The Asymptotic Distribution of Nagar's Bias-Adjusted TSLS Estimator under Partial Identification

Giovanni Forchini[1][2]

May 23, 2008

Abstract

The asymptotic distribution of the Nagar bias-adjusted two-stage-least-squares estimator is studied under the assumption of partial identification when the number of instruments increases at the same rate as the sample size. We find that the estimator of the identified parameters is consistent but has a non-standard asymptotic distribution and the estimator of the unidentified parameters has a non-degenerated distribution. The results have the same structure as those of Phillips (1989) and Choi and Phillips (1992) for the two stage least squares estimators for a fixed number of instruments.

1. Introduction

Recent literature on structural equations modelling has showed interest in Nagar (1959)'s bias-adjusted two-stage-least-squares (BATSLS) estimators (e.g., among others, Donald and Newey (2001), Hahn, Hausman and Kuersteiner (2004), Andrews, Moreira and Stock (2007)). It is easy to compute, consistent and asymptotically equivalent to the two-stage-least-squares (TSLS) estimator under standard asymptotics. In contrast to the TSLS estimator, it is also consistent when the number of instruments grows at the same rate as the sample size.

Partial identification in the sense of Phillips (1989) and Choi and Phillips (1992) can, in some cases, drastically affect the distribution of the estimator of interest in fundamental ways. Precisely, Phillips (1989) and Choi and Phillips (1992) show that the TSLS estimator of the identified parameters is consistent but has a non-standard asymptotic distribution while the estimator of the unidentified parameters converges in probability to a non-degenerated distribution. Forchini (2008) shows that, for a fixed number of instruments, the asymptotic distribution of the limited information maximum likelihood (LIML) estimator of both the

identified and the unidentified structural parameters does not have integer moments even if the LIML estimator is asymptotically normal when all the structural parameters are identified.

This paper derives the asymptotic distribution of the BATSLS estimator under partial identification when the number of instruments grows at the same rate as the sample size. We find that

1. The BATSLS estimator of the identified parameters is consistent but has a complicated distribution that is mean- and covariance matrix- mixed normal;

2. The estimator of the unidentified parameters converges in probability to a non-degenerate distribution which is covariance matrix-mixed normal and depends on the identified parameters; and

3. The asymptotic distribution of the BATSLS estimator of both the identified and the unidentified parameters are very similar to the asymptotic distribution of the TSLS estimator for a fixed number of instruments as derived by Phillips (1989) and Choi and Phillips (1992).

The structure of the paper is as follows. Section 2 specifies the model and some preliminary results. Section 3 deals with the consistency and the asymptotic distribution for the BATSLS estimator of both the identified and the unidentified coefficients. Section 4 concludes. Proofs are in the appendix. We use the same notation as Phillips (1989) and Choi and Phillips (1992) whereby $\Rightarrow$, $\rightarrow^{P}$ and $\equiv$ denote, respectively, weak convergence, convergence in probability and equality in distribution.

2. The model and the assumptions

Consider a linear structural equation of the form

(1)
$$\underset{(T\times 1)}{y} = \underset{(T\times n_1)}{X_1}\, b_1 + \underset{(T\times n_2)}{X_2}\, b_2 + u$$

with corresponding reduced form

(2)
$$(y, X_1, X_2) = \underset{(T\times k)}{Z}\left(\Pi b_1, \underset{(k\times n_1)}{\Pi}, \underset{(k\times n_2)}{0}\right) + (v, V_1, V_2).$$

Notice that the over-identifying restrictions have already been imposed on the reduced form, and that b_1 is regarded as being identified and b_2 as being unidentified.

The Nagar's BATSLS has the form

(3)
$$\begin{pmatrix} \hat{b}_1 \\ \hat{b}_2 \end{pmatrix} = \left[(X_1, X_2)'(P_Z - \hat{a}I_T)(X_1, X_2)\right]^{-1}\left[(X_1, X_2)'(P_Z - \hat{a}I_T)y\right],$$

where $\hat{a} = k/T$ and $P_Z = Z(Z'Z)^{-1}Z'$. Donald and Newey (2001) have suggested a similar modification of the TSLS estimator with $\hat{a} = (k - n - 1)/T$. Since the term $(n+1)/T$ tends

to zero as T tends to infinity, Nagar's and Donald and Newey's estimators must be asymptotically equivalent when T is large under both classical and large number of instruments asymptotics.

In order to simplify some results later on we let $M_Z = I_T - P_Z$ and

$$\hat{A} = \frac{1}{T}(y, X_1, X_2)'(P_Z - \hat{a} I_T)(y, X_1, X_2)$$

(4)

$$= \hat{a}(1 - \hat{a})(W - S),$$

where $W = (1/k)(y, X_1, X_2)' P_Z (y, X_1, X_2)$ and $S = (1/(T - k))(y, X_1, X_2)' M_Z (y, X_1, X_2)$. Partitioning $\hat{A}$ as

(5)

$$\hat{A} = \begin{pmatrix} \underset{(1\times1)}{\hat{a}} & \hat{b}_1' & \hat{b}_2' \\ \underset{(n_1\times1)}{\hat{b}_1} & \underset{(n_1\times n_1)}{\hat{A}_{11}} & \hat{A}_{21}' \\ \underset{(n_2\times1)}{\hat{b}_2} & \underset{(n_2\times n_1)}{\hat{A}_{21}} & \underset{(n_2\times n_2)}{\hat{A}_{22}} \end{pmatrix},$$

after some straightforward algebra one finds that the BATSLS estimators for the identified and the unidentified parameters are

(6)

$$\hat{b}_1 = \left(\hat{A}_{11} - \hat{A}_{21}' \hat{A}_{22}^{-1} \hat{A}_{21}\right)^{-1} \left(\hat{b}_1 - \hat{A}_{21}' \hat{A}_{22}^{-1} \hat{b}_2\right)$$

$$\hat{b}_2 = \left(\hat{A}_{22} - \hat{A}_{21} \hat{A}_{11}^{-1} \hat{A}_{21}'\right)^{-1} \left(\hat{b}_2 - \hat{A}_{21} \hat{A}_{11}^{-1} \hat{b}_1\right).$$

We make the following assumptions:

Assumptions

a) $\hat{a} = k/T \to a$ *where* $0 < a < 1$;

b) $(1/T)\Pi' Z' Z \Pi \to^P Q$ *where* Q *is a positive definite matrix of dimension* $(n_1 \times n_1)$.

c)   *The matrix* Z *has full column rank* k.

d)   *The rows of* (v, V_1, V_2) *are i.i.d. normal with mean vector zero and covariance matrix* Ω;

e) $\Omega = I_{n+1}, n = n_1 + n_2$.

