LECTURE 6
WEAK INSTRUMENTS

In this lecture, we discuss the IV regression model with weak instruments. The discussion follows Staiger and Stock's 1997 paper in *Econometrica*.

Model

Consider the following regression model with a single endogenous regressor and a number of exogenous regressors:

$$y_1 = y_2\gamma + Z_2\beta + u,$$

where y_1 is the n-vector of observations on the dependent variable, y_2 is the n-vector of observations on the endogenous regressor, Z_2 is the $n \times l_2$ matrix of l_2 exogenous regressors, u is the n-vector of residuals, $\gamma \in R$, and $\beta \in R^{l_2}$. We have the reduced form equation for y_2:

$$y_2 = Z_1\pi_1 + Z_2\pi_2 + v,$$

where Z_1 is the $n \times l_1$ matrix of l_1 instruments for y_2, $\pi_j \in R^{l_j}$ for $j = 1, 2$, and v is the n-vector of observations on the reduced form residuals.

If $\pi_1 = 0$, Z_1 is not a valid instrument for y_2 and the model is not identified. Here we consider the case when the model is identified, however, the relationship between Z_1 and y_2 (for a given Z_2) is weak, i.e. Z_1 is the matrix of weak instruments. Weak instruments are defined by the following assumption.

Assumption 1 *$\pi_1 = \pi_1(n) = n^{-1/2}C$, where $C \in R^{l_1}$ and fixed.*

Assumption 1 defines weak, but different from zero relationship between the endogenous regressor y_2 and its instruments Z_1 (after controlling for the effect of Z_2). We will rely on large n approximation for the distribution of the estimators, and, therefore, weakness of the relationship between y_2 and Z_1 has to be modelled in terms of the sample size n. This is because any fixed (independent of n) π_1, will be "large" when $n \to \infty$ as long as it is different from zero. Therefore, we assume that $\pi_1 = \pi_1(n) \to 0$ as $n \to \infty$. The rate of convergence is chosen in a such way so that small correlations captured by non-zero C's will appear in the limit.

In addition, we make the following assumptions.

Assumption 2 *(a) $\{(y_{1i}, y_{2i}, Z_{1i}, Z_{2i}) : i = 1, \ldots, n\}$ are iid.*

(b) $E\begin{pmatrix} Z_{1i} \\ Z_{2i} \end{pmatrix}\begin{pmatrix} u_i & v_i \end{pmatrix} = 0.$

(c) $E\left(\begin{pmatrix} u_i \\ v_i \end{pmatrix}\begin{pmatrix} u_i \\ v_i \end{pmatrix}' | Z_{1i}, Z_{2i}\right) = \begin{pmatrix} \sigma_u^2 & \sigma_{uv} \\ \sigma_{uv} & \sigma_v^2 \end{pmatrix} = \Sigma$, *a finite and positive definite matrix.*

(d) $E\begin{pmatrix} Z_{1i} \\ Z_{2i} \end{pmatrix}\begin{pmatrix} Z_{1i} \\ Z_{2i} \end{pmatrix}' = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{12}' & Q_{22} \end{pmatrix} = Q$, *a finite and positive definite matrix.*

Part (b) in the above assumption says that the instruments are exogenous. Part (c) says that the errors u and v are homoskedastic. It also introduces a partition on their second moment matrix. Part (d) says that the instrument have finite second moments. Not that part (d) also implies that Q_{11} and Q_{22} are positive definite.

Define

$$Z = \begin{pmatrix} Z_1 & Z_2 \end{pmatrix}.$$

It follows from Assumption 2(a),(d) and the WLLN that

$$
\begin{aligned}
n^{-1}Z'Z &= \begin{pmatrix} \frac{Z_1'Z_1}{n} & \frac{Z_1'Z_2}{n} \\ \frac{Z_2'Z_1}{n} & \frac{Z_2'Z_2}{n} \end{pmatrix} \\
&= \begin{pmatrix} \frac{\sum_{i=1} Z_{1i}Z_{1i}'}{n} & \frac{\sum_{i=1} Z_{1i}Z_{2i}'}{n} \\ \frac{\sum_{i=1} Z_{2i}Z_{1i}'}{n} & \frac{\sum_{i=1} Z_{2i}Z_{2i}'}{n} \end{pmatrix} \\
&\to_p \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{12}' & Q_{22} \end{pmatrix} \\
&= Q.
\end{aligned}
$$

Further, the CLT and Assumption 2 imply that

$$
\begin{aligned}
n^{-1/2} \begin{pmatrix} Z'u \\ Z'v \end{pmatrix} &= \begin{pmatrix} \frac{Z_1'u}{\sqrt{n}} \\ \frac{Z_2'u}{\sqrt{n}} \\ \frac{Z_1'v}{\sqrt{n}} \\ \frac{Z_2'v}{\sqrt{n}} \end{pmatrix} \\
&= n^{-1/2} \sum_{i=1}^{n} \begin{pmatrix} u_i \\ v_i \end{pmatrix} \otimes \begin{pmatrix} Z_{1i} \\ Z_{2i} \end{pmatrix} \\
&\to_d N\left(0, \Sigma \otimes Q\right) \\
&= \begin{pmatrix} \Phi_1 \\ \Phi_2 \\ \Psi_1 \\ \Psi_2 \end{pmatrix}. \qquad (1)
\end{aligned}
$$

Thus, for example, the asymptotic distribution of $Z_1'u/\sqrt{n}$ is given by the distribution of $\Phi_1 \sim N\left(0, \sigma_u^2 Q_{11}\right)$. It is important for further results that convergence in distribution is joint. For example, the asymptotic covariance of $Z_1'u/\sqrt{n}$ and $Z_2'v/\sqrt{n}$ is given by the covariance of Φ_1 and Ψ_2, which equals $\sigma_{uv}Q_{12}$.

