1(\xi_\tau(x))^2 f_{Y,X}(Q_Y[\tau],x)dydx\\
&= \frac{1}{nh^3}\int\int K'(u)^2\beta_1(\xi_\tau(x))^2 f_{Y,X}(Q_Y[\tau]+hu,x)dudx\\
&= \frac{1}{nh^3}\int\int K'(u)^2\beta_1(\xi_\tau(x))^2\left[f_{Y,X}(Q_Y[\tau],x)+huf_{Y,X}^{(1)}(\tilde{Q}_Y[\tau],x)\right]dudx\\
&= \frac{1}{nh^3}\int\int K'(u)^2\beta_1(\xi_\tau(x))^2 f_{Y,X}(Q_Y[\tau],x)dudx+o(n^{-1}h^{-2}).
\end{aligned}
$$

For the other term, we have

$$
\begin{aligned}
\frac{1}{nh^4}E\left[K'\left(\frac{Y-Q_Y[\tau]}{h}\right)\beta_1(\xi_\tau(X))\right]^2 &= \frac{1}{nh^4}\left[\int\int K'\left(\frac{y-Q_Y[\tau]}{h}\right)\beta_1(\xi_\tau(x))f_{Y,X}(y,x)dydx\right]^2\\
&= \frac{1}{nh^2}\left[\int\int K'(u)\beta_1(\xi_\tau(x))f_{Y,X}(Q_Y[\tau]+hu,x)dudx\right]^2\\
&= O(n^{-1}h^{-2}).
\end{aligned}
$$

Combining the bias and variance results, we obtain

$$E\left[\frac{1}{n}\sum_{i=1}^{n}\frac{\partial K_h(Y_i-q)}{\partial q}\bigg|_{q=Q_Y[\tau]}\cdot\beta_1(\xi_\tau(X_i))\right]-\int\beta_1(\xi_\tau(x))f^{(1)}_{Y,X}(Q_Y[\tau],x)dx=O_p(n^{-1/2}h^{-3/2}+h^r).$$

For the remaining term in (A.7) we use the fact that by Assumption 2, the second derivative of the kernel is bounded:

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\frac{\partial^2K_h(y_i-q)}{\partial^2 q}\bigg|_{q=\tilde{q}}\cdot\beta_1(\xi_\tau(x_i))&=\frac{1}{nh^3}\sum_{i=1}^{n}K''\left(\frac{y_i-\tilde{q}}{h}\right)\cdot\beta_1(\xi_\tau(x_i))\\&\leq\frac{\sup_{u\in\mathbb{R}}|K''(u)|}{h^3}\cdot\frac{1}{n}\sum_{i=1}^{n}\beta_1(\xi_\tau(x_i))\\&=O_p(h^{-3}),\end{aligned}$$

since $\frac{1}{n}\sum_{i=1}^{n}\beta_1(\xi_\tau(x_i))=O_p(1)$.

Now we show that T_2 in (A.6) satisfies

$$T_2:=\frac{1}{n}\sum_{i=1}^{n}K_h(y_i-Q_Y[\tau])\cdot\left[\hat{\beta}_1(\hat{\xi}_\tau(x_i))-\beta_1(\xi_\tau(x_i))\right]=o_p(n^{-1/2}h^{-1/2}).$$

We use the following decomposition, similar to the one in Theorem 1:

$$\hat{\beta}_1(\hat{\xi}_\tau(x))-\beta_1(\xi_\tau(x))=\hat{\beta}_1(\hat{\xi}_\tau(x))-\beta_1(\hat{\xi}_\tau(x))+\beta_1(\hat{\xi}_\tau(x))-\beta_1(\xi_\tau(x)).$$

We have

$$\begin{aligned}T_2&:=\frac{1}{n}\sum_{i=1}^{n}K_h(y_i-Q_Y[\tau])\cdot\left[\hat{\beta}_1(\hat{\xi}_\tau(x_i))-\beta_1(\xi_\tau(x_i))\right]\\&=\frac{1}{n}\sum_{i=1}^{n}K_h(y_i-Q_Y[\tau])\cdot\left[\hat{\beta}_1(\hat{\xi}_\tau(x_i))-\beta_1(\hat{\xi}_\tau(x_i))\right]+\frac{1}{n}\sum_{i=1}^{n}K_h(y_i-Q_Y[\tau])\cdot\left[\beta_1(\hat{\xi}_\tau(x_i))-\beta_1(\xi_\tau(x_i))\right]\\&\leq\sup_{\eta\in[\epsilon,1-\epsilon]}|\hat{\beta}(\eta)-\beta(\eta)|\frac{1}{n}\sum_{i=1}^{n}K_h(y_i-Q_Y[\tau])+\frac{1}{n}\sum_{i=1}^{n}K_h(y_i-Q_Y[\tau])\cdot\left[\beta_1(\hat{\xi}_\tau(x_i))-\beta_1(\xi_\tau(x_i))\right].\end{aligned}$$