Assumptions *a*, *b* and *c* are fairly standard. Assumption *a* allows the number of instruments to grow at the same rate as the sample size. Assumption *b* requires the matrix $(1/T)\Pi' Z' Z \Pi$ to converge as T increases and assumption *c* states that the ordinary least squares estimator of the reduced form parameters exists for any sample size.

Assumption *d* requires the reduced form error to be i.i.d. normal as in Bekker (1994). Essentially it allows us to derive asymptotic results that depend only on the mean of (y, X_1, X_2) and Ω. Violations of this assumption would make the covariance matrices of the

asymptotic distributions presented below depend on the fourth moments of the reduced form errors.

Assumption e is just a normalization as in Phillips (1989) and Choi and Phillips (1992), and requires that standardizing transformations and rotations of coordinates to isolate identified and unidentified parameters have already been performed.

Given Assumptions d and e, $(v, V_1, V_2): N(0, I_T \otimes I_{n+1})$ so that

$$(T-k)S: W_{n+1}(T-k, I_{n+1}), \quad kW: W_{n+1}\left(k, \begin{pmatrix} (b_1, I_{n_1})'\Pi'Z'Z\Pi(b_1, I_{n_1}) & 0 \\ 0 & 0 \end{pmatrix}\right) \text{ and they are}$$

independent.

3. The consistency and asymptotic distributions

As one would expect from existing results on the asymptotic properties of estimators in partially identified model, the BATSLS estimator of the identified parameters is consistent but that of the unidentified parameters converges in distribution to a non-standard non-degenerate distribution. Some intermediate results allow us to prove this.

Lemma 1. *If the assumptions a-e are satisfied then*

(7)
$$\hat{A} \to^P (1-a)\begin{pmatrix} b_1'Qb_1 & b_1'Q & 0 \\ Qb_1 & Q & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and

(8)

$$\sqrt{T}vec\left[\hat{A} - (1-\hat{a})\begin{pmatrix} (b_1, I_{n_1})'\left[\frac{1}{T}\Pi'Z'Z\Pi\right](b_1, I_{n_1}) & 0 \\ 0 & 0 \end{pmatrix}\right] \to^D N\left(0, a^2(1-a)\Sigma_1 + a(1-a)^2\Sigma_2\right)$$

where

$$\Sigma_1 = \left(I_{(n+1)^2} + K_{n+1}\right)\left(I_{n+1} \otimes I_{n+1}\right)$$

$$\Sigma_2 = \left(I_{(n+1)^2} + K_{n+1}\right)\left[\left(I_{n+1} \otimes I_{n+1}\right) + \frac{1}{a}I_{n+1} \otimes \left\{(b_1, I_{n_1}, 0)'Q(b_1, I_{n_1}, 0)\right\}\right.$$

$$\left. \frac{1}{a}\left\{(b_1, I_{n_1}, 0)'Q(b_1, I_{n_1}, 0)\right\} \otimes I_{n+1}\right]\right)$$

and K_{n+1} denotes the commutation matrix (e.g. Magnus and Neudecker (1988)).

The proof of Lemma 1 can be adapted to deal with non-normality provided conditions for the validity of a central limit theorem hold (i.e. the existence of the fourth moments of (v, V_1, V_2)). In this case one bounds from above the variances of S and W and shows that the bound tends to zero as the sample size grows. So, equation (7) would still hold but the covariance matrix in (8) would depend on the fourth moments of the rows of (v, V_1, V_2) and would have a more complex structure.

Consistency of Nagar's BATSLS estimator of the identified parameters follows from Lemma 1 and the continuous mapping theorem:

(9)
$$\hat{b}_1 = \left(\hat{A}_{11} - \hat{A}_{21}{}' \left(\sqrt{T}\hat{A}_{22} \right)^{-1} \left(\sqrt{T}\hat{A}_{21} \right) \right)^{-1} \left(\hat{b}_1 - \hat{A}_{21}{}' \left(\sqrt{T}\hat{A}_{22} \right)^{-1} \left(\sqrt{T}\hat{b}_2 \right) \right)$$
$$\rightarrow^P \left((1-a)Q \right)^{-1} \left((1-a)Qb_1 \right) = b_1.$$

For the estimator of the unidentified parameters one has

$$\hat{b}_2 = \left(\sqrt{T}A_{22} - \left(\sqrt{T}A_{21} \right) A_{11}^{-1} A_{21}{}' \right)^{-1} \left(\sqrt{T}b_2 - \left(\sqrt{T}A_{21} \right) A_{11}^{-1} b_1 \right)$$
$$= \left(\sqrt{T}A_{22} - \left(\sqrt{T}A_{21} \right) A_{11}^{-1} A_{21}{}' \right)^{-1} \left(\sqrt{T} \left(b_2 - A_{21}b_1 \right) - \left(\sqrt{T}A_{21} \right) \left(A_{11}^{-1} b_1 - b_1 \right) \right)$$
$$= \left(\sqrt{T}A_{22} \right)^{-1} \left(\sqrt{T}b_2 - \left(\sqrt{T}A_{21} \right) b_1 \right) + o_p(1)$$
$$= \left(\sqrt{T}A_{22} \right)^{-1} \sqrt{T} \left(b_2 - A_{21}b_1 \right) + o_p(1).$$

Lemma 1 implies that both $\sqrt{T}A_{22}$ and $\sqrt{T}\left(b_2 - A_{21}b_1 \right)$ have an asymptotic non-degenerated distribution so that $\hat{b}_2$ converges to a non-standard non-degenerate distribution. In order to find this distribution, one has to express (8) in a clearer form. What we need is the following result.

