Tests for Endogeneity and Instrument Suitability*

PHOEBUS J. DHRYMES
Columbia University

June 2003

Abstract

This paper deals with two alternatives to the so-called Hausman test for the exogeneity of instruments, in the context of a model where one or more explanatory variables are possibly correlated with the structural error. These two alternatives are at least as good or better than the Hausman test and are much simpler to carry out.

A small Monte Carlo study illustrates this result.

(JEL C10, C30, C43)

Key Words: Hausman test, instrumental variables, residual correlation test.

An earlier version of this paper was delivered as the ASSET Lecture in the annual meetings of the Association of Southern European Economic Theorists, Paphos Cyprus, November 1 and 2, 2002.
I would like to thank Timothy McKenna for his excellent research assistance in programming and executing the Monte Carlo study.

1 Introduction

In this paper we deal with the issue of how to determine whether in the following simple model

$$y_i = x_{i\cdot}\beta + u_i, \quad x_{ik} = z_{i\cdot}\gamma + v_i, \quad i = 1, 2, \ldots n, \tag{1}$$

x_{ik} is or is not **endogenous** or, more precisely, is or is not correlated with the error term u_i. In Eq. (1), $x_{i\cdot}$ is a k-element **row** vector whose last element, x_k, is suspected to be endogenous, in the sense that it is correlated with the structural error u_i; similarly, $z_{i\cdot}$ is an m-element row vector of "instruments", which are asserted to be independent of the structural errors u_i and v_i. The vectors $w_{i\cdot} = (u_i, v_i)$, $i = 1, 2, \ldots, n$ are asserted to be independent identically distributed (i.i.d.) with mean zero and an unrestricted positive definite matrix, i.e.

$$Ew_i' = 0, \quad \mathrm{Cov}(w_i') = \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} > 0. \tag{2}$$

The literature on this model has a long history because the possibility that one or more of the explanatory variables in a regression model is correlated with the error term can arise for a number of reasons. The early literature, Wald (1940), dealt with the simplest version of the model above as an error in variables model. It was inspired by, and connected to, physical science issues such as the Heisenberg uncertainty principle which introduces, in scientific measurements, errors of observation. Wald solved this problem by inventing the notion of instrumental variables although, at the time, he was not aware of this concept *per se* and its other potential uses. Wald's solution[1] noted that **if** one could rank the observations on the **independent** variable by their **true** value, then one could consistently estimate the parameters of the (regression) line, because a line is determined by two points. Using this additional information one could select the two points in such a way that the slope is well determined and converges to the true slope, at least in probability. Berkson (1950) noted that OLS estimators

[1] For a discussion of Wald's paper see Dhrymes (1978), pp. 245-47.

in such models would be consistent and unbiased **if** one assumes that the scientist **sets** the values of the explanatory variable(s)-in fact, ensuring that the standard assumptions of the general linear model hold. Durbin (1954) notes that of course such a rationale may be fine in the natural sciences but not in economics. He correctly saw the problem in its contemporary setting, i.e. as an issue in simultaneous equations theory. In fact, he explicitly introduced what has come to be known as the "Hausman" test. However, at the time it was not well understood how to handle "excess" instruments, and his "Hausman" statistic is not as efficient as one could make it. Wu (1973) extended Durbin's work and produced what has become known as the "Hausman" test. Hausman (1978) has given a broader rationale for such a test and gave several examples of its possible applications.[2] Oddly enough even though Durbin, Wu and Hausman all note that the crucial aspect of this problem is the relationship between the error terms in the two equations, none make this the focal point of their analysis. By contrast, this paper concentrates entirely on this facet and motivates tests on the basis of this relationship **only**.

Faced with the model of Eq. (1), practitioners are likely to estimate the parameters of the (first equation of the) model by "2SLS".

The 2SLS procedure customarily employed in the applied literature regresses, in the first stage, $x_{\cdot k}$ on $Z = (z_{i\cdot})$ to obtain

$$\hat{\gamma} = (Z'Z)^{-1}Z'x_{\cdot k}, \quad \hat{x}_{\cdot k} = Z\hat{\gamma}, \quad \hat{v} = x_{\cdot k} - \hat{x}_{\cdot k}. \tag{3}$$

One then defines $\hat{X} = (X_1, \hat{x}_{\cdot k})$, where $X_1 = (x_{\cdot 1}, x_{\cdot 2}, \ldots, x_{\cdot k-1})$ and

$$y = \hat{X}\beta + u + \beta_k \hat{v}; \tag{4}$$

in the second stage one obtains the 2SLS estimator of the structural parameter β as

$$\hat{\beta} = (\hat{X}'\hat{X})^{-1}\hat{X}'y = \beta + (\hat{X}'\hat{X})^{-1}\hat{X}'[u + \beta_k \hat{v}]. \tag{5}$$

[2]For a comment on the Hausman test, and its unreliability when applied to testing prior restrictions in structural (simultaneous) equations models, see Dhrymes (1994a); for a discussion of its general nature see Dhrymes (1994b), pp. 52-60 .

To justify, motivate, or compel estimation of parameters by "2SLS", it is a widespread practice in this literature to test for the endogeneity (or exogeneity) of x_k by the Hausman test.

It is also a widespread practice to divide the matrix of the "instruments", Z, as $Z = (Z_1, Z_2)$, where the two constituent matrices are of dimension $n \times m_1$ and $n \times m_2$, respectively. It is asserted that Z_1 contains instruments that are independent of the error term, but there is some doubt as to whether Z_2 has the same properties. The suitability of the instruments in Z_2 is also judged by means of the Hausman test, i.e. one obtains the estimators of Eq. (1) by using Z and Z_1 alone, and then obtains the "Hausman statistic" based on the difference $\hat{\beta}(Z) - \hat{\beta}(Z_1)$. One or more of these procedures is recommended in widely used textbooks such as Greene (2000), Chapters 9 and 14, and Davidson and Mackinnon (1993), Chapter 7, pp. 240ff.

This author, Dhrymes (1994a), has pointed out that, in the context of a complete system of equations for which the Hausman test was originally devised, results obtained by the application of such test are model dependent, and often unreliable. The problem is that the Hausman test does not have a specific parametric hypothesis to test, the results are difficult to interpret and in some instances prove to be quite unreliable when the null is false.

In this paper we shall devise appropriate tests to address both questions noted above.

2 Testing for Exogeneity

2.1 Structured Models

Proceeding somewhat formally, notice that if we consider the equations of Eq. (1) a complete model, a necessary and sufficient condition to obtain consistent and efficient estimators of the underlying parameters by **least squares** is that the system be **simply recursive**. If that holds and, in addition, the distribution of the errors is **jointly normal** the resulting estimators are also maximum likelihood (ML) estimators and will possess the property of sufficiency.

The model as formulated above is simply recursive if and only if

$$\sigma_{12} = 0. \tag{6}$$

Hence, the hypothesis of exogeneity may be formulated as

$H_0: \quad \sigma_{12} = 0$
as against the alternative
$H_1: \quad \sigma_{12} \neq 0\,.$

Under **the assumption of normality** the likelihood function is given by

$$L_n^*(\beta, \gamma, \Sigma) = (2\pi)^{-n} |\Sigma|^{-(n/2)} e^{-(1/2)\mathrm{tr}\Sigma^{-1}S}, \quad S = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix}, \tag{7}$$

where $s_{11} = (y - X\beta)'(y - X\beta)\,,\ s_{12} = (y - X\beta)'(x_{\cdot k} - Z\gamma)\,,\ s_{21} = (s_{12})'\,,$ $s_{22} = (x_{\cdot k} - Z\gamma)'(x_{\cdot k} - Z\gamma)\,.$ There are two ways this test may be implemented. We can employ the **likelihood ratio** test (LRT), or we can devise a test based **entirely** on the null.

