Quantile Regressions via Method of Moments with multiple fixed effects

Fernando Rios-Avila[a,*], Leonardo Siles, Gustavo Canavire-Bacarreza[b]

[a] *Levy Economics Institute of Bard College,*

[b] *World Bank,*

Abstract

This paper proposes a new method to estimate quantile regressions with multiple fixed effects. The method expands on the strategy proposed by Machado and Santos Silva (2019), allowing for multiple fixed effects, and providing various alternatives for the estimation of Standard errors. We provide Monte Carlo simulations to show the finite sample properties of the proposed method in the presence of two sets of fixed effects. Finally, we apply the proposed method to estimate **something interesting**

Keywords: Fixed effects, Linear heteroskedasticity, Location-scale model, Quantile regression

1. Introduction

Quantile regression (QR), introduced by Koenker and Bassett (1978), is an estimation strategy used for modeling the relationships between explanatory variables X and the conditional quantiles of the dependent variable $q_\tau(y|x)$. Using QR one can obtain richer characterizations of the relationships between dependent and independent variables, by accounting for otherwise unobserved heterogeneity.

A relatively recent development in the literature has focused on extending quantile regressions analysis to include individual fixed effects in the framework of panel data. However, as described in Neyman and Scott (1948), and Lancaster (2000), when individual fixed effects are included in quantile regression analysis it generates an incident parameter problem. While many strategies have been proposed for estimating this type of model (see Galvão and Kato, 2018 for a brief review), neither has become standard because of their restrictive assumptions in regards to the individual effects, the computational complexity, and implementation.

*Corresponding author

Email addresses: `friosavi@levy.org` (Fernando Rios-Avila), `leo@gmail.com` (Leonardo Siles), `gcanavire@worldbank.org` (Gustavo Canavire-Bacarreza)

More recently, Machado and Santos Silva (2019) (MSS hereafter) proposed a methodology based on a conditional location-scale model similar to the one described in He (1997) and Zhao (2000), for the estimation of quantile regressions models for panel data via a method of moments. This method allows individual fixed effects allowing to have heterogeneous effects on the entire conditional distribution of the outcome, rather constraining their effect to be a location shift only as in Canay (2011), Koenker (2004), and Lancaster (2000).

In principle, under the assumption that data generating process behind the data is based on a multiplicative heteroskedastic process that is linear in parameters (Cameron and Trivedi (2005), Machado and Santos Silva (2019), He (1997) and Zhao (2000)), the effect of a variable X on the q_th quantile can be derived as the combination of a location effect, and scale effect moderated by the quantile of an underlying i.i.d. error. For statistical inference, MSS derives the asymptotic distribution of the estimator, suggesting the use of bootstrap standard errors, as well.

While this methodology is not meant to substitute the use of standard quantile regression analysis, given the assumptions required for the identification of the model, it provides a simple and fast alternative for the estimation of quantile regression models with individual fixed effects.

In this framework, our paper expands on Machado and Santos Silva (2019), following some of the suggestions by the authors regarding further research. First, making use of the properties of GMM estimators, we derive various alternatives for the estimation of standard errors based on the empirical Influence functions of the estimators. Second, we reconsider the application of Frisch–Waugh–Lovell (FWL) theorem (Frisch and Waugh (1933) and Lovell (1963)) to extend the MSS estimator to allow for the inclusion of multiple fixed effects, for example, individual and year fixed effects.

The rest of the paper is restructured as follows. Section 2 presents the basic setup of the location-Scale model described in He (1997) and Zhao (2000), tying the relationship between the standard quantile regression model, and the location and scale model. It also revisits MSS methodology, proposing alternative estimators for the standard errors based on the properties of GMM estimators and the empirical influence functions. It also shows that FWL theorem can be used to control for multiple fixed effects. Section 3 presents the results of a small simulation study and Section 4 illustrates the application of the proposed methods with two empirical examples. Seccion 5 concludes.

2. Methodology

2.1. Quantile Regression: Location-Scale model

Quantile regressions are used to identify relationships between the explanatory variables x and the conditional quantiles of the dependent variable $Q(y|\tau, X)$. This relationship is commonly assumed to follow a linear functional form:

$$q(Y|X, \tau) = X\beta(\tau) \tag{1}$$

This allows for nonlinearities in the effect of X on Y across all values of τ. This formulation can also be related to a random coefficient model, where all coefficients are assumed to be some nonlinear function of τ, where τ follows a random uniform distribution.

