Discontinuities of Weak Instrument Limiting Distributions

Jinyong Hahn* Guido Kuersteiner†

March 11, 2005

Abstract

We consider two stage least squares (2SLS) estimators of a simple simultaneous equations model when identification fails asymptotically. We investigate how the limiting distribution of the estimator changes as we vary our parametrization to allow for increasing degrees of nonidentification.

Keywords: weak instruments, two stage least squares, near non-identification

JEL: C13,C31

In this note we consider a simple simultaneous equations model under conditions that imply asymptotically vanishing identification. Staiger and Stock (1997) considered models of this type under a specific specification of weak instruments. We generalize their specification to a continuum of parametrizations under which identification becomes weak, thus allowing for varying degrees of weakness. We analyze the limiting distribution of the 2SLS estimator for our model and describe how it changes as we change our assumptions about the severity of the identification problem.

*Brown University, Dept. of Economics, Box B, Providence, RI 02912, email:Jinyong_Hahn@brown.edu

†Corresponding Author: MIT, Dept. of Economics, 50 Memorial Drive E52-371A, Cambridge, MA 02142, phone: 617 253 2118, fax: 617 253 1330, email:gkuerste@mit.edu

We consider the following model

$$\begin{aligned} y_i &= x_i \cdot \beta + \varepsilon_i \\ x_i &= z_i'\pi + v_i = z_i'\left(n^{-\delta}\mu\right) + v_i \end{aligned}$$

where

$$0 < \delta < \infty$$

and μ is a $k \times 1$ vector of nonstochastic constants. Staiger and Stock (1997) considered the case where $\delta = \frac{1}{2}$ to examine properties of IV estimators with weak instruments. Because we allow for $\delta < 1/2$ our instruments are stronger in that case and we call the instruments "nearly weak". If $\delta > 1/2$ then the instruments are weaker than in the case considered by Staiger and Stock and we refer to that situation as the "near-nonidentified" case.

We observe a sample $\{x_i, y_i, z_i\}_{i=1}^n$ and organize the observations as $y = (y_1, ..., y_n)'$, $x = (x_1, ..., x_n)'$ and $z = (z_1', ..., z_n')$ such that x and y are $n \times 1$ vectors and z is a $n \times k$ matrix. For simplicity, we will assume that z_i are nonstochastic such that $n^{-1}z'z$ is fixed at M, and (ε_i, v_i) is i.i.d. bivariate normal with zero mean.

1 First Order Asymptotics

Consider the distribution of 2SLS

$$\begin{aligned} b &= \frac{x'Py}{x'Px} = \beta + \frac{x'P\varepsilon}{x'Px} \\ &= \beta + \frac{x'z\,(z'z)^{-1}\,z'\varepsilon}{x'z\,(z'z)^{-1}\,z'x} \end{aligned} \tag{1}$$

Here, $P = z\,(z'z)^{-1}\,z$ denotes the usual projection matrix. We need to consider three components to analyze the asymptotic distribution of b: $x'z$, $z'z$, and $z'\varepsilon$. Define the following quantities

$$\left(Z_\varepsilon', Z_v'\right)' \equiv \left(n^{-1/2}\varepsilon' z, n^{-1/2}\sum_{i=1}^n z_i' v_i\right)' \sim N\left(0, \Sigma \otimes M\right)$$

with

$$\Sigma = \begin{bmatrix} \sigma_\varepsilon^2 & \sigma_{\varepsilon v} \\ \sigma_{\varepsilon v} & \sigma_v^2 \end{bmatrix}.$$

Lemma 1 *For* $0 < \delta < 1/2$

$$n^{-1+\delta} z' x \sim M\mu + n^{-1/2+\delta} \cdot Z_v = M\mu + O_p\left(n^{-1/2+\delta}\right),$$

for $\delta = 1/2$

$$n^{-1/2} z' x \sim M\mu + Z_v = M\mu + O_p(1),$$

and for $1/2 < \delta < \infty$

$$n^{-1/2} z' x \sim O\left(n^{-\delta+1/2}\right) + Z_v = N\left(0, \sigma_v^2 M\right) + o(1).$$

Proof. For $0 < \delta < 1/2$, the assertion can be proved by noting that

$$\begin{aligned} n^{-1+\delta} z' x &= n^{-1+\delta} \sum_{i=1}^n z_i x_i = n^{-1+\delta} n^{-\delta} \left(\sum_{i=1}^n z_i z_i' \right) \mu + n^{-1+\delta} \sum_{i=1}^n z_i v_i \\ &= M\mu + n^{-1/2+\delta} \left(n^{-1/2} \sum_{i=1}^n z_i v_i \right). \end{aligned}$$