The LRT statistic is obtained as

$$\lambda = \frac{\max_{H_0} L_n^*}{\max_{H_1} L_n^*}. \tag{8}$$

As is well known, **under the null**, the maximum likelihood (ML) estimator of the parameters in question is the OLS estimator,

$$\hat{\beta}_{OLS} = (X'X)^{-1}X'y, \quad \hat{\gamma} = (Z'Z)^{-1}Z'x_{\cdot k}. \tag{9}$$

Consequently,

$$\hat{\sigma}_{11} = \frac{\hat{u}'\hat{u}}{n}, \quad \hat{\sigma}_{22} = \frac{\hat{v}'\hat{v}}{n}, \hat{u} = y - X\hat{\beta}, \quad \hat{v} = x_{\cdot k} - Z\hat{\gamma}, \quad \text{so that}$$

$$\max_{H_0} L_n^* = [\hat{\sigma}_{11}\hat{\sigma}_{22}]^{-(n/2)}][(2\pi)^{-n}e^{-n}]. \tag{10}$$

Under the **alternative**, the ML estimator of the underlying parameters is the **3SLS** estimator

$$\hat{\delta}_{3SLS} = \begin{pmatrix} \hat{\beta} \\ \hat{\gamma} \end{pmatrix} = \begin{bmatrix} \hat{\sigma}^{11}\hat{X}'\hat{X} & \hat{\sigma}^{12}X'Z \\ \hat{\sigma}^{21}Z'X & \hat{\sigma}^{22}Z'Z \end{bmatrix}^{-1} \begin{bmatrix} \hat{\sigma}^{11}\hat{X}'y + \hat{\sigma}^{12}\hat{X}'x_{\cdot k} \\ \hat{\sigma}^{21}Z'y + \hat{\sigma}^{22}Z'x_{\cdot k} \end{bmatrix}. \tag{11}$$

This is so "because" in this case the Jacobian of the transformation from (u_i, v_i) to (y_i, x_{ik}) is **unity**.[3]

Thus, **under the alternative**

$$\max_{H_1} L_n^* = [\hat{\sigma}_{11}\hat{\sigma}_{22}(1-r_{12}^2)]^{-(n/2)}][(2\pi)^{-n}e^{-n}], \quad r_{12}^2 = \frac{\hat{\sigma}_{12}^2}{\hat{\sigma}_{11}\hat{\sigma}_{22}} \tag{12}$$

and, consequently,

$$\lambda = \left(\frac{(\hat{\sigma}_{11}\hat{\sigma}_{22})_{OLS}}{(\hat{\sigma}_{11}\hat{\sigma}_{22})_{3SLS}(1-r_{12}^2)}\right)^{-(n/2)} \sim (1-r_{12}^2)^{n/2}, \quad \text{or} \quad r_{12}^2 \sim 1-\lambda^{2/n}, \tag{13}$$

which shows that the square of the correlation coefficient between the 3SLS residuals of the first and second equation is a **likelihood ratio ststistic**.

What is suggested by the discussion above is that we may base the test of exogeneity on the 3SLS residuals from the first equation and the generalized least squares residuals from the second, using the sample covariance. It can be easily shown that

$$\frac{\hat{v}'_{GLS}\hat{u}_{3SLS}}{\sqrt{n}} \sim \frac{v'u}{\sqrt{n}} - \sigma_{22}(\hat{\beta}-\beta)_{k(3SLS)}, \tag{14}$$

and moreover that

$$\frac{\hat{v}'_{OLS}\hat{u}_{2SLS}}{\sqrt{n}} \sim \frac{v'u}{\sqrt{n}} - \sigma_{22}(\hat{\beta}-\beta)_{k(2SLS)}. \tag{15}$$

Although, as seen from Eqs. (14) and (15) the results are somewhat different when we use 2SLS residuals from the first equation and OLS residuals from the second, we shall examine the test based on the **two stage least squares** residuals from the first equation and the OLS residuals from the second in order to conform with much of the empirical practice, even

[3] Actually, the simplification induced by the unit Jacobian is overstated above. The ML estimator in this case is the 3SLS estimator **iterated to convergence**, so that the entities $\hat{\sigma}^{ij}$, (normally) the 2SLS estimators, should be understood to be the 3SLS estimators of the covariance parameters. But this aspect plays no appreciable role in the discussion to follow. For a more general discussion of the relation between 3SLS and ML estimators see Dhrymes (1973).

though the LRT requires the use of 3SLS residuals[4] from the first and GLS residuals from the second equation, i.e. we consider

$$\frac{\hat{u}\tilde{v}}{\sqrt{n}} = \frac{1}{\sqrt{n}} u'[I_n - X(\hat{X}'\hat{X})^{-1}\hat{X}'][I_n - Z(Z'Z)^{-1}Z']v, \tag{16}$$

where $\hat{X} = (X_1, \hat{x}_{\cdot k})$, $\hat{x}_{\cdot k} = Z(Z'Z)^{-1}Z'x_{\cdot k}$. As it will turn out, the test statistic in Eq. (16) has precisely the same limiting distribution as that in Eq. (17) below, viz.

$$\frac{1}{\sqrt{n}} x'_{\cdot k}\hat{u} = x'_{\cdot k}[I_n - X(\hat{X}'\hat{X})^{-1}\hat{X}']u. \tag{17}$$

Alternatively, we may estimate the model under the null, in which case we find the OLS estimates of β and γ, obtain the residuals and then carry out the test. The first procedure is a **conformity** test, i.e. we estimate the parameters without imposing the restriction of the null and then we ask whether the results conform to the requirements of the null.[5] In the second alternative we operate entirely under the null, i.e. we estimate all parameters assuming the null to be true, and then ask whether the null is supported by the evidence. We shall establish the properties of both tests.

2.2 Derivation and Properties of the Test Statistics

In this and subsequent sections certain matrices will recur frequently and so we shall employ the following notation for ease of exposition. A matrix of the form $Q(Q'Q)^{-1}Q'$, will be routinely denoted by

$$P_q = Q(Q'Q)^{-1}Q', \tag{18}$$

it being a projection matrix,[6] i.e. for any suitably dimensioned vector, y, $P_q y$ gives the projection of y on the space spanned by the columns of Q.

[4]Notice that **under the null** 2SLS and 3SLS are asymptotically equivalent. See also the discussion in Appendix II.

[5]For reasons unclear to me such tests are referred to in the literature as Wald tests.

[6]For a discussion of the projection theorem see, e.g. Dhrymes (1998), Chapter 2.

Because in what follows a certain problem will recur frequently we examine the issue in question before we proceed. Consider the entities

$$v'P_{\hat{x}}u, \quad v'P_z u, \quad v'P_x u, \quad v'X(\hat{X}'\hat{X})^{-1}\hat{X}'u, \quad v'\hat{X}(\hat{X}'\hat{X})^{-1}X'u.$$

Under the null, all of these entities upon division by $\sqrt{n}$, converge to zero **at least in probability**. Under the alternative, however, their behavior is quite **different**. Thus,

$$\begin{aligned}
\frac{1}{\sqrt{n}}v'P_{\hat{x}}u &\xrightarrow{\mathrm{P}} 0,\\
\frac{1}{\sqrt{n}}v'P_z u &\xrightarrow{\mathrm{P}} 0,\\
\frac{1}{\sqrt{n}}v'P_x u &\xrightarrow{\mathrm{d}} \pm\infty,\\
\frac{1}{\sqrt{n}}v'X(\hat{X}'\hat{X})^{-1}\hat{X}'u &\xrightarrow{\mathrm{d}} \sigma_{12}e'_{\cdot k}\hat{B}\frac{\hat{X}'u}{\sqrt{n}}, \quad \hat{B} = \left(\frac{\hat{X}'\hat{X}}{n}\right)^{-1}\\
\frac{1}{\sqrt{n}}v'\hat{X}(\hat{X}'\hat{X})^{-1}X'u &\xrightarrow{\mathrm{d}} \sigma_{12}e'_{\cdot k}\hat{B}\frac{\hat{X}'v}{\sqrt{n}}, \qquad (19)
\end{aligned}$$

where $e_{\cdot k}$ is a k-element column vector all of whose elements are zero except for the last, which is unity.

The residuals from the 2SLS procedure are given by

$$\begin{aligned}
\hat{u} &= y - X\hat{\beta} = u - X(\hat{\beta}-\beta)\\
&= u - X(\hat{X}'\hat{X})^{-1}\hat{X}'[u+\beta_k\hat{v}]\\
&= [I - X(\hat{X}'\hat{X})^{-1}\hat{X}']u. \qquad (20)
\end{aligned}$$

The last equality follows because

$$Z = (X_1, P), \quad X_1 = ZI^*_{k-1}, \; I^*_{k-1} = (I_{k-1}, 0)', \; \hat{X} = Z(I^*_{k-1}, \hat{\gamma}),$$

so that $\hat{X}'\hat{v} = 0$.