An alternative formulation of quantile regressions is the location-scale model. This approach assumes that the conditional quantile of Y given X and τ can be expressed as a combination of two models: the location model, which describes the central tendency of the conditional distribution, and the scale model, which describes deviations from the central tendency:

$$q(Y|X,\tau) = X\beta + X\gamma(\tau) \tag{2}$$

Here, the location parameters β are typically identified using a linear regression model (as in Machado and Santos Silva (2019)), or a median regression (as in Melly (2005)), and the scale parameters $\gamma(\tau)$ can be estimated using standard approaches.

Both the standard quantile regression (Equation 1) and the location-scale specification (Equation 2) can be estimated as the solution to a weighted minimization problem:

$$\hat{\beta}(\tau) = \underset{\beta}{\operatorname{argmin}} \left(\sum_{i \in y_i \geq x_i'\beta} \tau(y_i - x_i'\beta) - \sum_{i \in y_i < x_i'\beta} (1-\tau)(y_i - x_i'\beta) \right) \tag{3}$$

One characteristic of this estimator is that the $\beta(\tau)$ coefficients are identified locally, and thus the estimated quantile coefficients will exhibit considerable variation when analyzed across τ. It is also implicit that if one requires an analysis of the entire distribution, it would be necessary to estimate the model for each quantile.[1]

One insightful extension to the location-scale parameterizations suggested by He (1997), Cameron and Trivedi (2005), and Machado and Santos Silva (2019) is to assume that the data-generating process (DGP) can be written as a linear model with a multiplicative heteroskedastic process that is linear in parameters.[2]

$$\begin{aligned} y_i &= x_i'\beta + \nu_i \\ \nu_i &= \varepsilon_i \times x_i'\gamma \end{aligned} \tag{4}$$

Under the assumption that ε is an independent and identically distributed (iid) unobserved random variable that is independent of X, the conditional quantile of Y given X and τ can be written as:

[1]There are other estimators that provide smoother estimates for the quantile regression coefficients using a kernel local weighted approach (Kaplan and Sun, 2017), as well as identifying the full set of quantile coefficients simultaneously assuming some parametric functional forms (Frumento and Bottai, 2016).

[2]Machado and Santos Silva (2019) also discuss a model where heteroskedasticity can be an arbitrary nonlinear function $\sigma(x_i'\gamma)$, but develop the estimator for the linear case, i.e., when $\sigma()$ is the identity function.

$$q(Y|X,\tau) = X\beta + q(\varepsilon|\tau) \times X\gamma \tag{5}$$

In this setup, the traditional quantile coefficients are identified as the location model coefficients, plus the scale model coefficients moderated by the $\tau_t h$ unconditional quantile of the standardized error ε.

$$\beta(\tau) = \beta + q(\varepsilon|\tau) \times \gamma \tag{6}$$

While this specification imposes a strong assumption on the DGP, it has two advantages over the standard quantile regression model. First, because the location and scale model can be identified globally, with only a single parmater ($q(\varepsilon|\tau)$) requiring local estimation, this estimation approach would be more efficient than the standard quantile regression model (Zhao (2000)). Second, under the assumption that $X\gamma$ is strictly possitive, the model would produce quantile coefficients that do not cross.

Following MSS, the quantile regression model defined by Equation 5 can be estimated using a method of moments approach. And while its possible to identify all coefficients ($\beta, \gamma, q(\varepsilon|\tau)$) simultaneously, we describe and use the implementation approach advocated by MSS which identifies each set of coefficients separately.

1. The location model can be estimated using a standard linear regression model, where the dependent variable is the outcome Y, and the independent variables are the explanatory variables X (including a constant) with an error u, which is by definition heteroskedastic. In this case, the location model coefficients are identified under the following condition:

$$\begin{aligned} y_i &= x_i'\beta + \nu_i \\ E[x_i \nu_i] &= 0 \end{aligned} \tag{7}$$

2. After the location model is estimated, the scale coefficients can be identified by modeling heteroskedasticity as a linear function of characteristics X. For this we use the absolute value of the errors from the location model u as dependent variable, which would allow us to estimate the conditional standard deviation (rather than conditional variance) of the errors. In this case, the coefficients are identified under the following condition:

$$\begin{aligned} |\nu_i| &= x_i'\gamma + \omega_i \\ E[x_i \omega_i] &= 0 \\ E[x_i(|\nu_i| - x_i'\gamma)] &= 0 \end{aligned} \tag{8}$$

3. Finally, given the location and scale coefficients, the τ_{th} quantile of the error e can be estimated using the following condition:

$$
\begin{aligned}
E\left[\mathbb{1}\left(x_i'(\beta+\gamma q(\varepsilon|\tau))\geq y_i\right)-\tau\right]&=0\\
E\left[\mathbb{1}\left(q(\varepsilon|\tau)\geq \frac{y_i-x_i'\beta}{x_i'\gamma}\right)-\tau\right]&=0
\end{aligned}
\tag{9}
$$

Where one identifies the quantile of the error ε using standardized errors $\frac{y_i-x_i'\beta}{x_i'\gamma}$, or by finding the values that identify the overall quantile coefficients $\beta(\tau)=\beta+\gamma q(\varepsilon|\tau)$. Afterwords, the conditional quantile coefficients is simply defined as the combination of the location and scale coefficients.