Define

$$
\begin{aligned}
X &= \begin{pmatrix} y_2 & Z_2 \end{pmatrix}, \\
P &= Z\left(Z'Z\right)^{-1} Z'.
\end{aligned}
$$

Since the errors are homoskedastic, we can focus on the 2SLS estimator:

$$
\begin{pmatrix} \widehat{\gamma} \\ \widehat{\beta} \end{pmatrix} = \left(X'PX\right)^{-1} X'Py_1.
$$

A simple regression example

Suppose that $\beta = \pi_2 = 0$ and $l_1 = 1$. This is a just identified case, and 2SLS estimator of γ reduces to the IV estimator:

$$
\begin{aligned}
\widehat{\gamma} &= \frac{Z_1' y_1}{Z_1' y_2} \\
&= \gamma + \frac{Z_1' u}{Z_1' (Z_1 \pi_1 + v)} \\
&= \gamma + \frac{Z_1' u}{Z_1' (Z_1 C / \sqrt{n} + v)} \\
&= \gamma + \frac{Z_1' u / \sqrt{n}}{Z_1' Z_1 / n C + Z_1' v / \sqrt{n}} \\
&\to_d \gamma + \frac{\Phi_1}{Q_{11} C + \Psi_1} \\
&= \gamma + \Delta,
\end{aligned} \tag{2}
$$

where the distribution of the random variable Δ is given by

$$\Delta = \frac{\Phi_1}{Q_{11} C + \Psi_1},$$

and (2) follows from the joint convergence in (1) and by the Continuous Mapping Theorem (CMT).

This simple example illustrates the problem with IV estimation in presence of weak instruments. First, the 2SLS (IV) estimator is inconsistent. Second, instead of usual converge in probability, $\widehat{\gamma}$ converges in distribution to a random variable. The asymptotic distribution of $\widehat{\gamma}$ is non-standard and depends on the ratio of two normal random variables. The "bias" term Δ has an inverse relationship with C.

Suppose that the econometrician ignores the fact that Z_1 is a weak instrument, and for testing $H_0 : \gamma = \gamma_0$, he uses the usual t statistic based on $\widehat{\gamma}$:

$$t = \frac{\widehat{\gamma} - \gamma_0}{\sqrt{\widehat{AsyVar} (\widehat{\gamma}) / n}},$$

where $\widehat{AsyVar} (\widehat{\gamma})$ is an estimator of the usual asymptotic variance of the IV estimator when the instruments are strong:

$$
\begin{aligned}
\widehat{AsyVar} (\widehat{\gamma}) &= \widehat{\sigma}_u^2 \frac{Z_1' Z_1 / n}{(Z_1' y_2 / n)^2}, \\
\widehat{\sigma}_u^2 &= \| y_1 - y_2 \widehat{\gamma} \|^2 / n,
\end{aligned}
$$

where $\|c\|^2 = c'c$ is the Euclidean norm. Under H_0, we have

$$t = \frac{Z_1' u}{\sqrt{\widehat{\sigma}_u^2 Z_1' Z_1}}. \tag{3}$$

In order to derive the asymptotic distribution of t under H_0, first, note that

$$\frac{Z_1' u / \sqrt{n}}{\sqrt{Z_1' Z_1 / n}} \to_d \frac{\Phi_1}{\sqrt{Q_{11}}}. \tag{4}$$

Next,

$$
\begin{aligned}
\widehat{\sigma}_u^2 &= \left\| y_2 \left(\widehat{\gamma} - \gamma \right) - u \right\|^2 / n \\
&= \left\| \left(Z_1 C / \sqrt{n} + v \right) \left(\widehat{\gamma} - \gamma \right) - u \right\|^2 / n \\
&= \left(\widehat{\gamma} - \gamma \right)^2 \left(C^2 \frac{Z_1' Z_1}{n^2} + \frac{v'v}{n} + 2C \frac{Z_1' v}{n\sqrt{n}} \right) - 2 \left(\widehat{\gamma} - \gamma \right) \left(C \frac{Z_1' u}{n\sqrt{n}} + \frac{v'u}{n} \right) + \frac{u'u}{n} \\
&\to_d \sigma_v^2 \Delta^2 - 2\sigma_{vu} \Delta + \sigma_u^2,
\end{aligned}
\tag{5}
$$

