On the equivalence of instrumental variables estimators for linear models

Antonio F. Galvao*
Gabriel Montes-Rojas†

Abstract

This note shows the equivalence of different instrumental variables estimators to solve the endogeneity problem in linear models when valid instruments are available. We demonstrate that the exclusion restriction estimator proposed by Chernozhukov and Hansen (2006) is equivalent to the two-stage least squares and the control function estimators for linear models.

Key Words: Instrumental variables; least-squares; control function; quantile regression

JEL Classification: C14, C23

*Department of Economics, University of Iowa, W284 PBB, 21 E. Market Street, Iowa City, IA 52242. E-mail: antonio-galvao@uiowa.edu

†CONICET-Universidad de San Andrés. Email: gmontesrojas@udesa.edu.ar

1 Introduction

The problems of endogeneity and causality occupy a substantial amount of research in theoretical and applied econometrics. One of the most popular approaches to solve endogeneity and achieve causal interpretation is the instrumental variables (IV) estimator, (see, e.g., Hausman, 1983; Angrist, Imbens, and Rubin, 1996; Angrist and Krueger, 2001; Angrist and Pischke, 2009). This note shows that the exclusion restriction estimator proposed by Chernozhukov and Hansen (2006) (CH) to solve endogeneity in quantile regression (QR) models also applies to estimate conditional average models. Using an analogue least-squares (LS) version of the CH estimator, we demonstrate that the CH estimator for the endogeneous variables coefficients is equal to the standard two-stage LS (2SLS) estimator for linear models. Given that for linear models, the control function (CF) approach is also equivalent to the 2SLS, this note thus illustrates the equivalence of the three estimators: 2SLS, CF and CH. We also show that the equivalence does not hold in general for the exogeneous variables coefficients, unless the model is exactly identified with the same number of instruments as endogenous variables.

Chernozhukov and Hansen (2005, 2006) proposed a variant of the IV approach, called the *inverse* approach, for QR models. Their estimator uses the exclusion restriction imposed by the IV, which does not belong to the outcome equation of the structural model but has a key role in constructing the conditioning set. CH estimator is one of an extensive list of estimators for semiparametric QR structural models with endogeneity. This method has been extensively used in the literature to accommodate different problems of endogeneity in QR models, e.g., Chernozhukov and Hansen (2008) develop robust inference procedures, Jun (2008) develops a testing procedure that is robust to identification quality, Galvao (2011) extends the QR IV to dynamic panel data models, Su and Yang (2011) propose a spatial quantile autoregression model, among others.

Although we show that the estimators are equivalent for conditional average models with endogeneity, the result does not extend trivially to conditional quantile models. As argued by Lee (2007) different estimators (i.e. IV, CF, CH, etc.) are in general only suitable for a particular QR structural model. Depending on the assumptions on a given structural model, different alternative approaches have been proposed for identification in structural QR models, and the corresponding estimators are not equivalent.[1]

[1]Three major approaches can be considered. First, an IV approach as those proposed by Hong and

2 The linear two-stage model and CH estimator

Suppose we have a linear structural model given by

$$\begin{aligned} y &= d\alpha + X\beta + u, \\ d &= X\delta_X + Z\delta_Z + v, \end{aligned} \tag{1}$$

where y is an outcome scalar variable, d is a scalar endogenous variable, X a vector of k_X exogenous covariates, and Z is a vector of k_Z IV. Moreover, u is a scalar random variable that aggregates unobserved factors in the outcome equation, and v is a scalar containing unobserved disturbances determining d and correlated with u. Due to the dependence of v and u, d is also sampled depending on u, and thus is endogeneous. We assume that Z is independent of both u and v. This is a particular case of the structural models listed in Lee (2007) and CH.

Consider the following QR model for (1),

$$Q_y(\tau|d,X) = d\alpha(\tau) + X\beta(\tau), \quad \tau \in (0,1). \tag{2}$$

Note that *without* endogeneity equation (2) can be written as

$$y = d\alpha(w) + X\beta(w), \quad w|d,X \sim U(0,1),$$

where $U(0,1)$ is a standard uniform distribution.

