1-1-2007

Consistent Estimation with Weak Instruments in Panel Data

Chihwa Kao
Syracuse University

Long Liu
Syracuse University

ISSN: 1525-3066

Center for Policy Research
Working Paper No. 95

CONSISTENT ESTIMATION WITH WEAK INSTRUMENTS IN PANEL DATA

Chihwa Kao and Long Lui

Center for Policy Research
Maxwell School of Citizenship and Public Affairs
Syracuse University
426 Eggers Hall
Syracuse, New York 13244-1020
(315) 443-3114 | Fax (315) 443-1081
e-mail: ctrpol@syr.edu

May 2007

$5.00

CENTER FOR POLICY RESEARCH – Spring 2007

Timothy Smeeding, Director
Professor of Economics & Public Administration

Associate Directors

Margaret Austin
Associate Director
Budget and Administration

Douglas Wolf
Professor of Public Administration
Associate Director, Aging Studies Program

John Yinger
Professor of Economics and Public Administration
Associate Director, Metropolitan Studies Program

SENIOR RESEARCH ASSOCIATES

Badi Baltagi .... Economics
Pablo Beramendi............................Political Science
Dan Black.. Economics
William Duncombe .... Public Administration
Gary Engelhardt ....Economics
Deborah Freund ....Public Administration
Madonna Harrington Meyer ....Sociology
Christine Himes.......................................Sociology
William C. Horrace ....Economics
Duke Kao ....Economics
Eric Kingson .... Social Work
Thomas Kniesner ....Economics
Jeffrey Kubik ....Economics
Andrew London.......................................Sociology
Len Lopoo ....Public Administration
Jerry Miner ....Economics
Jan Ondrich...Economics
John Palmer ....Public Administration
Lori Ploutz-Snyder.......................Exercise Science
David Popp............................Public Administration
Grant Reeher .... Political Science
Christopher Rohlfs ....Economics
Stuart Rosenthal ....Economics
Ross Rubenstein....................Public Administration
Margaret Usdansky ....Sociology
Michael Wasylenko ....Economics
Janet Wilmoth ....Sociology

GRADUATE ASSOCIATES

Javier Baez ....Economics
Sonali Ballal ....Public Administration
Jesse Bricker ....Economics
Maria Brown.....................................Social Science
Mike Eriksen ....Economics
Qu Feng ....Economics
Katie Fitzpatrick.....................................Economics
Alexandre Genest ....Public Administration
Julie Anna Golebiewski...........................Economics
Nadia Greenhalgh-Stanley ....Economics
Tamara Hafner.......................Public Administration
Yue Hu ....Economics
Sung Hyo Hong...................................... Economics
Joseph Marchand................................... Economics
Neelakshi Medhi .... Social Science
Larry Miller .... Public Administration
Wendy Parker .... Sociology
Emily Pas .... Economics
Shawn Rohlin .... Economics
Cynthia Searcy........................ Public Administration
Jeff Thompson .... Economics
Coady Wing .... Public Administration
Daniel Yanulavich .... Public Administration
Ryan Yeung .... Public Administration

STAFF

Kelly Bogart ....Administrative Secretary
Martha Bonney.. ...Publications/Events Coordinator
Karen Cimilluca.…………..Administrative Secretary
Kitty Nasto.........................Administrative Secretary
Candi Patterson......................Computer Consultant
Mary Santy……..……………Administrative Secretary
Tammy Tanner…………Librarian/Office Coordinator

Abstract

This note analyzes the asymptotic distribution for instrumental variables regression for panel data when the available instruments are weak. We show that consistency can be established in panel data.

JEL classification: C33

Key Words: Weak Instrument; Two Stage Least Squares; Panel Data; Concentration Parameter.

Consistent Estimation with Weak Instruments in Panel Data

Chihwa Kao*, Long Liu†
Syracuse University

April 20, 2007

Abstract

This note analyzes the asymptotic distribution for instrumental variables regression for panel data when the available instruments are weak. We show that consistency can be established in panel data.

