Advanced Econometrics II
TA Session Problems No. 2

Agnieszka Borowska

13.01.2015

Note: this is only a draft of the problems discussed on Tuesday and might contain some typos or more or less imprecise statements. If you find some, please let me know.

Restrictions testing

Remark: with the IV estimators, we focus on their **asymptotic** distributions, i.e. large-sample test. The reason for this is that the **finite sample** properties the IV estimators are usually unknown, so that exact tests are unavailable.

The model:

$$y = X\beta_0 + u, \qquad \mathbb{E}(uu^T) = \sigma_0^2 \mathbb{I}_n, \quad \mathbb{E}(u_t|X_t) \neq 0, \quad \mathbb{E}(u_t|W_t) = 0.$$

Consider the partition of X: $X = [X_1 \; X_2]$ and the corresponding partition of β: $\beta = [\beta_1 \; \beta_2]$, where X_1 is $n \times k_1$ and X_2 is $n \times k_2$. We wish to test

$$\begin{array}{ll} H_0: & \beta_2 = \beta_{20}, \\ H_1: & \beta_2 \neq \beta_{20}. \end{array} \tag{1}$$

t test

When β_2 is $n \times 1$, so we have a single restriction, we will call it β_i. Then, the test statistic is given by

$$t_{\beta_i} = \frac{\hat{\beta}_i - \beta_{i0}}{\left(\widehat{\mathrm{Var}}(\hat{\beta}_i)\right)^{-1}} \overset{a}{\sim} \mathcal{N}(0,1). \tag{8.47}$$

Wald test

When β_2 is $n \times k_2$, so we have k_2 restrictions, the Wald statistic has the form

$$W_{\beta_2} = \left(\hat{\beta}_2 - \beta_{20}\right)^T \left(\widehat{\mathrm{Var}}(\hat{\beta}_2)\right)^{-1} \left(\hat{\beta}_2 - \beta_{20}\right) \overset{a}{\sim} \chi^2(k_2) \tag{8.48}$$

IVGNR based tests

Alternatively, we can use the IV variant of the **Gauss-Newton (artificial) regression**[1] (IVGNR) to test for (1). Recall that the IVGNR has the form

$$\underbrace{y - X\beta}_{\text{regressand}} = \underbrace{P_W X}_{\text{regressor}} b + \text{ residuals},$$

[1]The details regarding the Gauss-Newton regression can be found in Chapter 6.5 in DM.

where now β is a *parameter* vector at which the regressand is evaluated, so that it is b which is the parameter estimated in this regression.

If we test for *linear* restrictions, then without loss of generality we can take $\beta_{20} = 0$, so that (1) can be written as

$$H_0 : y = X_1\beta_1 + u,$$
$$H_1 : y = X_1\beta_1 + X_2\beta_2 + u.$$

Important: use the same instrument matrix W for both, H_0 and H_1.

Then, the IVGNR corresponding to H_0 and H_1 are given by

$$\text{IVGNR}_0: \; y - X_1\acute{b}_1 = P_W X_1 b_1 + \text{residuals}, \tag{8.53}$$
$$\text{IVGNR}_1: \; y - X_1\acute{b}_1 = P_W X_1 b_1 + P_W X_2 b_2 + \text{residuals}, \tag{8.54}$$

Important: evaluate both IVGNRs at the same parameter values $[\acute{b}_1 \; \acute{b}_2]$, which **must satisfy the null**, so that $\acute{b}_2 = 0$. Moreover, $\acute{b}$ needs to be a consistent estimator under the null.

Then, the asymptotically valid test of H_0 against H_1 is provided by the **artificial F statistic**, obtained from running the two above IVGNRs, i.e.

$$F = \frac{(\text{SSR}_0 - \text{SSR}_1)/k_2}{\text{SSR}_1/(n-k)}, \tag{8.55}$$

which, when multiplied by k_2, under the null asymptotically follows $\chi^2(k_2)$ distribution.

