Available online at www.sciencedirect.com
ScienceDirect
Economics Letters 95 (2007) 272–277
economics letters
www.elsevier.com/locate/econbase

Inference approaches for instrumental variable quantile regression

Victor Chernozhukov [a], Christian Hansen [b,*], Michael Jansson [c]

[a] *Massachusetts Institute of Technology, United States*
[b] *University of Chicago, Graduate School of Business, 5807 S. Woodlawn Ave., Chicago, IL 60637, United States*
[c] *University of California-Berkeley, United States*

Received 7 February 2006; received in revised form 13 September 2006; accepted 25 October 2006

Abstract

We consider asymptotic and finite sample confidence bounds in instrumental variables quantile regressions of wages on schooling with relatively weak instruments. We find practically important differences between the asymptotic and finite sample interval estimates.

Keywords: Quantile regression; Instrumental variables; Schooling

JEL classification: C10; J31

1. Introduction

In this note, we outline three approaches to obtaining inference statements for quantile regression (QR) models with endogeneity: an asymptotic approximation provided in Chernozhukov and Hansen (2006) that extends the results of Koenker and Bassett (1978) for the QR model with all exogenous variables to a model with endogeneity when instruments are available, an alternate approach to inference presented in Chernozhukov and Hansen (2006) that will be asymptotically valid under weak identification, and an approach to obtaining finite sample confidence regions for parameters of a model defined by quantile restrictions developed in Chernozhukov, Hansen, and Jansson (2005). We compare

* Corresponding author. Tel.: +1 773 834 1702.
E-mail address: chansen1@chicagogsb.edu (C. Hansen).

0165-1765/$ - see front matter
doi:10.1016/j.econlet.2006.10.016

the three approaches in an example due to Card (1995) which uses college proximity as an instrument for years of education in a quantile regression of wages on education. We find substantial differences between the three approaches that appear to be due to the weakness of the instruments. The findings suggest caution should be used when relying on asymptotic approximations in QR models with endogeneity.

2. Inference approaches

We consider a random coefficient model with structural equation given by

$$Y = D'\alpha(U) + X'\beta(U), \tag{2.1}$$

where D and U may be statistically dependent,

A1. $D'\alpha(U)+X'\beta(U)$ is strictly increasing in U for almost every D and X,
A2. U–$U(0, 1)$, and
A3. U is independent of X and Z variables that do not enter the structural equation.

That is, D is a vector of potentially endogenous variables with random coefficients $\alpha(U)$; X is a vector of exogenous variables that enter the structural equation with random coefficients $\beta(U)$, and Z is a vector of exogenous variables that are excluded from the structural equation where it is assumed that dim $(Z) \geq \dim(D)$. This model incorporates the conventional linear QR model of Koenker and Bassett (1978) where $\alpha(U) \equiv 0$ and $Z=X$.

Under conditions A1.–A3., the problem of dependence between U and D is overcome through the presence of instruments, Z, that affect D but are independent of U. From Eq. (2.1) and the monotonicity assumed in A1., the event $\{Y \leq D'\alpha(\tau)+X'\beta(\tau)\}$ is equivalent to the event $\{U \leq \tau\}$. It then follows under A2. and A3. that

$$P[Y \leq D'\alpha(\tau) + X'\beta(\tau)|Z,X] = \tau. \tag{2.2}$$

Eq. (2.2) provides a moment restriction that can be used to estimate the structural parameters $\alpha(\tau)$ and $\beta(\tau)$. For example, one may use this set of moment restrictions to form a GMM estimator; see Pakes and Pollard (1989).