Key Words: *Weak Instrument; Two Stage Least Squares; Panel Data; Concentration Parameter.*

1 Motivation and Results

In recent year, economists have been concerned with the problem of weak instruments or partial identification, see Stock, Wright and Yogo (2002) for an excellent survey. Economists found that the first stage F statistic in the two stage least squares (2SLS) regression is often low, say, less than 10. In this case, the usual asymptotic normal approximations can be quite poor, even if the number of observations is large. To provide better asymptotic approximations in this case, Staiger and Stock (1997) derive the weak-instrument asymptotics for instrumental variables estimators. Staiger and Stock show that the 2SLS is inconsistent (i.e., converges to a random variable) and has a nonstandard limiting distribution. In this note we study the asymptotics of 2SLS with weak instruments in panel models. We show that the consistency of 2SLS can be established in panel data. We use $(n,T) \overset{\text{seq}}{\to} \infty$ to denote the sequential limit, i.e., $n \to \infty$ followed by $T \to \infty$.

Consider the following panel linear IV regression model with a single endogenous regressor

$$y_t = Y_t \beta + u_t \tag{1}$$

and

$$Y_t = Z_t \Pi + v_t \tag{2}$$

$t = 1, 2, \cdots, T$, where y_t and Y_t are $n \times 1$ vectors of observations on endogenous variables, Z_t is a $n \times k$ matrix of instruments, Π is a $k \times 1$ coefficient vector, and u_t and v_t are $n \times 1$ vectors of disturbance terms.

*Address correspondence to: Chihwa Kao, Center for Policy Research, 426 Eggers Hall, Syracuse University, Syracuse, NY 13244-1020; e-mail: cdkao@maxwell.syr.edu.

†Long Liu: Economics Department, 110 Eggers Hall, Syracuse University, Syracuse, NY 13244-1020; e-mail: loliu@maxwell.syr.edu.

Let u_{it} and v_{it} be ith element of u_t and v_t respectively. The errors $(u_{it}, v_{it})'$ are assumed to be iid $N(0, \Sigma)$, where the elements of Σ are σ_u^2, σ_{uv} and σ_v^2, and let $\rho = \sigma_{uv}/(\sigma_u \sigma_v)$. In this note, the errors are assumed be iid for simplicity. This assumption can be relaxed to weak dependence across time series and cross-section at the expense of complicated notations and will be studied in a different paper. Equation (1) is the structural equation and β is the scalar parameter of interest. The reduced-form equation (2) relates the endogenous regressor to the instruments. As proposed by Staiger and Stock (1997), the following assumption is used to describe the nature of weak instruments.

Assumption 1 $\Pi = C/\sqrt{n}$, *where* C *is a* $k \times 1$ *constant matrix.*

The strength of instrument can be measured by the concentration parameter λ_t^2, which is defined as

$$\lambda_t^2 = \Pi' Z_t' Z_t \Pi / \sigma_v^2.$$

Under Assumption 1 we obtain

$$\lambda_t^2 = \Pi' Z_t' Z_t \Pi / \sigma_v^2 = C' Z_t' Z_t C / \left(n \sigma_v^2\right) \xrightarrow{p} C' Q C / \sigma_v^2 \equiv \lambda^2 \tag{3}$$

as $n \to \infty$ where we assume for a give t,

$$\frac{1}{n} Z_t' Z_t \xrightarrow{p} Q = E\left(Z_t' Z_t\right).$$

Then the panel 2SLS estimator is

$$\hat{\beta}_{2SLS} = \sum_{t=1}^{T} \left(Y_t' P_{Z_t} y_t\right) / \sum_{t=1}^{T} \left(Y_t' P_{Z_t} Y_t\right)$$

where $P_{Z_t} = Z_t \left(Z_t' Z_t\right)^{-1} Z_t'$.