The OLS residuals from the first stage are evidently given by

$$\hat{v} = [I - P_z]v. \tag{21}$$

It may be shown by direct computation that, under the null,

$$\frac{1}{\sqrt{n}}\hat{v}'\hat{u} = \frac{1}{\sqrt{n}}v'[I - P_z - (0, \hat{v})(\hat{X}'\hat{X})^{-1}\hat{X}']u, \tag{22}$$

because $P_z X = \hat{X}$. Moreover,

$$\frac{1}{n}v'(I - P_z)v \xrightarrow{\mathrm{P}} \sigma_{22},$$

so that the quantity in the right member of Eq. (11) behaves like

$$\frac{1}{\sqrt{n}}\hat{v}'\hat{u} \quad \sim \quad \frac{1}{\sqrt{n}}(v - s)'u, \quad \text{where} \tag{23}$$

$$s' \quad = \quad \sigma_{22}(\hat{B}_{21}, \hat{B}_{22})\hat{X}', \quad \begin{bmatrix} \hat{B}_{11} & \hat{B}_{12} \\ \hat{B}_{21} & \hat{B}_{22} \end{bmatrix} = \hat{B}. \tag{24}$$

Under the null, this a sequence of independent non-identically distributed random variables with mean zero and variance

$$\phi_i = E[(v_i - s_i)u_i]^2 = E[v_i^2 u_i^2 - 2s_i v_i u_i^2 + s_i^2 u_i^2]$$

$$= E[v_i^2 u_i^2] + E[s_i^2 u_i^2] = \sigma_{11}\sigma_{22} + \sigma_{11}s_i^2;$$

$$\sigma_{11}\omega = \lim_{n\to\infty} \sum_{i=1}^{n} \frac{1}{n}\phi_i = \sigma_{11}\sigma_{22}[1 + \sigma_{22}\hat{B}_{22}]. \tag{25}$$

The derivation of the results above assumes that the joint distribution of u and v is symmetric (which ensures that odd moments vanish) and normal which, under the null, ensures that u and v are mutually independent.

Moreover, the Lindeberg condition is satisfied, and hence by the Lindeberg CLT, see Dhrymes (1989), p. 271, we conclude that under the null

$$\frac{1}{\sqrt{n}}\hat{v}'\hat{u} \xrightarrow{\mathrm{d}} N(0, \sigma_{11}\omega), \ \omega = \sigma_{22}[1 + \sigma_{22}(\hat{R}_z - \hat{R}_{x_1})^{-1}]. \tag{26}$$

The entity in square brackets in the last equation of Eq. (14)is found as follows:

$$\begin{aligned}\operatorname*{plim}_{n\to\infty}\frac{1}{n}s's &= \operatorname*{plim}_{n\to\infty}\frac{1}{n}\left(\sigma_{22}^2(\hat{B}_{21},\hat{B}_{22})\left(\frac{\hat{X}'\hat{X}}{n}\right)(\hat{B}_{21},\hat{B}_{22})'\right)\\ &= \sigma_{22}^2\operatorname*{plim}_{n\to\infty}\left(\frac{1}{n}[\hat{x}'_{\cdot k}\hat{x}_k-\hat{x}_{\cdot k}P_{x_1}\hat{x}_{\cdot k}]\right)^{-1} = \sigma_{22}^2(\hat{R}_z-\hat{R}_{x_1})^{-1},\end{aligned}$$

because $P_{x_1}P_z = P_{x_1}$, and R_z, R_{x_1} are, respectively, the mean sum of regression squares in the regression of $x_{\cdot k}$ on Z and X_1. Thus, the statistic[7]

$$t_n = \frac{\hat{v}'\hat{u}}{\sqrt{n\hat{\sigma}_{11}\hat{\sigma}_{22}[1+\hat{\sigma}_{22}\hat{B}_{22}]}} \stackrel{\mathrm{d}}{\to} N(0,1), \tag{27}$$

may be used to test the hypothesis of exogeneity, where

$$\hat{\sigma}_{11} = \frac{1}{n}\hat{u}'\hat{u},\ \ \hat{\sigma}_{22} = \frac{1}{n}\hat{v}'\hat{v},$$

$$\hat{B}_{22}^{-1} = \frac{1}{n}\hat{x}'_{\cdot k}\hat{x}_{\cdot k} - \frac{\hat{x}'_{\cdot k}X_1}{n}\left(\frac{X_1'X_1}{n}\right)^{-1}\frac{X_1'\hat{x}_{\cdot k}}{n} = \hat{R}_z-\hat{R}_{x_1}, \tag{28}$$

where $\hat{R}_z,\ \hat{R}_{x_1}$ are, respectively, the means of the regression sum of squares in the regression of $x_{\cdot k}$ on Z and X_1.

A variant of this test, i.e. a direct test that x_k is correlated with the structural error, is obtained by considering the entity

$$\frac{1}{\sqrt{n}}x'_{\cdot k}\hat{u}. \tag{29}$$

This test is asymptotically and numerically identical with the previous test, as will be readily demonstrated by noting that

$$x_{\cdot k} = \hat{x}_{\cdot k}+\hat{v},\quad \hat{x}'_{\cdot k}\hat{u} = 0,\ \text{so that}\ \ x'_{\cdot k}\hat{u} = \hat{v}'\hat{u}.$$

[7]Note that, to avoid notational clutter, we use the symbol(s) $\hat{B}_{ij},\ i,j=1,2$ to mean both the elements of the block matrix $(\hat{X}'\hat{X}/n)^{-1}$, as well as their respective limits. Similarly for the symbols B_{ij} and their relation to $(X'X/n)^{-1}$.

Hence, a test can be carried out through the statistic above, viz.

$$t_n = \frac{x'_{\cdot k}\hat{u}}{\sqrt{n[\hat{\sigma}_{11}\hat{\sigma}_{22}(1+\hat{\sigma}_{22}\hat{B}_{22})]}} \xrightarrow{\mathrm{d}} N(0,1). \tag{30}$$

The alternative test, based entirely on the null, uses the residuals of the OLS regressions from both equations, viz. $\tilde{u} = (I - P_x)u$ and $\hat{v} = (I - P_z)v$ to form the statistic

$$\frac{1}{\sqrt{n}}\hat{v}'\tilde{u} = \frac{1}{\sqrt{n}}v'(I - P_z)(I - P_x)u. \tag{31}$$

Employing precisely the same argument as above, we establish that

$$\frac{1}{\sqrt{n}}\hat{v}'\tilde{u} \sim \frac{1}{\sqrt{n}}(v - s^*)'u = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(v_i - s_i^*)u_i,$$

$$s^{*\prime} = \sigma_{22}e'_{\cdot k}BX', \qquad B = \left(\frac{X'X}{n}\right)^{-1}. \tag{32}$$

To determine its limiting distribution we note that

$$s_i^* = \sigma_{22}e'_{\cdot k}B\bar{x}'_{i\cdot} + \sigma_{22}B_{22}v_i, \quad \bar{X} = (X_1, Z\gamma) \tag{33}$$

and, consequently,

$$v_i - s_i^* = (1 - \sigma_{22}B_{22})v_i - \sigma_{22}e'_{\cdot k}B\bar{x}'_i$$

$$\phi_i = E[(v_i - s_i^*)u_i]^2 = (1 - \sigma_{22}B_{22})^2\sigma_{11}\sigma_{22} + \sigma_{11}\sigma_{22}^2 e'_{\cdot k}B\bar{x}'_{i\cdot}\bar{x}_{i\cdot}Be_{\cdot k}$$

$$\sigma_{11}\omega_1 = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{\infty}\phi_i = \sigma_{11}\sigma_{22}[1 - \sigma_{22}B_{22}]. \tag{34}$$

By the Lindeberg CLT it follows that

$$\frac{1}{\sqrt{n}}\hat{v}'\tilde{u} \xrightarrow{\mathrm{d}} N(0, \sigma_{11}\omega_1), \tag{35}$$

where

$$\omega_1 = \sigma_{22}[1 - \sigma_{22}B_{22}], \quad \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}, \quad B_{22}^{-1} = \frac{1}{n}[x'_{\cdot k}(I - P_{x_1})x_{\cdot k}]. \tag{36}$$

Therefore a test of the hypothesis of exogeneity can be carried out through the test statistic

$$t_n^{(1)} = \frac{\hat{v}'\tilde{u}}{\sqrt{n\hat{\sigma}_{11}\hat{\sigma}_{22}(1-\hat{\sigma}_{22}B_{22})}} \xrightarrow{\mathrm{d}} N(0,1). \tag{37}$$

Remark 1. Notice that the estimator $\hat{\sigma}_{11}^2$ is based on the OLS resdiuals from the structural equation and that the test statistic above is not of the same form as that of Eq. (27), due to the fact that $\hat{B}_{22} \geq B_{22}$. Notice further that B_{22}^{-1} is **the mean sum of the squared residuals** in the regression of $x_{\cdot k}$ on X_1.