When the variable d is endogenous and valid instrument Z is available, the structural model is still given by (2), but it can be represented as

$$y = d\alpha(w) + X\beta(w), \quad w|X,Z \sim U(0,1). \tag{3}$$

Thus, the valid instrument Z is not part of the structural model, but it allows one to write the random coefficients representation of the QR model as in (3).

Consider now the sample counterpart of the random variables (y,d,X,Z) of model (1) given by a sample of size n, $\{y_i, d_i, X_i, Z_i\}_{i=1}^n$. Given the intuition above, CH proposed the

Tamer (2003), Chen, Linton, and van Keilegom (2003), and Honoré and Hu (2004). Second a "fitted value" approach builds on the papers by Amemiya (1982) and Powell (1983), in which a first step fits the endogeneous variable(s) as a function of exogeneous covariates and IV, and this is then plugged in a second-stage. Third, Lee (2007) adopts a two-step CF approach where in first step consists of estimation of the residuals of the reduced-form equation for the endogenous explanatory variable. Abadie, Angrist, and Imbens (2002) and Chernozhukov and Hansen (2005) develop IV estimators of quantile treatment effects estimators.

following estimator based on the exclusion restriction given by Z,

$$\widehat{\alpha}_{CH}(\tau) = \arg\min_{\alpha} \|\widehat{\gamma}(\alpha,\tau)\|_{A(\tau)}, \tag{4}$$

$$(\widehat{\beta}(\alpha,\tau), \widehat{\gamma}(\alpha,\tau)) = \arg\min_{(\beta,\gamma)} \frac{1}{n}\sum_{i=1}^{n} \rho_\tau(y_i - d_i\alpha - X_i\beta(\tau) - Z_i\gamma(\tau)), \tag{5}$$

where $\|x\|_A = \sqrt{x'Ax}$, $A(\tau)$ is any uniformly positive definite matrix, and $\rho_\tau(u) = u(\tau - 1[u < 0])$ the usual QR check function. The final inverse or exclusion restriction estimator is thus $(\widehat{\alpha}_{CH}(\tau), \widehat{\beta}_{CH}(\widehat{\alpha}_{CH}(\tau), \tau))$. The intuition behind this estimator is that Z does not belong to the structural (outcome equation) model. However, by including Z, this plays a crucial role in the conditioning set used to estimate the structural parameter $\alpha(\tau)$. Note that an attractive feature of this model, which shares with the some semiparametric IV estimators, is that it does not need to specify a 'first-stage'. The estimator is implemented in the following way: (i) Consider a grid $\mathcal{A}$ for possible values of α. (ii) for each value $\alpha \in \mathcal{A}$ model (5) is estimated. (iii) Then $\widehat{\alpha}_{CH}$ is selected by finding the minimum of (4) over the pre-specified grid.

The analogue estimator for the conditional mean model (1) is then defined as

$$\widehat{\alpha}_{CHLS} = \arg\min_{\alpha} \widehat{\gamma}(\alpha)'\Omega\widehat{\gamma}(\alpha), \tag{6}$$

$$(\widehat{\beta}(\alpha), \widehat{\gamma}(\alpha)) = \arg\min_{(\beta,\gamma)} \frac{1}{n}\sum_{i=1}^{n} (y_i - d_i\alpha - X_i\beta - Z_i\gamma)^2, \tag{7}$$

where Ω is a positive definite matrix (also quadratic and with the corresponding dimensions of γ). We show that the estimator defined in (6)-(7), which we define as CHLS, is equivalent to the 2SLS estimator when valid instruments are available and when the optimal Ω is used.