Lemma 2. *If the assumptions a-e are satisfied then*

$$(10)\ \sqrt{T} \begin{pmatrix} b_1 - A_{11}b_1 \\ b_2 - A_{21}b_1 \\ vec(A_{21}) \\ vec(A_{22}) \end{pmatrix} \rightarrow^D N \left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \underset{(n_1 \times n_1)}{(1-a)\Omega_{11}} & 0 & 0 & 0 \\ 0 & \underset{(n_2 \times n_2)}{a(1-a)\Omega_{22}} & a(1-a)\Omega_{32}{}' & 0 \\ 0 & \underset{(n_1 n_2 \times n_2)}{a(1-a)\Omega_{32}} & \underset{(n_1 n_2 \times n_1 n_2)}{(1-a)\Omega_{33}} & 0 \\ 0 & 0 & 0 & \underset{(n_2{}^2 \times n_2{}^2)}{a(1-a)\Omega_{44}} \end{pmatrix} \right)$$

where

$$\Omega_{11} = (1 + b_1'b_1)\left[a I_{n_1} + (1-a)Q\right] + ab_1b_1'$$

$$\Omega_{22} = (1 + b_1'b_1)I_{n_2}$$

(11)
$$\Omega_{33} = a\left(I_{n_1} \otimes I_{n_2}\right) + (1-a)\left(Q \otimes I_{n_2}\right)$$

$$\Omega_{32} = -\left(b_1 \otimes I_{n_2}\right)$$

$$\Omega_{44} = \left(I_{n_2^2} + K_{n_2}\right)\left(I_{n_2} \otimes I_{n_2}\right).$$

It follows from Lemma 2 that

$$\sqrt{T}\frac{1}{\sqrt{a(1-a)}}(b_2 - A_{21}b) \to^D N(0, \Omega_{22})$$

$$\sqrt{T}\frac{1}{\sqrt{a(1-a)}}vec(A_{22}) \to^D vec(M) \equiv N(0, \Omega_{44}),$$

where M is an $(n_2 \times n_2)$ matrix, and they are asymptotically independent. Therefore, since

$$\hat{b}_2 = \left(\sqrt{T}A_{22}\right)^{-1}\sqrt{T}(b_2 - A_{21}b_1) + o_p(1), \text{ we have}$$

$$\hat{b}_2 \mid \frac{1}{\sqrt{a(1-a)}}\sqrt{T}A_{22} \to^D \left[\frac{1}{\sqrt{a(1-a)}}\sqrt{T}A_{22}\right]^{-1} N(0, \Omega_{22})$$

so that removing the conditioning

(12)
$$\hat{b}_2 = \int N\left(0, (1 + b_1'b_1)M^{-2}\right)dM.$$

The estimator of the unidentified parameters has an asymptotic non-degenerate distribution that is a normal covariance-matrix mixture with singular mixing distribution. In the totally unidentified case these would not be there and $\hat{b}_2 = \int N(0, M^{-2})dM$ with $vec(M) \equiv N\left(0, (I_{n^2} + K_n)(I_n \otimes I_n)\right)$. Notice that (12) has the same structure as the density of the TSLS estimator of the unidentified parameters in Corollary 3.1(b) of Choi and Phillips (1992) with some obvious changes since in our case the number of instruments also tends to infinity. In contrast to the analogous results of Choi and Phillips (1992) the BATSLS estimator of the unidentified parameters depends on the identified parameters.

We now investigate the asymptotic distribution of the estimator of the identified parameters

$$\sqrt{T}\left(\hat{b}_1 - b_1\right) = \sqrt{T}\left(\hat{A}_{11} - \hat{A}_{21}{}'\hat{A}_{22}^{-1}\hat{A}_{21}\right)^{-1}\left(\hat{b}_1 - \hat{A}_{21}{}'\hat{A}_{22}^{-1}\hat{b}_2\right) - \sqrt{T}b_1$$

$$= \sqrt{T}\left(\hat{A}_{11} - \hat{A}_{21}{}'\hat{A}_{22}^{-1}\hat{A}_{21}\right)^{-1}\left(\hat{b}_1 - \hat{A}_{21}{}'\hat{A}_{22}^{-1}\hat{b}_2 - \hat{A}_{11}b_1 + \hat{A}_{21}{}'\hat{A}_{22}^{-1}\hat{A}_{21}b_1\right)$$

$$= \left(\hat{A}_{11} - \hat{A}_{21}{}'\left(\sqrt{T}\hat{A}_{22}\right)^{-1}\left(\sqrt{T}\hat{A}_{21}\right)\right)^{-1}\left(\sqrt{T}\left(\hat{b}_1 - \hat{A}_{11}b_1\right) - \left(\sqrt{T}\hat{A}_{21}\right){}'\left(\sqrt{T}\hat{A}_{22}\right)^{-1}\sqrt{T}\left(\hat{b}_2 - \hat{A}_{21}b_1\right)\right).$$

Using the continuous mapping theorem and Lemma 1

$$\hat{A}_{11} - \hat{A}_{21}{}'\left(\sqrt{T}\hat{A}_{22}\right)^{-1}\left(\sqrt{T}\hat{A}_{21}\right) \to^{P} (1-a)Q.$$