2.3 Unstructured Models

Suppose, now, that in Eq. (1) we do not have a **complete** model but we suspect that, either through measurement error, omitted variable, or simultaneity, the variable x_k may be correlated with the structural error u. The difference between this case and the earlier one is that a precisely specified set of instruments is not available, and there is no precise specification for the error vector v. Rather, there is under consideration one (or more) matrix of instruments, $W_1 = (X_1, P_1)$, whose elements are asserted to be independent of the structural error u, and the structural parameter vector β is to be estimated by instrumental variables (IV) methods. Following the procedure given initially in Dhrymes (1969) and applied subsequently in similar contexts by Amemiya (1974), Jorgenson and Laffont (1974), Hansen (1982), *inter alia*, put

$$R_1 R_1' = W_1' W_1 \tag{38}$$

and consider

$$R_1^{-1} W_1' y = R_1^{-1} W_1' X\beta + R_1^{-1} W_1' u. \tag{39}$$

As is well known, see Dhrymes (1970), pp. 302-303, 2SLS (OLS) is the optimal IV estimator in the context of the transformed model of Eq. (28), when the admissible class of instruments is $W_1 A$, where A is an arbitrary nonsingular matrix. The OLS estimator of β in the context of the transformed model in Eq. (28) is

$$\hat{\beta} = [X'P_{w_1}X]^{-1}X'P_{w_1}y = \beta + [X'P_{w_1}X]^{-1}X'P_{w_1}u. \tag{40}$$

On the assumption that

$$\frac{1}{\sqrt{n}}W_1'u \xrightarrow{\mathrm{d}} N(0, \sigma_{11}\Omega_{11}), \quad \Omega_{11} = \operatorname*{plim}_{n\to\infty} \frac{W_1'W_1}{n},$$

it may be easily shown that

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{\mathrm{d}} N(0, \sigma_{11}\Psi^{-1}), \quad \Psi = \operatorname*{plim}_{n\to\infty} \frac{X'P_{w_1}X}{n} = [M_{xw_1}\Omega_{11}^{-1}M_{w_1x}], \tag{41}$$

where the generic notation M_{xy} denotes the limit of second moment (or second cross moment) matrices.

To test the hypothesis that x_k is correlated with the structural error, we consider the entity

$$\frac{1}{\sqrt{n}}x_{\cdot k}'\hat{u}, \ \hat{u} = y - X\hat{\beta} = Cu, \ C = I - X[X'P_{w_1}X]^{-1}X'P_{w_1}, \tag{42}$$

which may be written more extensively as

$$\frac{1}{\sqrt{n}}x_{\cdot k}'\hat{u} = \frac{1}{\sqrt{n}}x_{\cdot k}'u - \frac{x_{\cdot k}'X}{n}\left(\frac{X'P_{w_1}X}{n}\right)^{-1}\frac{X'W_1}{n}\left(\frac{W_1'W_1}{n}\right)^{-1}\frac{1}{\sqrt{n}}W_1'u. \tag{43}$$

Under the null, the rightmost member of the equation above obeys the conditions of the Lindeberg CLT and, consequently,

$$\frac{1}{\sqrt{n}}x_{\cdot k}'\hat{u} \xrightarrow{\mathrm{d}} N(0, \sigma_{11}\omega_2). \tag{44}$$

This result is established as follows:

$$x_{\cdot k} = P_{w_1}x_{\cdot k} + (I - P_{w_1})x_{\cdot k}$$

$$P_{w_1}CC' = P_{w_1} - P_{\hat{x}}, \quad \hat{X} = P_{w_1}X, \ P_{\hat{x}} = \hat{X}(\hat{X}'\hat{X})^{-1}\hat{X}'$$

$$P_{w_1}CC'(I - P_{w_1}) = 0$$

$$(I-P_{w_1})X=(0,v), \quad v=(I-P_{w_1})x_{\cdot k}$$

$$(I-P_{w_1})CC'(I-P_{w_1})=I-P_{w_1}+(I-P_{w_1})X(\hat{X}'\hat{X})^{-1}X'(I-P_{w_1})$$

$$P_{w_1}CC'P_{w_1}=P_{w_1}-P_{\hat{x}}, \quad \text{so that}$$

$$\omega_2=\operatorname*{plim}_{n\to\infty}\frac{1}{n}x'_{\cdot k}(P_{w_1}-P_{\hat{x}})x_{\cdot k}+\sigma^*_{22}+\hat{B}_{22}\sigma^{*2}_{22},$$

$$\sigma^*_{22}=\operatorname*{plim}_{n\to\infty}\frac{x'_{\cdot k}(I-P_{w_1})x_{\cdot k}}{n}. \tag{45}$$

To show that, *mutatis mutandis*, ω_2 is the "same" as ω in the first two tests, note that

$$x'_{\cdot k}(P_{w_1}-P_{\hat{x}})x_{\cdot k}=0, \tag{46}$$

and further note that

$$x'_{\cdot k}X=\hat{x}'_{\cdot k}\hat{X}+(0,\ \hat{v}^{*\prime}\hat{v}^*),$$

because

$$x'_{\cdot k}(P_{w_1}-P_{\hat{x}})=\hat{x}'_{\cdot k}-\hat{x}'_{\cdot k}\hat{X}(\hat{X}'\hat{X})^{-1}\hat{X}'=\hat{x}'_{\cdot k}-e'_{\cdot k}\hat{X}'=0, \tag{47}$$

and identify σ_{22} with σ^*_{22}. It follows, then, that

$$t^{(2)}_n=\frac{x'_{\cdot k}\hat{u}}{\sqrt{n\hat{\sigma}_{11}\hat{\sigma}^*_{22}[1+\hat{\sigma}^*_{22}\hat{B}_{22}]}}\to N(0,1) \tag{48}$$

is a suitable test statistic for testing the hypothesis that $x_{\cdot k}$ is exogenous, when a complete model s not specified. In the preceding

$$\hat{\sigma}_{11}=\frac{\hat{u}'\hat{u}}{n}, \quad \hat{\sigma}^*_{22}=\frac{x'_{\cdot k}(I-P_{w_1})x_{\cdot k}}{n}.$$

In this section we have proved the following:

Proposition 1. Consider the model in Eq. (1), a complete model under the assumptions

- A1 $w_{i\cdot}=(u_i,v_i)$ is a sequence of i.i.d. random vectors with mean zero and covariance matrix $\Sigma>0$.

A2 The variables in the vector $z_{i\cdot}$, as well as the first $k-1$ elements of the vector $x_{i\cdot}$, are independent of the vector $w_{i\cdot}$, but the last element of $x_{i\cdot}$, x_k , may be correlated with u_i , i.e. it may be **endogenous**.

A test of the hypothesis that x_k is **exogenous** is equivalent to the test that $\sigma_{12}=0$. A test of this hypothesis may be carried out using the test statistic(s) in Eq. (27), or, equivalently, Eq. (30). An alternative test, based entirely on the null, may be carried out based on the statistic in Eq. (37).