Here we use $\sup_{\eta\in[\epsilon,1-\epsilon]}|\hat{\beta}(\eta)-\beta(\eta)|=O_p(n^{-1/2})$ and $\frac{1}{n}\sum_{i=1}^{n}K_h(y_i-Q_Y[\tau])=O_p(1)$, to conclude that the first term is $o_p(n^{-1/2}h^{-1/2})$. By the Cauchy-Schwarz inequality, the second term is bounded by

$$\frac{1}{n}\sum_{i=1}^{n}K_h(y_i-Q_Y[\tau])\cdot\left[\beta_1(\hat{\xi}_\tau(x_i))-\beta_1(\xi_\tau(x_i))\right]\leq$$
$$\frac{1}{\sqrt{h}}\left[\frac{1}{nh}\sum_{i=1}^{n}K\left(\frac{y_i-Q_Y[\tau]}{h}\right)^2\right]^{1/2}\cdot\left[\frac{1}{n}\sum_{i=1}^{n}\left[\beta_1(\hat{\xi}_\tau(x_i))-\beta_1(\xi_\tau(x_i))\right]^2\right]^{1/2}.$$

Here

$$\frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{y_i - Q_Y[\tau]}{h}\right)^2 = O_p(1).$$

For the other term, we use Markov's inequality

$$\begin{aligned}
\Pr\left[\frac{1}{n}\sum_{i=1}^{n}[\beta_1(e(X_i)) - \beta_1(\xi_\tau(X_i))]^2 > \varepsilon\right]\Big|_{e=\hat{\xi}_\tau} &\leq \frac{1}{n\varepsilon}E\left|\sum_{i=1}^{n}[\beta_1(e(X_i)) - \beta_1(\xi_\tau(X_i))]^2\right|\Big|_{e=\hat{\xi}_\tau} \\
&\leq \frac{1}{\varepsilon}E\left[[\beta_1(e(X_i)) - \beta_1(\xi_\tau(X_i))]^2\right]\big|_{e=\hat{\xi}_\tau}.
\end{aligned}$$

Since by Assumption 5, $E\left[\left(n^{1/4}\left[\beta_1(e(X_i)) - \beta_1(\xi_\tau(X_i))\right]\right)^2\right]\big|_{e=\hat{\xi}_\tau} = o_p(1)$, then

$$T_2 := \frac{1}{n}\sum_{i=1}^{n} K_h(y_i - Q_Y[\tau]) \cdot \left[\beta_1(\hat{\xi}_\tau(x_i)) - \beta_1(\xi_\tau(x_i))\right] = o_p(n^{-1/2}).$$

Hence,

$$T_2 := \frac{1}{n}\sum_{i=1}^{n} K_h(y_i - Q_Y[\tau]) \cdot \left[\beta_1(\hat{\xi}_\tau(x_i)) - \beta_1(\xi_\tau(x_i))\right] = o_p(n^{-1/2}h^{-1/2}).$$

For T_3 in (A.6), we have

$$\begin{aligned}
T_3 &:= \frac{1}{n}\sum_{i=1}^{n}\left[K_h(y_i - \hat{Q}_Y[\tau]) - K_h(y_i - Q_Y[\tau])\right] \cdot \left[\hat{\beta}_1(\hat{\xi}_\tau(x_i)) - \beta_1(\xi_\tau(x_i))\right] \\
&\leq \sup_{\eta\in[\epsilon,1-\epsilon]}|\hat{\beta}(\eta) - \beta(\eta)|\frac{1}{n}\sum_{i=1}^{n}\left[K_h(y_i - \hat{Q}_Y[\tau]) - K_h(y_i - Q_Y[\tau])\right] \\
&+ \frac{1}{n}\sum_{i=1}^{n}\left[K_h(y_i - \hat{Q}_Y[\tau]) - K_h(y_i - Q_Y[\tau])\right] \cdot \left[\beta_1(\hat{\xi}_\tau(x_i)) - \beta_1(\xi_\tau(x_i))\right].
\end{aligned}$$

By (A.5) we have that

$$\begin{aligned}
\sup_{\eta\in[\epsilon,1-\epsilon]}|\hat{\beta}(\eta) - \beta(\eta)|\frac{1}{n}\sum_{i=1}^{n}\left[K_h(y_i - \hat{Q}_Y[\tau]) - K_h(y_i - Q_Y[\tau])\right] &= O_p(n^{-1/2})o_p(n^{-1/2}h^{-1/2}) \\
&= o_p(n^{-1}h^{-1/2}).
\end{aligned}$$