3 Two-stage least squares equivalence

In the 2SLS procedure, the mean regresion parameters of the endogeneous variables are estimated by imposing a linear 'first-stage', where the conditional expectation of d on (X, Z) is,

$$\mathrm{E}[d|X,Z] = X\delta_X + Z\delta_Z.$$

Consider now matrix $(\boldsymbol{y}, \boldsymbol{d}, \boldsymbol{X}, \boldsymbol{Z})$ of dimension $n \times (1, 1, k_X, k_Z)$. Let $\widehat{\boldsymbol{d}} \equiv \boldsymbol{X}\widehat{\delta}_X + \boldsymbol{Z}\widehat{\delta}_Z$, where $\widehat{\delta}_X = (\boldsymbol{X}'\boldsymbol{M}_{\boldsymbol{Z}}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{M}_{\boldsymbol{Z}}\boldsymbol{d}$ and $\widehat{\delta}_Z = (\boldsymbol{Z}'\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{M}_{\boldsymbol{X}}\boldsymbol{d}$, with $\boldsymbol{M}_{\cdot} = \boldsymbol{I} - \cdot(\cdot'\cdot)^{-1}\cdot'$

the standard residual projection (square and idempotent) matrix on the space spanned by the variables $\cdot$. Then, the 2SLS IV estimator is obtained from the 'second-stage' regression model

$$\boldsymbol{y} = \widehat{\boldsymbol{d}}\widehat{\alpha}_{2SLS} + \boldsymbol{X}\widehat{\beta}_{2SLS} + \boldsymbol{r_{2SLS}},$$

where $\boldsymbol{r_{2SLS}}$ means simply the difference between the regressand and the rest of the right side of the regression.

The CF approach applied to the linear two-stage model (1) uses the same 'first-stage' but a different 'second-stage'. Define $\widehat{\boldsymbol{v}} \equiv \boldsymbol{d} - \widehat{\boldsymbol{d}}$ as the 'first-stage' residual. Then in the CF 'second-stage', this residual is included as an additional regressor, together with $\boldsymbol{d}$ and $\boldsymbol{X}$. That is,

$$\boldsymbol{y} = \widehat{\boldsymbol{d}}\widehat{\alpha}_{CF} + \boldsymbol{X}\widehat{\beta}_{CF} + \widehat{\boldsymbol{v}}\widehat{\eta}_{CF} + \boldsymbol{r_{CF}},$$

where $\boldsymbol{r_{CF}}$ is the resulting regression residual.

A well known result is that $\widehat{\alpha}_{CF} = \widehat{\alpha}_{2SLS}$ and $\widehat{\beta}_{CF} = \widehat{\beta}_{2SLS}$. The simple proof follows from the Frisch-Waugh-Lovell theorem in which if we run a regression of $(\boldsymbol{d}, \boldsymbol{X})$ on $\widehat{\boldsymbol{v}}$ we obtain $(\boldsymbol{X}, \widehat{\boldsymbol{d}})$. See the illustrative notes by Imbens and Wooldridge (2007).

The main result of this note is that $\widehat{\alpha}_{CF} = \widehat{\alpha}_{2SLS} = \widehat{\alpha}_{CHLS}$ thus completing the equivalence between the CHLS and 2SLS IV and CF estimators for linear models with endogeneity. The following proposition formalizes the result.

Proposition 1 *Under model in equation (1), let Z be a set of valid instruments for d, i.e. $\boldsymbol{d}'\boldsymbol{M_X Z}(\boldsymbol{Z}'\boldsymbol{M_X Z})^{-1}\boldsymbol{Z}'\boldsymbol{M_X d}$ and $\boldsymbol{Z}'\boldsymbol{M_X Z}$ are invertible. If $\Omega = \boldsymbol{Z}'\boldsymbol{M_X Z}$,*
(i) the 2SLS IV and CHLS estimators for α are equivalent, and
(ii) the estimators for β are not equal unless the model is exactly identified, i.e. $k_Z = 1$ in which case $\widehat{\gamma}(\widehat{\alpha}_{CHLS}) = \boldsymbol{0}$.

Proof. Note that from the 2SLS IV model we have

$$\begin{aligned}
\widehat{\alpha}_{2SLS} &= (\widehat{\boldsymbol{d}}'\boldsymbol{M_X}\widehat{\boldsymbol{d}})^{-1}\widehat{\boldsymbol{d}}'\boldsymbol{M_X y} \\
&= \left((\boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{M_X Z})^{-1}\boldsymbol{Z}'\boldsymbol{M_X d})'\boldsymbol{M_X}(\boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{M_X Z})^{-1}\boldsymbol{Z}'\boldsymbol{M_X d})\right)^{-1}(\boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{M_X Z})^{-1}\boldsymbol{Z}'\boldsymbol{M_X d})'\boldsymbol{M_X y} \\
&= \left(\boldsymbol{d}'\boldsymbol{M_X Z}(\boldsymbol{Z}'\boldsymbol{M_X Z})^{-1}\boldsymbol{Z}'\boldsymbol{M_X d}\right)^{-1}\left(\boldsymbol{d}'\boldsymbol{M_X Z}(\boldsymbol{Z}'\boldsymbol{M_X Z})^{-1}\boldsymbol{Z}'\boldsymbol{M_X y}\right),
\end{aligned}$$