Theorem 1 *Under assumption 1,*

1.

$$\hat{\beta}_{2SLS} \xrightarrow{p} \beta + \frac{\sigma_u \rho k}{\sigma_v \left(\lambda^2 + k\right)},$$

2.

$$\sqrt{T}\left(\hat{\beta}_{2SLS} - \beta - \frac{\sigma_u \rho k}{\sigma_v \left(\lambda^2 + k\right)}\right) \xrightarrow{d} N\left(0, \frac{\sigma_u^2 \left[\lambda^2 + k\left(1 + \rho^2\right)\right]}{\sigma_v^2 \left(\lambda^2 + k\right)^2}\right)$$

as $(n, T) \overset{\text{seq}}{\rightarrow} \infty$.

2 Proof and Discussion

Define $\zeta_{1tn} = Y_t' P_{Z_t} u_t$, $\zeta_{2tn} = Y_t' P_{Z_t} Y_t$, $\xi_{1Tn} = \frac{1}{T}\sum_{t=1}^{T} \zeta_{1tN}$, and $\xi_{2Tn} = \frac{1}{T}\sum_{t=1}^{T} \zeta_{2tn}$. Hence

$$\hat{\beta}_{2SLS} - \beta = \frac{\sum_{t=1}^{T} \left(Y_t' P_{Z_t} u_t\right)}{\sum_{t=1}^{T} \left(Y_t' P_{Z_t} Y_t\right)} = \frac{\frac{1}{T}\sum_{t=1}^{T} \zeta_{1tn}}{\frac{1}{T}\sum_{t=1}^{T} \zeta_{2tn}} = \frac{\xi_{1Tn}}{\xi_{2Tn}}.$$

We first present the following lemma:

Lemma 1 *Under Assumption 1*

1. $\zeta_{1tn} = \sigma_v \sigma_u \left(\lambda z_{tu} + s_{tuv}\right) + o_p(1)$ as $n \to \infty$, where $z_{tu} = \frac{\Pi' Z_t' u_t}{\sigma_u \sqrt{\Pi' Z_t' Z_t \Pi}}$ and $s_{tuv} = \frac{v_t' P_{Z_t} u_t}{\sigma_v \sigma_u}$,

2. $\zeta_{2tn} = \sigma_v^2 \left(\lambda^2 + 2\lambda z_{tv} + s_{tvv}\right) + o_p(1)$ as $n \to \infty$, where $z_{tv} = \frac{\Pi' Z_t' v_t}{\sigma_v \sqrt{\Pi' Z_t' Z_t \Pi}}$ and $s_{tvv} = \frac{v_t' P_{Z_t} v_t}{\sigma_v^2}$,

3. $\xi_{1Tn} \xrightarrow{p} \sigma_v \sigma_u \rho k$ as $(n, T) \overset{\text{seq}}{\to} \infty$,

4. $\xi_{2Tn} \xrightarrow{p} \sigma_v^2 \left(\lambda^2 + k\right)$ as $(n, T) \overset{\text{seq}}{\to} \infty$.

Proof. Consider (1) and (2). Following Rothenberg (1984), we know

$$\begin{aligned}
\zeta_{1tn} &= Y_t' P_{Z_t} u_t = \Pi' Z_t' u_t + v_t' P_{Z_t} u_t \\
&= \sigma_v \sigma_u \left(\sqrt{\frac{\Pi' Z_t' Z_t \Pi}{\sigma_v^2}}\right) \left(\frac{\Pi' Z_t' u_t}{\sigma_u \sqrt{\Pi' Z_t' Z_t \Pi}}\right) + \left(\sigma_v \sigma_u\right) \left(\frac{v_t' P_{Z_t} u_t}{\sigma_v \sigma_u}\right) \\
&= \sigma_v \sigma_u \lambda_t z_{tu} + \left(\sigma_v \sigma_u\right) s_{tuv} \\
&= \sigma_v \sigma_u \left(\lambda_t z_{tu} + s_{tuv}\right)
\end{aligned}$$

