école d'économie d'aix-marseille
aix-marseille school of economics

Working Papers / Documents de travail

A Particular Form of Non-Constant Effect in Two-Stage Quantile Regression

Tae-Hwan Kim
Christophe Muller

WP 2015 - Nr 22

A Particular Form of Non-Constant Effect in Two-Stage Quantile Regression

Tae-Hwan Kim [a] and Christophe Muller [b]

[a] Department of Economics, Yonsei University, Seoul 120-749, Korea
tae-hwan.kim@yonsei.ac.kr

[b] Aix-Marseille University (Aix-Marseille School of Economics), CNRS & EHESS
14, Boulevard Jules Ferry, Aix-en-Provence, France
christophe.muller@univ-amu.fr

May 2015

Abstract: We study the fitted-value approach to quantile regression in the presence of endogeneity under a weakened form of the IV condition. In this context, we exhibit the possibility of a particular form of non-constant effect models with the fitted-value approach, a situation often believed to be ruled out. However, only the constant effect coefficients of the model can be consistently estimated. Finally, we discuss practical examples where this approach can be useful to avoid misspecification of quantile models.

JEL Codes: C13, C21, C31.

Key Words: Two-Stage Estimation, Quantile Regression, Fitted-Value Approach.

This work was carried out thanks to the support of the A*MIDEX project (No. ANR-11-IDEX-0001-02) funded by the "Investissements d'Avenir" French Government program, managed by the French National Research Agency (ANR). We are grateful to P. Lavergne for a useful discussion and to participants in a seminar at the Toulouse School of Economics for their comments. Remaining errors are ours.

1 Introduction

This paper considers the estimation of a structural linear equation using quantile regression in the presence of endogeneity problems. Since the seminal work by Koenker and Bassett (1978), the literature on quantile regressions has expanded a great deal. There are two trends in the literature about quantile regression with endogeneity. The first one, denoted the 'structural approach,' corresponds to models directly specified in terms of the conditional quantile of the structural equation of interest. In that case, semiparametric restrictions needed for identification are directly imposed on the structural errors terms or on the structural model if there are no separable errors. The second trend, denoted the 'fitted-value approach,' is based on the conditional quantile of the reduced-form equation. Accordingly, the restrictions are imposed on the reduced-form errors when they can be separated. In this latter approach, the analysts substitute the endogenous regressors with the fitted-values obtained from ancillary equations based on some other exogenous variables. In this paper, we exhibit a particular case of non-constant effect (i.e., quantile-dependent coefficients) for linear quantile regression with the fitted-value approach, a situation sometimes believed to be ruled out.

The literature on the structural approach for linear quantile regressions is abundant,[1] but it meets computation difficulties for the correction of endogeneity issues.[2] On the other hand, the fitted-value approach corresponds to a simple two-step quantile regression procedure that is analogous to the 2SLS method. Such a procedure

[1] See for example: Abadie et al. (2002), Chen et al. (2003), Chesher (2003), Hong and Tamer (2003), Chernozhukov and Hansen (2005, 2006, 2008), Imbens and Newey (2006), Ma and Koenker (2006), Chernozhukov, Imbens and Newey (2007), Lee (2007).

[2] As documented in Kim and Muller (2013).

has been employed by empirical researchers.[3] Partial theoretical results had been obtained by Amemiya (1982) and Powell (1983), who analyse the two-stage least-absolute-deviations estimators in simple settings, and by redefining the dependent variable. Chen and Portnoy (1996) and Kim and Muller (2004, 2012) investigate such two-stage quantile regressions with diverse first-step estimators (LS, LAD, and trimmed least squares estimators). However, all these authors deal with constant effect specifications. In contrast, our focus in this paper is the occurrence of non-constant effect with the fitted-value approach.

The paper is organized as follows. Section 2 presents the model and the assumptions. In Section 3, we exhibit and analyse a case of non-constant effect for the fitted-value approach. Finally, we conclude the paper in Section 4.

2 The Model

Assume that our interest lies in the parameter vector $\alpha_{0\theta} = (\beta'_{0\theta}, \gamma'_{0\theta})'$ in the following linear equation for T observations and an arbitrary quantile index $\theta \in (0, 1)$:

$$\begin{aligned} y_t &= x'_{1t}\beta_{0\theta} + Y'_t\gamma_{0\theta} + u_{t\theta} \qquad (1) \\ &= z'_t\alpha_{0\theta} + u_{t\theta}, \end{aligned}$$

where $[y_t, Y'_t]$ is a $(G+1)$-row vector of endogenous variables, x'_{1t} is a K_1-row vector of exogenous variables, $z_t = [x'_{1t}, Y'_t]'$ and $u_{t\theta}$ is an error term.

Since we wish to study non-constant effect models, we emphasize that the coefficient vector and the errors may vary with the considered quantile index θ. We denote by x'_{2t} the row vector of the K_2 exogenous variables excluded from (1), and we assume $K_2 \geq G$. We further assume that Y_t can be linearly predicted from the exogenous

[3] Arias et al. (2001), Garcia et al. (2001), Chortareas et al. (2012) and Chepatrakul et al. (2012).

variables through the following equation, which we assume to be correctly specified.

$$Y_t' = x_t'\Pi_0 + V_t', \tag{2}$$

where $x_t' = [x_{1t}', x_{2t}']$ is an unbounded K-rows vector with $K = K_1+K_2$, Π_0 is a $K \times G$ matrix of unknown parameters, and V_t' is a G-row vector of unknown error terms. To avoid absurdities, we assume that the columns in x_t are linearly independent. Using (1) and (2), y_t can also be expressed as:

$$y_t = x_t'\pi_{0\theta} + v_{t\theta}, \tag{3}$$

where

$$\begin{aligned} \pi_{0\theta} &= H(\Pi_0)\alpha_{0\theta} \text{ with } H(\Pi_0) = \left[\begin{pmatrix} I_{K_1} \\ 0 \end{pmatrix}, \Pi_0\right] \\ \text{and } v_{t\theta} &= u_{t\theta} + V_t'\gamma_{0\theta}. \end{aligned} \tag{4}$$