where we are using the fact that $\boldsymbol{M_X X} = \boldsymbol{0}$. Moreover,

$$\widehat{\beta}_{2SLS} = (\boldsymbol{X}'\boldsymbol{M}_{\hat{\boldsymbol{d}}}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{M}_{\hat{\boldsymbol{d}}}\boldsymbol{y} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'(\boldsymbol{y} - \boldsymbol{d}\widehat{\alpha}_{2SLS}).$$

The second equality determines that $\widehat{\beta}_{2SLS}$ can be obtained as an OLS estimator of a regression of $(\boldsymbol{y} - \boldsymbol{d}\widehat{\alpha}_{2SLS})$ on $\boldsymbol{X}$.

The CHLS estimator in (6)-(7) is implemented, for a given α, by considering the auxiliary regression

$$\boldsymbol{y} - \boldsymbol{d}\alpha = \boldsymbol{X}\widehat{\beta}(\alpha) + \boldsymbol{Z}\widehat{\gamma}(\alpha) + \boldsymbol{r}(\alpha),$$

where $(\widehat{\beta}(\alpha), \widehat{\gamma}(\alpha))$ is the OLS estimator of running a regression of $\boldsymbol{y} - \boldsymbol{d}\alpha$ on $(\boldsymbol{X}, \boldsymbol{Z})$ and $\boldsymbol{r}(\alpha)$ the resulting regression residual. Then, notice that

$$\begin{aligned}
\widehat{\gamma}(\alpha) &= \left(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{Z}\right)^{-1}\left(\boldsymbol{Z}'\boldsymbol{M_X}(\boldsymbol{y} - \boldsymbol{d}\alpha)\right), \\
\widehat{\beta}(\alpha) &= \left(\boldsymbol{X}'\boldsymbol{M_Z}\boldsymbol{X}\right)^{-1}\left(\boldsymbol{X}'\boldsymbol{M_Z}(\boldsymbol{y} - \boldsymbol{d}\alpha)\right).
\end{aligned}$$

The CHLS estimator is then given by

$$\begin{aligned}
\widehat{\alpha}_{CHLS} &= \arg\min_{\alpha}\left[\left(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{Z}\right)^{-1}\left(\boldsymbol{Z}'\boldsymbol{M_X}(\boldsymbol{y} - \boldsymbol{d}\alpha)\right)\right]'\Omega\left[\left(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{Z}\right)^{-1}\left(\boldsymbol{Z}'\boldsymbol{M_X}(\boldsymbol{y} - \boldsymbol{d}\alpha)\right)\right] \\
&= \arg\min_{\alpha}\Big\{\left(\boldsymbol{y}'\boldsymbol{M_X}\boldsymbol{Z}\right)\left(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{Z}\right)^{-1}\Omega\left(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{Z}\right)^{-1}\left(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{y}\right) \\
&\quad +(\alpha'\boldsymbol{d}'\boldsymbol{M_X}\boldsymbol{Z})\left(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{Z}\right)^{-1}\Omega\left(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{Z}\right)^{-1}\left(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{d}\alpha\right) \\
&\quad -2(\boldsymbol{y}'\boldsymbol{M_X}\boldsymbol{Z})\left(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{Z}\right)^{-1}\Omega\left(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{Z}\right)^{-1}\left(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{d}\alpha\right)\Big\}.
\end{aligned}$$

Solving the minimization with respect to α, we obtain from the first order condition

$$\begin{aligned}
\widehat{\alpha}_{CHLS} &= \left[\left(\boldsymbol{d}'\boldsymbol{M_X}\boldsymbol{Z}\right)\left(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{Z}\right)^{-1}\Omega\left(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{Z}\right)^{-1}\left(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{d}\right)\right]^{-1} \\
&\quad \left(\boldsymbol{d}'\boldsymbol{M_X}\boldsymbol{Z}\right)\left(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{Z}\right)^{-1}\Omega\left(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{Z}\right)^{-1}\left(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{y}\right).
\end{aligned}$$

Now note that if $\Omega = \boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{Z}$ we obtain that $\widehat{\alpha}_{CHLS} = \widehat{\alpha}_{2SLS}$, as stated in (i).