Using Lemma 2,

$$\sqrt{T}\left(\hat{b}_1 - \hat{A}_{11}b_1\right) \to^{D} N\left(0,(1-a)\Omega_{11}\right)$$

and is independent of everything else,

$$\sqrt{T}vec\left(A_{22}\right) \to^{D} \sqrt{a(1-a)}vec(M) \equiv \sqrt{a(1-a)}N\left(0,\Omega_{44}\right)$$

and is also independent of everything else. Finally,

$$\sqrt{T}\begin{pmatrix} b_2 - A_{21}b_1 \\ vec\left(A_{21}\right) \end{pmatrix} \to^{D} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} a(1-a)\Omega_{22} & a(1-a)\Omega_{32}{}' \\ a(1-a)\Omega_{32} & (1-a)\Omega_{33} \end{pmatrix}\right)$$

so that

$$\sqrt{T}\left(b_2 - A_{21}b_1\right) \mid \sqrt{T}vec\left(A_{21}\right) \to^{D} N\left(a\sqrt{1-a}\,\Omega_{32}{}'\Omega_{33}^{-1}vec(L), a(1-a)\Omega_{22.3}\right)$$

$$\sqrt{T}vec\left(A_{21}\right) \to^{D} (1-a)^{1/2}vec(L) \equiv (1-a)^{1/2}N\left(0,\Omega_{33}\right)$$

where $\Omega_{22.3} = \Omega_{22} - a\Omega_{32}{}'\Omega_{33}^{-1}\Omega_{32}$ and L is an $\left(n_2 \times n_1\right)$ matrix. Therefore

$$\sqrt{T}\left(\hat{b}_1 - b_1\right) \mid \left(\sqrt{T}\hat{A}_{21}\right),\left(\sqrt{T}\hat{A}_{22}\right) \to^{D}$$

$$(1-a)^{-1}Q^{-1}N\left(0,(1-a)\Omega_{11}\right)$$

$$- \frac{1}{(1-a)\sqrt{a}}Q^{-1}L'M^{-1}N\left(a\sqrt{1-a}\,\Omega_{32}{}'\Omega_{33}^{-1}vec(L), a(1-a)\Omega_{22.3}\right)$$

$$\equiv N\left(0, \frac{1}{1-a}Q^{-1}\Omega_{11}Q^{-1}\right)$$

$$- N\left(\sqrt{\frac{a}{1-a}}Q^{-1}L'M^{-1}\Omega_{32}{}'\Omega_{33}^{-1}vec(L), \frac{1}{1-a}Q^{-1}L'M^{-1}\Omega_{22.3}\left(M^{-1}\right)'LQ^{-1}\right).$$

Removing the conditioning, we see that $\sqrt{T}\left(\hat{b}_1 - b_1\right)$ has a very complicated non-standard asymptotic distribution that is mean- and covariance matrix- mixed normal and one of the mixing densities is singular

(13)

$$\sqrt{T}\left(\hat{b}_1 - b_1\right) \to N\left(0, \frac{1}{1-a}Q^{-1}\Omega_{11}Q^{-1}\right) -$$

$$\int N\left(\sqrt{\frac{a}{1-a}}Q^{-1}L'M^{-1}\Omega_{32}{}'\Omega_{33}{}^{-1}vec(L), \frac{1}{1-a}Q^{-1}L'M^{-1}\Omega_{22.3}\left(M^{-1}\right)'LQ^{-1}\right)dLdM.$$

Notice also that if we knew that b_2 were not identified we could simplify (1) and (2) to eliminate X_2 and consider

(14)
$$y = X_1 b_1 + u$$

with corresponding reduced form

(15)
$$\left(y, X_1\right) = Z\left(\Pi b_1, \Pi\right) + \left(v, V_1\right).$$

In this case the asymptotic distribution of the Nagar' BATSLS estimator is

(16)
$$\sqrt{T}\left(\hat{b}_1 - b_1\right) \to N\left(0, \frac{1}{1-a}Q^{-1}\Omega_{11}Q^{-1}\right),$$

so that the second term in (13) captures the fact the we have unidentified structural parameters.

4. Conclusions

This paper has studied the asymptotic properties of Nagar's BATSLS when the number of instruments increases at the same rate as the sample size under partial identification.

We have found similar results to those of Phillips (1989) and Choi and Phillips (1992) in the sense that the BATSLS estimator of the identified structural parameters is consistent, but the estimator of the unidentified parameters has a non-degenerate non-standard limit distribution which depends on the identified parameters. The asymptotic distribution of the BATSLS estimator of the identified parameters is a mean- and covariance-matrix mixed normal and has a very complex structure.

Appendix

Proof of Lemma 1

Since $\left(T-k\right)S : W_{n+1}\left(T-k, I_{n+1}\right)$, standard results (e.g. Muirhead (1982) p. 90) imply that

$$E(S) = I_{n+1} \quad \text{and} \quad \text{cov}\left(vec(S)\right) = \frac{1}{T-k}\left(I_{(n+1)^2} + K_{n+1}\right)\left(I_{n+1} \otimes I_{n+1}\right) \to 0 \quad \text{as} \quad T \to \infty, \quad \text{so that}$$

$S \to^P I_{n+1}$. Similarly, since $kW : W_{n+1}\left(k, \begin{pmatrix} (b_1, I_{n_1})'\Pi'Z'Z\Pi(b_1, I_{n_1}) & 0 \\ 0 & 0 \end{pmatrix}\right)$ standard results