For other term we have

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\left[K_h(y_i-\hat{Q}_Y[\tau])-K_h(y_i-Q_Y[\tau])\right]\cdot\left[\beta_1(\hat{\xi}_\tau(x_i))-\beta_1(\xi_\tau(x_i))\right]\\
&=(\hat{Q}_Y[\tau]-Q_Y[\tau])\frac{1}{nh^2}\sum_{i=1}^{n}K'\left(\frac{y_i-\tilde{Q}_Y[\tau]}{h}\right)\cdot\left[\beta_1(\hat{\xi}_\tau(x_i))-\beta_1(\xi_\tau(x_i))\right]\\
&\leq(\hat{Q}_Y[\tau]-Q_Y[\tau])\frac{1}{h}\left[\frac{1}{nh^2}\sum_{i=1}^{n}K'\left(\frac{y_i-\tilde{Q}_Y[\tau]}{h}\right)^2\right]^{1/2}\cdot\left[\frac{1}{n}\sum_{i=1}^{n}\left[\beta_1(\hat{\xi}_\tau(x_i))-\beta_1(\xi_\tau(x_i))\right]^2\right]^{1/2}\\
&=O_p(n^{-1/2})\frac{1}{h}o_p(n^{-1/4})\\
&=o_p(n^{-3/4}h^{-1}),
\end{aligned}
$$

since

$$
\frac{1}{nh^2}\sum_{i=1}^{n}K'\left(\frac{y_i-\tilde{Q}_Y[\tau]}{h}\right)^2=O_p(1).
$$

Now, $n^{1/2}h^{1/2}n^{-3/4}h^{-1}=n^{-1/4}h^{-1/2}$, so we need $nh^2\to\infty$, which is satisfied by Assumption 4 which states $nh^5\to\infty$. Therefore, T_3 is $o_p(n^{-1/2}h^{-1/2})$. This means that

$$
\hat{m}_1(\hat{Q}_Y[\tau],\hat{\beta}_1,\hat{\xi}_\tau)-\hat{m}_1(Q_Y[\tau],\beta_1,\xi_\tau)=o_p(n^{-1/2}h^{-1/2}). \tag{A.8}
$$

Therefore, (A.5) and (A.8) imply that (A.3) is actually

$$
\widehat{UQPE}_{X_1}(\tau)-\widetilde{UQPE}_{X_1}(\tau)=o_p(n^{-1/2}h^{-1/2}).
$$

A.4 Proof of Corollary 1

Recall that by equation (11), we have that

$$
UQPE_{X_1}(\tau)=E\left[\beta_1(\xi_\tau(X))|Y=Q_Y[\tau]\right],
$$

and that we defined

$$
U_\tau:=\beta_1(\xi_\tau(X))-E\left[\beta_1(\xi_\tau(X))|Y\right]
$$

and, by construction, $E[U_\tau|Y]=0$ a.s. and, by Assumption 6, $E[U_\tau^2|Y=y]=\sigma_\tau^2(y)<\infty$ for every y in the support of Y. Moreover, for (x_i,y_i), we have

$$
u_{\tau,i}:=\beta_1(\xi_\tau(x_i))-E\left[\beta_1(\xi_\tau(X))|Y=y_i\right]. \tag{A.9}
$$

We focus on

$$\begin{aligned}
\sqrt{nh}\left(\widetilde{UQPE}_{X_1}(\tau) - UQPE_{X_1}(\tau)\right) &= \sqrt{nh}\left(\frac{\sum_{i=1}^n K_h(y_i - Q_Y[\tau])\cdot \beta_1(\xi_\tau(x_i))}{\sum_{i=1}^n K_h(y_i - Q_Y[\tau])} - UQPE_{X_1}(\tau)\right) \\
&= \sqrt{nh}\left(\frac{\sum_{i=1}^n K_h(y_i - Q_Y[\tau])\cdot [\beta_1(\xi_\tau(x_i)) - UQPE_{X_1}(\tau)]}{\sum_{i=1}^n K_h(y_i - Q_Y[\tau])}\right) \\
&= \sqrt{nh}\left(\frac{\sum_{i=1}^n K_h(y_i - Q_Y[\tau])\cdot u_{\tau,i}}{\sum_{i=1}^n K_h(y_i - Q_Y[\tau])}\right) \\
&+ \sqrt{nh}\left(\frac{\frac{1}{n}\sum_{i=1}^n K_h(y_i - Q_Y[\tau])\cdot [E\left[\beta_1(\xi_\tau(X))|Y = y_i\right] - UQPE_{X_1}(\tau)]}{\frac{1}{n}\sum_{i=1}^n K_h(y_i - Q_Y[\tau])}\right).
\end{aligned} \tag{A.10}$$