Regarding (ii), by substituting the above equation into $\widehat{\beta}(\alpha)$ we obtain

$$\begin{aligned}
\widehat{\beta}_{CHLS} &= \left(\boldsymbol{X}'\boldsymbol{M_Z}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{M_Z}(\boldsymbol{y} - \boldsymbol{d}\widehat{\alpha}_{CHLS}) \\
&= \left(\boldsymbol{X}'\boldsymbol{M_Z}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{M_Z}(\boldsymbol{y} - \boldsymbol{d}\widehat{\alpha}_{2SLS}),
\end{aligned}$$

which is not necessarily equal to $(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'(\boldsymbol{y} - \boldsymbol{d}\widehat{\alpha}_{2SLS}) = \widehat{\beta}_{2SLS}$.

Finally, note that when the model is exactly identified (same number of instruments as endogenous variables) then

$$\begin{aligned}
&\widehat{\gamma}(\widehat{\alpha}_{CHLS}) = \left(\boldsymbol{X}'\boldsymbol{M_Z}\boldsymbol{X}\right)^{-1}\left(\boldsymbol{X}'\boldsymbol{M_Z}(\boldsymbol{y} - \boldsymbol{d}\widehat{\alpha}_{CHLS})\right) \\
= &\left(\boldsymbol{X}'\boldsymbol{M_Z}\boldsymbol{X}\right)^{-1}\left\{\boldsymbol{X}'\boldsymbol{M_Z}\left(\boldsymbol{y} - \boldsymbol{d}\left[\left(\boldsymbol{d}'\boldsymbol{M_X}\boldsymbol{Z}\right)\left(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{Z}\right)^{-1}\left(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{d}\right)\right]^{-1}\left(\boldsymbol{d}'\boldsymbol{M_X}\boldsymbol{Z}\right)\left(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{Z}\right)^{-1}\left(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{y}\right)\right)\right. \\
= &\left(\boldsymbol{X}'\boldsymbol{M_Z}\boldsymbol{X}\right)^{-1}\left\{\boldsymbol{X}'\boldsymbol{M_Z}\left(\boldsymbol{y} - \boldsymbol{d}(\boldsymbol{d}'\boldsymbol{M_X}\boldsymbol{Z})^{-1}\left(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{Z}\right)\left(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{d}\right)^{-1}\left(\boldsymbol{d}'\boldsymbol{M_X}\boldsymbol{Z}\right)\left(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{Z}\right)^{-1}\left(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{y}\right)\right)\right\} \\
= &\left(\boldsymbol{X}'\boldsymbol{M_Z}\boldsymbol{X}\right)^{-1}\left\{\boldsymbol{X}'\boldsymbol{M_Z}\left(\boldsymbol{y} - \boldsymbol{d}(\boldsymbol{d}'\boldsymbol{M_X}\boldsymbol{Z})^{-1}(\boldsymbol{Z}'\boldsymbol{M_X}\boldsymbol{y})\right)\right\} \\
= &\ \boldsymbol{0}.
\end{aligned}$$

Then $\widehat{\beta}_{CHLS}$ can be obtained from the regression

$$\begin{aligned}\boldsymbol{y} - \boldsymbol{d}\widehat{\alpha}_{2SLS} &= \boldsymbol{X}\widehat{\beta}_{CHLS} + \boldsymbol{Z}\widehat{\gamma}_{CHLS} + \boldsymbol{r}(\widehat{\alpha}_{CHLS}) \\ &= \boldsymbol{X}\widehat{\beta}_{CHLS} + \boldsymbol{Z}\mathbf{0} + \boldsymbol{r}(\widehat{\alpha}_{CHLS}) = \boldsymbol{X}\widehat{\beta}_{2SLS} + \boldsymbol{r_{2SLS}}.\end{aligned}$$