As noted in Blundell and Powell (2006): "The reduced form for y_t may be of interest if the values of IVs are control variables for the policy maker." However, in the fitted-value approach, the reduced form can also be viewed as an intermediary stage for calculations. Amemiya (1974) pointed out that, while substitution of fitted values in nonlinear structural functions generally yields inconsistent estimates of the structural parameters, consistent estimation methods that substitute fitted values into the structural function rely on linearity of the regression, where the model is based on the reduced form error $v_t = u_t + V_t'\gamma_0$ with similar stochastic properties to the structural error u_t. We first consider the following quantile restriction on the reduced-form errors, typically used in the fitted-value approach, for a given quantile θ. Later on, we will also discuss the situation when this assumption is imposed over the whole quantile process altogether.

Assumption 1: *$E(\psi_{\theta_0}(v_{t\theta_0})|x_t) = 0$, for an arbitrary given $\theta_0 \in (0,1)$, where $\psi_\theta(z) = \theta - 1_{[z \leq 0]}$, and $1_{[.]}$ is the indicator function.*

Assumption 1 imposes that zero is the given θ_0^{th}-quantile of the conditional distribution of $v_{t\theta_0}$. This assumption allows the identification of the coefficients of the quantile regression model centered in quantile θ_0. It is associated with the fitted-value approach in which, first, the conditional quantile restriction is placed on the reduced-form error $v_{t\theta_0}$, and second, the information set used for the conditional restriction exclusively consists of the exogenous variables x_t. It has been used in Amemiya (1982), Powell (1983), Chen and Portnoy (1996), and Kim and Muller (2004, 2012). In particular, Kim and Muller (2012) provide an analysis of asymptotic properties based on this assumption and with broad conditions on the model. One issue is how the quantile restrictions for different quantiles come together to correspond to a unique quantile process for a single model; that is, all these restrictions are compatible. The analysis we pursue will clarify this point, which is at the core of understand the possible non-constancy of effects. In contrast, in the structural approaches, the conditional quantile restrictions directly characterize the structural error term $u_{t\theta}$.

Although directly imposing the conditional quantile restriction on the structural error term is more popular, the fitted value approach is an alternative way of exploring the conditional distribution of the dependent variable y_t through the reduced-form error. In that case, the structural effect is recovered through the second stage of the estimation. The link of structural and reduced-form parameters are described by (4). Identification is allowed by the following assumption.

Assumption 2: *$H(\Pi_0)$ is of full column rank.*

Assumption 2 is the usual necessary condition for IV regressions, for example, that obtained through usual exclusion restrictions. This is an identification condition for simultaneous linear equations models. Similar conditions can be obtained when other estimators are used. In the fitted-value approach, a first-stage estimator of Π_0 in (2)[4] $\hat{\Pi}$ allows the construction of the fitted value $\hat{Y}_t' = x_t'\hat{\Pi}$, which is substituted for Y_t' in the final estimation stage. In this paper, we are not only interested in an estimator of the structural parameter $\alpha_{0\theta}$, but also in an estimator of the reduced-form parameter $\pi_{0\theta}$, denoted $\hat{\pi}_\theta$.

Let us now define two-stage quantile regressions based on the fitted-value approach in our setting. These two-stage estimators help to deal with endogeneity problems in quantile regression. For any quantile $\theta \in (0,1)$, we define $\rho_\theta(z) = z\psi_\theta(z)$. If the orthogonality condition, $E(z_t\psi_\theta(u_{t\theta})) = 0$, is satisfied, then the quantile regression estimator (QR) is consistent. However, when $u_{t\theta}$ and Y_t (a sub-vector of z_t) are statistically linked under weak endogeneity of Y_t, these orthogonality conditions might not be satisfied. In that case, the QR of $\alpha_{0\theta}$ is generally not consistent, and mere quantile regressions cannot be used.

As usual in this literature, we define, for any quantile θ, the two-stage quantile regression estimator $\hat{\alpha}_\theta$ of $\alpha_{0\theta}$ as a solution to the following program:

$$\min_{\alpha} \; S_T(\alpha, \hat{\Pi}, \theta) = \sum_{t=1}^{T} \rho_\theta(y_t - x_t' H(\hat{\Pi})\alpha). \tag{5}$$

The estimator is straightforward to calculate with usual quantile regression algorithms. In contrast, structural methods based on conditional quantiles of $u_{t\theta}$ involve computational complications such as grid search for simulation techniques (Cher-

[4] We could allow this estimation stage to depend on the considered quantile θ, for example when a quantile regression is used for the estimation, as implemented in Kim and Muller (2004). However, to stick to usual procedures, the estimation of the prediction of Y_t is here rather assumed to be independent of θ, e.g., by OLS.

nozhukov and Hansen, 2006) or preliminary nonparametric estimation plagued by the 'multidimensionality curse' (Abadie et al., 2002, Chen et al., 2003, Lee, 2007). Currently, these methods can hardly be used in practice with large data sets when more than very few conditioning variables occur in the model, due to excessive computation burden, especially when grid search is included.[5]

However, the fitted-value approach has a one shortcoming: It is often believed that the fitted-value approach used with quantile regressions exclusively implies the constant effect. We show in the next section that, provided a relatively weak independence assumption, some kind of non-constant effect in the structural model with the fitted-value approach can be exhibited.