Similarly, note that for $1/2 < \delta < \infty$ we have

$$\begin{aligned} n^{-1/2} z' x &= n^{-1/2} \sum_{i=1}^n z_i x_i = n^{-1} n^{-\delta+1/2} \left(\sum_{i=1}^n z_i z_i' \right) \mu + n^{-1/2} \sum_{i=1}^n z_i v_i \\ &= n^{-\delta+1/2} M\mu + n^{-1/2} \sum_{i=1}^n z_i v_i. \end{aligned}$$

■

Lemma 2

$$n^{-1/2} z' \varepsilon = Z_\varepsilon = O_p(1)$$

Using the preceding lemmas, for the case of nearly weak instruments we can conclude that

$$n^{1/2-\delta}(b-\beta)=\frac{\left(n^{-1+\delta}x'z\right)\left(n^{-1}z'z\right)^{-1}\left(n^{-1/2}z'\varepsilon\right)}{\left(n^{-1+\delta}x'z\right)\left(n^{-1}z'z\right)^{-1}\left(n^{-1+\delta}z'x\right)}=O_p(1)$$

This makes sense because it gives a continuity in the rate of convergence as $\delta\uparrow\frac{1}{2}$. As $\delta\uparrow\frac{1}{2}$, we would like to have the situation where $b-\beta$ approaches $O_p(1)$ in order to maintain some continuity to Staiger and Stock's (1997) analysis. On the other hand, when $\delta\geq 1/2$ we find by applying the lemmas and the same argument as before that

$$b-\beta=O_p(1).$$

In other words for all fixed values of $\delta\geq 1/2$ the 2SLS estimator is inconsistent and as we will show later has some limiting distribution.

We now analyze the asymptotic distribution itself. Let $Z_1=\sigma_v^{-1}M^{-1/2}Z_v$, $Z_2=\sigma_\varepsilon^{-1}M^{-1/2}Z_\varepsilon$, and $\bar{\mu}=\sigma_v^{-1}M^{1/2}\mu$. Again, the preceding lemmas imply the following result.

Theorem 1 *Let b be defined in (1). For $0<\delta<1/2$, it follows that*

$$n^{1/2-\delta}(b-\beta)\rightarrow N\left(0,\sigma_\varepsilon^2\left(\mu'M\mu\right)^{-1}\right).$$

For $\delta=1/2$,

$$b-\beta\rightarrow\frac{\sigma_\varepsilon}{\sigma_v}\frac{\left(\bar{\mu}+Z_1\right)'Z_2}{\left(\bar{\mu}+Z_1\right)'\left(\bar{\mu}+Z_1\right)},$$

and for $\delta>1/2$,

$$b-\beta\rightarrow\frac{\sigma_\varepsilon}{\sigma_v}\frac{Z_1'Z_2}{Z_1'Z_1}.$$

In other words, nearly weak IV asymptotics roughly predict that

$$b-\beta\approx N\left(0,\frac{1}{n}\sigma_\varepsilon^2\left(\pi'M\pi\right)^{-1}\right),$$

which is the same prediction as implied by the "usual" first order asymptotic theory. An immediate consequence of this result is that the nearly weak limit distribution does not reflect the type of finite sample moments usually associated with the 2SLS estimator while it was shown by Chao and Swanson (2000) that the weak instrument limit of Staiger and Stock (1997) preserves the exact finite sample moments of 2SLS under the conditions imposed in this paper.

Finally, for the near-nonidentified case some of the features of the finite sample moments such as the dependence of the mean squared error and bias on the correlation between ε and v are preserved while others such as the dependence of the bias on the number of overidentifying restrictions are lost.

The limiting distributions near the point of nonidentification are therefore discontinuous in the sense that their implied moments change discontinuously when $\delta \uparrow 1/2$ and $\delta \downarrow 1/2$. While the nearly weak limit does not seem to be of much interest in capturing small sample distortions of 2SLS the distinction between $\delta = 1/2$ and $\delta > 1/2$ is more delicate and the relevance of the corresponding limit distribution for the actual small sample distribution in general will depend on sample size and parametrization of the model.

Testable implications of the weak instrument and near-nonidentification limits can be obtained by considering the first two moments of the respective limit distributions.