Combining the results in (3)-(5), and using the CMT,

$$
\begin{aligned}
t &\to_d \frac{\Phi_1 / \sqrt{Q_{11}}}{\sqrt{\sigma_v^2 \Delta^2 - 2\sigma_{vu} \Delta + \sigma_u^2}} \\
&= \frac{\Phi_1 / \sqrt{\sigma_u^2 Q_{11}}}{\sqrt{\frac{\sigma_v^2}{\sigma_u^2} \Delta^2 - 2 \frac{\sigma_{vu}}{\sigma_u^2} \Delta + 1}}.
\end{aligned}
$$

Note that while $\Phi_1 / \sqrt{\sigma_u^2 Q_{11}} \sim N(0,1)$, the denominator in the above expression is a random variable, which depends on the numerator through Φ_1. Furthermore, the limiting distribution of t depends on the unknown nuisance parameter C, which cannot be estimated consistently. Thus, under the null, the t statistic does not have the usual standard normal distribution. Consequently, a test that rejects the null hypothesis in favor of the alternative, $H_1 : \gamma \neq \gamma_0$, when $|t| > z_{1-\alpha/2}$ is invalid, since its asymptotic size is different from α. Similarly, the usual confidence intervals based on $\widehat{\gamma}$ are invalid, since they are constructed by inverting the t test. Let

$$
CI_\alpha = \left[\widehat{\gamma} - z_{1-\alpha/2} \sqrt{\widehat{AsyVar} \left(\widehat{\gamma} \right) / n}, \widehat{\gamma} + z_{1-\alpha/2} \sqrt{\widehat{AsyVar} \left(\widehat{\gamma} \right) / n} \right],
$$

and let γ_0 be the true value of γ. Then,

$$
\begin{aligned}
\lim_{n\to\infty} P \left(\gamma_0 \in CI_a \right) &= \lim_{n\to\infty} P \left(|t| < z_{1-\alpha/2} \right) \\
&= P \left(\left| \frac{\Phi_1 / \sqrt{\sigma_u^2 Q_{11}}}{\sqrt{\frac{\sigma_v^2}{\sigma_u^2} \Delta^2 - 2 \frac{\sigma_{vu}}{\sigma_u^2} \Delta + 1}} \right| < z_{1-\alpha/2} \right).
\end{aligned}
$$

Large sample properties of the 2SLS estimator of γ

The 2SLS estimator can be written as

$$
\begin{pmatrix} \widehat{\gamma} \\ \widehat{\beta} \end{pmatrix} = \left(\widehat{X}' \widehat{X} \right)^{-1} \widehat{X}' y_1,
$$

where

$$
\begin{aligned}
\widehat{X} &= \begin{pmatrix} \widehat{y}_2 & Z_2 \end{pmatrix}, \\
\widehat{y}_2 &= Z_1 \widehat{\pi}_1 + Z_2 \widehat{\pi}_2,
\end{aligned}
$$

and $\widehat{\pi}_1$ and $\widehat{\pi}_2$ are the LS coefficients from the regression of y_2 against Z_1 and Z_2. Define

$$
\begin{aligned}
M_2 &= I_n - P_2, \\
P_2 &= Z_2 \left(Z_2' Z_2 \right)^{-1} Z_2'.
\end{aligned}
$$

Then we can write

$$\begin{aligned}\widehat{\gamma} &= \frac{\widehat{y}_2' M_2 y_1}{\widehat{y}_2' M_2 \widehat{y}_2} \\ &= \gamma + \frac{\widehat{y}_2' M_2 u}{\widehat{y}_2' M_2 \widehat{y}_2} \\ &= \gamma + \frac{\widehat{\pi}_1' Z_1' M_2 u}{\widehat{\pi}_1' Z_1' M_2 Z_1 \widehat{\pi}_1}.\end{aligned} \tag{6}$$

Next,

$$\widehat{\pi}_1 = \left(Z_1' M_2 Z_1\right)^{-1} Z_1' M_2 y_2.$$

Combining the last two equations,

$$\widehat{\gamma} - \gamma = \left(\frac{y_2' M_2 Z_1}{\sqrt{n}} \left(\frac{Z_1' M_2 Z_1}{n}\right)^{-1} \frac{Z_1' M_2 y_2}{\sqrt{n}}\right)^{-1} \frac{y_2' M_2 Z_1}{\sqrt{n}} \left(\frac{Z_1' M_2 Z_1}{n}\right)^{-1} \frac{Z_1' M_2 u}{\sqrt{n}}. \tag{7}$$

We have

$$\begin{aligned}Z_1' M_2 Z_1 / n &= \frac{Z_1' Z_1}{n} - \frac{Z_1' Z_2}{n} \left(\frac{Z_2' Z_2}{n}\right)^{-1} \frac{Z_2' Z_1}{n} \\ &\to_p Q_{11} - Q_{12} \left(Q_{22}^{-1}\right) Q_{12}' \\ &= Q_{1\cdot 2}.\end{aligned} \tag{8}$$

$$\begin{aligned}Z_1' M_2 y_2 / \sqrt{n} &= Z_1' M_2 \left(Z_1 \pi_1 + v\right) / \sqrt{n} \\ &= Z_1' M_2 \left(Z_1 C / \sqrt{n} + v\right) / \sqrt{n} \\ &= \frac{Z_1' M_2 Z_1}{n} C \\ &+ \frac{Z_1' v}{\sqrt{n}} - \frac{Z_1' Z_2}{n} \left(\frac{Z_2' Z_2}{n}\right)^{-1} \frac{Z_2' v}{\sqrt{n}} \\ &\to_d Q_{1\cdot 2} C + \Psi_1 - Q_{12} Q_{22}^{-1} \Psi_2 \\ &= Q_{1\cdot 2} C + \Psi_{1\cdot 2},\end{aligned} \tag{9}$$