3 Non-constant Effect in Fitted-Value Approaches

Let us start again with Equations (1) and (2), with possible non-constant effect in structural and reduced form equations, *without imposing a priori Assumption 1.* In order to deal with unique quantile values so as to simplify the discussion, we make the following continuity and monotonicity assumption.

Assumption 3: *For a given quantile θ_0, the cdf of $v_{t\theta_0}$ conditional on x_t, denoted $F_{v_t|x_t}$, the cdf of $v_{t\theta_0}$ conditional on x_{1t}, denoted $F_{v_t|x_{1t}}$, and the marginal cdf of x_{2t}, denoted $F_{x_{2t}}$, are continuous and strictly increasing.*

First, note that, under Assumption 3, an inverse cdf term can always be isolated in the reduced-form equation by denoting: $y_t = x_t'\pi_{0\theta} + v_{t\theta} = F^{-1}_{v_t|x_t}(\theta) + x_t'\pi_{01\theta} + v^*_{t\theta}$ and $v^*_{t\theta} = v_{t\theta} - F^{-1}_{v_t|x_t}(\theta)$. Then, by construction, $v^*_{t\theta}$ is characterized by the conditional

[5] See the computation times in Kim and Muller (2013).

quantile restriction: $E(\psi_\theta(v_{t\theta}^*)|x_t) = \theta - P\left[v_{t\theta}^* \leq 0|x_t\right] = \theta - P\left[v_{t\theta} \leq F_{v_{t|x_t}}^{-1}(\theta)|x_t\right] = \theta - \theta = 0$. As a consequence, one can always obtain the quantile regression restriction at θ for the reduced-form, provided one accepts the introduction of a possible nuisance bias term $F_{v_{t|x_t}}^{-1}(\theta)$ that can affect all the coefficients of the model. We here investigate the conditions in which this bias takes simplified and useful forms.

In order to have models for different quantiles belonging to the same additive quantile regression process, one must have $v_{t\theta'} = v_{t\theta} - F_{v_{t|x_t}}^{-1}(\theta')$, where $F_{v_{t|x_t}}$ is the cdf of $v_{t\theta}$ with the quantile index θ chosen as a given starting value. In these conditions, if we have $E(\psi_\theta(v_{t\theta})|x_t) = 0$, then it is easy to prove that we have also $E(\psi_{\theta'}(v_{t\theta'})|x_t) = 0$ for any other $\theta' \in (0,1)$. Indeed, since for any v, $\psi_{\theta'}(v) = \theta' - I_{[v \leq 0]}$, we have $E(\psi_{\theta'}(v)|x_t) = \theta' - P(v \leq 0|x_t)$. Then, $E(\psi_{\theta'}(v_{t\theta'})|x_t) = \theta' - P(v_{t\theta'} \leq 0|x_t) = \theta' - P(v_{t\theta} - F_{v_{t|x_t}}^{-1}(\theta') \leq 0|x_t) = \theta' - P(v_{t\theta} \leq F_{v_{t|x_t}}^{-1}(\theta')|x_t)$. Under Assumption 3, this is equal to $\theta' - F_{v_{t|x_t}}\left(F_{v_{t|x_t}}^{-1}(\theta')\right) = \theta' - \theta' = 0$.

We can also verify that $F_{v_{t|x_t}}^{-1}(\theta) = 0$, for our chosen value θ_0 of θ. Indeed, $E(\psi_\theta(v_{t\theta})|x_t) = 0$ can be written as $0 = \theta - P(v_{t\theta} \leq 0|x_t) = \theta - F_{v_{t|x_t}}(0)$, which implies $F_{v_{t|x_t}}^{-1}(\theta) = 0$ under Assumption 3.

3.1 Generating a non-constant effect

Let us now return to our maintained model, but *instead of Assumption 1*, we now consider the following weaker restriction.

Assumption 4: *For a given quantile θ_0 indexing the cdfs, as mentioned before:*

$$F_{v_t|x_t}^{-1}(\theta) = F_{v_t|x_{1t}}^{-1}(\theta), \tag{6}$$

for all $\theta \in (0,1)$.

Proposition 1: *Under Assumption 3, Assumption 4 is equivalent to that $v_{t\theta_0}$ is independent on x_{2t}, conditionally on x_{1t}, where θ_0 is a chosen value of θ for defining $F_{v_t|x_t}$.*

Proof: Let us first consider Assumption 4 without the variables x_{1t}, that is, $F^{-1}_{v_t|x_{2t}}(\theta) = F^{-1}_{v_t}(\theta)$, for all θ, which we wish to show to be equivalent to: $v_{t\theta_0}$ is independent of x_{2t}. By definition, under Assumption 3, the latter statement means that $f_{v_t,x_{2t}}(v, x_2) = f_{v_t}(v)\ f_{x_{2t}}(x_2)$, for all v and x_2, and where $f_{v_t,x_{2t}}$ is the joint pdf of $v_{t\theta_0}$ and x_{2t}, f_{v_t} is the marginal pdf of $v_{t\theta_0}$, and $f_{x_{2t}}$ is the marginal pdf of x_{2t}. Since, under Assumption 3, there is no level of x_{2t} with vanishing marginal pdf, we can rewrite the equality as $f_{v_t}(v) = f_{v_t,x_{2t}}(v, x_2)/f_{x_{2t}}(x_2)$, for all v and x_2. Then, by integration with respect to v on both sides of the equality, we obtain $F_{v_t}(v) = F_{v_t|x_{2t}}(v\ |x_{2t} = x_2)$, for all x_2. By inversion, thanks to Assumption 3, this is equivalent to $F^{-1}_{v_t|x_{2t}}(\theta) = F^{-1}_{v_t}(\theta) = v$, for all v, i.e., for all $\theta = F_{v_t}(v)$. Introducing variables x_{1t} and conditioning on them is straightforward by considering the different cdfs for each level of x_{1t}. QED.