For convenience, denote

$$Z = P_W X,$$
$$Z_1 = P_W X_1,$$
$$Z_2 = P_W X_2.$$

Recall that in the OLS setting

$$y = X\beta + u$$

the SSR can be expressed as

$$\text{SSR} = y^T M_X y,$$

where for an orthogonal projection matrix P_A, the matrix $M_A = \mathbb{I} - P_A$ is a matrix of the complementary projection. Similarly, we have now

$$\text{SSR}_0 = \left(y - X_1\acute{b}_1\right)^T M_{Z_1} \left(y - X_1\acute{b}_1\right),$$
$$\text{SSR}_1 = \left(y - X_1\acute{b}_1\right)^T M_{Z} \left(y - X_1\acute{b}_1\right),$$

Hence, k_2 times the numerator in (8.55) is given by

$$\begin{aligned} \text{SSR}_0 - \text{SSR}_1 &= \left(y - X_1\acute{b}_1\right)^T (M_{Z_1} - M_Z) \left(y - X_1\acute{b}_1\right) \\ &= \left(y - X_1\acute{b}_1\right)^T (P_Z - P_{Z_1}) \left(y - X_1\acute{b}_1\right), \end{aligned}$$

If we evaluate the nominator at $\acute{b} = \beta_0$, the true parameter value[2], then $y - X\beta_0 = u$, so the above expression simplifies to

$$\text{SSR}_0 - \text{SSR}_1 = u^T (P_Z - P_{Z_1})\, u, \tag{8.58}$$

i.e. a **quadratic form** in u, the vector of error terms, and the difference of two projection matrices, which is here also an orthogonal projection matrix, projecting on to a space of dimension $k - k_1 = k_2$.

[2] We can do this as the value of the parameter at which we evaluate the IVGNR regressands does not influence the *difference* between both SSRs.

What is the **distribution** of (8.58)? If we *assume normality* of u and fix X and W, then by Thm 4.1.2[3] we obtain

$$\frac{u^T\left(P_Z - P_{Z_1}\right)u}{\sigma_0^2} \sim \chi^2(k_2)$$

Before moving to the main exercise, recall the **Frisch-Waugh-Lovell theorem**.

Thm. 2.1. *Consider the following two regressions*:

$$y = X_1\beta_1 + X_2\beta_2 + u,$$
$$M_1 y = M_1 X_2 \beta_2 + \text{ residuals.}$$

The OLS estimates of β_2 from both regressions are numerically identical. Also, the residuals from both regressions are numerically identical.

DM 8.18 (part of)

Show that k_2 times the artificial F statistic from the pair of IVGNRs (8.53) and (8.54) is ***asymptotically equal*** *to the Wald statistic (8.48). Why are these two statistics* ***not numerically identical****?*

- First, consider the the artificial F statistic (8.55).

  To start with, let's deal with its *denominator*, given by

  $$\text{SSR}_1/(n-k) = \frac{1}{n-k}\left(y - X_1\acute{b}_1\right)^T M_Z \left(y - X_1\acute{b}_1\right).$$

  This is simply the estimate of the error variance from the IVGNR (8.54), we will denote it by $\acute{\sigma}^2$. It can be shown[4] that it consistently estimates the error variance, i.e. it tends to σ_0^2 with $n \to \infty$.

  Next, analyse its *nominator* multiplied by k_2. Application of the FWL theorem to IVGNR_1 means that SSR_1 is the same as the one from the FWL regression

  $$M_{Z_1}\left(y - X_1\acute{b}_1\right) = M_{Z_1} Z_2 b_2 + \text{ residuals.} \tag{2}$$

  Notice, that the OLS estimate of b_2 from (2) is given by

  $$\hat{b}_2 = \left(Z_2^T M_{Z_1} Z_2\right)^{-1} Z_2^T M_{Z_1} M_{Z_1}\left(y - X_1\acute{b}_1\right).$$

  Next, recall the SSR from OLS was can be expresses as

  $$y^T y - y^T X\hat{\beta}.$$

  Similarly, the SSR from (2) takes the form

  $$\underbrace{\left(y - X_1\acute{b}_1\right)^T M_{Z_1}\left(y - X_1\acute{b}_1\right)}_{\text{SSR}_0} - \underbrace{\left(y - X_1\acute{b}_1\right)^T M_{Z_1} Z_2 \left(Z_2^T M_{Z_1} Z_2\right)^{-1} Z_2^T M_{Z_1}\left(y - X_1\acute{b}_1\right)}_{(*)}.$$