2.2. Standard Errors: GLS, Robust, Clustered

As discussed in the previous section, the estimation of quantile regression coefficients using the location-scale model with heteroskedstic linear errors can be estimated using a the following set of moments, which fits in the Generalized Method of Moments framework:

$$
\begin{aligned}
E[x_i\nu_i]=E[h_{1,i}]&=0\\
E[x_i(|\nu_i|-x_i\gamma)]=E[h_{2,i}]&=0\\
E\left[\mathbb{1}\left(q(\varepsilon|\tau)\geq \frac{y_i-x_i'\beta}{x_i'\gamma}\right)-\tau\right]=E[h_{3,i}]&=0
\end{aligned}
\tag{10}
$$

Under the conditions described in Newey and McFadden (1994) (see section 7), Cameron and Trivedi (2005) (see chapter 6.3.9) or as shown in Machado and Santos Silva (2019), the location, scale and residual quantile coefficients are asymptotically normal.[3]

Call $\theta=[\beta' \quad \gamma' \quad q(\varepsilon|\tau)']'$ the set of coefficients that are identified by the modement conditions in Equation 10, a just identified model. And the function h_i is a vector function that stacks all the moments at the individual level described in Equation 10. Then $\hat{\theta}$ follows a normal distribution with mean θ and variance-covariance matrix $V(\theta)$ that is estimated as:

$$
\hat{V}(\hat{\theta})=\frac{1}{N}\bar{G}(\hat{\theta})^{-1}\left(\frac{1}{N}\sum_{i=1}^{N}h_ih_i'\Big|_{\theta=\hat{\theta}}\right)\bar{G}(\hat{\theta})^{-1}
$$

Which is equivalent to the Eicker-White Heteroskedastic-Consistent estimator for least-squares estimators.

Here, the inner product is the moment covariance matrix, and $\bar{G}(\theta)$ is the Jacobian matrix of the moment equations (Equation 10) evaluated at $\hat{\theta}$.

$$
\bar{G}(\theta)=-\frac{1}{N}\sum_{i=1}\frac{\partial h_i}{\partial\theta'}\Big|_{\theta=\hat{\theta}}
$$

[3] Zhao (2000) also shows that the quantile coefficients for the location-scale model also follows a normal distribution, but uses the assumption that the location model is derived using a least absolute deviation approach (median regression).

In this framework, the quantile regression coefficients, a combination of the location-scale-quantile estimates, will follow a normal distribution with mean $\beta(\tau) = \beta + q(\varepsilon|\tau)\gamma$ and variance-covariance matrix equal to:

$$\hat{V}(\beta(\tau)) = \Xi \hat{V}(\hat{\theta}) \Xi'$$

where Ξ is a $k \times (2k+1)$ matrix defined as:

$$\Xi = [I(k), \hat{q}(\varepsilon|\tau) \times I(k), \hat{\gamma}]$$

with $I(k)$ being an identity matrix of dimension k (number of explanatory variables in X including the constant).

While it is possible to estimate the variance-covariance matrix using simultaneous model estimation, for a just identified model, it is more efficient to estimate each set of coefficients separately. The variance-covariance matrix can be estimated using the empirical influence functions of the estimators (see Jann (2020) for an overview of the application, and Hampel et al. (2005) for an in-depth review).

1. Need to add the general condition Var=(inf (theta) * inf(theta))

Thatn inf(theta) = inf beta gamma and q)tay

and finally define what the IFs are (see appendix)

The different types of Standard errors estimation, thus, depend on the assumptions imposed for the estimation of $V(\theta)$.

2.2.1. Robust Standard Errors

Cross product of IFs

2.2.2. Clustered Standard Errors

Sun then cross product

2.2.3. GLS Standard Errors

Original paper;

2.3. Multiple Fixed Effects: Expanding on Machado and Santos Silva (2019)

Using the setup described in the previous section, Machado and Santos Silva (2019) proposes an extension to the model proposed by He (1997) that would allow for the estimation of quantile regression models with panel data, allowing for the inclusion of individual fixed effects. However, as the authors suggest, the methodology can be generalized to allow for the inclusion of multiple fixed effects. This type of analysis may be useful when considering data such as employer-employee linked data (Abowd et al., 2006), or teacher-student linked data (Harris and Sass, 2011). Or, in the most common case, allowing to control for both individual and time fixed effects.