If a complete model is not available, so that we deal **only** with the first equation of Eq. (1), but we have an instrumental matrix $W_1=(X_1,S_1)$, a test of the exogeneity of x_k may be based on the test statistic of Eq. (48).

2.4 Testing for the Exogeneity of Instruments Actually Used

Having obtained the estimator dealt with immediately above, we may wish to test whether the matrix W_1 contains elements that are indeed independent of the structural error $\hat{u}$. To this end consider the entity

$$\frac{1}{\sqrt{n}}W_1'\hat{u}=G\frac{1}{\sqrt{n}}W_1'u, \quad G=I-\frac{W_1'X}{n}\left(\frac{X'P_{w_1}X}{n}\right)^{-1}\frac{X'W_1}{n}\left(\frac{W_1'W_1}{n}\right)^{-1}. \tag{49}$$

Under the null, the entity above obeys the conditions of the Lindeberg CLT and thus

$$\frac{1}{\sqrt{n}}W_1'\hat{u}\xrightarrow{\mathrm{d}}N(0,\sigma_{11}\Phi_1), \quad \Phi_1=\operatorname*{plim}_{n\to\infty}G\frac{W_1'W_1}{n}G'. \tag{50}$$

The expression above can be considerably simplified by direct manipulation to show that

$$G\frac{W_1'W_1}{n}G'=\frac{W_1'W_1}{n}-\frac{W_1'X}{n}\left(\frac{X'P_{w_1}X}{n}\right)^{-1}\frac{X'W_1}{n}. \tag{51}$$

To show that the distribution is or is not **degenerate**, we need to show that, for all $n\geq k+m$, the matrix in Eq. (50) is positive semi-definite or positive definite, respectively. To this end apply the simultaneous decomposition theorem, see Dhrymes (2000), p. 87 and note that the difference may be expressed as

$$G\frac{W_1'W_1}{n}G' = R_1\begin{bmatrix}0 & 0\\ 0 & I_{m-1}\end{bmatrix}R_1', \quad \text{where} \quad \frac{W_1'W_1}{n} = R_1R_1'. \tag{52}$$

This is so because W_1 is of dimension $n \times k+m-1$, and the characteristic roots of the second matrix of the difference (in Eq. (50), **in the metric** of the first matrix of the difference consist of k unities and $m-1$ zeros. Thus, the matrix Φ_1 is positive **semi-definite** of rank $m-1$.

Remark 2. The result above means that we **cannot** test for the exogeneity of **all** instruments; it is similar, *mutatis mutandis*, to our inability to test for the validity of all prior restrictions in a complete simultaneous equations system, see Dhrymes (1994a).

Let S_j be a **selection** matrix, i.e. a matrix that contains $m_1-1 \geq j$ of the columns of the $k+m_1-1$- dimensioned identity matrix, I_{k+m_1-1} . Then,

$$\frac{S_j'W_1'\hat{u}}{n} \xrightarrow{\text{d}} N(0,\ \sigma_{11}S_j'\Phi_1S_j). \tag{53}$$

Consequently, we can test the hypothesis using the statistic below,

$$t_j = \frac{[S_j'\hat{u}]'[S_j'G(W_1'W_1/n)G'S_j]^{-1}[S_j'\hat{u}]}{\hat{u}'\hat{u}/n} \xrightarrow{\text{d}} \chi^2_{m-1}. \tag{54}$$

This statistic is simple to compute since all entities, save the selection matrix, are routinely computed in the process of obtaining the estimator, and the test is the precise analog of the test for (some or all) over-identifying restrictions in the general linear structural econometric model (GLSEM) given in Dhrymes (1994).

Remark 3. The fact that Φ_1 is of rank $m-1$ means that at least one $m-1$ dimensional sub-matrix is nonsingular. But there may very well be several. This means that more than one group of $m-1$ instruments may be tested for exogeneity. From the point of view of the researcher it may not be desired to test for the exogeneity of variables in the matrix X_1, only for those in P_1. However, in the context we have created above it is certainly possible to do so for both, provided $n-k \geq k-1$ and $n-k \geq m_1 - 1$.

2.5 Testing for the Exogeneity of Contemplated Instruments

In the context of the discussion of the previous section, suppose we obtained the IV estimator using the instrumental matrix W_1, but now we wish to test for the exogeneity of possible instruments we did not use, say those contained in the matrix P_2, where P_2 is $n \times m_2$ of rank m_2.

This problem may be solved by the same method as above, assuming that we are satisfied that W_1 is an admissible instrumental matrix, and that the resulting estimator is consistent. To this end consider the entity

$$\frac{1}{\sqrt{n}} P_2' \hat{u} = \frac{1}{\sqrt{n}} P_2' u - \frac{P_2' X}{n} \left(\frac{X' P_{w_1} X}{n} \right)^{-1} \frac{X' W_1}{n} \left(\frac{W_1' W_1}{n} \right)^{-1} \frac{W_1' u}{\sqrt{n}} = \frac{1}{\sqrt{n}} P_2' C u, \tag{55}$$

where $C = I - X(X' P_{w_1} X/n)^{-1} X' P_{w_1}$.

By the Lindeberg CLT

$$\frac{1}{\sqrt{n}} P_2' \hat{u} \stackrel{\mathrm{d}}{\to} N(0,\ \sigma_{11} \Phi_2), \quad \text{where} \tag{56}$$

$$\Phi_2 = \operatorname*{plim}_{n\to\infty} \frac{1}{n} P_2' C C' P_2.$$

To evaluate Φ_2, we note that

$$P_2 = P_{w_1} P_2 + (I - P_{w_1}) P_2 \tag{57}$$

$$P_{w_1} C C' = P_{w_1} - P_{\hat{x}}, \quad \hat{X} = P_{w_1} X, \ P_{\hat{x}} = \hat{X}(\hat{X}'\hat{X})^{-1}\hat{X}'$$

$$P_{w_1} C C' (I - P_{w_1}) = 0$$

$$(I - P_{w_1}) X = (0, v), \quad v = (I - P_{w_1}) x_{\cdot k}$$

$$(I - P_{w_1})CC'(I - P_{w_1}) = I - P_{w_1} + (I - P_{w_1})X(\hat{X}'\hat{X})^{-1}X'(I - P_{w_1})$$

$$P_{w_1}CC'P_{w_1} = P_{w_1} - P_{\hat{x}}, \quad \text{so that}$$

$$\Phi_2 = \operatorname*{plim}_{n\to\infty} \frac{1}{n} P_2'(P_{w_1} - P_{\hat{x}})P_2 + \hat{B}_{22}\xi\xi', \quad \xi = \operatorname*{plim}_{n\to\infty} \frac{P_2'v}{n}.$$

Since $\hat{B}_{22}\xi\xi'$ is a positive semi-definite matrix of rank 1, and $P_{\hat{x}}$ is an idempotent matrix of rank $n-k-m+1$ it follows that the rank of Φ_2 is m_2, provided $n-k-m+1 \geq m_2$. In such a case, all additional instruments (i.e. the variables contained in P_2) can be tested for instrument suitability, by means of the statistic

$$t_n = \frac{1}{n}(P_2'\hat{u})'\tilde{\Phi}_2^{-1}(P_2'\hat{u}) \xrightarrow{\mathrm{d}} \chi^2_{m_2}, \tag{58}$$

where $\tilde{\Phi}_2$ is a consistent estimator of Φ_2. Any subset of the instruments may be tested by the same method, using a suitably dimensioned selection matrix