Moment conditions (2.2) also suggest a different procedure for estimating $\alpha(\tau)$ and $\beta(\tau)$. For a given value of the structural parameter, say α, run the ordinary QR of $Y — D'\alpha$ on X and Z to obtain $(\hat{\beta}(\alpha,\tau), \hat{\gamma}(\alpha,\tau))$ where $\hat{\gamma}(\alpha,\tau)$ are the estimated coefficients on the instruments Z. Then note that Eq. (2.2) implies that 0 is the τth conditional quantile of $Y-D'\alpha(\tau)-X'\beta(\tau)$ given Z and X, so one may estimate $\alpha(\tau)$ by finding a value for α that makes the coefficient on the instrumental variable $\hat{\gamma}(\alpha,\tau)$ as close to 0 as possible. That is,

$$\hat{\alpha}(\tau) = \arg \inf_{\alpha \in \mathcal{A}} n[\hat{\gamma}(\alpha,\tau)']\hat{A}(\alpha)[\hat{\gamma}(\alpha,\tau)], \tag{2.3}$$

where $\mathcal{A}$ is the parameter space for α, $\hat{A}(\alpha)=A(\alpha)+o_p(1)$ and $A(\alpha)$ is positive definite uniformly in $\alpha \in \mathcal{A}$. It is convenient to set $A(\alpha)$ equal to the inverse of the asymptotic co-variance matrix of $\sqrt{n}(\hat{\gamma}(\alpha,\tau)-\gamma(\alpha,\tau))$ in which case $W_n(\alpha)$ is the Wald statistic for testing $\gamma(\alpha,\tau)=0$. The parameter estimates are then given by $(\hat{\alpha}(\tau), \hat{\beta}(\tau))=(\hat{\alpha}(\tau), \hat{\beta}(\hat{\alpha}(\tau),\tau))$. Chernozhukov and Hansen (2006) provide additional details, verify that the above procedure consistently estimates $\alpha(\tau)$ and $\beta(\tau)$, and find the limiting distribution from which conventional asymptotic inference immediately follows.

The above procedure may be modified to obtain confidence intervals that will be asymptotically valid when there is only a weak relationship between D and Z. The idea is to base inference on the Wald statistic $W_n(\alpha)$ for testing whether the coefficients on the instruments are zero (i.e. whether $\gamma(\alpha,\tau)=0$). The intuition for this approach is that at the true $\alpha(\tau)$ the structural Eq. (2.1) implies that $\gamma(\alpha,\tau)=0$. Thus, fixing a value for $\alpha(\tau)$, $\gamma(\alpha,\tau)$ should be near zero if this value is consistent with the structural equation. More formally, when $\alpha=\alpha(\tau), W_n(\alpha) \xrightarrow{d} \chi^2(\dim(\gamma))$. Thus, a valid confidence region for $\alpha(\tau)$ can be based on the inversion of this Wald statistic: $CR_p[\alpha(\tau)]:=\{\alpha: W_n(\alpha)<c_p\}$ contains $\alpha(\tau)$ with probability approaching p, where c_p is the p-percentile of a $\chi^2(\dim(\gamma))$ distribution. Chernozhukov and Hansen (2006) show that this procedure provides a valid confidence region for $\alpha(\tau)$ without putting restrictions on the dependence between D and Z.

The final approach to inference we consider is based on the observation that $P[Y \le D'\alpha(\tau)+X'\beta(\tau)|Z,X]=\tau$ implies that the event $\{Y \le D'\alpha(\tau)+X'\beta(\tau)\}$ is distributed exactly as a Bernoulli(τ) conditional on X and Z regardless of the sample size. Thus, any test statistic that depends only on this event, X, and Z will have a distribution that does not depend on any unknown parameters in finite samples and so can be used to construct valid finite sample inference statements. This basic approach extends work on finite sample inference for unconditional quantiles; see Walsh (1960) and MacKinnon (1964).

Chernozhukov et al. (2005) make use of the aforementioned distributional fact and consider inference based on the GMM objective function:

$$L_n(\alpha,\beta)=\frac{1}{2}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} m_i(\alpha,\beta)\right)' W_n\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} m_i(\alpha,\beta)\right), \tag{2.4}$$

where $m_i(\alpha,\beta)=[\tau-1(Y_i \le D_i'\alpha-X_i'\beta)]\ (Z_i',X_i')'$. In this expression, W_n is a positive definite weight matrix, which is fixed conditional on $(X_1,Z_1),\ldots,(X_n,Z_n)$.

$$W_n=\frac{1}{\tau(1-\tau)}\left(\frac{1}{n}\sum_{i=1}^{n}(Z_i',X_i')'(Z_i',X_i')\right)^{-1},$$

which equals the inverse of the variance of $n^{-1/2}\sum_{i=1}^{n} m_i(\alpha(\tau),\beta(\tau))$ conditional on $(X_1,Z_1),\ldots,(X_n,Z_n)$, is a natural choice of W_n.

