Quantile Regressions via Method of Moments with multiple fixed effects

Fernando Rios-Avila Leonardo Siles
Gustavo Canavire-Bacarreza

2023-10-23

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**

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.

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 (**zhao2020?**), 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); (**zhao2020?**)), the effect of a variable X on the $q_t h$ 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 (**zhao2020?**), 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.

Methodology

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 **??**) and the location-scale specification (Equation **??**) 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]

$$y_i = x_i'\beta + \varepsilon_i \times x_i'\gamma(\tau) \tag{4}$$

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.

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

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.

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

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 **??** 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:

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

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:

$$\begin{aligned} |u_i| &= x_i'\gamma + v_i \\ E[(|u| - X\gamma)X] &= 0 \end{aligned} \tag{8}$$

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.

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:

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

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 in Equation **??**, a just identified model. And the functions $G_{k,l}$ and $S_{k.l}$ equal to:

$$
\begin{aligned}
G_{k,l}=N^{-1}\sum_{i=1}^{N}\frac{\partial h_{k,i}}{\partial \theta_l}\Big|_{\hat{\theta}} \ \forall k,l \in [1,2,3] \\
S_{k,l}=N^{-1}\sum_{i=1}^{N} h_{k,i}h_{l,i}'\Big|_{\hat{\theta}} \ \forall k,l \in [1,2,3]
\end{aligned}
$$

where $\theta_1=\beta$, $\theta_2=\gamma$ and $\theta_3=Q(\varepsilon,\tau)$.

Then

$$
\sqrt{N}(\hat{\theta}-\theta) \xrightarrow{d} N(0, G^{-1}SG^{-1})
$$

[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).