Consider the first term:

$$\sqrt{nh}\left(\frac{\sum_{i=1}^n K_h(y_i - Q_Y[\tau])\cdot u_{\tau,i}}{\sum_{i=1}^n K_h(y_i - Q_Y[\tau])}\right) = \frac{1}{\hat{f}_Y(Q_Y[\tau])}\sum_{i=1}^n \frac{1}{\sqrt{nh}}K\left(\frac{y_i - Q_Y[\tau]}{h}\right)\cdot u_{\tau,i}.$$

For this term we use the Lindberg-Feller CLT. First we compute the variance of the sum.

$$\begin{aligned}
Var\left[\sum_{i=1}^n \frac{1}{\sqrt{nh}}K\left(\frac{Y_i - Q_Y[\tau]}{h}\right)\cdot u_{\tau,i}\right] &= \frac{1}{h}E\left[K\left(\frac{Y - Q_Y[\tau]}{h}\right)^2 \sigma_\tau^2(Y)\right] \\
&= \frac{1}{h}\int K\left(\frac{y - Q_Y[\tau]}{h}\right)^2 \sigma_\tau^2(y) f_Y(y)dy \\
&= \int K(u)^2\sigma^2(Q_Y[\tau] + hu) f_Y(Q_Y[\tau] + hu)du \\
&\to \sigma_\tau^2(Q_Y[\tau]) f_Y(Q_Y[\tau]) \int K(u)^2 du.
\end{aligned}$$

because $\sigma_\tau^2(y)$ and $f_Y(y)$ are continuous, and are bounded. The conclusion follows from the dominated convergence theorem. We write

$$\sigma_{\tau,n}^2 := Var\left[\sum_{i=1}^n \frac{1}{\sqrt{nh}}K\left(\frac{Y_i - Q_Y[\tau]}{h}\right)\cdot u_{\tau,i}\right] \to \sigma_{\tau,0}^2.$$

To apply the Lindberg-Feller CLT, we define

$$\omega_{in} := \frac{1}{\sqrt{nh}}K\left(\frac{Y_i - Q_Y[\tau]}{h}\right).$$

We need to show that, for some $\delta > 0$,

$$\lim_{n\to\infty}\sum_{i=1}^n E\left|\frac{\omega_{in}U_{\tau,i}}{\sigma_{\tau,n}}\right|^{2+\delta} = 0.$$

We have

$$\sum_{i=1}^{n} E\left|\frac{\omega_{in}U_{\tau,i}}{\sigma_{\tau,n}}\right|^{2+\delta} = \sum_{i=1}^{n}\left|\frac{\sigma_{\tau,0}}{\sigma_{\tau,n}}\right|^{2+\delta} E\left|\frac{\omega_{in}u_{\tau,i}}{\sigma_{\tau,0}}\right|^{2+\delta}.$$

It will be sufficient to focus on

$$\begin{aligned}
\sum_{i=1}^{n} E\left|\frac{\omega_{in}U_{\tau,i}}{\sigma_{0,\tau}}\right|^{2+\delta} &= \frac{n}{|\sigma_{0,\tau}|^{2+\delta}} E\left|\omega_{in}U_{\tau,i}\right|^{2+\delta} \\
&= \frac{n}{|\sigma_{0,\tau}|^{2+\delta}} E\left[|\omega_{in}|^{2+\delta} E[|U_{\tau,i}|^{2+\delta}|Y_i]\right] \\
&\leq C\frac{n}{|\sigma_{0,\tau}|^{2+\delta}} E\left[|\omega_{in}|^{2+\delta}\right] \\
&= \frac{Cn}{|\sigma_{0,\tau}|^{2+\delta}}(nh)^{-1-\delta/2}\int K\left(\frac{y-Q_Y[\tau]}{h}\right)^{2+\delta} f_Y(y)dy \\
&= \frac{Cn}{|\sigma_{0,\tau}|^{2+\delta}}(nh)^{-1-\delta/2}h\int K(u)^{2+\delta} f_Y(Q_Y[\tau]+hu)du \\
&= \frac{c}{(nh)^{\delta/2}}\int K(u)^{2+\delta} f_Y(Q_Y[\tau]+hu)du,
\end{aligned}$$

which goes to 0 since $(nh)^{1/2} \to \infty$ and $f_Y(y)$ is continuous at $y = Q_Y[\tau]$. We have used the bound of the higher order conditional expectation of U_τ: $E[|U_{\tau,i}|^{2+\delta}|Y_i] < C$ a.s., and that $\int |K(u)|^{2+\delta}du < \infty$. Therefore,

$$\frac{1}{\hat{f}_Y(Q_Y[\tau])}\sum_{i=1}^{n}\frac{1}{\sqrt{nh}}K\left(\frac{y_i-Q_Y[\tau]}{h}\right)\cdot u_{\tau,i} \xrightarrow{d} N\left(0, \sigma_\tau^2(Q_Y[\tau])f_Y(Q_Y[\tau])^{-1}\int K(u)^2du\right) \quad \text{(A.11)}$$

since $\hat{f}_Y(Q_Y[\tau]) = f_Y(Q_Y[\tau]) + o_p(1)$.