(e.g. Muirhead (1982) p. 442 and 518) yield

$$E(W) = I_{n+1} + \frac{1}{\hat{a}}(b_1, I_{n_1}, 0)'\left(\frac{1}{T}\Pi'Z'Z\Pi\right)(b_1, I_{n_1}, 0) \to^P I_{n+1} + \frac{1}{a}(b_1, I_{n_1}, 0)'Q(b_1, I_{n_1}, 0)$$

and

$$\mathrm{cov}(vec(W)) = \frac{1}{k}(I_{n+1} + K)\left[\frac{k}{k}(I_{n+1} \otimes I_{n+1}) + I_{n+1} \otimes \left\{(b_1, I_{n_1}, 0)'\left[\frac{T}{k}\frac{1}{T}\Pi'Z'Z\Pi\right](b_1, I_{n_1}, 0)\right\}\right.$$
$$\left. + \left\{(b_1, I_{n_1}, 0)'\left[\frac{T}{k}\frac{1}{T}\Pi'Z'Z\Pi\right](b_1, I_{n_1}, 0)\right\} \otimes I_{n+1}\right]$$
$$\to^P 0$$

as $T \to \infty$, so that $W \to^P I_{n+1} + \frac{1}{a}(b_1, I_{n_1}, 0)'Q(b_1, I_{n_1}, 0)$. Then equation (7) follows from

$$\hat{A} = \hat{a}(1-\hat{a})(W - S)$$
$$\to^P a(1-a)\left[I_{n+1} + \frac{1}{a}(b_1, I_{n_1}, 0)'Q(b_1, I_{n_1}, 0) - I_{n+1}\right]$$
$$= (1-a)(b_1, I_{n_1}, 0)'Q(b_1, I_{n_1}, 0).$$

To prove the second part of the lemma notice that since $(T-k)S : W_{n+1}(T-k, I_{n+1})$, standard results imply that

$$\sqrt{T-k}\left(vec(S) - vec(I_{n+1})\right) \to^D N(0, \Sigma_1)$$

(e.g. Muirhead (1982) p. 90) and one can prove that

$$\sqrt{k}\left(vec(W) - vec\left(I_{n+1} + \frac{1}{\hat{a}}\begin{pmatrix} (b_1, I_{n_1})'\left[\frac{1}{T}\Pi'Z'Z\Pi\right](b_1, I_{n_1}) & 0 \\ 0 & 0 \end{pmatrix}\right)\right) \to^d N(0, \Sigma_2)$$

where

$$\Sigma_2 = \left(I_{(n+1)^2} + K_{n+1}\right)\lim_{T \to \infty}\left[(I_{n+1} \otimes I_{n+1}) + I_{n+1} \otimes \left\{(b_1, I_{n_1}, 0)'\left[\frac{T}{k}\frac{1}{T}\Pi'Z'Z\Pi\right](b_1, I_{n_1}, 0)\right\}\right.$$
$$\left. + \left\{(b_1, I_{n_1}, 0)'\left[\frac{T}{k}\frac{1}{T}\Pi'Z'Z\Pi\right](b_1, I_{n_1}, 0)\right\} \otimes I_{n+1}\right]$$
$$= \left(I_{(n+1)^2} + K_{n+1}\right)\left[(I_{n+1} \otimes I_{n+1}) + \frac{1}{a}I_{n+1} \otimes \left\{(b_1, I_{n_1}, 0)'Q(b_1, I_{n_1}, 0)\right\}\right.$$
$$\left. \frac{1}{a}\left\{(b_1, I_{n_1}, 0)'Q(b_1, I_{n_1}, 0)\right\} \otimes I_{n+1}\right]).$$

Then

$$\sqrt{T}\,vec\left[\hat{A}-(1-\hat{a})\begin{pmatrix}(b_1,I_{n_1})'\left[\dfrac{1}{T}\Pi'Z'Z\Pi\right](b_1,I_{n_1}) & 0 \\ 0 & 0\end{pmatrix}\right]$$

$$=\hat{a}(1-\hat{a})\sqrt{T}\,vec\left[(W-S)-\dfrac{1}{\hat{a}}\begin{pmatrix}(b_1,I_{n_1})'\left[\dfrac{1}{T}\Pi'Z'Z\Pi\right](b_1,I_{n_1}) & 0 \\ 0 & 0\end{pmatrix}\right]$$

$$=\hat{a}(1-\hat{a})\sqrt{T}\,vec\left[W-\dfrac{1}{\hat{a}}\begin{pmatrix}(b_1,I_{n_1})'\left[\dfrac{1}{T}\Pi'Z'Z\Pi\right](b_1,I_{n_1}) & 0 \\ 0 & 0\end{pmatrix}-S\right]$$

$$=\hat{a}(1-\hat{a})\left\{\sqrt{\dfrac{T}{k}}\sqrt{k}\,vec\left[W-\dfrac{1}{\hat{a}}\begin{pmatrix}(b_1,I_{n_1})'\left[\dfrac{1}{T}\Pi'Z'Z\Pi\right](b_1,I_{n_1}) & 0 \\ 0 & 0\end{pmatrix}\right]-\sqrt{\dfrac{T}{T-k}}\sqrt{T-k}\,vec(S)\right\}$$

$$\to^{D} a(1-a)\left\{\sqrt{\dfrac{1}{a}}N(0,\Sigma_2)-\sqrt{\dfrac{1}{1-a}}N(0,\Sigma_1)\right\}\equiv N\left(0,a^2(1-a)\Sigma_1+a(1-a)^2\Sigma_2\right).$$

Proof of Lemma 2

Let $\hat{Q}=(1/T)\Pi'Z'Z\Pi$ and write

$$\bar{A}=\hat{A}-(1-\hat{a})\begin{pmatrix}(b_1,I_{n_1})'Q(b_1,I_{n_1}) & 0 \\ 0 & 0\end{pmatrix}$$

$$=\begin{pmatrix}\hat{a}-(1-\hat{a})b_1'Qb_1 & (\hat{b}_1-(1-\hat{a})Qb_1)' & \hat{b}_2' \\ \hat{b}_1-(1-\hat{a})Qb_1 & \hat{A}_{11}-(1-\hat{a})Q & \hat{A}_{21}' \\ \hat{b}_2 & \hat{A}_{21} & \hat{A}_{22}\end{pmatrix}.$$