Our condition is akin to the ones in the control function literature, only here the control function is observable and equal to x_{1t}. We do not require complete exogeneity, as we will discuss later. It is also related to the notion of 'conditional exogeneity' in White and Chalak (2010). Note that the restriction in Assumption 4 implies that the OLS in the reduced form is inconsistent in the case where the x_{1t} are endogenous. One may also have $E(\psi_\theta(v_{t\theta})|x_{1t}) \neq 0$ under this hypothesis. This means that x_{1t} is endogenous in the sense of the quantile regressions of quantile θ for equation (3). In such situation, equation (3) no longer characterizes a genuine 'reduced form' based only on exogenous regressors, and we instead denote it the 'pseudo reduced-form equation.' In the rest of this section, we show how a non-constant effect can be obtained for conditional quantiles of the pseudo reduced-form and then conveyed to the conditional quantiles of the structural form.

An example of such a setting is a pseudo reduced-form equation for a labour market study in which the dependent variable is the logarithm of the wage rate for a sample of workers, while the two independent variables in this equation are the industrial sector (x_{1t}) and the birth quarter (x_{2t}). The birth quarter is used as an instrument for education that is assumed to be the sole endogenous independent variable in the corresponding assumed structural model (for example, like in Angrist and Krueger, 1991). In such an empirical problem, one typically expects the wage rate to be positively correlated with the capital of the firm, which is itself assumed to be an omitted variable incorporated in the error v_t and should be typically correlated with the industrial sector. Then, x_{1t} and v_t should be correlated, while x_{2t} and v_t should be independent according to the usual reasons justifying the use of quarter of birth as an instrument in wage equations.[6]

Alternatively, Assumption 4 can be imposed only for a limited range of quantiles θ of interest, still for a given θ_0 indexing the cdfs. Indeed, as Koenker states, instruments can be valid for some quantiles only (p. 285, Koenker, 2005): "When instruments are only effective over a limited quantile range, then mean effects are unidentified unless one makes stringent assumptions about the homogeneity of effects across quantiles." In addition, Kim and Muller (2013), who provide a test of exogeneity of regressors for quantile regression, show by applying it to UK consumption data that exogeneity may, in practice, be satisfied only for a limited range of quantiles, even in the constant

[6] Conditioning on the industrial sector might make the hypothesis of independence of the birth quarter and the error more plausible if the sector was a common determinant of the latter two variables, although it is unclear why this should be the case.

effect case. Since $v_{t\theta} = u_{t\theta} + V'_{t\theta}\gamma_{0\theta}$, Assumption 4 for all θ can also be obtained, for example, by assuming that $u_{t\theta}$ and V_t are independent of x_{2t}, conditional on x_{1t}. Thus, this condition is also connected to a natural instrumental variable characterization of x_{2t} for the structural model.

We now show that Assumption 4 implies that there are constant effects in the quantile regressions of the pseudo reduced-form equation of the coefficients of the variables in x_{2t}, but not necessarily of the coefficients of the variables in x_{1t}. Indeed, consider another quantile index θ different from the previously chosen θ_0 and let us impose Assumption 1 to it, that is: the typical quantile restriction for a quantile θ. Then, we examine how this conditional quantile restriction interacts with Assumption 4 assumed with the chosen θ_0.

We have $P\left[v_{t\theta} \leq 0 \middle| \, x_t\right] = \theta$, i.e., $P\left[y_t \leq x'_t\pi_{0\theta} \mid x_t\right] = \theta$, which implies that $P\left[x'_{1t}\pi_{0K_1\theta_0} + x'_{2t}\pi_{0K_2\theta_0} + v_{t\theta_0} \leq x'_{1t}\pi_{0K_1\theta} + x'_{2t}\pi_{0K_2\theta} \mid x_t\right] = \theta$, where $\pi_{0K_1\theta_0}, \pi_{0K_2\theta_0}, \pi_{0K_1\theta}, \pi_{0K_2\theta}$ are the respective components of $\pi_{0\theta_0}$ and $\pi_{0\theta}$ according to the partition of x_t into x_{1t} and x_{2t}. By regrouping, we obtain

$$P\left[v_{t\theta_0} \leq x'_{1t}\left(\pi_{0K_1\theta} - \pi_{0K_1\theta_0}\right) + x'_{2t}\left(\pi_{0K_2\theta} - \pi_{0K_2\theta_0}\right) \mid x_t\right] = \theta,$$

which in turn implies that $F_{v_t|x_t}\left[x'_{1t}\left(\pi_{0K_1\theta} - \pi_{0K_1\theta_0}\right) + x'_{2t}\left(\pi_{0K_2\theta} - \pi_{0K_2\theta_0}\right)\right] = \theta$. Finally, under Assumption 3, we have

$$x'_{1t}\left(\pi_{0K_1\theta} - \pi_{0K_1\theta_0}\right) + x'_{2t}\left(\pi_{0K_2\theta} - \pi_{0K_2\theta_0}\right) = F^{-1}_{v_t|x_t}(\theta) = F^{-1}_{v_t|x_{1t}}(\theta), \tag{7}$$

where the latter equality is obtained using Assumption 4.

Since $F^{-1}_{v_t|x_{1t}}(\theta)$ does not depend on x_{2t}, equation (7) implies $\pi_{0K_2\theta} = \pi_{0K_2\theta_0}$, that is: constant effect for the variables x_{2t}. As a consequence, we drop the dependence of $\pi_{0K_2\theta}$ on θ, which becomes π_{0K_2}. In contrast, there is no restriction on the effect of the variables x_{1t}, which may vary with θ in that case: there might be some non-constant effect. To be consistent with (7), we must also assume:

Assumption 5: $F^{-1}_{v_t|x_{1t}}(\theta)$ must be linear in x_{1t}.