Since statistic $L_n(\alpha,\beta)$ defined by Eq. (2.4) depends only on X, Z, and $\{Y \le D'\alpha+X'\beta\}$ it is conditionally pivotal in finite samples when evaluated at $(\alpha,\beta)=(\alpha(\tau)', \beta(\tau)')'$; that is, its distribution at the true parameter values does not depend on any unknown parameters and so exact finite sample critical values for this statistic may be obtained. Let $c_n(p)$ denote the pth quantile of the distribution of the statistic given in Eq. (2.4), and note that $c_n(p)$ may be obtained easily through simulation by, for $j=1,\ldots,J$, 1. drawing $(B_{i,j}, i \le n)$ as iid Bernoulli(τ) random variables, 2. computing $L_{n,j}=\frac{1}{2}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} m_{i,j}\right)' W_n\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} m_{i,j}\right)$ where $m_{i,j}=[\tau-B_{i,j}]\ (Z_i', X_i')'$, and 3. obtaining $c_n(\alpha)$ as the α-quantile of the sample $(L_{n,j}, j=1,\ldots,J)$ for a large number J.

Table 1
Schooling coefficient

Estimation method	τ=0.25	τ=0.50	τ=0.75
A. Quantile regression (no instruments)			
$\alpha(\tau)$	0.074	0.074	0.079
Quantile regression (asymptotic)	(0.064,0.083)	(0.065,0.083)	(0.071,0.087)
Finite sample	(0.047,0.100)	(0.050,0.101)	(0.057,0.098)
B. IV Quantile regression (proximity to 2 and 4 year college as instruments)			
$\alpha(\tau)$	0.175	0.033	0.103
Inverse quantile regression (asymptotic)	(0.081,0.269)	(−0.077,0.142)	(0.025,0.180)
Weak instrument (asymptotic)	(0.018,0.500]	(−0.053,0.500]	(0.068,0.483)
Finite sample	[−0.100,0.500]	[−0.100,0.500]	[−0.100,0.500]

Note: 95% level confidence interval estimates for the schooling coefficient in a simple wage model. Schooling is treated as exogenous in Panel A and as endogenous in Panel B. The first row in each panel reports the point estimate, and the second row reports the interval estimate obtained using the asymptotic approximation. In Panel A, the third row reports an interval estimate obtained using the finite sample approach. In Panel B, the third row reports a weak-instrument robust interval, and the fourth row reports the finite sample interval.

Given $c_n(p)$, a valid p-level confidence region for $(\alpha(\tau)', \beta(\tau)')'$ may then be obtained as the set $CR(p) \equiv \{(\alpha', \beta')': L_n(\alpha, \beta) \leq c_n(p)\}$. $CR(p)$ defines a joint confidence set for all of the model's parameters from which one can define a confidence bound for a real-valued functional $\psi(\alpha(\tau), \beta(\tau), \tau)$ as $CR(p,\psi) = \{\psi(\alpha, \beta, \tau), (\alpha', \beta')' \in CR(p)\}$. It is worth noting that the joint confidence region $CR(p)$ is not conservative, but confidence bounds for functionals may be. Also, forming confidence regions by inverting a test-statistic will generally be computationally demanding. Chernozhukov et al. (2005) presents various feasible algorithms for computing confidence regions and bounds in this context.

3. Case study: Card's (1995) schooling example

We illustrate the different approaches to inference by estimating a wage model that allows for heterogeneity in the effect of schooling on wages as well as other heterogeneity in the effects of other control variables. Because we are concerned that the level of schooling and earnings may be jointly determined, we instrument for schooling. We employ the same data and identification strategy as Card (1995).