where

$$\Psi_{1\cdot 2} = \Psi_1 - Q_{12} Q_{22}^{-1} \Psi_2.$$

Note that

$$\Psi_{1\cdot 2} \sim N\left(0, \sigma_v^2 Q_{1\cdot 2}\right).$$

Lastly,

$$\begin{aligned}Z_1' M_2 u / \sqrt{n} &= \frac{Z_1' u}{\sqrt{n}} - \frac{Z_1' Z_2}{n} \left(\frac{Z_2' Z_2}{n}\right)^{-1} \frac{Z_2' u}{\sqrt{n}} \\ &\to_d \Phi_1 - Q_{12} Q_{22}^{-1} \Phi_2 \\ &= \Phi_{1\cdot 2},\end{aligned} \tag{10}$$

where

$$\Phi_{1\cdot 2} = \Phi_1 - Q_{12} Q_{22}^{-1} \Phi_2 \sim N\left(0, \sigma_u^2 Q_{1\cdot 2}\right).$$

Combining (7)-(10), we obtain

$$\begin{aligned}\widehat{\gamma} &\to_d \gamma + \frac{\left(Q_{1\cdot 2} C + \Psi_{1\cdot 2}\right)' Q_{1\cdot 2}^{-1} \Phi_{1\cdot 2}}{\left(Q_{1\cdot 2} C + \Psi_{1\cdot 2}\right)' Q_{1\cdot 2}^{-1} \left(Q_{1\cdot 2} C + \Psi_{1\cdot 2}\right)} \\ &= \gamma + \Delta,\end{aligned}$$

where

$$\Delta = \frac{\left(Q_{1\cdot 2}C + \Psi_{1\cdot 2}\right)' Q_{1\cdot 2}^{-1}\Phi_{1\cdot 2}}{\left(Q_{1\cdot 2}C + \Psi_{1\cdot 2}\right)' Q_{1\cdot 2}^{-1}\left(Q_{1\cdot 2}C + \Psi_{1\cdot 2}\right)}.$$

We conclude that the 2SLS estimator of γ is inconsistent. Note that the denominator in the expression for Δ has a noncentral χ^2 distribution with l_1 degrees of freedom and the noncentrality parameter given by $C'Q_{1\cdot 2}C$. Similarly to the simple example discussed in the previous section, the usual tests and confidence intervals for γ are invalid due to the weak IVs problem.

Large sample properties of the 2SLS estimator of β

The 2SLS estimator of β can be written as

$$\begin{aligned} \widehat{\beta} &= \left(Z_2'Z_2\right)^{-1} Z_2'\left(y_1 - \widehat{y}_2\widehat{\gamma}\right) \\ &= \beta + \left(Z_2'Z_2\right)^{-1} Z_2'\left(y_2\gamma - \widehat{y}_2\widehat{\gamma}\right) + \left(Z_2'Z_2\right)^{-1} Z_2'u. \end{aligned}$$

From (6),

$$\begin{aligned} y_2\gamma - \widehat{y}_2\widehat{\gamma} &= \left(y_2 - \widehat{y}_2\right)\widehat{\gamma} - y_2\left(\widehat{\gamma} - \gamma\right) \\ &= \left(I_n - Z\left(Z'Z\right)^{-1} Z'\right) v\widehat{\gamma} - y_2\left(\widehat{\gamma} - \gamma\right) \\ &= Mv\widehat{\gamma} - y_2\left(\widehat{\gamma} - \gamma\right), \end{aligned}$$

where

$$\begin{aligned} M &= I_n - Z\left(Z'Z\right)^{-1} Z' \\ &= I_n - P, \end{aligned}$$

a projection matrix onto the space orthogonal to Z's. Note that $Z_2'M = 0$. We have

$$\begin{aligned} \widehat{\beta} - \beta &= \left(Z_2'Z_2\right)^{-1} Z_2'\left(Mv\widehat{\gamma} - y_2\left(\widehat{\gamma} - \gamma\right)\right) + \left(Z_2'Z_2\right)^{-1} Z_2'u \\ &= -\left(Z_2'Z_2\right)^{-1} Z_2'y_2\left(\widehat{\gamma} - \gamma\right) + \left(Z_2'Z_2\right)^{-1} Z_2'u \\ &= -\left(Z_2'Z_2\right)^{-1} Z_2'\left(Z_1C/\sqrt{n} + Z_2\pi_2 + v\right)\left(\widehat{\gamma} - \gamma\right) + \left(Z_2'Z_2\right)^{-1} Z_2'u \\ &= -\left(\frac{Z_2'Z_2}{n}\right)^{-1} \frac{Z_2'Z_1}{n\sqrt{n}} C\left(\widehat{\gamma} - \gamma\right) - \pi_2\left(\widehat{\gamma} - \gamma\right) - \left(\frac{Z_2'Z_2}{n}\right)^{-1} \frac{Z_2'v}{n}\left(\widehat{\gamma} - \gamma\right) + \left(\frac{Z_2'Z_2}{n}\right)^{-1} \frac{Z_2'u}{n}. \end{aligned}$$

Using the results of the previous sections,

$$\widehat{\beta} - \beta \to_d -\pi_2\Delta.$$

Thus, $\widehat{\beta}$ is inconsistent, if π_2 is fixed and different from zero. In particular, the 2SLS estimator of the coefficients on exogenous regressors is consistent if $\pi_2 = 0$, i.e. the exogenous regressors Z_2 are uncorrelated with the endogenous regressor y_2 after controlling for the instruments Z_1.