The second term in the expansion of $\sqrt{nh}\left(\widetilde{UQPE}_{X_1}(\tau) - UQPE_{X_1}(\tau)\right)$ is a bias term. We now find its rate of convergence. We start with the numerator.

$$\begin{aligned}
&E\left[\frac{1}{n}\sum_{i=1}^{n} K_h(Y_i - Q_Y[\tau])\cdot[E[\beta_1(\xi_\tau(X))|Y=Y_i] - UQPE_{X_1}(\tau)]\right] \\
&= \frac{1}{h}\int_{\mathcal{Y}} K\left(\frac{y-Q_Y[\tau]}{h}\right)[E[\beta_1(\xi_\tau(X))|Y=y] - UQPE_{X_1}(\tau)]\, f_Y(y)dy \\
&= \int K(u)\,[E[\beta_1(\xi_\tau(X))|Y=Q_Y[\tau]+hu] - UQPE_{X_1}(\tau)]\, f_Y(Q_Y[\tau]+hu)du \\
&= \int K(u)\,[E[\beta_1(\xi_\tau(X))|Y=Q_Y[\tau]+hu] - UQPE_{X_1}(\tau)]\, f_Y(Q_Y[\tau]+hu)du \\
&= \int K(u)E[\beta_1(\xi_\tau(X))|Y=Q_Y[\tau]+hu]\, f_Y(Q_Y[\tau]+hu)du \\
&- UQPE_{X_1}(\tau)\int_{\mathcal{Y}} K(u)f_Y(Q_Y[\tau]+hu)du.
\end{aligned}$$

We do a Taylor expansion on the density and the conditional expectation and we use the fact that $\int K(u)du = 1$, $\int u^j K(u)du = 0$ when $j = 1, \ldots, r-1$, and $\int u^r K(u)du < \infty$. Let $(Ef_Y)^{(j)}(q)$ denote the j-derivative with respect to y of the product $E\left[\beta_1(\xi_\tau(X))|Y = y\right] f_Y(y)$ evaluated at $y = q$. The first term is

$$\begin{aligned}
&\int K(u)E\left[\beta_1(\xi_\tau(X))|Y = Q_Y[\tau] + hu\right] f_Y(Q_Y[\tau] + hu)du \\
&= UQPE_{X_1}(\tau) + \int K(u) \sum_{j=1}^{r-1} \frac{h^j u^j (Ef_Y)^{(j)}(Q_Y[\tau])}{j!} du \\
&+ \frac{h^r}{r!} \int K(u)u^r (Ef_Y)^{(j)}(\tilde{Q}_Y[\tau])du \\
&= UQPE_{X_1}(\tau) + O(h^r),
\end{aligned}$$

since the derivatives are uniformly bounded. Now, for the other term we do a similar expansion of the density.

$$\begin{aligned}
&UQPE_{X_1}(\tau) \int_{\mathcal{Y}} K(u) f_Y(Q_Y[\tau] + hu)du \\
&= UQPE_{X_1} + \int_{\mathcal{Y}} K(u) \sum_{j=1}^{r-1} \frac{h^j u^j f_Y^{(j)}(Q_Y[\tau])}{j!} du + \frac{h^r}{r!} \int K(u)u^r f_Y^{(r)}(\tilde{Q}_Y[\tau])du \\
&= UQPE_{X_1}(\tau) + O(h^r).
\end{aligned}$$

Therefore, we obtain that the bias is of order $O(h^r)$:

$$E\left[\frac{1}{n}\sum_{i=1}^{n} K_h(Y_i - Q_Y[\tau]) \cdot \left[E\left[\beta_1(\xi_\tau(X))|Y = Y_i\right] - UQPE_{X_1}(\tau)\right]\right] = O(h^r).$$