Theorem 2 *Let* $\psi_1 = \frac{\sigma_\varepsilon}{\sigma_v}\left(\bar{\mu} + Z_1\right)' Z_2 / \left(\bar{\mu} + Z_1\right)'\left(\bar{\mu} + Z_1\right)$ *and* $\psi_2 = \frac{\sigma_\varepsilon}{\sigma_v} Z_1' Z_2 / Z_1' Z_1$*. Then*

$$
\begin{aligned}
E\left[\psi_1\right] &= \sigma_{\varepsilon v}\sigma_v^{-2} e^{-\bar{\mu}'\bar{\mu}/2}\ {}_1F_1\left(\frac{k}{2}-1, \frac{k}{2}, \frac{\bar{\mu}'\bar{\mu}}{2}\right), \\
E\left[\psi_1^2\right] &= \sigma_\varepsilon^{-1}\sigma_{\varepsilon v}^2\sigma_v^{-3}\left(\frac{\sigma_\varepsilon^2\sigma_v^2}{\sigma_{\varepsilon v}^2\left(k-2\right)}\ {}_1F_1\left(\frac{k}{2}-1, \frac{k}{2}, \frac{\bar{\mu}'\bar{\mu}}{2}\right) + \frac{k-3}{k-2}\ {}_1F_1\left(\frac{k}{2}-2, \frac{k}{2}, \frac{\bar{\mu}'\bar{\mu}}{2}\right)\right)
\end{aligned}
$$

where ${}_1F_1(a, b, x) = \sum_{j=0}^{\infty}(a)_j/(b)_j x^j/j!$ *is the confluent hypergeometric function with* $(a)_j =$

$\Gamma(a+j)/\Gamma(a)$ *and*

$$
\begin{aligned}
E\left[\psi_2\right] &= \sigma_{\varepsilon v}\sigma_v^{-2}, \\
E\left[\psi_2^2\right] &= \frac{\sigma_\varepsilon \sigma_v^{-1}}{k-2} + \sigma_\varepsilon^{-1}\sigma_{\varepsilon v}^2 \sigma_v^{-3}\frac{k-3}{k-2}.
\end{aligned}
$$

Proof. For a proof of the first result see Chao and Swanson (2000). The second result follows from $\exp(0) = 1$ and ${}_1F_1(.,.,0) = 1$. ■

The theorem reveals a discontinuity in the limiting distribution when $\delta \downarrow 1/2$ with μ fixed. At the same time, and for a given sample size n, one can represent the degree of non-identification by an appropriate choice of $\bar{\mu}'\bar{\mu}$. In other words, as $\bar{\mu}'\bar{\mu} \to 0$ the moments of ψ_1 predict the same behavior of the 2SLS as under the near-nonidentified limit distribution, namely that the bias becomes insensitive to the number of instruments.

2 Higher Order Asymptotics

The results in the previous section show that the near weak instrument approximation does not lead to asymptotic distributions that differ in their predictions form the usual first order asymptotics with strong identification. In particular the asymptotic bias of the estimator is zero. We now consider "higher order" asymptotics to investigate if the near weak instrument case leads to different higher order approximations to bias and mean squared error.

For this purpose we write

$$
\begin{aligned}
\left(n^{-1+\delta}x'z\right)\left(n^{-1}z'z\right)^{-1}\left(n^{-1/2}z'\varepsilon\right) &= \left(M\mu + n^{-\gamma}Z_v\right)' M^{-1} Z_\varepsilon \\
&= \mu' Z_\varepsilon + n^{-\gamma} Z_v' M^{-1} Z_\varepsilon
\end{aligned}
$$

and

$$\begin{aligned} \left(n^{-1+\delta}x'z\right)\left(n^{-1}z'z\right)^{-1}\left(n^{-1+\delta}z'x\right) &= \left(M\mu+n^{-\gamma}Z_v\right)' M^{-1}\left(M\mu+n^{-\gamma}Z_v\right) \\ &= \mu' M\mu+2n^{-\gamma}\mu' Z_v+n^{-2\gamma}Z_v' M^{-1}Z_v \end{aligned}$$

where

$$\gamma \equiv \frac{1}{2}-\delta$$

denotes the rate of convergence of 2SLS. Let

$$\begin{aligned} \Lambda_1 &= \mu' Z_\varepsilon \\ \Lambda_2 &= \mu' M\mu \\ a &= Z_v' M^{-1} Z_\varepsilon \\ b_1 &= 2\mu' Z_v \\ b_2 &= Z_v' M^{-1} Z_v \\ s &= n^{-\gamma} \end{aligned}$$

We may then write $n^\gamma \left(b-\beta\right)$ as

$$\frac{\Lambda_1+as}{\Lambda_2+b_1 s+b_2 s^2},$$

which has a power series expansion

$$\frac{\Lambda_1}{\Lambda_2}+\left(\frac{a}{\Lambda_2}-\frac{\Lambda_1 b_1}{\Lambda_2^2}\right)s+\left(-\frac{\Lambda_1 b_2}{\Lambda_2^2}-\frac{b_1 a}{\Lambda_2^2}+\frac{b_1^2\Lambda_1}{\Lambda_2^3}\right)s^2+O\left(s^3\right)$$