Reconsider the original model, and assume there are sets of unobserved heterogeneity that are assumed to be constant across observations, if they belong to common groups. In panel data, the groups would be the individual fixed effects and the time fixed effects. Without loss of generality, we can assume that the data generating process is as follows:

$$y_i = x_i'\beta + \delta_{g1} + \delta_{g2} + \nu_i$$
$$\nu_i = \varepsilon_i \times (x_i'\gamma + \zeta_{g1} + \zeta_{g2})$$

where we assume x_i vary across groups g_1 and g_2, thus are not collinear, and that $\delta's$ and $\zeta's$ are the location and scale effects associated with groups fixed effects.[4]

If the dimension of groups g_k is low, this model could be estimated using a dummy inclussion approach following Section 2.1, and standard errors obtained as discussed in Section 2.2. However, if the dimensionality of g_k is high, the dummy inclusion approach may not be feasible. Instead, a more feasible approach is to apply the Frisch-Waugh-Lovell (FWL) theorem, and partial out the impact of the group fixed effects on the control variables x_i, the outcome of interest y_i, with a similar approach for the identification of $\sigma(x)$. In the case of unbalanced setups, with multiple groups, the estimation involves iterative processes for which various approaches have been suggested and implemented (see Correia (2016), Gaure (2013), Rios-Avila (2015), among others).

When applying the partialing out approach, some modifications to the approach described in Section 2.1 are needed.

1. For all dependent and independent variables in the model ($w = y, x$), we partial out the group fixed effects, and obtain the centered-residualized variables:

$$w_i = \delta_{g1}^w + \delta_{g2}^w + u_i^w$$
$$w_i^{rc} = E(w_i) + \hat{u}_i^w$$

2. We estimate the location model using the centered-residualized variables:

$$y_i^{rc} = x_i^{rc'}\beta + nu_i$$

3. Since $|\hat{\nu}_i|$ is the dependent variable for the scale model, we apply the partialling out and recentering to this expression ($|\hat{\nu}_i|^{rc}$), and use that to estimate the model:

$$|\hat{\nu}_i|^{rc} = x_i^{rc'}\gamma + e_i$$

4. Finally the standardized residuals ε_i can be obtained as follows

$$\hat{\varepsilon}_i = \frac{\nu_i}{|\hat{\nu}_i| - \hat{e}_i}$$

[4]We could just as well consider multiple sets of fixed effects

where $|\hat{\nu}_i| - \hat{e}_i$ is the prediction for the conditional standard deviation $\sigma(x_i) = x_i'\gamma + \zeta_{g1} + \zeta_{g2}$

The τ_{th} quantile of the error ε can be estimated as usual, and the variance-covariance matrices obtained in the same way as before (see Section 2.2), but using x_i^{rc} instead of x_i when estimating the influence functions for all estimated coefficients.

3. Monte Carlo Simulations

4. Application: Something interesting

5. Conclusions

6. Appendix

6.1. Model Identification

The estimation of quantile regression via moments assumes that the DGP is linear in parameters, with an heteroskedastic error term that is also linear function of parameters:

$$y = x\beta + \nu$$
$$\nu = \varepsilon \times x\gamma$$

where ε is an unobserved i.i.d. random variable that is independent of x, and such that $x\gamma$ is larger than zero for any x.

In this case, the τ_{th} conditional quantile model can be written as:

$$q(y|\tau, X) = x(\beta + q(\varepsilon|\tau) \times \gamma)$$

This model is identified under the following conditions:

$$\begin{aligned} E[(y_i - x_i'\beta)x_i] = E[h_{1,i}] = 0 \\ E[(|y_i - x_i'\beta| - x_i'\gamma)x_i] = E[h_{2,i}] = 0 \\ E\left[\mathbb{1}\left(q(\varepsilon|\tau)x_i'\gamma + x_i'\beta \geq y_i\right) - \tau\right] = E[h_{3,i}] = 0 \end{aligned}$$

For simplicity, for the rest of the appendix, I will use q_τ^ε to represent $q(\varepsilon|\tau)$.