3 Power of Tests: Limiting Distribution under the Alternative

3.1 Structured Models

We first examine the tests exhibited in Eqs. (27) and (37). In view of Eq. (8), we have

$$\begin{aligned}
\frac{\hat{v}'\hat{u}}{\sqrt{n}} &= \frac{1}{\sqrt{n}} v'[I - X(\hat{X}'\hat{X})^{-1}\hat{X}']u - \frac{1}{\sqrt{n}}(v'P_z u - v'P_{\hat{x}}u) \\
&\sim \frac{1}{\sqrt{n}} v'[I - X(\hat{X}'\hat{X})^{-1}\hat{X}']u \\
&= \frac{1}{\sqrt{n}} v'(I - P_{\hat{x}})u - \frac{v'v}{n} e'_{\cdot k}\left(\frac{\hat{X}'\hat{X}}{n}\right)^{-1} \frac{1}{\sqrt{n}}\hat{X}'u \\
&\sim \frac{1}{\sqrt{n}}(v - s)'u = \frac{1}{\sqrt{n}}\sum_i^n (v_i - s_i)u_i. \qquad (59)
\end{aligned}$$

where $s_i = \sigma_{22} e'_{\cdot k} \hat{B} \hat{x}'_{i\cdot}$. However, the summands, $(v_i - s_i)u_i$ **do not** have mean zero under the alternative; their mean is σ_{12}. Thus, we need to consider

$$\frac{\hat{v}'\hat{u}}{\sqrt{n}} - \sqrt{n}\sigma_{12} \sim \frac{1}{\sqrt{n}} \sum_{i=1}^{n} [(v_i - s_i)u_i - \sigma_{12})], \tag{60}$$

which obeys the conditions of the Lindeberg CLT. Consequently, using the same arguments as before, we conclude that

$$\frac{\hat{v}'\hat{u}}{\sqrt{n}} - \sqrt{n}\sigma_{12} \xrightarrow{\mathrm{d}} N(0, \omega^*), \quad \omega^* = \sigma_{11}\sigma_{22}[1 + \sigma_{22}\hat{B}_{22}] + \sigma_{12}^2.^{8} \tag{61}$$

Thus, **under the alternative**

$H_1 : \qquad \sigma_{12} = \frac{\phi^*}{\sqrt{n}}$,

the test statistic(s) in Eqs. (27) and (37) obey

$$t_n^2 \quad \xrightarrow{\mathrm{d}} \quad \chi^2(\lambda), \qquad \lambda = \frac{\phi^{*2}}{2\sigma_{11}\sigma_{22}(1 + \sigma_{22}\hat{B}_{22})}. \tag{62}$$

Turning now to the test (statistic) of Eq. (37) and repeating the steps leading to Eqs. (31) through (34) we see that **under the alternative**

$$E(v_i - s_i^*)u_i = (1 - \sigma_{22}B_{22})\sigma_{12}. \tag{63}$$

Thus,

$$\frac{1}{\sqrt{n}}\hat{v}'\tilde{u} - (1 - \sigma_{22}B_{22})\sqrt{n}\sigma_{12} \sim \frac{1}{\sqrt{n}} \sum_{i=1}^{n} [(1 - \sigma_{22}B_{22})(v_i u_i - \sigma_{12}) - \sigma_{22} e'_{\cdot k} B \bar{x}'_{i\cdot} u_i], \tag{64}$$

and consequently

$$\frac{1}{\sqrt{n}}\hat{v}'\tilde{u} - (1 - \sigma_{22}B_{22})\sqrt{n}\sigma_{12} \xrightarrow{\mathrm{d}} N(0, \sigma_{11}\omega). \tag{65}$$

[8] In fact, the last term, σ_{12}^2, should not be there because it results when we sum, divide by n and take the limit of the variances of the summands. Since under the **alternative** $\sigma_{12} = \phi^*/\sqrt{n}$, it follows that upon taking limits $\omega^* = \sigma_{11}\omega$ as in Eq. (26).

But this implies that the test statistic in Eq. (37) obeys **under the alternative**

$$t_n^{(1)} = \frac{\hat{v}'\tilde{u}}{\sqrt{n\hat{\sigma}_{11}\hat{\sigma}_{22}(1-\hat{\sigma}_{22}B_{22})}} \xrightarrow{\mathrm{d}} \chi^2(\lambda_1), \quad \lambda_1 = \frac{\phi^{*2}(1-\sigma_{22}B_{22})}{2\sigma_{11}\sigma_{22}} \tag{66}$$

Turning now to the entity in Eq. (42) and the test (statistic) in Eq. (48), we consider the limiting distribution under the alternative.

Reamrk 4. In dealing with this issue under the null, we had concealed the fact that the standard assumptions made in this context are not sufficient to enable formal derivations. It is not enough merely to say that we have some instruments which are independent of the structural error u. More is required as will become apparent when the argument unfolds.

Rewrite Eq. (43) as

$$\frac{1}{\sqrt{n}}x'_{\cdot k}\hat{u} = \frac{1}{\sqrt{n}}x'_{\cdot k}u - \frac{1}{\sqrt{n}}\left(\frac{x_{\cdot k}X}{n}\right)\left(\frac{X'P_{w_1}X}{n}\right)^{-1} X'P_{w_1}u, \tag{67}$$

and note that

$$x'_{\cdot k}X = x'_{\cdot k}P_{w_1}X + x'_{\cdot k}(I-P_{w_1})X = x'_{\cdot k}P_{w_1}X + e'_{\cdot k}v'v, \tag{68}$$

where $v = (I-P_{w_1})x_{\cdot k}$. Thus, the statistic of Eq. (48) behaves like (upon taking a certain probability limit)

$$\frac{1}{\sqrt{n}}x'_{\cdot k}\hat{u} \sim \frac{1}{\sqrt{n}}v'u - \frac{1}{\sqrt{n}}du, \quad d = \sigma^*_{22}e'_{\cdot k}\left(\frac{X'P_{w_1}X}{n}\right)^{-1} X'P_{w_1}. \tag{69}$$

Remark 5. At this stage the need for more precise specification becomes evident, because what we have assumed until now is that v is the **orthogonal complement** of the projection of $x_{\cdot k}$ on the space spanned by W_1. To treat the elements of v as random variables with the properties (once removed) required for the invocation of a CLT is almost tantamount to stating that $x_{\cdot k} = W_1\gamma + v$, in which case we should have a complete model, as in the case of the first test we considered. In fact, this must be the implicit assumption made by practitioners who employ such techniques.

Be that as it may, we note that **under the alternative**

$$\frac{1}{\sqrt{n}} x'_{\cdot k} \hat{u} \sim \frac{1}{\sqrt{n}} \sum_{i=1}^{n} [v_i u_i - d_i u_i], \tag{70}$$

and the individual terms of the sum **do not have mean zero**. Thus, if we mimic the previous derivations we are led to consider

$$\frac{1}{\sqrt{n}} x'_{\cdot k} \hat{u} - \sqrt{n}\sigma_{12} \sim \frac{1}{\sqrt{n}} \sum_{i=1}^{n} [(v_i u_i - \sigma_{12}) - d_i u_i], \tag{71}$$

which will obey the conditions of the Lindeberg CLT, whence we conclude

$$\frac{1}{\sqrt{n}} x'_{\cdot k} \hat{u} - \sqrt{n}\sigma_{12} \xrightarrow{\mathrm{d}} N(0, \sigma_{11}\sigma^{*}_{22}[1 + \sigma^{*}_{22}\hat{B}_{22}]), \tag{72}$$

$$[t_n^{(2)}]^2 \xrightarrow{\mathrm{d}} \chi^2(\lambda_2), \quad \lambda_2 = \frac{\phi^{*2}}{2\sigma_{11}\sigma^{*}_{22}[1 + \sigma^{*}_{22}\hat{B}_{22}]}. \tag{73}$$

3.2 Unstructured Models

Here we examine, under the alternative, the limiting distribution of the entity in Eq. (38) and, hence, that of the test statistic in Eq. (43).

If we consider the vector $\xi_{i\cdot} = (u_i, w_{i1\cdot})$, where its second component it the ith row of W_1–and thus a $k+m$-dimensioned random vector with mean, say zero for convenience, and covariance matrix

$$\mathrm{Cov}(\xi_{i\cdot}) = \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{1\cdot} \\ \sigma_{\cdot 1} & \Sigma_{22} \end{bmatrix}, \tag{74}$$

we may, with little loss of relevance, assume that

$$\Sigma_{22} = \operatorname*{plim}_{n\to\infty} \frac{1}{n} W_1' W_1 \tag{75}$$

and state the hypotheses under consideration

$H_0: \quad \sigma_{\cdot 1} = 0$
as against the alternative $H_1: \quad \sigma_{\cdot 1} = \frac{1}{\sqrt{n}}\phi^{*}_{3}$,

where ϕ_1^* is a **non-zero** vector.