  Hence, k_2 times $(*)$ is the nominator of the artificial F statistic. Notice that we can simplify $(*)$ as

  $$\begin{aligned} Z_2^T M_{Z_1} X_1 &= Z_2^T\left(\mathbb{I} - P_{Z_1}\right)X_1 \\ &= Z_2^T X_1 - Z_2^T P_{Z_1} X_1 \\ &= Z_2^T X_1 - Z_2^T Z_1\left(Z_1^T Z_1\right)^{-1} Z_1^T X_1 \\ &= (P_W X_2)^T X_1 - (P_W X_2)^T P_W X_1\left(X_1^T P_W X_1\right)^{-1} X_1^T P_W X_1 \\ &= (P_W X_2)^T X_1 - X_2^T P_W X_1 \\ &= \mathbb{O}. \end{aligned}$$

  Then, $(*)$ becomes

  $$\underbrace{y^T M_{Z_1} Z_2\left(Z_2^T M_{Z_1} Z_2\right)^{-1} Z_2^T M_{Z_1} y}_{(**)} \tag{3}$$

[3] **Thm 4.1.2.** If P is a projection matrix with rank r and z is an n-vector that is distributed as $\mathcal{N}(0, \mathbb{I})$, then the quadratic form $z^T P z$ is distributed as $\chi^2(r)$.

[4] Cf. DM, Ex. 8.16.

- Second, consider the Wald W statistic (8.48)

$$W_{\beta_2} = \hat{\beta}_2^T \left(\widehat{\text{Var}}(\hat{\beta}_2)\right)^{-1} \hat{\beta}_2.$$

It is a quadratic form in vector $\hat{\beta}_2$ and the inverse of the covariance matrix of that vector. To find the formula for $\hat{\beta}_2$, consider the second-stage 2SLS regression

$$\begin{aligned} y &= P_W X_1 \beta_1 + P_W X_2 \beta_2 + \text{ residuals} \\ &= Z_1 \beta_1 + Z_2 \beta_2 + \text{ residuals}. \end{aligned}$$

Then, the application of the FWL theorem yields

$$M_{Z_1} y = M_{Z_1} Z_2 \beta_2 + \text{ residuals},$$

so that

$$\hat{\beta}_2 = \left(Z_2^T M_{Z_1} Z_2\right)^{-1} Z_2^T M_{Z_1} y,$$

with the corresponding estimate of the covariance matrix

$$\hat{\sigma}^2 \left(Z_2^T M_{Z_1} Z_2\right)^{-1},$$

where

$$\hat{\sigma}^2 = \frac{||y - X\hat{\beta}_{IV}||^2}{n}$$

is the IV estimate of σ_0^2. Combining these two results, we can express the Wald statistic as

$$\begin{aligned} W_{\beta_2} &= \left(\left(Z_2^T M_{Z_1} Z_2\right)^{-1} Z_2^T M_{Z_1} y\right)^T \left(\hat{\sigma}^2 \left(Z_2^T M_{Z_1} Z_2\right)^{-1}\right)^{-1} \left(Z_2^T M_{Z_1} Z_2\right)^{-1} Z_2^T M_{Z_1} y \\ &= \frac{1}{\hat{\sigma}^2} y^T M_{Z_1} Z_2 \left(Z_2^T M_{Z_1} Z_2\right)^{-1} Z_2^T M_{Z_1} Z_2 \left(Z_2^T M_{Z_1} Z_2\right)^{-1} Z_2^T M_{Z_1} y \\ &= \frac{1}{\hat{\sigma}^2} \underbrace{y^T M_{Z_1} Z_2 \left(Z_2^T M_{Z_1} Z_2\right)^{-1} Z_2^T M_{Z_1} y}_{(**)} \end{aligned} \tag{4}$$

- Finally, notice that the terms denoted with (∗∗) in (3) and (4) are identical, so that

$$\begin{aligned} k_2 F &= \frac{(**)}{\acute{\sigma}^2}, \\ W_{\beta_2} &= \frac{(**)}{\hat{\sigma}^2}. \end{aligned}$$

Thus, both quantities differ only wrt their denominators. They are not the same because SSR_1 used in k_2F is **not** the same as the SSR from the IV estimation of the unrestricted model used in W_{β_2}. And it is this difference in the denominators which makes both quantities **not numerically identical**.

- However, since the denominator of the artificial F statistic is asymptotically equal to σ_0^2 and the IV estimator for the variance of the error terms $\hat{\sigma}^2$ is consistent (which means it is also asymptotically equal to σ_0^2), we can conclude that k_2F and W_{β_2} are indeed **asymptotically equal**.