6.2. Estimation of the variance-covariance matrix

In this model, to estimate the variance-covariance matrix the set of coefficients $\theta' = [\beta' \ \gamma' \ q_\tau^\varepsilon]$, we need to obtain the influence functions of all coefficients, which are defined as:

$$\lambda_i = \bar{G}(\theta)^{-1} \begin{bmatrix} h_{1,i} \\ h_{2,i} \\ h_{3,i} \end{bmatrix}$$

where the Jacobian matrix $\bar{G}(\theta)$ is defined as:

$$\bar{G}(\theta) = \begin{bmatrix} \bar{G}_{11} & G_{12} & G_{13} \\ \bar{G}_{21} & \bar{G}_{22} & G_{23} \\ \bar{G}_{31} & \bar{G}_{32} & \bar{G}_{13} \end{bmatrix}$$

with

$$\bar{G}_{j,k} = -\frac{1}{N}\sum_{i=1}^{N}\frac{\partial h_{j,i}}{\partial \theta_k'} \;\; \forall j,k \in 1,2,3$$

6.2.1. First Moment Condition: Location Model

$$h_{1,i} = x_i(y_i - x_i'\beta)$$

$$\begin{aligned} \bar{G}_{1,1} &= -\frac{1}{N}\sum_{i=1}^{N}\frac{\partial h_{1,i}}{\partial \beta'} \\ &= -\frac{1}{N}\sum_{i=1}^{N}(-x_i x_i') \\ &= N^{-1}X'X \end{aligned}$$

$$\bar{G}_{1,2} = \bar{G}_{1,3} = 0$$

6.2.2. Second Moment Condition: Scale model

$$h_{2,i} = x_i(|y_i - x_i'\beta| - x_i'\gamma)$$

$$\begin{aligned} \bar{G}_{2,1} &= -\frac{1}{N}\sum\frac{\partial h_{2,i}}{\partial \beta'} \\ &= \frac{1}{N}\sum x_i x_i' \frac{y_i - x_i'\beta}{|y_i - x_i'\beta|} \\ \frac{y_i - x_i'\beta}{|y_i - x_i'\beta|} &= sign(y_i - x_i'\beta) \end{aligned}$$

Under the assumption $\varepsilon_i \times x\gamma$, or in this case $y_i - x_i'\beta$, is uncorrelated with x, we can simplify the expression as:

$$\begin{aligned} \bar{G}_{2,1} &= N^{-1}\left(N^{-1}\sum sign(y_i - x_i'\beta)\right)\sum x_i x_i' \\ &= N^{-1}E[sign(y_i - x_i'\beta)]X'X \end{aligned}$$

$$
\begin{aligned}
\bar{G}_{2,2} &= -\frac{1}{N}\sum \frac{\partial h_{2,i}}{\partial \gamma'} \\
&= \frac{1}{N}\sum x_i x_i' \qquad = N^{-1}X'X
\end{aligned}
$$

$$
\bar{G}_{2,3} = 0
$$

6.2.3. Third Moment Condition: Quantile of Standardized Residual

$$
\begin{aligned}
h_{3,i} &= \mathbb{1}\left(q_\tau^\varepsilon x_i'\gamma + x_i'\beta - y_i \geq 0\right) - \tau \text{ or} \\
h_{3,i} &= \mathbb{1}\left(q_\tau^\varepsilon \geq \frac{y_i - x_i'\beta}{x_i'\gamma}\right) - \tau = \mathbb{1}(q_\tau^\varepsilon \geq \varepsilon) - \tau
\end{aligned}
$$

Because the indicator function $\mathbb{1}()$ is not differentiable, we borrow from the non-parametric literature, and approximate the indicator function with the integral of a kernel density function $I()$. This function $I()$, is monotonic and symetrical function around zero, with a domain over the real numbers, and a range between 0 and 1.

With an arbirarily small bandwidth h, this function will approximate the indicator function:

$$
\lim_{h \to 0} I\left(\frac{z}{h}\right) \approx \mathbb{1}(z \geq 0)
$$

Thus the function $h_{3,i}$ can be approximated as:

$$
h_{3,i} \approx I\left(q_\tau^\varepsilon x_i'\gamma + x_i'\beta - y_i\right) - \tau
$$

Now, we can obtain the Jacobian matrix $\bar{G}_{3,1}$ as:

$$
\begin{aligned}
\bar{G}_{3,1} &= -\frac{1}{N}\sum \frac{\partial h_{3,i}}{\partial \beta'} \\
&= -N^{-1}\sum K_h(q_\tau^\varepsilon x_i'\gamma + x_i'\beta - y_i)x_i'
\end{aligned}
$$

Under the assumption that we have enough observations within each combination of x, and of multiplicative heteroskedasticity, we have:

$$
\begin{aligned}
E(K_h(q_\tau^\varepsilon x_i'\gamma + x_i'\beta - y_i)|X) &= f_{y|X}(q_\tau^\varepsilon x_i'\gamma + x_i'\beta) \\
&= \frac{1}{x_i'\gamma} f_\varepsilon(q_\tau^\varepsilon)
\end{aligned}
$$

where $f_{y|X}$ is the conditional probability density function of y given X, and f_ε is the unconditional distribution of the standardized error. With this, we can rewrite the Jacobian matrix as:

$$\bar{G}_{3,1} = -N^{-1} f_\varepsilon(q_\tau^\varepsilon) \sum \frac{x_i'}{x_i'\gamma}$$