and

$$\begin{aligned}
\zeta_{2tn} &= Y_t' P_{Z_t} Y_t = \Pi' Z_t' Z_t \Pi + 2\Pi' Z_t' v_t + v_t' P_{Z_t} v_t \\
&= \sigma_v^2 \lambda_t^2 + 2\sigma_v^2 \left(\sqrt{\frac{\Pi' Z_t' Z_t \Pi}{\sigma_v^2}}\right) \left(\frac{\Pi' Z_t' v_t}{\sigma_v \sqrt{\Pi' Z_t' Z_t \Pi}}\right) + \sigma_v^2 \left(\frac{v_t' P_{Z_t} v_t}{\sigma_v^2}\right) \\
&= \sigma_v^2 \lambda_t^2 + 2\sigma_v^2 \lambda_t z_v + \sigma_v^2 s_{vv} \\
&= \sigma_v^2 \left(\lambda_t^2 + 2\lambda_t z_v + s_{tvv}\right)
\end{aligned}$$

where

$$z_{tu} = \frac{\Pi' Z_t' u_t}{\sigma_u \sqrt{\Pi' Z_t' Z_t \Pi}},$$

$$z_{tv} = \frac{\Pi' Z_t' v_t}{\sigma_v \sqrt{\Pi' Z_t' Z_t \Pi}},$$

$$s_{tuv}=\frac{v_t' P_{Z_t} u_t}{\sigma_v \sigma_u},$$

and

$$s_{tvv}=\frac{v_t' P_{Z_t} v_t}{\sigma_v^2}.$$

As stated in Rothenberg (1984), the $(z_{tu}, z_{tv})'$ is bivariate normal with zero means, unit variances, and correlation coefficient ρ. The random variable s_{tuv} has mean $k\rho$ and variance $k\left(1+\rho^2\right)$ and s_{tvv} has mean k and variance $2k$.

It is clear that

$$\begin{aligned}
\zeta_{1tn} &= \sigma_v \sigma_u \left(\lambda_t z_{tu} + s_{tuv}\right) \\
&= \sigma_v \sigma_u \left(\lambda z_{tu} + s_{tuv}\right) + o_p(1) \\
&= \zeta_{1t} + o_p(1)
\end{aligned}$$

and

$$\begin{aligned}
\zeta_{2tn} &= \sigma_v^2 \left(\lambda_t^2 + 2\lambda_t z_{tv} + s_{tvv}\right) \\
&= \sigma_v^2 \left(\lambda^2 + 2\lambda z_{tv} + s_{tvv}\right) + o_p(1) \\
&= \zeta_{2t} + o_p(1)
\end{aligned}$$

as $n \to \infty$ where $\zeta_{1t} = \sigma_v \sigma_u \left(\lambda z_{tu} + s_{tuv}\right)$ and $\zeta_{2t} = \sigma_v^2 \left(\lambda^2 + 2\lambda z_{tv} + s_{tvv}\right)$.

Consider (3). As $T \to \infty$,

$$\xi_{1Tn} = \frac{1}{T}\sum_{t=1}^{T} \zeta_{1tn} \xrightarrow{p} E\left(\zeta_{1tn}\right) = E\left(\sigma_v \sigma_u \left(\lambda z_{tu} + s_{tuv}\right)\right) = \sigma_v \sigma_u \rho k$$

by a law of large numbers (LLN) and $E\left(z_{tu}\right)=0$ and $E\left(s_{tuv}\right)=\rho k$.

Similarly,

$$\xi_{2Tn} = \frac{1}{T}\sum_{t=1}^{T} \zeta_{2tn} \xrightarrow{p} E\left(\zeta_{2tn}\right) = E\left[\sigma_v^2 \left(\lambda^2 + 2\lambda z_{tv} + s_{tvv}\right)\right] = \sigma_v^2 \left(\lambda^2 + k\right)$$

proving (4). ∎

Proof of Theorem 1

Proof. For (1), recall that $\hat{\beta}_{2SLS} - \beta = \frac{\xi_{1Tn}}{\xi_{2Tn}}$. Using the Lemmas 1.3 and 1.4, we obtain

$$\hat{\beta}_{2SLS} = \beta + \frac{\xi_{1Tn}}{\xi_{2Tn}} \xrightarrow{p} \beta + \frac{\sigma_v \sigma_u \rho k}{\sigma_v^2 \left(\lambda^2 + k\right)} = \beta + \frac{\sigma_u \rho k}{\sigma_v \left(\lambda^2 + k\right)}$$

as $(n,T) \overset{\text{seq}}{\to} \infty$.