Then consider

$$\tilde{A}=\begin{pmatrix}1 & -b_1' & 0 \\ 0 & I_{n_1} & 0 \\ 0 & 0 & I_{n_2}\end{pmatrix}\bar{A}\begin{pmatrix}1 & 0 & 0 \\ -b_1 & I_{n_1} & 0 \\ 0 & 0 & I_{n_2}\end{pmatrix}$$

$$=\begin{pmatrix}a-2b_1'b_1+b_1'A_{11}b_1 & (b_1-A_{11}b_1)' & (b_2-A_{21}b_1) \\ b_1-A_{11}b & A_{11}-(1-\hat{a})\hat{Q} & A_{21}' \\ b_2-A_{21}b_1 & A_{21} & A_{22}\end{pmatrix}.$$

This matrix contains $b_1-A_{11}b_1$, $b_2-A_{21}b_1$, A_{21} and A_{22} which appear in the statement of the Lemma. Then using the expression above and Theorem 2 p. 30 of Magnus and Neudecker (1988) followed by the continuous mapping theorem we obtain

$$\sqrt{T}\,vec\left(\overset{\%}{A}\right)=\left(\begin{pmatrix}1&-\mathrm{b}_1{}'&0\\0&I_{n_1}&0\\0&0&I_{n_2}\end{pmatrix}\otimes\begin{pmatrix}1&-\mathrm{b}_1{}'&0\\0&I_{n_1}&0\\0&0&I_{n_2}\end{pmatrix}\right)\sqrt{T}\,vec\left(\overline{A}\right)$$

$$\rightarrow^{D}\left(\begin{pmatrix}1&-\mathrm{b}_1{}'&0\\0&I_{n_1}&0\\0&0&I_{n_2}\end{pmatrix}\otimes\begin{pmatrix}1&-\mathrm{b}_1{}'&0\\0&I_{n_1}&0\\0&0&I_{n_2}\end{pmatrix}\right)N\left(0,a\,(1-a)^2\,\Sigma_2+a^2\,(1-a)\,\Sigma_1\right)$$

$$\equiv N\left(0,\Sigma\right),$$

so that $\sqrt{T}\left((b_1-A_{11}\mathrm{b}_1)',(b_2-A_{21}\mathrm{b}_1)',vec(A_{21})',vec(A_{22})\right)$ is asymptotically normal with mean vector zero. We now have to focus on the covariance matrix. Notice that we do not need the whole covariance matrix in the latter display. We need just some parts of it.

The covariance matrix Σ is

(17)

$$(1-a)^2\left(I_{(n+1)^2}+K_{n+1}\right)\left(\begin{pmatrix}1+\mathrm{b}_1{}'\mathrm{b}_1&-\mathrm{b}_1{}'&0\\-\mathrm{b}_1&I_{n_1}&0\\0&0&I_{n_2}\end{pmatrix}\otimes\begin{pmatrix}0&0&0\\0&Q&0\\0&0&0\end{pmatrix}+\begin{pmatrix}0&0&0\\0&Q&0\\0&0&0\end{pmatrix}\otimes\begin{pmatrix}1+\mathrm{b}_1{}'\mathrm{b}_1&-\mathrm{b}_1{}'&0\\-\mathrm{b}_1&I_{n_1}&0\\0&0&I_{n_2}\end{pmatrix}\right).$$

Now we need to work out the variance and covariance matrices of $\sqrt{T}\left((b_1-A_{11}\mathrm{b}_1)',(b_2-A_{21}\mathrm{b}_1)',vec(A_{21})',vec(A_{22})\right)$. We do this by selecting submatrices of $\overset{\%}{A}$ by operation having the form $F\overset{\%}{A}B$ this will transform $\sqrt{T}\,vec\left(\overset{\%}{A}\right)$ to $\sqrt{T}\,vec\left(F\overset{\%}{A}B\right)=(B'\otimes F)\sqrt{T}\,vec\left(\overset{\%}{A}\right)$ so that this will have asymptotic distribution $(B'\otimes F)N(0,\Sigma)\equiv N\left(0,(B'\otimes F)\Sigma(B\otimes F')\right)$.

Part 1: $b_1-A_{11}\mathrm{b}$ and $b_2-A_{21}\mathrm{b}_1$

Consider

$$\begin{pmatrix}b_1-A_{11}\mathrm{b}\\b_2-A_{21}\mathrm{b}_1\end{pmatrix}=vec\begin{pmatrix}b_1-A_{11}\mathrm{b}\\b_2-A_{21}\mathrm{b}_1\end{pmatrix}$$

$$=vec\left(\begin{pmatrix}0&I_{n_1}&0\\0&0&I_{n_2}\end{pmatrix}\overset{\%}{A}\begin{pmatrix}1\\0\\0\end{pmatrix}\right)$$

$$\rightarrow^{D}N\left(0,\left((1\;\;0\;\;0)\otimes\begin{pmatrix}0&I_{n_1}&0\\0&0&I_{n_2}\end{pmatrix}\right)\Sigma\left(\begin{pmatrix}1\\0\\0\end{pmatrix}\otimes\begin{pmatrix}0&0\\I_{n_1}&0\\0&I_{n_2}\end{pmatrix}\right)\right).$$