Let's make the following assumption.

Assumption 3 *$\pi_2 = \pi_2(n) = n^{-1/2}D$, where $D \in R^{l_2}$ and fixed.*

Using the last assumption, we have

$$\begin{aligned} \sqrt{n}\left(\widehat{\beta} - \beta\right) &= -\left(\frac{Z_2'Z_2}{n}\right)^{-1} \frac{Z_2'Z_1}{n} C\left(\widehat{\gamma} - \gamma\right) - D\left(\widehat{\gamma} - \gamma\right) - \left(\frac{Z_2'Z_2}{n}\right)^{-1} \frac{Z_2'v}{\sqrt{n}}\left(\widehat{\gamma} - \gamma\right) + \left(\frac{Z_2'Z_2}{n}\right)^{-1} \frac{Z_2'u}{\sqrt{n}} \\ &\to_d -Q_{22}^{-1}\left(Q_{12}'C\Delta + \Psi_2\Delta - \Phi_2\right) - D\Delta. \end{aligned}$$

In the case of weak correlation between Z_2 and y_2, the 2SLS estimator of β is consistent. However, its asymptotic distribution is nonstandard and depends on unknown nuisance parameters C and D which cannot be estimated consistently.

Inference on γ: a simple regression

Suppose we are interested in testing $H_0 : \gamma = \gamma_0$ against $H_1 : \gamma \neq \gamma_0$. Consider first a model with no exogenous regressors: $\beta = \pi_2 = 0$, and $l_1 \geq 1$ IVs:

$$\begin{aligned} y_1 &= \gamma y_2 + u, \\ y_2 &= Z_1 \pi_1 + v \end{aligned}$$

The null restricted residuals are given by $y_1 - y_2\gamma_0$. Under the null,

$$\begin{aligned} Z_1' \left(y_1 - y_2\gamma_0\right)/\sqrt{n} &= Z_1' u/\sqrt{n} \\ &\to_d \Phi_1 \\ &= N\left(0, \sigma_u^2 Q_{11}\right). \end{aligned} \tag{11}$$

The variance term $\sigma_u^2 Q_{11}$ can be estimated consistently by $\widetilde{\sigma}_u^2 Z_1' Z_1/n$, where the estimator $\widetilde{\sigma}_u^2$ is based on the null restricted residuals:

$$\widetilde{\sigma}_u^2 = \left(y_1 - y_2\gamma_0\right)' \left(y_1 - y_2\gamma_0\right)/n.$$

Under the null,

$$\widetilde{\sigma}_u^2 \to_p \sigma_u^2. \tag{12}$$

Let

$$P_1 = Z_1 \left(Z_1' Z_1\right)^{-1} Z_1',$$

and consider the following statistic:

$$\begin{aligned} \widetilde{AR}_n\left(\gamma_0\right) &= \left(y_1 - y_2\gamma_0\right)' P_1 \left(y_1 - y_2\gamma_0\right)/\widetilde{\sigma}_u^2 \\ &= \frac{\left[\left(y_1 - y_2\gamma_0\right)' Z_1/\sqrt{n}\right] \left(Z_1' Z_1/n\right)^{-1} \left[Z_1' \left(y_1 - y_2\gamma_0\right)/\sqrt{n}\right]}{\widetilde{\sigma}_u^2}. \end{aligned} \tag{13}$$

When the null hypothesis is true, (11) and (12) imply that

$$\widetilde{AR}_n\left(\gamma_0\right) \to_d \chi_{l_1}^2,$$

which holds regardless of the strength of the instruments. Hence, a test that rejects the null when $\widetilde{AR}_n\left(\gamma_0\right) > \chi_{l_1,1-\alpha}^2$, where $\chi_{l_1,1-\alpha}^2$ is the $1 - \alpha$ quantile of the $\chi_{l_1}^2$, has the asymptotic size α.

Suppose that $l_1 = 1$. Then, under H_0, the statistic in (13) can be written as

$$\widetilde{AR}_n\left(\gamma_0\right) = \left(\frac{Z_1' u}{\sqrt{\widetilde{\sigma}_u^2 Z_1' Z_1}}\right)^2.$$

By comparing the above equation with (3), we observe that $\widetilde{AR}_n\left(\gamma_0\right)$ is a null-restricted version of the t-statistic (squared): we replace the usual estimator of σ^2 with its null-restricted version. Thus, in this case we can solve the problem of weak identification by using a null-restricted t-statistic.