Consider (2). First we write

$$v_t = E(v_t|u_t) + \eta_t = \frac{\sigma_{uv}}{\sigma_u^2} u_t + \eta_t$$

where η_t and u_t are independent. Then

$$\begin{aligned} & Var(\zeta_{1t}) \\ = \; & Var(\sigma_v \sigma_u (\lambda z_{tu} + s_{tuv})) \\ = \; & \sigma_v^2 \sigma_u^2 \left[\lambda^2 Var(z_{tu}) + Var(s_{tuv}) + 2Cov(z_{tu}, s_{tuv})\right] \\ = \; & \sigma_v^2 \sigma_u^2 \left[\lambda^2 + k\left(1+\rho^2\right)\right] \end{aligned}$$

because $Var(z_{tu}) = 1$, $Var(s_{tuv}) = k\left(1+\rho^2\right)$, and

$$\begin{aligned} & Cov(z_{tu}, s_{tuv}) \\ = \; & E(z_{tu} s_{tuv}) - E(z_{tu}) E(s_{tuv}) \\ = \; & E\left(z_{tu} \frac{v_t' P_{Z_t} u_t}{\sigma_v \sigma_u}\right) \\ = \; & E\left[z_{tu}\left(\rho \frac{u_t' P_{Z_t} u_t}{\sigma_u^2} + \frac{\eta_t' P_{Z_t} u_t}{\sigma_v \sigma_u}\right)\right] \\ = \; & E\left[\frac{\Pi' Z_t' u_t}{\sigma_u \sqrt{\Pi' Z_t' Z_t \Pi}} \left(\rho \frac{u_t' P_{Z_t} u_t}{\sigma_u^2}\right)\right] + E\left[\frac{\Pi' Z_t' u_t}{\sigma_u \sqrt{\Pi' Z_t' Z_t \Pi}} \left(\frac{\eta_t' P_{Z_t} u_t}{\sigma_v \sigma_u}\right)\right] \\ = \; & 0 \end{aligned}$$

since

$$E\left[E\left[\frac{\Pi' Z_t' u_t}{\sigma_u \sqrt{\Pi' Z_t' Z_t \Pi}} \left(\frac{\eta_t' P_{Z_t} u_t}{\sigma_v \sigma_u}\right)|u_t\right]\right] = 0$$

since $E(\eta_t|u_t) = 0$ and

$$E\left[\frac{\Pi' Z_t' u_t}{\sigma_u \sqrt{\Pi' Z_t' Z_t \Pi}} \left(\rho \frac{u_t' P_{Z_t} u_t}{\sigma_u^2}\right)\right] = 0$$

e.g., Corollary 10.9.2 in Graybill (1983).

By a central limit theorem we have

$$\sqrt{T}(\xi_{1T} - \sigma_v \sigma_u \rho k) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} (\zeta_{1t} - \sigma_v \sigma_u \rho k) \xrightarrow{d} N\left(0, \sigma_v^2 \sigma_u^2 \left[\lambda^2 + \lambda\left(1+\rho^2\right)\right]\right)$$

as $T \to \infty$.