This concludes statement (ii) $\widehat{\beta}_{CHLS} = \widehat{\beta}_{2SLS}$ for $k_Z = 1$. ■

As a final important remark note that Ω weights the coefficient of each instrument in the quadratic form $\widehat{\gamma}(\alpha)'\Omega\widehat{\gamma}(\alpha)$ of equation (6). Given that 2SLS is efficient the optimal weighting scheme for CHLS thus requires Ω to be given by the variance of Z net of X. When the model is exactly identified, i.e. $k_Z = 1$, the result holds with the 2SLS being the simple IV estimator. This follows immediately from Proposition 1. Note that if the number of columns in $\boldsymbol{d}$ is the same as in $\boldsymbol{Z}$, then $\boldsymbol{d'M_XZ}(\boldsymbol{Z'M_XZ})^{-1}\boldsymbol{Z'M_Xd} = \boldsymbol{d'(Z'M_XZ)^{-1}d}$ and $\boldsymbol{d'M_XZ}(\boldsymbol{Z'M_XZ})^{-1}\boldsymbol{Z'M_Xy} = \boldsymbol{d'(Z'M_XZ)^{-1}y}$, and we can simplify the 2SLS estimator to $\widehat{\alpha}_{2SLS} = (\boldsymbol{Z'M_Xd})^{-1}\boldsymbol{Z'M_Xy}$.

References

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

Amemiya, T. (1982): "Two stage least absolute deviations estimators," *Econometrica*, 50, 689–711.

Angrist, J. D., G. W. Imbens, and D. B. Rubin (1996): "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, 91, 444–455.

Angrist, J. D., and A. Krueger (2001): "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments," *Journal of Economic Perspectives*, 15, 69–85.

Angrist, J. D., and J.-S. Pischke (2009): *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, Princeton, New Jersey.

Chen, X., O. Linton, and I. van Keilegom (2003): "Estimation of Semiparametric Models When the Criterion Function is Not Smooth," *Econometrica*, 71, 1591–1608.

Chernozhukov, V., and C. Hansen (2005): "An IV Model of Quantile Treatment Effects," *Econometrica*, 73, 245–261.

——— (2006): "Instrumental Quantile Regression Inference for Structural and Treatment Effects Models," *Journal of Econometrics*, 132, 491–525.

——— (2008): "Instrumental Variable Quantile Regression: A Robust Inference Approach," *Journal of Econometrics*, 142, 379–398.

Galvao, A. F. (2011): "Quantile Regression for Dynamic Panel Data with Fixed Effects," *Journal of Econometrics*, 164(1), 142–157.

Hausman, J. A. (1983): "Specification and Estimation of Simultaneous Equation Models," in *Handbook of Econometrics*, ed. by Z. Griliches, and M. D. Intriligator, vol. II, pp. 391–448. North-Holland Press.

Hong, H., and E. Tamer (2003): "Inference in censored models with endogenous regressors," *Econometrica*, 71, 905–932.

Honoré, B. E., and L. Hu (2004): "On the performance of some robust instrumental variables estimators," *Journal of Business and Economic Statistics*, 22(1), 30–39.

Imbens, G., and J. Wooldridge (2007): "What's New in Econometrics? Lecture 6," NBER Summer Courses 2007 `http://www.nber.org/WNE/lect_6_controlfuncs.pdf`.

Jun, S. J. (2008): "Weak Identification Robust Tests in an Instrumental Quantile Model," *Journal of Econometrics*, 144, 118–138.

Lee, S. (2007): "Endogeneity in Quantile Regression Models: A Control Function Approach," *Journal of Econometrics*, 101(2), 1131–1158.

Powell, J. L. (1983): "The asymptotic normality of two-stage least absolute deviations estimators," *Econometrica*, 51, 1569–1576.

Su, L., and Z. Yang (2011): "Instrumental Variable Quantile Estimation of Spatial Autoregressive Models," Manuscript, Singapore Management University.