Asymptotically, however, the expression $\sum \frac{a_i}{b_i}$ can be approximated using taylor expansions by $N\frac{\bar{a}}{\bar{b}}$. [5] Thus, we can rewrite the last term as:

$$\bar{G}_{3,1} = -f_\varepsilon(q_\tau^\varepsilon) \frac{\bar{x}_i'}{\bar{x}_i'\gamma}$$

The Jacobian for the second matrix $\bar{G}_{3,2}$ can be derived similarly:

$$\begin{aligned}
\bar{G}_{3,2} &= -\frac{1}{N} \sum \frac{\partial h_{3,i}}{\partial \gamma'} \\
&= -N^{-1} \sum K_h(q_\tau^\varepsilon x_i'\gamma + x_i'\beta - y_i) q_\tau^\varepsilon x_i' \\
&= -N^{-1} \sum f_{y|x}(q_\tau^\varepsilon x_i'\gamma + x_i'\beta) q_\tau^\varepsilon x_i' \\
&= -N^{-1} f_\varepsilon(q_\tau^\varepsilon) q_\tau^\varepsilon \sum \frac{x_i'}{x_i'\gamma} \\
&= -f_\varepsilon(q_\tau^\varepsilon) q_\tau^\varepsilon \frac{\bar{x}_i'}{\bar{x}_i'\gamma}
\end{aligned}$$

and the Jacobian for the third matrix $\bar{G}_{3,3}$ is:

$$\begin{aligned}
\bar{G}_{3,3} &= -\frac{1}{N} \sum \frac{\partial h_{3,i}}{\partial q_\tau^\varepsilon} \\
&= -\frac{1}{N} \sum K_h(q_\tau^\varepsilon x_i'\gamma + x_i'\beta - y_i) x_i'\gamma \\
&= -\frac{1}{N} \sum f_{y|X}(q_\tau^\varepsilon x_i'\gamma + x_i'\beta) x_i'\gamma \\
&= -\frac{1}{N} \sum f_\varepsilon(q_\tau^\varepsilon) \frac{x_i'\gamma}{x_i'\gamma} \\
&= -f_\varepsilon(q_\tau^\varepsilon)
\end{aligned}$$

6.3. Influence functions

6.3.1. Location coefficients

$$\lambda_i(\beta) = \bar{G}_{1,1}^{-1} h_{1,i} = N(X'X)^{-1}(x_i(y_i - x_i'\beta)) = N(X'X)^{-1}(x_i \nu_i)$$

Which can also be written as a function of the standardized residuals:

$$\lambda_i(\beta) = \bar{G}_{1,1}^{-1} h_{1,i} = N(X'X)^{-1}(x_i(y_i - x_i'\beta)) = N(X'X)^{-1}(x_i(x_i'\gamma \times \varepsilon))$$

[5] This approximation will be useful when we consider the estimation of the influence functions.

*6.3.2. Scale Coefficients coefficients***

$$\begin{aligned}
\lambda_i(\gamma) &= \bar{G}_{2,2}^{-1}\Big(h_{2,i} - \bar{G}_{2,1}\lambda_i(\beta)\Big) \\
&= N(X'X)^{-1}\Big(x_i(|\nu_i| - x_i'\gamma) - N^{-1}E[sign(\nu_i)]X'X[N(X'X)^{-1}(x_i\nu_i)]\Big) \\
&= N(X'X)^{-1}\Big(x_i(|\nu_i| - x_i'\gamma) - E[sign(\nu_i)](x_i\nu_i)\Big) \\
&= N(X'X)^{-1}x_i\Big(|\nu_i| - E[sign(\nu_i)]\nu_i - x_i'\gamma\Big)
\end{aligned}$$

However,

$$\begin{aligned}
|\nu_i| &= \nu_i \times \mathbb{1}(\nu_i \geq 0) - \nu_i \times \mathbb{1}(\nu_i < 0) \\
|\nu_i| &= \nu_i \times \mathbb{1}(\nu_i \geq 0) - \nu_i \times [1 - \mathbb{1}(\nu_i \geq 0)] \\
|\nu_i| &= 2\nu_i \times \mathbb{1}(\nu_i \geq 0) - \nu_i
\end{aligned}$$