Similarly,

$$\xi_{2Tn} = \frac{1}{T} \sum_{t=1}^{T} \zeta_{2tn} \xrightarrow{p} E(\zeta_{2tn}) = E\left[\sigma_v^2\left(\lambda^2 + 2\lambda z_{tv} + S_{tvv}\right)\right] = \sigma_v^2\left(\lambda^2 + k\right)$$

by a LLN. Then we have

$$\begin{aligned}
\sqrt{T}\left(\hat{\beta}_{2SLS}-\beta-\frac{\sigma_v\sigma_u\rho k}{\sigma_v^2\left(\lambda^2+k\right)}\right) &= \frac{\sqrt{T}\left[\xi_{1Tn}-\sigma_v\sigma_u\rho k\right]}{\xi_{2Tn}} \\
&\xrightarrow{d} \frac{N\left(0,\sigma_v^2\sigma_u^2\left[\lambda^2+k\left(1+\rho^2\right)\right]\right)}{\sigma_v^2\left(\lambda^2+k\right)} \\
&= N\left(0,\frac{\sigma_u^2\left[\lambda^2+k\left(1+\rho^2\right)\right]}{\sigma_v^2\left(\lambda^2+k\right)^2}\right)
\end{aligned}$$

$(n,T)\overset{\text{seq}}{\rightarrow}\infty$ proving the theorem. ■

Since $\sigma_u, \sigma_v, \lambda^2$ and k are all positive, the sign of the bias term of panel 2SLS, $\hat{\beta}_{2SLS}$, is determined by ρ which is the sample correlation of disturbance terms u_t and v_t. When $\rho = 0$, i.e., Y_t is uncorrelated with u_t, $\hat{\beta}_{2SLS}$ becomes consistent; when ρ has the same sign as β, $\hat{\beta}_{2SLS}$ is overestimated; when ρ has a different sign from β, $\hat{\beta}_{2SLS}$ is underestimated.

Once σ_u, σ_v, ρ and λ^2 are consistently estimated, the bias can be corrected by using a bias-corrected estimator. For example, the bias-corrected estimator can be constructed as

$$\tilde{\beta}_{BC2SLS} = \hat{\beta}_{2SLS} - \frac{\widehat{\sigma}_u\widehat{\rho}k}{\widehat{\sigma}_v\left(\widehat{\lambda}^2+k\right)} \tag{4}$$

where $\hat{\sigma}_u, \hat{\sigma}_v$ and $\hat{\rho}$ can be estimated from the residuals $\hat{u}_{it}$ of 2SLS regression and $\hat{v}_{it}$ of the first stage regression. A consistent estimator of λ can be constructed by following Moon and Phillips (2000).

In the cross-sectional case, when the concentration parameter stays constant as the sample size grows, the signal of the model is too weak comparing to the noise. Hence the model is weakly identified, i.e., the two stage least square estimator is inconsistent, and more importantly, 2SLS converge to a random variable. However, in the panel set-up, if time series dimension is large, the weak signal can be strengthened by repeating regression across the time series dimension. It is, in spirit, similar to the argument of establishing the consistency in the panel spurious regression, e.g., Phillips and Moon (1999) and Kao (1999).

References

[1] Graybill, F. A. (1983) *Matrices with Applications in Statistics.* Wadsworth.

[2] Kao, C. (1999) Spurious Regression and Residual-Based Tests for Cointegration in Panel Data. *Journal of Econometrics* 90, 1-44.

[3] Moon, H. R., and P. C. B. Phillips (2000) Estimation of Autoregressive Roots Near Unity Using Panel Data. *Econometric Theory* 16, 927-997.

[4] Phillips, P. C. B., and H. R. Moon (1999) Linear Regression Limit Theory for Nonstationary Panel Data. *Econometrica* 67, 1057-1111.

[5] Rothenberg, T. J. (1984) Approximating the Distributions of Econometric Estimators of Test Statistics, in *Handbook of Econometrics.* Vol. II, ed. by Z. Griliches and M.D. Intriligator. Amsterdam: North-Holland, 881-936.

[6] Staiger, D., and J. H. Stock (1997) Instrumental Variables Regression with Weak Instruments. *Econometrica* 65, 557-587.

[7] Stock, H. J., H. J. Wright, and M. Yogo (2002) A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments. *Journal of Business and Economic Statistics* 20, 518-29.