We will work out the covariance matrix in detail for this case, but for the cases considered below we will skip the details. The asymptotic covariance matrix above is

$$\left(\begin{pmatrix}1 & 0 & 0\end{pmatrix}\otimes\begin{pmatrix}0 & I_{n_1} & 0\\ 0 & 0 & I_{n_2}\end{pmatrix}\right)\Sigma\left(\begin{pmatrix}1\\ 0\\ 0\end{pmatrix}\otimes\begin{pmatrix}0 & 0\\ I_{n_1} & 0\\ 0 & I_{n_2}\end{pmatrix}\right)$$

$$=a\left(1-a\right)\left(\begin{pmatrix}1 & 0 & 0\end{pmatrix}\otimes\begin{pmatrix}0 & I_{n_1} & 0\\ 0 & 0 & I_{n_2}\end{pmatrix}\right)\left(I_{(n+1)^2}+K_{n+1}\right)\left(\begin{pmatrix}1+b_1{}'b_1\\ -b_1\\ 0\end{pmatrix}\otimes\begin{pmatrix}-b_1{}' & 0\\ I_{n_1} & 0\\ 0 & I_{n_2}\end{pmatrix}\right)$$

$$+\left(1-a\right)^2\left(\begin{pmatrix}1 & 0 & 0\end{pmatrix}\otimes\begin{pmatrix}0 & I_{n_1} & 0\\ 0 & 0 & I_{n_2}\end{pmatrix}\right)\left(I_{(n+1)^2}+K_{n+1}\right)\left(\begin{pmatrix}1+b_1{}'b_1\\ -b_1\\ 0\end{pmatrix}\otimes\begin{pmatrix}0 & 0\\ Q & 0\\ 0 & 0\end{pmatrix}\right).$$

Now we focus on the first term:

$$\left(\begin{pmatrix}1 & 0 & 0\end{pmatrix}\otimes\begin{pmatrix}0 & I_{n_1} & 0\\ 0 & 0 & I_{n_2}\end{pmatrix}\right)\left(I_{(n+1)^2}+K_{n+1}\right)\left(\begin{pmatrix}1+b_1{}'b_1\\ -b_1\\ 0\end{pmatrix}\otimes\begin{pmatrix}-b_1{}' & 0\\ I_{n_1} & 0\\ 0 & I_{n_2}\end{pmatrix}\right)$$

$$=\left(\begin{pmatrix}1 & 0 & 0\end{pmatrix}\otimes\begin{pmatrix}0 & I_{n_1} & 0\\ 0 & 0 & I_{n_2}\end{pmatrix}\right)\left(\begin{pmatrix}1+b_1{}'b_1\\ -b_1\\ 0\end{pmatrix}\otimes\begin{pmatrix}-b_1 & 0\\ I_{n_1} & 0\\ 0 & I_{n_2}\end{pmatrix}\right)$$

$$+\left(\begin{pmatrix}1 & 0 & 0\end{pmatrix}\otimes\begin{pmatrix}0 & I_{n_1} & 0\\ 0 & 0 & I_{n_2}\end{pmatrix}\right)K_{n+1}\left(\begin{pmatrix}1+b_1{}'b_1\\ -b_1\\ 0\end{pmatrix}\otimes\begin{pmatrix}-b_1{}' & 0\\ I_{n_1} & 0\\ 0 & I_{n_2}\end{pmatrix}\right).$$

Using Theorem 9 in Chapter 3 of Magnus and Neudecker (1988) we obtain

$$\left(1+b_1{}'b_1\right)I_n+\left(\begin{pmatrix}1 & 0 & 0\end{pmatrix}\otimes\begin{pmatrix}0 & I_{n_1} & 0\\ 0 & 0 & I_{n_2}\end{pmatrix}\right)\left(\begin{pmatrix}-b_1{}' & 0\\ I_{n_1} & 0\\ 0 & I_{n_2}\end{pmatrix}\otimes\begin{pmatrix}1+b_1{}'b_1\\ -b_1\\ 0\end{pmatrix}\right)$$

$$=\left(1+b_1{}'b_1\right)I_n+\left(\begin{pmatrix}-b_1 & 0\end{pmatrix}\otimes\begin{pmatrix}-b_1{}'\\ 0\end{pmatrix}\right)$$

$$=\left(1+b_1{}'b_1\right)I_n+\begin{pmatrix}-b_1\\ 0\end{pmatrix}\begin{pmatrix}-b_1{}' & 0\end{pmatrix}$$

$$=\begin{pmatrix}\left(1+b_1{}'b_1\right)I_{n_1}+b_1b_1{}' & 0\\ 0 & \left(1+b_1{}'b_1\right)I_{n_2}\end{pmatrix}.$$

Similarly we can show that the second term equals

$$\left(1+b_1{}'b_1\right)\otimes\begin{pmatrix}Q & 0\\ 0 & 0\end{pmatrix}.$$

Part 2: $b_2-A_{21}b_1$, $vec\left(A_{21}\right)$ and $vec\left(A_{22}\right)$

We now consider

$$\sqrt{T}\begin{pmatrix} b_2 - A_{21}b_1 \\ vec(A_{21}) \\ vec(A_{22}) \end{pmatrix} = \sqrt{T}\,vec\begin{pmatrix} b_2 - A_{21}b_1 & A_{21} & A_{22} \end{pmatrix}$$

$$= \sqrt{T}\,vec\left(\begin{pmatrix} 0 & 0 & I_{n_2} \end{pmatrix} \hat{A}\right) \to^{D} N\left(0, \left(I_{n+1} \otimes \begin{pmatrix} 0 & 0 & I_{n_2} \end{pmatrix}\right)\Sigma\left(I_{n+1} \otimes \begin{pmatrix} 0 \\ 0 \\ I_{n_2} \end{pmatrix}\right)\right).$$

Proceeding as before we find that the asymptotic covariance matrix is

$$\begin{pmatrix} a(1-a)(1+b_1{}'b_1)I_{n_2} & -a(1-a)(b_1{}'\otimes I_{n_2}) & 0 \\ -a(1-a)(b_1 \otimes I_{n_2}) & a(1-a)(I_{n_1}\otimes I_{n_2})+(1-a)^2 Q \otimes I_{n_2} & 0 \\ 0 & 0 & a(1-a)(I_{n_2^2}+K_{n_2})(I_{n_2}\otimes I_{n_2}) \end{pmatrix}.$$