And

$$\begin{aligned}
E[sign(\nu_i)] &= E[\mathbb{1}(\nu_i \geq 0)] - E[\mathbb{1}(\nu_i < 0)] \\
E[sign(\nu_i)] &= E[\mathbb{1}(\nu_i \geq 0)] - E[(1 - \mathbb{1}(\nu_i \geq 0))] \\
E[sign(\nu_i)] &= 2E[\mathbb{1}(\nu_i \geq 0)] - 1
\end{aligned}$$

Thus,

$$\begin{aligned}
\lambda_i(\gamma) &= N(X'X)^{-1}x_i\Big(2\nu_i \times \mathbb{1}(\nu_i \geq 0) - \nu_i - (2E[\mathbb{1}(\nu_i \geq 0)] - 1)\nu_i - x_i'\gamma\Big) \\
&= N(X'X)^{-1}x_i\Big(2\nu_i \times \mathbb{1}(\nu_i \geq 0) - 2E[\mathbb{1}(\nu_i \geq 0)]\nu_i - x_i'\gamma\Big) \\
&= N(X'X)^{-1}x_i\Big(2\nu_i \times [\mathbb{1}(\nu_i \geq 0) - E[\mathbb{1}(\nu_i \geq 0)]] - x_i'\gamma\Big) \\
&= N(X'X)^{-1}x_i\Big(\tilde{\nu}_i - x_i'\gamma\Big)
\end{aligned}$$

This last expression is the equivalent simplification used in Machado and Santos Silva (2019) and Im (2000). If the scale function is strictly possitive, it also follows that $\mathbb{1}(\nu_i \geq 0) = \mathbb{1}(\varepsilon_i \geq 0)$. Thus, it can be simplified as:

$$\lambda_i(\gamma) = N(X'X)^{-1}x_i(x_i'\gamma) \times (\tilde{\varepsilon}_i - 1)$$

6.3.3. Quantile of standardized residual

$$\begin{aligned}
\lambda_i(q_\tau^\varepsilon) &= \bar{G}_{3,3}^{-1}\Big(h_{3,i} - \bar{G}_{3,1}\lambda_i(\beta) - \bar{G}_{3,2}\lambda_i(\gamma)\Big) \\
&= -\frac{1}{f_\varepsilon(q_\tau^\varepsilon)} \times \Bigg(\Big(\mathbb{1}(q_\tau^\varepsilon \geq \varepsilon) - \tau\Big) \\
&+ f_\varepsilon(q_\tau^\varepsilon)\frac{\bar{x}_i'}{\bar{x}_i'\gamma} N(X'X)^{-1}x_i(x_i'\gamma \times \varepsilon) \\
&+ f_\varepsilon(q_\tau^\varepsilon)q_\tau^\varepsilon\frac{\bar{x}_i'}{\bar{x}_i'\gamma} N(X'X)^{-1}x_i(\tilde{\nu}_i - x_i'\gamma)\Bigg) \\
&= \frac{\tau - \mathbb{1}(q_\tau^\varepsilon \geq \varepsilon)}{f_\varepsilon(q_\tau^\varepsilon)} - \frac{x_i'\gamma \times \varepsilon_i}{\bar{x}_i'\gamma} - q_\tau^\varepsilon\frac{\tilde{\nu}_i - x_i'\gamma}{\bar{x}_i'\gamma}
\end{aligned}$$

References

Abowd, J.M., Kramarz, F., Roux, S., 2006. Wages, mobility and firm performance: Advantages and insights from using matched worker–firm data*. The Economic Journal 116, F245–F285. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-0297.2006.01099.x, doi:https://doi.org/10.1111/j.1468-0297.2006.01099.x, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-0297.2006.01099.x.

Cameron, A.C., Trivedi, P.K., 2005. Microeconometrics: methods and applications. Cambridge University Press, Cambridge ; New York.

Canay, I.A., 2011. A simple approach to quantile regression for panel data. The Econometrics Journal 14, 368–386. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1368-423X.2011.00349.x, doi:https://doi.org/10.1111/j.1368-423X.2011.00349.x, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1368-423X.2011.00349.x.

Correia, S., 2016. A Feasible Estimator for Linear Models with Multi-Way Fixed Effects URL: http://scorreia.com/research/hdfe.pdf.

Frisch, R., Waugh, F.V., 1933. Partial time regressions as compared with individual trends. Econometrica 1, 387–401. URL: http://www.jstor.org/stable/1907330.

Frumento, P., Bottai, M., 2016. Parametric modeling of quantile regression coefficient functions. Biometrics 72, 74–84. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/biom.12410, doi:https://doi.org/10.1111/biom.12410, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/biom.12410.

Gaure, S., 2013. OLS with multiple high dimensional category variables. Computational Statistics & Data Analysis 66, 8–18. URL: https://www.sciencedirect.com/science/article/pii/S0167947313001266, doi:https://doi.org/10.1016/j.csda.2013.03.024.

Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J., Stahel, W.A., 2005. Robust Statistics: The Approach Based on Influence Functions. Wiley Series in Probability and Statistics. 1 ed., Wiley. URL: https://onlinelibrary.wiley.com/doi/book/10.1002/9781118186435, doi:10.1002/9781118186435.

Harris, D.N., Sass, T.R., 2011. Teacher training, teacher quality and student achievement. Journal of Public Economics 95, 798–812. URL: https://www.sciencedirect.com/science/article/pii/S0047272710001696, doi:https://doi.org/10.1016/j.jpubeco.2010.11.009.

He, X., 1997. Quantile curves without crossing. The American Statistician 51, 186–192. URL: https://www.tandfonline.com/doi/abs/10.1080/00031305.1997.10473959, doi:10.1080/00031305.1997.10473959, arXiv:https://www.tandfonline.com/doi/pdf/10.1080/00031305.1997.10473959.

Im, K.S., 2000. Robustifying glejser test of heteroskedasticity. Journal of Econometrics 97, 179–188. URL: https://www.sciencedirect.com/science/article/pii/S0304407699000615, doi:https://doi.org/10.1016/S0304-4076(99)00061-5.

Jann, B., 2020. Influence functions continued. A framework for estimating standard errors in reweighting, matching, and regression adjustment. University of Bern Social Sciences Working Papers 35. University of Bern, Department of Social Sciences. URL: https://ideas.repec.org/p/bss/wpaper/35.html.

Kaplan, D.M., Sun, Y., 2017. Smoothed estimating equations for instrumental variables quantile regression. Econometric Theory 33, 105–157. URL: https://www.cambridge.org/core/journals/

econometric-theory/article/abs/smoothed-estimating-equations-for-instrumental-variables-quantile-regression/59EE6581C077098DDBB87F3E98F9BF90#, doi:10.1017/S0266466615000407. publisher: Cambridge University Press.

Koenker, R., 2004. Quantile regression for longitudinal data. Journal of Multivariate Analysis 91, 74–89. URL: https://www.sciencedirect.com/science/article/pii/S0047259X04001113, doi:https://doi.org/10.1016/j.jmva.2004.05.006. special Issue on Semiparametric and Nonparametric Mixed Models.

Koenker, R., Bassett, G., 1978. Regression quantiles. Econometrica 46, 33–50. URL: http://www.jstor.org/stable/1913643.

Lancaster, T., 2000. The incidental parameter problem since 1948. Journal of Econometrics 95, 391–413. URL: https://www.sciencedirect.com/science/article/pii/S0304407699000445, doi:https://doi.org/10.1016/S0304-4076(99)00044-5.

Lovell, M.C., 1963. Seasonal adjustment of economic time series and multiple regression analysis. Journal of the American Statistical Association 58, 993–1010. URL: https://www.tandfonline.com/doi/abs/10.1080/01621459.1963.10480682, doi:10.1080/01621459.1963.10480682, arXiv:https://www.tandfonline.com/doi/pdf/10.1080/01621459.1963.10480682.

Machado, J.A., Santos Silva, J., 2019. Quantiles via moments. Journal of Econometrics 213, 145–173. URL: https://www.sciencedirect.com/science/article/pii/S0304407619300648, doi:10.1016/j.jeconom.2019.04.009. annals: In Honor of Roger Koenker.

Melly, B., 2005. Decomposition of differences in distribution using quantile regression. Labour Economics 12, 577–590. URL: https://www.sciencedirect.com/science/article/pii/S0927537105000382, doi:https://doi.org/10.1016/j.labeco.2005.05.006.

Newey, W.K., McFadden, D., 1994. Chapter 36 Large sample estimation and hypothesis testing, in: Handbook of Econometrics. Elsevier. volume 4, pp. 2111–2245. URL: https://linkinghub.elsevier.com/retrieve/pii/S1573441205800054, doi:10.1016/S1573-4412(05)80005-4.

Neyman, J., Scott, E.L., 1948. Consistent estimates based on partially consistent observations. Econometrica 16, 1–32. URL: http://www.jstor.org/stable/1914288.

Rios-Avila, F., 2015. Feasible Fitting of Linear Models with N Fixed Effects. The Stata Journal 15, 881–898. URL: https://doi.org/10.1177/1536867X1501500318, doi:10.1177/1536867X1501500318. publisher: SAGE Publications.

Zhao, Q., 2000. Restricted Regression Quantiles. Journal of Multivariate Analysis 72, 78–99. URL: https://www.sciencedirect.com/science/article/pii/S0047259X99918493, doi:https://doi.org/10.1006/jmva.1999.1849.