The statistic in (13) was originally suggested by Anderson and Rubin (1949) in the following form. Let

$$M_1 = I_n - P_1.$$

The Anderson-Rubin statistic (AR) is given by

$$AR\left(\gamma_0\right) = \frac{\left(y_1 - y_2\gamma_0\right)' P_1 \left(y_1 - y_2\gamma_0\right)/l_1}{\left(y_1 - y_2\gamma_0\right)' M_1 \left(y_1 - y_2\gamma_0\right)/\left(n - l_1\right)}.$$

Note that under the null,

$$\begin{aligned}
&\frac{(y_1 - y_2\gamma_0)' M_1 (y_1 - y_2\gamma_0)}{n - l_1} \\
&= \frac{u'u}{n - l_1} - \frac{u'Z_1}{n}\left(\frac{Z_1'Z_1}{n}\right)^{-1}\frac{Z_1'u}{n}\frac{n}{n-l_1} \\
&\to_p \sigma_u^2.
\end{aligned}$$

Therefore, under the null,

$$AR(\gamma_0) \to_d \chi^2_{l_1}.$$

Thus, the two versions have the same asymptotic distribution under the null. However, under the null, when the disturbances are normally distributed, $AR(\gamma_0)$ has the $F_{l_1, n-l_1}$ distribution in finite samples. Assuming that $u|Z_1 \sim N\left(0, \sigma_u^2 I_n\right)$,

$$\begin{aligned}
u'P_1 u/\sigma_u^2 | Z_1 &\sim \chi^2_{\text{rank}(P_1)} \\
&= \chi^2_{l_1}. \\
u'M_1 u/\sigma_u^2 | Z_1 &\sim \chi^2_{\text{rank}(M_1)} \\
&= \chi^2_{n-l_1}.
\end{aligned}$$

Further, numerator and denominator are independent due to normality and orthogonality of P_1 and M_1. It follows that, under the null, $AR(\gamma_0) \sim F_{l_1, n-l_1}$.

While the test based on the AR statistic has correct size regardless of the power of the instruments, its power depends on the strength of the IVs. Consider a fixed alternative $H_1 : \gamma = \gamma_0 + \delta$. In this case,

$$\begin{aligned}
M_1 (y_1 - y_2\gamma_0) &= M_1 (y_2\delta + u) \\
&= M_1 ((Z_1\pi_1 + v)\delta + u) \\
&= M_1 (v\delta + u).
\end{aligned}$$

Hence,

$$(y_1 - y_2\gamma_0)' M_1 (y_1 - y_2\gamma_0) / (n - l_1) \to_p \sigma_v^2\delta^2 + \sigma_u^2 + 2\sigma_{uv}\delta.$$

Further,

$$\begin{aligned}
P_1 (y_1 - y_2\gamma_0) &= P_1 ((Z_1\pi_1 + v)\delta + u) \\
&= Z_1 \left(C/\sqrt{n}\right)\delta + P_1 (v\delta + u).
\end{aligned}$$

Next, suppose for simplicity that $l_1 = 1$.

$$\begin{aligned}
&(y_1 - y_2\gamma_0)' P_1 (y_1 - y_2\gamma_0) \\
&= \frac{Z_1'Z_1}{n}C^2\delta^2 + (v\delta + u)' P_1 (v\delta + u) + 2C\delta\frac{Z_1'(v\delta + u)}{\sqrt{n}} \\
&\to_d Q_{11}C^2\delta^2 + (\Psi_1\delta + \Phi_1)^2 Q_{11}^{-1} + 2C\delta(\Psi_1\delta + \Phi_1) \\
&= \left(C\delta\sqrt{Q_{11}} + \frac{\Psi_1\delta + \Phi_1}{\sqrt{Q_{11}}}\right)^2.
\end{aligned}$$

We conclude that under the fixed alternatives,

$$AR(\gamma_0) \to_d \left(C\delta\sqrt{\frac{Q_{11}}{\sigma_v^2\delta^2 + \sigma_u^2 + 2\sigma_{uv}\delta}} + \frac{\Psi_1\delta + \Phi_1}{\sqrt{Q_{11}\left(\sigma_v^2\delta^2 + \sigma_u^2 + 2\sigma_{uv}\delta\right)}}\right)^2.$$

Note that

$$\frac{\Psi_1\delta + \Phi_1}{\sqrt{Q_{11}\left(\sigma_v^2\delta^2 + \sigma_u^2 + 2\sigma_{uv}\delta\right)}} \sim N(0,1),$$

and, therefore, $AR(\gamma_0)$ has asymptotically noncentral χ_1^2 distribution with the noncentrality parameter

$$\frac{C^2\delta^2 Q_{11}}{\sigma_v^2\delta^2 + \sigma_u^2 + 2\sigma_{uv}\delta}.$$

Suppose that one rejects the null when $AR(\gamma_0) > \chi^2_{1,1-\alpha}$. First, note that the test is inconsistent against fixed alternatives. This is due to weakness of the IVs. Second, as usual the power of the test increases with the distance from the null δ. However, the test will have poor power if the instruments are very weak (small C). In particular, when the instruments and endogenous regressor are unrelated ($C = 0$), $AR(\gamma_0)$ has the same central χ_1^2 distribution regardless of the value of δ. Consequently, for all values of δ, the asymptotic power is equal to the size α.