Now, for the variance we have

$$\begin{aligned}
&Var\left[\frac{1}{n}\sum_{i=1}^{n} K_h(Y_i - Q_Y[\tau]) \cdot \left[E\left[\beta_1(\xi_\tau(X))|Y = Y_i\right] - UQPE_{X_1}(\tau)\right]\right] \\
&= \frac{1}{nh^2} Var\left[K\left(\frac{Y_i - Q_Y[\tau]}{h}\right)\left[E\left[\beta_1(\xi_\tau(X))|Y = Y_i\right] - UQPE_{X_1}(\tau)\right]\right] \\
&\leq \frac{1}{nh^2} E\left[K\left(\frac{Y_i - Q_Y[\tau]}{h}\right)^2 \left[E\left[\beta_1(\xi_\tau(X))|Y = Y_i\right] - UQPE_{X_1}(\tau)\right]^2\right] \\
&= \frac{1}{nh^2} \int K\left(\frac{y - Q_Y[\tau]}{h}\right)^2 \left[E\left[\beta_1(\xi_\tau(X))|Y = y\right] - UQPE_{X_1}(\tau)\right]^2 f_Y(y)dy \\
&= \frac{1}{nh} \int K(u)^2 \left[E\left[\beta_1(\xi_\tau(X))|Y = Q_Y[\tau] + hu\right] - UQPE_{X_1}(\tau)\right]^2 f_Y(Q_Y[\tau] + hu)du \\
&= \frac{1}{nh} \int K(u)^2 [huE^{(1)}\left[\beta_1(\xi_\tau(X))|Y = \tilde{Q}_Y[\tau]\right]]^2 \left[f_Y(Q_Y[\tau]) + huf_Y^{(1)}(\tilde{Q}_Y[\tau])\right] du.
\end{aligned}$$

This implies that

$$Var\left[\frac{1}{n}\sum_{i=1}^{n}K_h(Y_i - Q_Y[\tau])\cdot[E[\beta_1(\xi_\tau(X))|Y=Y_i] - UQPE_{X_1}(\tau)]\right] = O(n^{-1}h).$$

Therefore, we obtain that

$$\frac{1}{n}\sum_{i=1}^{n}K_h(Y_i - Q_Y[\tau])\cdot[E[\beta_1(\xi_\tau(X))|Y=Y_i] - UQPE_{X_1}(\tau)] = O_p(h^r + n^{-1/2}h^{1/2}).$$

Under Assumption 4, this term is $o_p(n^{-1/2}h^{-1/2})$, since

$$(nh)^{1/2}O_p(h^r + n^{-1/2}h^{1/2}) = O_p((nh)^{1/2}h^r + h^{1/4}) = o_p(1),$$

since $(nh)^{1/2}h^r \to 0$, and $h \to 0$ as $n \to \infty$. Therefore, the bias term is

$$\sqrt{nh}\left(\frac{\frac{1}{n}\sum_{i=1}^{n}K_h(y_i - Q_Y[\tau])\cdot[E[\beta_1(\xi_\tau(X))|Y=y_i] - UQPE_{X_1}(\tau)]}{\frac{1}{n}\sum_{i=1}^{n}K_h(y_i - Q_Y[\tau])}\right) = \frac{o_p(1)}{f_Y(Q_Y[\tau]) + o_p(1)} = o_p(1).$$

Combining this fact with (A.11), we obtain

$$\sqrt{nh}\left(\widetilde{UQPE}_{X_1}(\tau) - UQPE_{X_1}(\tau)\right) \xrightarrow{d} N\left(0, \sigma_\tau^2(Q_Y[\tau])f_Y(Q_Y[\tau])^{-1}\int K(u)^2du\right).$$

In view of the result in Theorem 2, we have

$$\sqrt{nh}\left(\widehat{UQPE}_{X_1}(\tau) - UQPE_{X_1}(\tau)\right) \xrightarrow{d} N\left(0, \sigma_\tau^2(Q_Y[\tau])f_Y(Q_Y[\tau])^{-1}\int K(u)^2du\right).$$

A.5 Inference for the matching function

By Theorem 1, we have that

$$\hat{\xi}_\tau(x) - \xi_\tau(x) = -\frac{1}{x'\dot{\beta}(\xi_\tau(x))}\frac{1}{n}\sum_{i=1}^{n}x'\Psi_i(\xi_\tau(x)) + \frac{1}{x'\dot{\beta}(\xi_\tau(x))}\frac{1}{n}\sum_{i=1}^{n}\psi_i(\tau) + o_p(n^{-1/2}).$$

We use this influence function representation to estimate the asymptotic variance. First, consider the term $x'\dot{\beta}(\xi_\tau(x))$. Assumptions 1.1 and 1.2, allow us to write $F_{Y|X}(x'\beta(\eta)|x) = \eta$, so that $x'\dot{\beta}(\eta) = f_{Y|X}(x'\beta(\eta)|x)^{-1}$, and hence

$$x'\dot{\beta}(\xi_\tau(x)) = \frac{1}{f_{Y|X}(x'\beta(\xi_\tau(x))|x)}.$$

The function $\eta \mapsto f_{Y|X}(x'\beta(\eta)|x)^{-1}$ is usually called the sparcity function. It plays a central role in inference for quantile regression, see Section 3.4.2 in Koenker (2005), but also in (conditional) mode estimation as in Ota, Kato, and Hara (2019). The estimator is

$$x'\widehat{\dot{\beta}(\xi_\tau(x))} = \left[\frac{x'\hat{\beta}(\hat{\xi}_\tau(x) + h_{KM}) - x'\hat{\beta}(\hat{\xi}_\tau(x) - h_{KM})}{2h_{KM}}\right]^{-1}.$$