This hypothesis, which we assume from now, is implied in particular by the often used model of the 'linear location-scale hypothesis' in the quantile regression literature on non-constant effect (e.g., Koenker, 2005, pp 104-105). In cases where Assumption 4 would be judged too unrealistic, this assumption can be easily relaxed by incorporating polynomial terms in x_{1t} in the model, as is usual for approximating nonlinear functions. Alternatively, one can specify a pseudo-reduced form (3) as partially linear in x_{2t}, and possibly nonlinear in x_{1t}, with one unknown nonlinear functional form. The above reasoning delivering constant effects for x_{2t} and unrestricted (nonlinear) effect for x_{1t} is valid. In that case, Assumption 5 would not be necessary. However, this would lead us toward nonparametric estimation methods, which is not what we discuss in this paper.

We now show that, under Assumptions 4-5, the bias in the θ^{th} conditional quantile regression estimator of the pseudo reduced-form model can be confined to the first K_1 variables. Indeed, we have by construction that

$$y_t = x'_{1t}\pi_{0K_1\theta} + x'_{2t}\pi_{0K_2} + F^{-1}_{v_t|x_{1t}}(\theta) + v^*_{t\theta},$$

where $\pi_{0K_1\theta}$ and π_{0K_2} denote the parameter vectors respectively associated with x_{1t} and x_{2t} in the reduced-form, and $v^*_{t\theta} = v_{t\theta} - F^{-1}_{v_t|x_{1t}}(\theta)$. Assuming that $F^{-1}_{v_t|x_{1t}}(\theta) = x'_{1t}\pi_{FK_1\theta}$ (the linear location-scale model), where $\pi_{FK_1\theta}$ is a parameter vector, we have

$$y_t = x'_{1t}\pi^*_{0K_1\theta} + x'_{2t}\pi_{0K_2} + v^*_{t\theta},$$

where $\pi^*_{0K_1\theta} = \pi_{0K_1\theta} + \pi_{FK_1\theta}$ denotes the translated parameter vector associated with x_{1t}. In that case, a standard estimation of the quantile regression of this 'translated' reduced-form, based on $E(\psi_\theta(v^*_{t\theta})|x_t) = 0$ that is satisfied by construction of $v^*_{t\theta}$ under

Assumption 4, would yield quantile regression estimators consistent to the true value of the K_2 last coefficients, π_{0K_2}. This can also be seen by rewriting the restriction as $E(\psi_\theta(y_t - x'_{1t}\pi^*_{0K_1\theta} - x'_{2t}\pi_{0K_2})|x_t) = 0$. Indeed, here the coefficients of x'_{2t} appear in their true value forms.

In these conditions, a non-constant effect is possible for the conditional quantiles of the pseudo reduced-form for $\pi^*_{0K_1\theta}$, although not for π_{0K_2} that has been shown not to vary with θ. For example, a non-constant effect quantile regression whose coefficients $\pi^*_{0K_1\theta}$ can be parameterized such as: $\pi^*_{0K_1\theta} = \pi^*_{0K_1} + \theta$. However, if the exhibited non-constant effect in the pseudo reduced-form allows for some generalization of the model as compared to constant effect cases and for a consistent estimation of the other parameters of the model, the non-constant effect cannot be estimated without the bias $\pi_{FK_1\theta}$.

3.2 Involving the whole quantile process

Often, authors dealing with quantile regressions assume Assumption 1 for the whole quantile process. In that case, there is no room for imposing Assumption 4 instead at one or several quantiles θ_0, and a constant effect is obtained for all coefficients, apart from the intercept if it exists. However, it is possible to impose Assumption 4 for all quantiles, instead of imposing Assumption 1. This would preserve the possibility of non-constant effect for some coefficients. Indeed, the consistency of the quantile estimation of some coefficients is still possible under this assumption, as shown in the following asymptotic representation.

Proposition 2: *Suppose that the following assumptions hold: Assumption 4 and*

(i): $T^{1/2}(\hat{\Pi} - \Pi_0) = O_p(1)$.

(ii): The sequence $\{(x'_t, u^*_{t\theta}, v^*_{t\theta})\}$ *is* m*-dependent.*

(iii) Let $f_t^*(.|x)$ be the conditional probability density function (PDF) of $v_{t\theta}^*$. The conditional PDF $f_t^*(\cdot|x)$ is assumed to be Lipschitz continuous for all x, strictly positive and bounded by a constant f_0 (i.e., $f_t^*(\cdot|x) < f_0$, for all x).

(iv) The matrices $Q = \lim_{T\to\infty} E\left[\frac{1}{T}\sum_{t=1}^{T} x_t x_t'\right]$ and $Q_0^* = \lim_{T\to\infty} E\left[\frac{1}{T}\sum_{t=1}^{T} f_t^*(0|x_t) x_t x_t'\right]$ are finite and positive definite.

(v) There exists a positive number $C > 0$ such that $E(\|x_t\|^3) < C < \infty$ for any t.

Then, the asymptotic representation for the 2SQR($\theta, 1$) estimator in Kim and Muller (2012), which is our estimator of interest here, is

$$T^{1/2}(\hat{\alpha}_\theta^* - \alpha_{0\theta}^*) = RT^{-1/2}\sum_{t=1}^{T} x_t \psi_\theta(v_{t\theta}^*) - RQ_0^* T^{1/2}(\hat{\Pi} - \Pi_0)\gamma_0 + o_p(1),$$

where $R = Q_{zz}^{*-1} H(\Pi_0)'$ and $Q_{zz}^* = H(\Pi_0)' Q_0^* H(\Pi_0)$.

Proof: Direct application of the result in Kim and Muller (2012) to the translated errors $v_{t\theta}^*$.