Proceeding as before, we consider

$$\frac{1}{\sqrt{n}} W_1' \hat{u} - \sqrt{n}\sigma_{\cdot 1} = \frac{1}{\sqrt{n}} [\sum_{i=1}^{n} (w_{i1\cdot}' u_i - \sigma_{\cdot 1}]. \tag{76}$$

The summands are a sequence of independent random vectors obeying the Lindeberg CLT conditions; moreover, the covariance matrix of the typical summand is

$$\begin{aligned} \mathrm{Cov}[w_{i1\cdot}' u_i - \sigma_{\cdot 1}] &= E\left(w_{i1\cdot}' u_i^2 w_{i1\cdot}' - [w_{i1\cdot}' u_i] \sigma_{\cdot 1}' - \sigma_{\cdot 1} [w_{i1\cdot}' u_i]' + \sigma_{\cdot 1} \sigma_{\cdot 1}' \right) \\ &= \sigma_{11} \Sigma_{22} + 2\sigma_{\cdot 1} \sigma_{\cdot 1}' - 2\sigma_{\cdot 1} \sigma_{\cdot 1}' + \sigma_{\cdot 1} \sigma_{\cdot 1}' \\ &= \sigma_{11} \Sigma_{22} + \sigma_{\cdot 1} \sigma_{\cdot 1}'. \end{aligned} \tag{77}$$

We therefore conclude that, under the alternative,

$$\frac{1}{\sqrt{n}} W_1' \hat{u} - \sqrt{n}\sigma_{\cdot 1} \xrightarrow{\mathrm{d}} N(0, \sigma_{11} \Phi_1), \tag{78}$$

and, consequently that **under the alternative**

$$\frac{1}{\sqrt{n}} W_1' \hat{u} \xrightarrow{\mathrm{d}} N(\phi_3^*, \sigma_{11} \Phi_1). \tag{79}$$

An immediate consequence of the preceding is that the test statistic of Eq. (54) converges in distribution to a **non-central** chi-square distribution, i.e. **under the alternative**

$$\begin{aligned} t_j &= \frac{[S_j' \hat{u}]' [S_j' G (W_1' W_1 / n) G' S_j]^{-1} [S_j' \hat{u}]}{\hat{u}' \hat{u} / n} \xrightarrow{\mathrm{d}} \chi^2_{m-1}((\lambda_3)), \\ \lambda_3 &= \frac{1}{2} \phi_3^{*\prime} (\sigma_{11} \Phi_1)^{-1} \phi_3^*. \end{aligned} \tag{80}$$

The discussion with respect to contemplated instruments is quite similar to that for instruments already employed, except that now we rule out correlation between the instruments already used and the structural error term. The only question is whether the contemplated instruments exhibit such a feature. For the current discussion it will be convenient if we recast Eq. (55) as

$$\frac{1}{\sqrt{n}}P_2'\hat{u} = \sum_{i=1}^{n}[p_{i1\cdot}' - \tilde{M}w_{i1\cdot}']u_i,$$

$$\tilde{M} = \frac{P_2'X}{n}\left(\frac{X'P_{w_1}X}{n}\right)^{-1}\frac{X'W_1}{n}\left(\frac{W_1'W_1}{n}\right)^{-1}\frac{W_1'u}{\sqrt{n}}. \tag{81}$$

Since $\tilde{M} \xrightarrow{\mathrm{P}} M$, we see that

$$\frac{1}{\sqrt{T}}P_2'\hat{u} \sim \sum_{i=1}^{n}[p_{i1\cdot}' - Mw_{i1\cdot}'u_i, \tag{82}$$

and the summands under the null are independent random vectors with mean

$$Ep_{i2\cdot}u_i = \sigma_{\cdot 1}, \quad \mathrm{Cov}(p_{i1\cdot}u_i) = \sigma_{11}\Sigma_{22} + \sigma_{\cdot 1}\sigma_{\cdot 1}'. \tag{83}$$

In deriving this result, we have defined the vector $\zeta_{i\cdot} = (ui,\ p_{i2\cdot})$ we have put

$$\mathrm{Cov}(\zeta_{i\cdot}) = \begin{bmatrix} \sigma_{11} & \sigma_{\cdot 1}' \\ \sigma_{\cdot 1} & \Sigma_{22} \end{bmatrix} \tag{84}$$

and assumed that $\Sigma_{22} = \mathrm{plim}_{n\to\infty}(P_2'P_2/n)$.

Since, under the alternative, $Ep_{i2}'u_i = \sigma_{\cdot 1} \neq 0$, we conclude that **under the alternative**

$H_1: \quad \sigma_{\cdot 1} = \frac{1}{\sqrt{n}}\phi^*4$,

$$\frac{1}{\sqrt{n}}P_2'\hat{u} \xrightarrow{\mathrm{d}} N(\phi_4^*, \sigma_{11}\phi_2) \tag{85}$$

and, consequently, the test statistic in Eq. (58) obeys,

$$t_n = \frac{1}{n}(P_2'\hat{u})'\tilde{\Phi}_2^{-1}(P_2'\hat{u}) \xrightarrow{\mathrm{d}} \chi^2_{m_2}(\lambda_4), \tag{86}$$

where

$$\lambda_4 = \frac{1}{2}\phi_4^{*\prime}(\sigma_{11}\Phi_2)^{-1}\phi_4^*.$$

4 Limiting Distribution of the Hausman Test

In the context of the model in Eq. (1) the Hausman test is based on the difference $\hat{\beta}_{\text{``}2SLS\text{''}} - \hat{\beta}_{OLS}$, where

$$\hat{\beta}_{\text{``}2SLS''} = \beta + (X'P_{w_1}X)^{-1}X'u, \quad \hat{\beta}_{OLS} = \beta + (X'X)^{-1}X'u. \tag{87}$$

Under the null, the difference obeys

$$\sqrt{n}(\hat{\beta}_{\text{``}2SLS''} - \hat{\beta}_{OLS}) \xrightarrow{\mathrm{d}} N(0, \Phi_h), \quad \Phi_h = \sigma_{11}(\hat{B} - B) \tag{88}$$

where, we recall

$$\hat{B} = \operatorname*{plim}_{n\to\infty}\left(\frac{X'P_{w_1}X}{n}\right)^{-1}, \quad B = \operatorname*{plim}_{n\to\infty}\left(\frac{X'X}{n}\right)^{-1}.$$

Since

$$\Phi_h \sim \hat{B}\left(\frac{X'X}{n} - \frac{X'P_{w_1}X}{n}\right)B = \hat{B}\left(\frac{1}{n}\begin{bmatrix} 0 & 0 \\ 0 & x'_{\cdot k}(I_n - P_{w_1})x_{\cdot k}\end{bmatrix}\right)B, \tag{89}$$

it follows that

$$\operatorname{rank}(\Phi_h) = 1. \tag{90}$$

The Hausman test statistic thus obeys, **under the null**,

$$t_h = n(\hat{\beta}_{\text{``}2SLS''} - \hat{\beta}_{OLS})'[\hat{\Phi}_h]_g(\hat{\beta}_{\text{``}2SLS\text{''}} - \hat{\beta}_{OLS}) \xrightarrow{\mathrm{d}} \chi^2_1, \tag{91}$$

i.e. it converges to a central chi square with one degree of freedom, where the notation $[\hat{\Phi}_h]_g$ indicates the **generalized inverse** of Φ_h.

Remark 6. There is a considerable degree of confusion in implementing this test. Typically, in computer packages and in textbooks one finds the statement

$$\hat{\Phi}_h = \left(\hat{\sigma}_{11(\text{``}2SLS'')}\frac{X'P_{w_1}X}{n}\right)^{-1} - \left(\hat{\sigma}_{11(OLS)}\frac{X'X}{n}\right)^{-1}$$

which in finite samples will be non-singular, although under the null this expression will converge to the singular matrix of rank one obtained earlier.

It would appear that, in the spirit of asymptotic theory, the expression above "ought" to be written as

$$\hat{\Phi}_h = \hat{\sigma}_{11(\text{``}2SLS'')} \left[\left(\frac{X' P_{w_1} X}{n} \right)^{-1} - \left(\frac{X'X}{n} \right)^{-1} \right], \tag{92}$$

which, mimics the asymptotic rank for all sample sizes and, in any event, converges to the desired entity both under the null and under the alternative.