In the case where either $\hat{\xi}_\tau(x) + h_{KM} > 1$ or $\hat{\xi}_\tau(x) - h_{KM} < 0$, the implementation is

$$x'\widehat{\dot{\beta}(\xi_\tau(x))} = \left[\frac{x'\hat{\beta}(\hat{\xi}_\tau(x) + \min\{h_{KM}, \tau_{\max} - \hat{\xi}_\tau(x)\}) - x'\hat{\beta}(\hat{\xi}_\tau(x) - \min\{h_{KM}, \hat{\xi}_\tau(x) - \tau_{\min}\})}{\min\{h_{KM}, \tau_{\max} - \hat{\xi}_\tau(x)\}) + \min\{h_{KM}, \hat{\xi}_\tau(x) - \tau_{\min}\}}\right]^{-1},$$

where h_{KM} is the bandwidth of Koenker and Machado (1999), $\tau_{\min} = \epsilon$, and $\tau_{\max} = 1 - \epsilon$ for some small $1/2 > \epsilon > 0$.

Now we focus on the term $x'\Psi_i(\xi_\tau(x))$, where

$$\Psi_i(\eta) = E\left[f_{Y|X}(X'\beta(\eta)|X)XX'\right]^{-1}\left(\eta - \mathbb{1}\{y_i \leq x_i'\beta(\eta)\}\right)x_i.$$

By Powell (1991), it can be estimated by

$$x'\widehat{\Psi_i(\xi_\tau(x))} = x'\left[\frac{1}{2nh_P}\sum_{j=1}^{n}\mathbb{1}\left\{|y_j - x_j'\hat{\beta}(\hat{\xi}_\tau(x))| \leq h_P\right\}x_jx_j'\right]^{-1}\left(\hat{\xi}_\tau(x) - \mathbb{1}\{y_i \leq x_i'\hat{\beta}(\hat{\xi}_\tau(x))\}\right)x_i.$$

where h_P is given in Section 3.4.2 in Koenker (2005). Finally,

$$\psi_i(\tau) = f_Y(Q_Y[\tau])^{-1}(\tau - \mathbb{1}\{y_i \leq Q_Y[\tau]\})$$

can be estimated by

$$\widehat{\psi_i(\tau)} = \hat{f}_Y(\hat{Q}_Y[\tau])^{-1}\left(\tau - \mathbb{1}\{y_i \leq \hat{Q}_Y[\tau]\}\right).$$

So, the estimator of the asymptotic variance is

$$\frac{1}{n}\frac{1}{x'\widehat{\dot{\beta}(\xi_\tau(x))}^2}\sum_{i=1}^{n}\left(-x'\widehat{\Psi_i(\xi_\tau(x))} + \widehat{\psi_i(\tau)}\right)^2.$$

References

ANDREWS, D. W. K. (1995): "Nonparametric Kernel Estimation for Semiparametric Models," *Econometric Theory*, 11, 560–596.

ANGRIST, J., V. CHERNOZHUKOV, AND I. FERNÁNDEZ-VAL (2006): "Quantile Regression under Misspecification, with an Application to the U.S. Wage Structure," *Econometrica*, 74, 539–563.

AUTOR, D. H., L. S. KATZ, AND M. S. KEARNEY (2005): "Rising Wage Inequality: The Role of Composition and Prices," NBER Working Paper 11628.

BERA, A., A. GALVAO, G. MONTES-ROJAS, AND S. PARK (2016): "Asymmetric Laplace Regression: Maximum Likelihood, Maximum Entropy and Quantile Regression," *Journal of Econometric Methods*, 5(1), 79–101.

BLUNDELL, R., X. CHEN, AND D. KRISTENSEN (2007): "Semi-nonparametric IV estimation of shape-invariant Engel curves," *Econometrica*, 75(6), 1613–1669.

CHAI, A., AND A. MONETA (2010): "Retrospectives: Engel curves," *Journal of Economic Perspectives*, 24(1), 225–40.

CHERNOZHUKOV, V., I. FERNÁNDEZ-VAL, AND A. GALICHON (2010): "Quantile and Probability Curves Without Crossing," *Econometrica*, 78(3), 1093–1125.

CHERNOZHUKOV, V., I. FERNÁNDEZ-VAL, AND A. E. KOWALSKI (2015): "Quantile regression with censoring and endogeneity," *Journal of Econometrics*, 186(1), 201–221.