One may also want to consider fully specified quantile regressions only at some quantiles. In that case, one may require Assumption 1 for these quantiles only, keeping Assumption 4 for other quantiles. Jun (2008) studies the variations across quantiles of identification through instrument variables, which suggests an interest in applying different semi-parametric IV restrictions at different quantiles. This would allow numerous distinct models with non-constant effect.

3.3 Transmitting the non-constant effect to the structural form

We now need to assess the consequences of the partial occurrence of non-constant effect in the pseudo reduced-form for the structural model. Assumption 4 allows us to rewrite the pseudo reduced-form model as follows.

$$y_t = x_t' \pi_{0\theta}^* + v_{t\theta}^*, \tag{8}$$

where $\pi_{0\theta}^*$ includes the bias term $\pi_{FK_1\theta}$. Eq. (8) can be considered as the proper reduced-form equation for the θ^{th} quantile regression since it satisfies $E(\psi_\theta(v_{t\theta}^*)|x_t) = 0$. This restriction indicates that both x_{1t} and x_{2t} can be here considered as exogenous in the quantile regression sense for the quantile index θ. Then, a quantile regression estimator $\hat{\pi}_{0\theta}^*$ corresponding to this equation would converge toward $\pi_{0\theta}^*$, under usual conditions. In that case, the corresponding estimator for the structural model is denoted $\hat{\alpha}_{0\theta}^*$ and converges to a value $\alpha_{0\theta}^* = \left(\beta_{0\theta}^{*'}, \gamma_{0\theta}^{*'}\right)'$ that may also incorporate a bias term.

To demonstrate this, we now decompose the link of the reduced-form and structural-form (possibly biased) parameters, by splitting system (4) into two blocks of equations, by partitioning $\Pi_0 = \begin{bmatrix} \Pi_{01} \\ \Pi_{02} \end{bmatrix} = [\Pi_{01}', \Pi_{02}']'$ according to the partition of $\pi_{0\theta}$. We obtain:

$$\pi_{0K_1\theta}^* = \beta_{0\theta}^* + \Pi_{01}\gamma_{0\theta}^*, \tag{9}$$

$$\pi_{0K_2\theta}^* = \Pi_{02}\gamma_{0\theta}^*, \tag{10}$$

where $\pi_{0K_2\theta}^* = \pi_{0K_2}$, which is unbiased and does not depend on θ, as seen above. If the system is exactly-identified (i.e., $K_2 = G$), then $\gamma_{0\theta}^*$ can be directly expressed in terms of Π_{02} and π_{0K_2}, which implies that $\gamma^*_{0\theta}$ does not depend on θ. For the other over-identifying case (i.e., $K_2 > G$), we have more information (or more equations) than is necessary to identify $\gamma^*_{0\theta}$. Similarly to 2SLS or GMM, the additional information does not perturb the identification of $\gamma^*_{0\theta}$. One can even enhance the finite sample performance of the resulting estimators by choosing an appropriate weighting matrix. In either the exactly-identifying case or the over-identifying case, $\gamma^*_{0\theta}$ is a function of Π_{02} and π_{0K_2} only. Hence, the property of unbiased constant effect in $\pi^*_{0K_2\theta}$ is conveyed to $\gamma^*_{0\theta} \equiv \gamma_0$. On the other hand, since $\gamma^*_{0\theta}$ is fully determined in (10), (9) shows that $\beta^*_{0\theta}$ incorporates the non-constant effect from $\pi^*_{0K_1\theta}$ and exactly the same bias as in $\pi^*_{0K_1\theta}$.

What about the relationship of estimators $\hat{\alpha}^*_{0\theta}$ and $\hat{\pi}^*_{0\theta}$ instead of that of the of parameters $\alpha^*_{0\theta}$ and $\pi^*_{0\theta}$? $\hat{\alpha}^*_{0\theta}$ can be computed from (5) and is therefore numerically identical to $\hat{\alpha}_{0\theta}$, except that it is now considered under Assumption 4 instead of Assumption 1. As a consequence, there is now a bias term incorporated in the estimator $\hat{\pi}^*_{0\theta}$ and in its limit, because of the endogeneity of x_{1t} in the pseudo reduced form. This bias also shows up in the associated errors $v^*_{t\theta}$, which are the limits of some associated residuals. However, asymptotically, the same linear relationship as between the limit values will hold for these estimators because the estimators converge respectively to $\alpha^*_{0\theta}$ and $\pi^*_{0\theta}$. The convergence of $\hat{\pi}^*_{0\theta}$ toward $\pi^*_{0\theta}$ is obtained by construction of the semiparametric conditional quantile restriction and by assuming that the usual conditions for convergence of quantile regressions are satisfied. The asymptotic development of $\hat{\alpha}^*_{0\theta}$ provides the characterisation of the convergence of this estimator toward a finite limit. This development can be easily adapted from Kim and Muller (2012), for example, under broad stochastic conditions, by using $v^*_{t\theta}$ instead of v_t in the corresponding empirical process and with minor adjustments. We discussed this in Section 3.2.

When a first-stage estimation is performed based on Eq. (2), the independent variables x_t consists of vectors x_{1t} and x_{2t}. If a non-zero asymptotic bias is present only in the coefficients of x_{1t} in the pseudo reduced-form estimator $\hat{\pi}_\theta$, which is the case on which we focus, then the non-zero asymptotic bias in the second-stage estimator $\hat{\alpha}_\theta = [\hat{\beta}_\theta', \hat{\gamma}_\theta']'$ is exclusively confined to the coefficients of x_{1t}; that is, only $\hat{\beta}_\theta$ is asymptotically biased with exactly the same bias as that of $\hat{\pi}_\theta$. Therefore, the parameter ($\gamma_{0\theta}$) for the endogenous variable Y_t in the structural equation in (1) can be consistently estimated. Then, in this setting, because π_{0K_2} is characterized by constant effect, and because $\gamma_{0\theta}$ is not connected to $\pi_{0K_1\theta}$, we have also constant effect for $\gamma_{0\theta}$ that can be estimated consistently. In contrast, a non-constant effect may occur for $\beta_{0\theta}$. However, $\beta_{0\theta}$ cannot be estimated without an asymptotic bias, which is exactly equal to the asymptotic bias in $\pi_{0K_1\theta}$.