Specifically, we suppose that the log of the wage is determined by the following linear quantile model:

$$Y = \alpha(U)S + X'\beta(U)$$

where Y is the log of the hourly wage; S is years of completed schooling; X consists of a constant and 14 demographic controls; and U is an unobservable normalized to follow a uniform distribution over (0,1). We might think of U as indexing unobserved ability, in which case $\alpha(\tau)$ may be thought of as the return to schooling for an individual with unobserved ability τ. Since we believe that years of schooling may be jointly determined with unobserved ability, we use an indicator for residence near a two year college in 1966 and an indicator for residence near a four year college in 1966 as instruments for schooling. Further details about the data sources, descriptive statistics, and arguments for the validity of the instruments are in Card (1995).

We summarize results for the schooling coefficient in Table 1. For quantile regression, we report 95%-level confidence intervals computed via the asymptotic approximation and finite sample approach. For instrumental variable quantile regression, we report 95%-level confidence intervals computed using the asymptotic approximation, the weak-instrument robust approach, and the finite sample approach. Both the finite sample and weak-instrument approach require that we invert a test-statistic for a range of potential values for $\alpha(\tau)$; in all cases, we consider values in the interval $[-.1,.5]$ which we feel covers essentially all plausible values for the return to an additional year of schooling.

The results have a number of interesting features. Looking first at the results in Panel A of Table 1 which treat schooling as exogenous, we see that the point estimates across the various quantiles are quite close, suggesting little heterogeneity in how the quantiles of wages vary with schooling. We also see that the finite sample and asymptotic intervals are qualitatively similar, though the finite sample intervals are wider than the asymptotic intervals.

In the estimates which treat schooling as endogenous, the differences between the inference approaches are striking. The point estimates suggest considerable heterogeneity in the returns to schooling, though even the usual asymptotic intervals are wide enough to make finding statistical evidence for this quite unlikely. Considering next the weak instrument robust intervals, we see that they are substantially wider than the usual asymptotic intervals in all cases. For each quantile we consider, the weak instrument robust intervals include essentially the entire parameter space. The weak instrument robust intervals do exclude 0 in the two cases where the usual intervals do. While the weak instrument robust intervals rely on less stringent assumptions than do the usual asymptotic intervals, they still rely on an asymptotic argument. Looking last at the finite sample intervals which are valid under minimal assumptions and do not require any asymptotic argument, we see that they cover the entire parameter space in every case.

The likely cause for the large discrepancies is the weakness of the relationship between the instruments and schooling. If one regresses education on the control variables and the two proximity instruments, the F-statistic on the excluded instruments is 8.32 which is in the range that many would consider weak in the usual linear IV model. The weakness of the instruments clearly explains the difference between the usual asymptotic and weak instrument asymptotic results. The differences between the weak instrument and finite sample intervals is then likely driven by the fact that with a large number of covariates and weak instruments the “effective” sample size is quite small. It would be interesting to explore these differences and develop formal procedures for judging differences between asymptotic and finite sample intervals. Overall, the difference between the asymptotic and finite sample inference statements calls the validity of the asymptotics into question in this example.

References

Card, D., 1995. Using geographic variation in college proximity to estimate the return to schooling. In: Christofides, L., Grant, E.K., Swindinsky, R. (Eds.), Aspects of Labour Economics: Essays in Honour of John Vanderkamp. University of Toronto Press.

Chernozhukov, V., Hansen, C., 2006. Instrumental quantile regression inference for structural and treatment effect models. J. Econom. 132 (2), 491–525.

Chernozhukov, V., Hansen, C., Jansson, M., 2005. “Finite-Sample Inference in Econo-metric Models via Quantile Restrictions,” mimeo.

Koenker, R., Bassett, G.S., 1978. Regression quantiles. Econometrica 46, 33–50.
MacKinnon, W.J., 1964. Table for both the sign test and distribution-free confidence intervals of the median for sample sizes to 1000. J. Am. Stat. Assoc. 59, 935–956.
Pakes, A., Pollard, D., 1989. Simulation and asymptotics of optimization estimators. Econometrica 57, 1027–1057.
Walsh, J.E., 1960. Nonparametric tests for median by interpolation from sign tests. Ann. Inst. Stat. Math. 11, 183–188.