CHERNOZHUKOV, V., I. FERNÁNDEZ-VAL, AND B. MELLY (2013): "Inference on Counterfactual Distributions," *Econometrica*, 81(6), 2205–2268.

ESCANCIANO, J. C., D. JACHO-CHÁVEZ, AND A. LEWBEL (2014): "Uniform Convergence of Weighted Sums of Non- and Semi-parametric Residuals for Estimation and Testing," *Journal of Econometrics*, 178, 426–443.

FIRPO, S., N. FORTIN, AND T. LEMIEUX (2009): "Unconditional Quantile Regression," *Econometrica*, 77(3), 953–973.

FORTIN, N., T. LEMIEUX, AND S. FIRPO (2011): "Decomposition Methods in Economics," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 4, pp. 1–12. Amsterdam: Elsevier.

HALL, P., AND J. HOROWITZ (2013): "A Simple Bootstrap Method For Constructing Nonparametric Confidence Bands For Functions," *Annals of Statistics*, 41, 1892–1921.

HÄRDLE, W., AND A. BOWMAN (1988): "Bootstrapping in Nonparametric Regression: Local Adaptive Smoothing and Confidence Bands," *Journal of the American Statistical Association*, 83, 1892–1921.

INOUE, A., T. LI, AND Q. XU (2021): "Two Sample Unconditional Quantile Effect," ARXIV: https://arxiv.org/pdf/2105.09445.pdf.

KOENKER, R. (2005): *Quantile Regression*. Cambridge University Press, Cambridge.

KOENKER, R., AND G. W. BASSETT (1978): "Regression Quantiles," *Econometrica*, 46, 33–49.

KOENKER, R., V. CHERNOZHUKOV, X. HE, AND L. PENG (2017): *Handbook of Quantile Regression*. Chapman and Hall/CRC.

KOENKER, R., AND K. HALLOCK (2001): "Quantile Regression," *Journal of Economic Perspecives*, 15 (4), 143–156.

KOENKER, R., AND J. A. F. MACHADO (1999): "Godness of Fit and Related Inference Processes for Quantile Regression," *Journal of the American Statistical Association*, 94, 1296–1310.

LEE, Y.-Y. (2021): "Nonparametric Weighted Average Quantile Derivative," *Econometric Theory*, pp. 1–39.

LEWBEL, A. (1997): "Consumer demand systems and household equivalence scales," *Handbook of Applied Econometrics*, 2, 167–201.

——— (2008): "Engel curves," *The New Palgrave Dictionary of Economics*, 2.

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

MAMMEN, E., C. ROTHE, AND M. SCHIENLE (2012): "Nonparametric Regression With Nonparametrically Generated Covariates," *Annals of Statistics*, 40, 1132–1170.

——— (2016): "Semiparametric Estimation with Generated Covariates," *Econometric Theory*, 32, 1140–1177.

MARTINEZ-IRIARTE, J. (2021): "Sensitivity Analysis in Unconditional Quantile Effects," Working Paper.

MARTINEZ-IRIARTE, J., G. MONTES-ROJAS, AND Y. SUN (2022): "Unconditional Effects of General Policy Interventions," Working Paper.

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

OTA, H., K. KATO, AND S. HARA (2019): "Quantile Regression Approach to Conditional Mode Estimation," *Electronic Journal of Statistics*, 13, 3120–3160.

PAGAN, A., AND A. ULLAH (1999): *Nonparametric Econometrics*. Cambridge University Press.

Powell, J. (1991): "Estimation of Monotonic Regression Models under Quantile Restrictions," in *Nonparametric and Semiparametric Models in Econometrics*, ed. by J. Powell, and G. Tauchen, pp. 1–12. Cambridge: Cambridge University Press.

Rios-Avila, F. (2020): "Recentered Influence Functions (RIFs) in Stata: RIF Regression and RIF Decomposition," *Stata Journal*, 20, 51–94.

Rothe, C. (2012): "Partial Distributional Policy Effects," *Econometrica*, 80(5), 2269–2301.

Sasaki, Y., T. Ura, and Y. Zhang (2022): "Unconditional Quantile Regression with High-Dimensional Data," *Quantitative Economics*, 13, 955–978.

Serfling, R. J. (1980): *Approximation Theorems of Mathematical Statistics*. New York: Wiley.

Song, K. (2008): "Uniform Convergence of Series Estimators over Function Spaces," *Econometric Theory*, 24, 1463–1499.

Sperlich, S. (2009): "A Note on Non-Parametric Estimation with Predicted Variables," *Econometrics Journal*, 12, 382–395.

van der Vaart, A. (1998): *Asymptotic Statistics*. Cambridge University Press, Cambridge.