Allowing a weakened IV condition has enabled us to introduce non-constant effect on the vector $\beta_{0\theta}$, even though this parameter is estimated with bias. This alone is a generalization of the *stricto sensu* constant effect structural quantile regression, which may be useful if the researcher's interest is concentrated on vector γ_0 that can be estimated consistently. Indeed, this approach avoids misspecification of the quantile regression in that case if the true DGP involves non-constant effect for $\beta_{0\theta}$ and constant effect for γ_0. We now discuss a few empirical examples of such a situation.

3.4 Empirical examples

An example of our model arises when evaluating a policy parameter that, first, is constant on an identifiable population, and second, affects an outcome variable that is shifted by a quantity proportional to the policy parameter. For example, one might wish to assess the impact of a cash transfer program on family earnings. Let y_t be the total earnings of family t and Y_t be the policy treatment, which is here a dummy variable describing whether or not the family is covered by the program. Assume that the treated families can be identified through observable characteristics, for example, because they live in a specific place or have given socio-demographic characteristics. For example, in France, the family allowances is a transfer program that is exclusively based on the number of children by age class in the family (ADECRI, 2008). The information on family composition is generally observed in household surveys. In that case, the treatment dummy variable Y_t can be observed, while it might be endogenous in an earnings model. Indeed, there might be unobserved characteristics that both determine income and household composition. For example, families with low work motivation might have both more children and less income. This is the case if having children is perceived by some families as a strategy to access social aid and compensate for insufficient incomes. However, assume that the cash amount transferred by such policy is unknown to the researcher. This amount would be here the constant parameter γ_0 to estimate.

Another similar context is to assess the impact on household earnings of taxes that are defined in terms of some identifiable categories of household. Assume that these unknown tax amounts are fixed within each household category. Then, in our setting with the household net income as a dependent variable, the taxes can be described by a vector of coefficients γ_0 that measures the role of variables describing the family categories. Another example is that of racketing extraction tariffs imposed by mafia organisations, which have the same property of being fixed for some given categories of businesses, while their actual amounts are generally unobservable by researchers. A last example is the impact of some unobserved spending for a discrete good or a discrete service on total family expenditure. For example, a given durable good with fixed characteristics might be observable, but not its fixed price. The estimation of the corresponding coefficient γ_0 would allow some inference about the ignored price.

In all of the above examples, we encounter the case of constant effect of the treatment variable of interest. Many variables x_{1t} that determine family earnings or expenditure - like location, activity types, education - can be included in such models. For the current discussion, assume that these variables are considered as exogenous, for example because they have been pre-determined. However, in the considered example their coefficients should generally correspond to a non-constant effect. Indeed, economic theory does not provide any reason for imposing constant effect for these variables. Such non-constantness is likely to perturb most estimation methods of quantile regressions that would wrongly assume constant coefficients for these variables. However, if the interest of the researcher is exclusively in the coefficient γ_0, the fact that β_0 has non-constant effect that cannot be estimated consistently is not an issue with the fitted-value approach.

Finally, some observed variability in the treatment can be incorporated by interacting the treatment with observed characteristics or by considering subpopulations defined in terms of these characteristics. In such situations, one can investigate treatment effects by following our estimation approach, while distinguishing several treatments with constant effect.

4 Conclusion

In this paper, we have shown how some particular causes of non-constant effect can be obtained with two-stage quantile regressions based on the fitted-value approach under endogeneity. However, we have established that only the coefficients of constant effect variables can be consistently estimated, even though well-specified non-constant effect have been introduced for the other variables. We have presented a few practical examples where our approach would be useful.

Our results are based on relatively little demanding instrumental variable conditions, for example that the (pseudo) reduced-form errors (or the structural errors) are independent of *SOME excluded variables,* conditionally on the other independent variables. Such weakening of the usual IV conditions is potentially useful since convincing instruments are typically difficult to find. Then, any reduction in the set of necessary instruments is valuable. This is also important because some instruments might be only valid over a limited range of quantiles.

Our approach is also interesting because a model with constant effect for one coefficient only (or a few coefficients only) is much more general that a general constant effect model, which is what is often estimated in the quantile regression literature. Thus, our approach corresponds to a specification intermediate between the constant effect quantile model and the fully non-constant effect quantile model. In addition, only allowing for biases and non-constant effect on the intercept plus on a single other coefficient of the included exogenous variables might already be sufficient in some cases to achieve consistent estimation of the coefficient of the endogenous variables of interest, under correct specification. As a matter of fact, there are a variety of more or less demanding independence conditions that can be used to allow for partially non-constant structural effects without giving up too many consistent estimates of the coefficients of the independent variables in the structural model, as long as constant effect can be assumed for some endogenous variables.[7]

Alternatively, one could assume instead that the (pseudo) reduced-form errors (or

[7] Allowing only a bias on the intercept is not useful to generate non-constant effect, as the original intercept is already used to center the additive error term in the reduced-form quantile regression.

the structural errors) are independent of some of the included structural regressors, conditional on the excluded independent variables. In that case, one would obtain constant effect for the structural coefficients of the exogenous regressors and possibly non-constant effect for the structural endogenous regressors, although the latter ones would not be estimable without bias. This approach could be useful in particular if the constant effect of interest is that of exogenous regressors.