Part 3: $b_1 - A_{11}b$ and $vec(A_{21})$

Consider

$$\sqrt{T}\begin{pmatrix} b_1 - A_{11}b \\ vec(A_{21}{}') \end{pmatrix} = \sqrt{T}\,vec\begin{pmatrix} (b_1 - A_{11}b & A_{21}{}') \end{pmatrix}$$

$$= \sqrt{T}\,vec\left(\begin{pmatrix} 0 & I_{n_1} & 0 \end{pmatrix}\hat{A}\begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & I_{n_2} \end{pmatrix}\right)$$

$$\to^{D} N\left(0, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & I_{n_2} \end{pmatrix}\otimes\begin{pmatrix} 0 & I_{n_1} & 0 \end{pmatrix}\Sigma\begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & I_{n_2} \end{pmatrix}\otimes\begin{pmatrix} 0 \\ I_{n_1} \\ 0 \end{pmatrix}\right).$$

The variance covariance matrix is

$$\begin{pmatrix} (1+b_1{}'b_1)\left(a(1-a)I_{n_1}+(1-a)^2 Q\right)+a(1-a)b_1b_1{}' & 0 \\ 0 & a(1-a)(I_{n_2}\otimes I_{n_1})+(1-a)^2(I_{n_2}\otimes Q) \end{pmatrix}.$$

Part 4: $b_1 - A_{11}b$ and A_{22}

Consider

13

$$\sqrt{T}\begin{pmatrix} b_1 - A_{11}\mathsf{b}_1 \\ b_2 - A_{21}\mathsf{b}_1 \\ vec\begin{pmatrix} A_{21} \\ A_{22} \end{pmatrix} \end{pmatrix} = \sqrt{T}\,vec\begin{pmatrix} b_1 - A_{11}\mathsf{b} & A_{21}{}' \\ b_2 - A_{21}\mathsf{b}_1 & A_{22} \end{pmatrix}$$

$$= \sqrt{T}\,vec\left(\begin{pmatrix} 0 & I_{n_1} & 0 \\ 0 & 0 & I_{n_2} \end{pmatrix} A \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & I_{n_2} \end{pmatrix} \right)$$

$$\to^D N\left(0, \left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & I_{n_2} \end{pmatrix} \otimes \begin{pmatrix} 0 & I_{n_1} & 0 \\ 0 & 0 & I_{n_2} \end{pmatrix} \right) \Sigma \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & I_{n_2} \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ I_{n_1} & 0 \\ 0 & I_{n_2} \end{pmatrix} \right) \right).$$

The asymptotic covariance matrix is

$$\mathsf{a}(1-\mathsf{a})\begin{pmatrix} (1+\mathsf{b}_1{}'\mathsf{b}_1)I_n & 0 \\ 0 & I_{n_2}\otimes I_n \end{pmatrix}$$

$$+\mathsf{a}(1-\mathsf{a})\begin{pmatrix} \begin{pmatrix} \mathsf{b}_1\mathsf{b}_1{}' \\ 0 \end{pmatrix} & 0 & -\begin{pmatrix} 0 \\ e_1\mathsf{b}_1{}' \end{pmatrix} & 0 & -\begin{pmatrix} 0 \\ e_2\mathsf{b}_1{}' \end{pmatrix} & 0 & \dots & -\begin{pmatrix} 0 \\ e_n\mathsf{b}_1{}' \end{pmatrix} & 0 \\ 0 & I_{n_2}\otimes\begin{pmatrix} -\mathsf{b}_1 \\ 0 \end{pmatrix} & 0 & I_{n_2}\otimes\begin{pmatrix} 0 \\ e_1 \end{pmatrix} & 0 & I_{n_2}\otimes\begin{pmatrix} 0 \\ e_2 \end{pmatrix} & \dots & 0 & I_{n_2}\otimes\begin{pmatrix} 0 \\ e_n \end{pmatrix} \end{pmatrix}$$

$$+(1-\mathsf{a})^2\left(\begin{pmatrix} 1+\mathsf{b}_1{}'\mathsf{b}_1 & 0 \\ 0 & I_{n_2} \end{pmatrix} \otimes \begin{pmatrix} Q & 0 \\ 0 & 0 \end{pmatrix} \right).$$

We are interested in the lower $n_2^2 \times n_1$ block which is 0.

References

Andrews, D. W. K., M. J. Moreira and J. H. Stock (2007). "Performance of conditional Wald tests in IV regression with weak instruments" *Journal of Econometrics* 139(1): 116-132.

Bekker, P. A. (1994). "Alternative Approximations to the Distributions of Instrumental Variable Estimators" *Econometrica* 62: 657-681.

Choi, I. and P. C. B. Phillips (1992). "Asymptotic and Finite Sample Distribution Theory for *IV* Estimators and Tests in Partially Identified Structural Equations" *Journal of Econometrics* 51: 113-150.

Donald, S. G. and W. K. Newey (2001). "Choosing the Number of Instruments" *Econometrica* 69(5): 1161-1191.

Forchini, G. (2008). "The Asymptotic Distribution of the LIML Estimator in a Partially Identified Structural Equation" *available at http://www-personal.buseco.monash.edu.au/~forchini/AsymLIML.pdf*.

Hahn, J., J. Hausman and G. Kuersteiner (2004). "Estimation with weak instruments: accuracy of higher-order bias and MSE approximations" *The Econometrics Journal* 7(1): 272-306.

Magnus, J. R. and H. Neudecker (1988). "Matrix Differential Calculus with Applications in Statistics and Econometrics". New York, John Wiley & Sons Ltd.

Muirhead, R. J. (1982). "Aspects of Multivariate Statistical Theory". New York, John Wiley and Sons, Inc.

Nagar, A. L. (1959). "The Bias and Moment Matrix of the General k-Class Estimators of the Parameters in Simultaneous Equations" *Econometrica* 27: 575-595.

Phillips, P. C. B. (1989). "Partially Identified Econometric Models" *Econometric Theory* 5: 181-240.