When $l_1 > 1$,

$$(y_1 - y_2\gamma_0)' P_1 (y_1 - y_2\gamma_0) = (y_1 - y_2\gamma_0)' Z_1/\sqrt{n}\,(Z_1'Z_1/n)^{-1} Z_1'(y_1 - y_2\gamma_0)/\sqrt{n}.$$

Next,

$$\begin{aligned} Z_1'(y_1 - y_2\gamma_0)/\sqrt{n} &= Z_1'\left(\left(Z_1C/\sqrt{n} + v\right)\delta + u\right)/\sqrt{n} \\ &\to_d Q_{11}C\delta + \Psi_1\delta + \Phi_1, \end{aligned}$$

and

$$\begin{aligned} AR(\gamma_0)\, l_1 &\to_d \frac{(Q_{11}C\delta + \Psi_1\delta + \Phi_1)' Q_{11}^{-1}(Q_{11}C\delta + \Psi_1\delta + \Phi_1)}{\sigma_v^2\delta^2 + \sigma_u^2 + 2\sigma_{uv}\delta} \\ &= \left\| \frac{Q_{11}^{1/2}C\delta}{\sqrt{\sigma_v^2\delta^2 + \sigma_u^2 + 2\sigma_{uv}\delta}} + \frac{Q_{11}^{-1/2}(\Psi_1\delta + \Phi_1)}{\sqrt{\sigma_v^2\delta^2 + \sigma_u^2 + 2\sigma_{uv}\delta}} \right\|^2. \end{aligned}$$

Note that

$$\Psi_1\delta + \Phi_1 \sim N\left(0, \left(\sigma_v^2\delta^2 + \sigma_u^2 + 2\sigma_{uv}\delta\right) Q_{11}\right).$$

Since Q_{11} is positive definite, it can be decomposed as $Q_{11} = H\Lambda H'$, where Λ is the diagonal matrix of positive eigenvalues and $H'H = HH' = I$. Define $\Lambda^{1/2}$ to be the diagonal matrix composed of the square roots of the elements of Λ,

$$Q_{11}^{1/2} = H\Lambda^{1/2},$$

and

$$Q_{11}^{-1/2} = \Lambda^{-1/2}H'.$$

Note that

$$\begin{aligned} Q_{11}^{-1/2} Q_{11} Q_{11}^{-1/2\prime} &= \Lambda^{-1/2}H'H\Lambda H'H\Lambda^{-1/2} \\ &= I_{l_1}. \end{aligned}$$

Thus,

$$\frac{Q_{11}^{-1/2}(\Psi_1\delta + \Phi_1)}{\sqrt{\sigma_v^2\delta^2 + \sigma_u^2 + 2\sigma_{uv}\delta}} \sim N(0, I_{l_1}),$$

or

$$\frac{Q_{11}^{1/2}C\delta}{\sqrt{\sigma_v^2\delta^2 + \sigma_u^2 + 2\sigma_{uv}\delta}} + \frac{Q_{11}^{-1/2}(\Psi_1\delta + \Phi_1)}{\sqrt{\sigma_v^2\delta^2 + \sigma_u^2 + 2\sigma_{uv}\delta}} \sim N\left(\frac{Q_{11}^{1/2}C\delta}{\sqrt{\sigma_v^2\delta^2 + \sigma_u^2 + 2\sigma_{uv}\delta}}, I_{l_1}\right),$$

and therefore, in the case of fixed alternatives and weak IVs, $AR(\gamma_0)$ has asymptotic non-central $\chi^2_{l_1}$ distribution with the noncentarlity parameter

$$\delta^2 \frac{C' Q_{11} C}{\sigma_v^2 \delta^2 + \sigma_u^2 + 2\sigma_{uv}\delta}.$$

The econometrician should reject H_0 when $AR(\gamma_0) l_1 > \chi^2_{l_1,1-\alpha}$. Again, power of the test depends on the strength of the IVs C. The test has no power if $C = 0$. Note also that there is wasting of degrees of freedom: while we have only one restriction under H_0 (γ is a scalar), the number of degrees of freedom $l_1 \geq 1$.

One can also show that in the case of strong IVs, the AR test has power against local alternatives $\gamma = \gamma_0 + \delta/\sqrt{n}$.

Inference on γ: regression with exogenous regressors

The joint hypotheses on γ and β can be tested in exactly the same way as in the simple regression case. However, often econometricians are interested in testing hypotheses on γ while leaving β unrestricted. This can be done by projecting the IVs Z_1 onto the space orthogonal to Z_2. In this case, the AR statistic takes the following form. Define

$$\begin{aligned} \widetilde{Z}_1 &= M_2 Z_1, \\ \widetilde{P}_1 &= \widetilde{Z}_1 \left(\widetilde{Z}_1' \widetilde{Z}_1 \right)^{-1} \widetilde{Z}_1'. \end{aligned}$$

The AR statistic is given by

$$AR(\gamma_0) = \frac{(y_1 - y_2\gamma_0)' \widetilde{P}_1 (y_1 - y_2\gamma_0)/l_1}{(y_1 - y_2\gamma_0)' M (y_1 - y_2\gamma_0)/(n - l_1 - l_2)}.$$