Therefore, we have

$$\begin{aligned}
& \frac{\mu' Z_\varepsilon + n^{-\gamma} Z_v' M^{-1} Z_\varepsilon}{\mu' M \mu + 2n^{-\gamma} \mu' Z_v + n^{-2\gamma} Z_v' M^{-1} Z_v} \\
= & \frac{\mu' Z_\varepsilon}{\mu' M \mu} \\
& + n^{-\gamma} \left(\frac{Z_v' M^{-1} Z_\varepsilon}{\mu' M \mu} - 2 \frac{(\mu' Z_\varepsilon)(\mu' Z_v)}{(\mu' M \mu)^2} \right) \\
& + n^{-2\gamma} \left(- \frac{(\mu' Z_\varepsilon)(Z_v' M^{-1} Z_v)}{(\mu' M \mu)^2} - \frac{(\mu' Z_v)(Z_v' M^{-1} Z_\varepsilon)}{(\mu' M \mu)^2} + \frac{(\mu' Z_v)^2 (\mu' Z_\varepsilon)}{(\mu' M \mu)^3} \right) \\
& + O_p\left(n^{-3\gamma}\right)
\end{aligned}$$

We therefore make the approximation

$$\begin{aligned}
& n^\gamma (b - \beta) \\
\approx & \frac{\mu' Z_\varepsilon}{\mu' M \mu} \\
& + n^{-\gamma} \left(\frac{Z_v' M^{-1} Z_\varepsilon}{\mu' M \mu} - 2 \frac{(\mu' Z_\varepsilon)(\mu' Z_v)}{(\mu' M \mu)^2} \right) \\
& + n^{-2\gamma} \left(- \frac{(\mu' Z_\varepsilon)(Z_v' M^{-1} Z_v)}{(\mu' M \mu)^2} - \frac{(\mu' Z_v)(Z_v' M^{-1} Z_\varepsilon)}{(\mu' M \mu)^2} + \frac{(\mu' Z_v)^2 (\mu' Z_\varepsilon)}{(\mu' M \mu)^3} \right)
\end{aligned}$$

Because the third moments of a normal vector are zero, we can say that

$$\begin{aligned}
E\left[n^\gamma (b - \beta)\right] \approx & E\left[\frac{\mu' Z_\varepsilon}{\mu' M \mu}\right] \\
& + n^{-\gamma} \left(E\left[\frac{Z_v' M^{-1} Z_\varepsilon}{\mu' M \mu}\right] - 2E\left[\frac{(\mu' Z_\varepsilon)(\mu' Z_v)}{(\mu' M \mu)^2}\right] \right)
\end{aligned}$$

But

$$\begin{aligned}
E\left[\frac{\mu' Z_\varepsilon}{\mu' M \mu}\right] &= 0 \\
E\left[\frac{Z_v' M^{-1} Z_\varepsilon}{\mu' M \mu}\right] &= \frac{K \sigma_{\varepsilon v}}{\mu' M \mu} \\
E\left[\frac{(\mu' Z_\varepsilon)(\mu' Z_v)}{(\mu' M \mu)^2}\right] &= \frac{2\sigma_{\varepsilon v}}{\mu' M \mu}
\end{aligned}$$

we have

$$E\left[n^\gamma (b - \beta)\right] \approx n^{-\gamma} \frac{(K-2)\sigma_{\varepsilon v}}{\mu' M \mu}$$

Therefore, the higher order asymptotics roughly predicts that

$$\begin{aligned}
E\left[b\right] &\approx \beta+n^{-2\gamma}\frac{\left(K-2\right)\sigma_{\varepsilon v}}{\mu' M\mu}=\beta+n^{-1}\frac{\left(K-2\right)\sigma_{\varepsilon v}}{\left(n^{-\delta}\mu\right)' M\left(n^{-\delta}\mu\right)} \\
&= \beta+n^{-1}\frac{\left(K-2\right)\sigma_{\varepsilon v}}{\pi' M\pi},
\end{aligned}$$

which is the same prediction coming from the "usual" higher order asymptotics. We therefore conclude that the near weak IV asymptotics is qualitatively the same as the standard asymptotics. (We skipped the mean squared error calculation, but it is expected to yield the same qualitative conclusion.)

References

[1] Chao, J., and N. Swanson, 2000, Bias and MSE of the IV Estimator under Weak Identification, unpublished working paper

[2] Staiger, D., and J. H. Stock, 1997, Instrumental Variables Regressions with Weak Instruments, Econometrica 65, 557 - 586.