Finally, all these results have consequences for the interpretation of the fitted-value approach. Namely, under our weakened IV hypothesis, the estimates based on the conditional quantile of the reduced-form can be used to recover at least the estimates of the coefficients of interest for constant effect variables for the conditional quantile of the non-constant effect structural form. Under such conditions, the conditional quantile restriction of the (pseudo) reduced-form can therefore be considered as useful, while partial, characterization of the structural conditional quantile distribution.

References

[1] Abadie, A., J. Angrist and G. Imbens (2002), "Instrumental Variables Estimates of the Effect of Subsidized Training on the Quantiles of Trainee Earnings," *Econometrica,* 70, 91-117.

[2] ADECRI (2008), "*The French Social System*," ADECRI, Paris.

[3] Amemiya, T. (1974), "The Nonlinear Two Stage Least Square Estimator," *Journal of Econometrics*, 2, 105-110.

[4] Amemiya, T. (1982), "Two Stage Least Absolute Deviations Estimators," *Econometrica*, 50, 689-711.

[5] Angrist, J. D., V. Chernozhukov and I. Fernandez-Val, "Quantile Regression under Misspecification, with an Application to the US Wage Structure," *Econometrica*, Vol. 77, N. 2, 539-569.

[6] Angrist, J. D. and A. B. Krueger (1991), "Does compulsory school attendance affect schooling and earnings?" *The Quarterly Journal of Economics*, November 1991, 106 (4), 979-1014.

[7] Arias, O., K.F. Hallock and W. Sosa-Escudero (2001), "Individual Heterogeneity in the Returns to Schooling: Instrumental Variables Quantile Regression Using Twins Data," *Empirical Economics*, 26, 7-40.

[8] Amemiya, T. (1974), "The Nonlinear Two-Stage Least-Square Estimator," J*ournal of Econometrics*, 2, 105-110.

[9] Blundell, R. and J. Powell (2006), "Endogeneity in Nonparametric and Semiparametric Regression Models" in M. Dewatripont, L.P. Hansen and S.J. Turnovsky,

eds., "*Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress, Vol. II*," Cambridge University Press, Cambridge.

[10] Chen, L.-A. and S. Portnoy (1996), "Two-Stage Regression Quantiles and Two-Stage Trimmed Least Squares Estimators for Structural Equation Models," *Commun. Statist.-Theory Meth.*, 25, 1005-1032.

[11] Chen, X., O. Linton and I. Van Keilegem (2003), "Estimation of Semi-Parametric Models When the Criterion Function is not Smooth," *Econometrica*, 71, 1591-1608.

[12] Chernozhukov, V. and C. Hansen (2005), "An IV Model of Quantile Treatment Effects," *Econometrica*, 73, 245-261.

[13] Chernozhukov, V. and C. Hansen (2006), "Instrumental Quantile Regression Inference for Structural and Treatment Effect Models," *Journal of Econometrics*, 132, 491-525.

[14] Chernozhukov, V. and C. Hansen (2008), "Instrumental Variable Quantile Regression: A Robust Inference Approach," *Journal of Econometrics*, 142, 379-398.

[15] Chernozhukov, V., G.W. Imbens and W.K. Newey (2007), "Instrumental Variable Estimation of Nonseparable Models," *Journal of Econometrics*, 139, 4-14.

[16] Chevapatrakul, T., T.-H. Kim and P. Mizen (2009), "The Taylor principle and monetary policy approaching a zero bound on nominal rates: quantile regression results for the United States and Japan," *Journal of Money, Credit and Banking*, 41, 1705-1723.

[17] Chesher, A. (2003), "Identification in Nonseparable Models," *Econometrica*, 71, 1405-1441.

[18] Chortareas, G., G. Magonis and T. Panagiotidis (2012), "The asymmetry of the New Keynesian Phillips Curve in the euro area," *Economics Letters*, 114, 161-163.

[19] Garcia, J., P.J. Hernandez and A. Lopez (2001), "How Wide is the Gap? An Investigation of Gender Wage Differences Using Quantile Regression," *Empirical Economics*, 26, 149-167.

[20] Hong, H. and E. Tamer (2003), "Inference in Censored Models with Endogenous Regressors," *Econometrica*, 71, 905-932.

[21] Imbens, G.W. and W.K. Newey (2006), "Identification and Estimation of Triangular Simultaneous Equations Models without Additivity," Discussion paper, MIT Department of Economics.

[22] Jun, S.J. (2008), "Weak identification robust tests in an instrumental quantile model," *Journal of Econometrics*, 144, 118-138.

[23] Kim, T. and C. Muller (2004), "Two-Stage Quantile Regressions when the First Stage is Based on Quantile Regressions," *The Econometrics Journal*, 18-46.

[24] Kim, T. and C. Muller (2012), "Bias Transmission and Variance Reduction in Quantile Regression," Working Paper Aix-Marseille School of Economics, N. 2012-21.

[25] Kim, T. and C. Muller (2013), "A Test of Exogeneity in Quantiles," Working Paper Aix-Marseille School of Economics, N. 2013-42.

[26] Koenker, R. (2005), "*Quantile Regression*," Cambridge University Press, New York.

[27] Koenker, R. and G. Bassett (1978), "Regression Quantiles," *Econometrica*, 46, 33-50.

[28] Lee, S. (2007), "Endogeneity in Quantile Regression Models: A Control Function Approach," *Journal of Econometrics*, 141, 1131-1158.

[29] Ma, L. and R. Koenker (2006), "Quantile Regression Methods for Recursive Structural Equation Models," *Journal of Econometrics,* 134, 471-506.

[30] Powell, J. (1983), "The Asymptotic Normality of Two-Stage Least Absolute Deviations Estimators," *Econometrica*, 51, 1569-1575.

[31] White, H. and K. Chalak (2010), "Testing a Conditional Form of Exogeneity," *Economics Letters*, 109, 88-90.