Consider again the fixed alternative $H_1 : \gamma = \gamma_0 + \delta$. We have

$$\begin{aligned} \widetilde{P}_1 (y_1 - y_2\gamma_0) &= \widetilde{P}_1 \left(\left(Z_1 C/\sqrt{n} + v \right) \delta + u \right) \\ &= \left(Z_1' M_2 Z_1 \right)^{-1} Z_1' M_2 \left(\left(Z_1 C/\sqrt{n} + v \right) \delta + u \right), \text{ and} \\ \frac{Z_1' M_2 \left(\left(Z_1 C/\sqrt{n} + v \right) \delta + u \right)}{\sqrt{n}} &\to_d Q_{1\cdot 2} C\delta + \Psi_{1\cdot 2}\delta + \Phi_{1\cdot 2}. \end{aligned}$$

Next,

$$\begin{aligned} \frac{(y_1 - y_2\gamma_0)' M (y_1 - y_2\gamma_0)}{n} &= \frac{(v\delta + u)' M (v\delta + u)}{n} \\ &\to_p \sigma_v^2 \delta^2 + \sigma_u^2 + 2\sigma_{uv}\delta. \end{aligned}$$

Combining the above results,

$$AR(\gamma_0) l_1 \to_d \frac{(Q_{1\cdot 2} C\delta + \Psi_{1\cdot 2}\delta + \Phi_{1\cdot 2})' Q_{1\cdot 2}^{-1} (Q_{1\cdot 2} C\delta + \Psi_{1\cdot 2}\delta + \Phi_{1\cdot 2})}{\sigma_v^2 \delta^2 + \sigma_u^2 + 2\sigma_{uv}\delta}.$$

Further,

$$\Psi_{1\cdot 2}\delta + \Phi_{1\cdot 2} \sim N\left(0, \left(\sigma_v^2 \delta^2 + \sigma_u^2 + 2\sigma_{uv}\delta\right) Q_{1\cdot 2}\right),$$

and, therefore,

$$\frac{Q_{1\cdot 2}^{-1/2} (Q_{1\cdot 2} C\delta + \Psi_{1\cdot 2}\delta + \Phi_{1\cdot 2})}{\sqrt{\sigma_v^2 \delta^2 + \sigma_u^2 + 2\sigma_{uv}\delta}} \sim N\left(\frac{Q_{1\cdot 2}^{1/2'} C\delta}{\sqrt{\sigma_v^2 \delta^2 + \sigma_u^2 + 2\sigma_{uv}\delta}}, I_{l_1} \right).$$

Hence, asymptotically $AR(\gamma_0)\,l_1$ has the noncentral $\chi^2_{l_1}$ distribution with the noncentrality parameter given by

$$\delta^2 \frac{C' Q_{1\cdot 2} C}{\sigma_v^2 \delta^2 + \sigma_u^2 + 2\sigma_{uv}\delta}.$$

Suppose that one rejects the null if $AR(\gamma_0)\,l_1 > \chi^2_{l_1,1-\alpha}$. When the null hypothesis is true ($\delta = 0$) or when the instruments are unrelated to the endogenous regressor ($C = 0$), $AR(\gamma_0)\,l_1$ has the central $\chi^2_{l_1}$ distribution, and

$$P\left(AR(\gamma_0)\,l_1 > \chi^2_{l_1,1-\alpha}\right) \to \alpha.$$

However, when the instruments are weak, and the null is false, the asymptotic power of the AR test exceeds α. The probability to reject the null increases with the magnitude of δ and C.

Confidence intervals for γ

Asymptotically valid confidence intervals for γ can be constructed by inverting the test based on the AR statistic (note that the usual confidence interval is the inverted t test). The robust to weak IVs confidence intervals are constructed as follows.

$$CI_{1-\alpha} = \left\{\gamma^* : AR(\gamma^*)\,l_1 < \chi^2_{l_1,1-\alpha}\right\}.$$

Thus, one collects all values of γ for which the null hypotheses $H_0 : \gamma = \gamma^*$ cannot be rejected. Let γ_0 be the true value of γ.

$$\begin{aligned} P(\gamma_0 \in CI_{1-\alpha}) &= P\left(AR(\gamma_0)\,l_1 < \chi^2_{l_1,1-\alpha}\right) \\ &\to 1-\alpha. \end{aligned}$$

Hence, such confidence intervals have a correct asymptotic coverage probability. The length of the confidence interval depends on the strength of the IVs. If the IVs and endogenous variable are unrelated, the AR test always accepts the null (with asymptotic probability $1-\alpha$). In this case, $CI_{1-\alpha}$ is infinite (contains the whole real line).

Multiple endogenous regressors

In the case of a vector of endogenous regressors ($Y_2 \in R^{n\times k}, \gamma \in R^k$), the AR statistic can be constructed in the same way as before:

$$AR(\gamma_0) = \frac{(y_1 - Y_2\gamma_0)' \widetilde{P}_1 (y_1 - Y_2\gamma_0)/l_1}{(y_1 - Y_2\gamma_0)' M (y_1 - Y_2\gamma_0)/(n - l_1 - l_2)}.$$

One should reject the null when $AR(\gamma_0)\,l_1 > \chi^2_{l_1,1-\alpha}$.