Under the alternative, the OLS estimator is inconsistent, and its inconsistency is given by

$$\operatorname*{plim}_{n\to\infty} (\hat{\beta}_{OLS} - \beta) = \sigma_{12} B e_{\cdot k}. \tag{93}$$

It may further be shown that **under the alternative**

$$\sqrt{n}(\hat{\beta}_{\text{``}2SLS''} - \hat{\beta}_{OLS}) - \sqrt{n}\sigma_{12} B e_{\cdot k} \xrightarrow{\mathrm{d}} N(0,\ \Phi_h), \tag{94}$$

or that

$$\sqrt{n}(\hat{\beta}_{\text{``}2SLS''} - \hat{\beta}_{OLS}) \xrightarrow{\mathrm{d}} N(\phi^* B e_{\cdot k}, \Phi_h) \tag{95}$$

and, consequently, that

$$t_h \xrightarrow{\mathrm{d}} \chi_1^2(\lambda_h), \quad \lambda_h = \frac{1}{2} \phi^{*2} e_{\cdot k}' B [\hat{\Phi}_h]_g B e_{\cdot k}. \tag{96}$$

The distributions of the other variants of the Hausman test, noted at the beginning, are similarly obtained and, thus, will not be pursued here.

BIBLIOGRAPHY

Amemiya, T. (1974), "The nonlinear least squares estimator", *Journal of Econometrics*, vol. 2, pp. 105-110.

Berkson, J. (1950), "Are there two regressions?" *Journal of the American Statistical Association*, vol. 47, pp.164-180

Davidson, R and J. MacKinnon (1993), *Estimation and Inference in Econometrics*, New York: Oxford University Press.

Dhrymes, P. J. (1969), "Alternative asymptotic tests of significance and related aspects of 2SLS and 3SLS estimated parameters", *The Review of Economic Studies*, vol. 36, pp. 213-236.

Dhrymes, P. J. (1970), *Econometrics: Statistical Foundations and Applications*, New York: Harper and Row.

Dhrymes, P. J. (1973), 'Small Sample and Asymptotic Relations between Maximum Likelihood and Three Stage Least Squares Estimators", **Econometrica**, vol. 41, pp. 357-364.

Dhrymes, P.J.(1978), *Introductory Econometrics*, New York: Springer Verlag.

Dhrymes, P. J. (1994a), "Specification tests in simultaneous equations systems", *Journal of Econometrics*, pp. 45-76.

Dhrymes, P. J. (1994b),*Topics in Advanced Econometrics: vol. II Linear and Nonlinear Simultaneous Equations*, New York: Springer Verlag.

Dhrymes, P. J. (1989), *Topics in Advanced Econometrics: Probability Foundations*, New York: Springer Verlag.

Dhrymes, P. J. (1998), *Time Series, Unit Roots and Cointegration*, San Diego: Academic Press.

Dhrymes, P. J. (2000), *Mathematics for Econometrics*, Third Edition, New York: Springer Verlag

Durbin, J. (1954), “Errors in Variables”, *Review of the International Statistical Insitute*, vol. 22, pp. 23-32.

Greene, W. H. (2000), *Econometric Analysis*, Fourth Edition, Upper Saddle River: Prentice Hall Inc.

Hansen, L. P. (1982),“Large Sample Properties of the Generalized Method of Moments”, *Econometrica*, vol. 50, pp. 1029-1054.

Hausman, J. A. (1978),“Specification Tests in Econometrics”, *Econometrica*, vol. 46, pp. 1251-1271.

Jorgenson, D. and J. Laffont *1974), “Efficient Estimation of Nonlinear Simultaneous Equations with Additive Disturbances”,*Annals of Economic and Social Measurement*, vol. 3, pp. 615-640.

Reiersol, O. (1950), “Identifiability of a linear relation between variables which are subject to error”, *Econometrica*, vol. 18, pp. 375-389.

Wald, A. (1940), “The fitting of straight lines if both variables are subject to error”, *Annals of Mathematical Statistics*, vol. 11, pp. 284-300.

Wu, D. (1973), “Alternative Tests of Independence between Stochastic Regressors and Disturbances”, *Econometrica*, vol. 41, pp.733-750.

APPENDIX I

In this appendix we derive an explicit expression for the generalized inverse required by the Hausman test(s). Consider the covariance matrix of the limiting distribution of the entity $\sqrt{n}(\hat{\beta}_{\text{“}2SLS\text{”}} - \hat{\beta}_{OLS})$,

$$C^* = \sigma_{11} C, \quad C = \operatorname*{plim}_{n \to \infty} \left[\left(\frac{X' P_{w_1} X}{n} \right)^{-1} - \left(\frac{X'X}{n} \right)^{-1} \right].$$

For ease of exposition, and without loss of relevance, in the argument below we shall dispense with the probability limit and the division by n and treat C as if it were $(X' P_{w_1} X)^{-1} - (X'X)^{-1}$. Write

$$\begin{aligned} C &= (X' P_{w_1} X)^{-1} - (X'X)^{-1} = (X' P_{w_1} X)^{-1} \begin{bmatrix} 0 & 0 \\ 0 & \hat{v}'\hat{v} \end{bmatrix} (X'X)^{-1} \\ &= \sigma_{22} (X' P_{w_1} X)^{-1} e_{\cdot k} e_{\cdot k}' (X'X)^{-1} = \tilde{B}_2 B_2', \quad \text{where} \\ \tilde{B} &= (X' P_{w_1} X)^{-1} = \begin{bmatrix} \tilde{B}_{11} & \tilde{B}_{12} \\ \tilde{B}_{21} & \tilde{B}_{22} \end{bmatrix} = (\tilde{B}_1, \tilde{B}_2), \\ B &= (X'X)^{-1} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = (B_1, B_2), \quad \text{and} \\ \tilde{B}_{12} &= -\tilde{B}_{22} (X_1' X)_1^{-1} X_1' x_{\cdot k}, \quad \tilde{B}_{22} = [\tilde{x}_{\cdot k}' \tilde{x}_{\cdot k} - x_{\cdot k}' P_{x_1} x_{\cdot k}]^{-1}, \\ B_{12} &= -B_{22} (X_1' X_1)^{-1} X_1' x_{\cdot k}, \quad B_{22} = [x_{\cdot k}' x_{\cdot k} - x_{\cdot k}' P_{x_1} x_{\cdot k}]^{-1}. \end{aligned}$$

Thus, we have the result

$$C = \tilde{B}_2 B_2',$$

where $\tilde{B}_2$ is $k \times 1$ of rank 1 and B_2' is $1 \times k$ of rank 1. This constitutes a **rank factorization** of the matrix C and, by Proposition 71 in Dhrymes (2000), the **generalized inverse** of C is given by

$$C_g = B_2 (B_2' B_2)^{-1} (\tilde{B}_2' \tilde{B}_2)^{-1} \tilde{B}_2'.$$

Carrying out the requisite operations and defining $\zeta = (X_1'X_1)^{-1}x_{\cdot k}$, we find

$$C_g = B_{22}^{-1}\frac{1}{1+\zeta'\zeta}\begin{bmatrix} \zeta\zeta' & -\zeta \\ -\zeta' & 1 \end{bmatrix}\frac{1}{1+\zeta'\zeta}\tilde{B}_{22}^{-1}, \quad \text{and thus}$$

$$C_g^* = \frac{1}{\sigma_{11}\sigma_{22}}\left(B_{22}^{-1}\frac{1}{1+\zeta'\zeta}\right)\begin{bmatrix} \zeta\zeta' & -\zeta \\ -\zeta' & 1 \end{bmatrix}\left(\frac{1}{1+\zeta'\zeta}\tilde{B}_{22}^{-1}\right).$$

Inserting this representation in the non-centrality parameter of Eq. (96) yields

$$\lambda_h = \frac{1}{2}\phi^{*2}e_{\cdot k}'B[\hat{\Phi}_h]_g Be_{\cdot k} = \frac{1}{2\sigma_{11}}\phi^{*2}\frac{x_{\cdot k}'(P_{w_1}-P_{x_1})x_{\cdot k}}{x_{\cdot k}'(I-P_{x_1})x_{\cdot k}}.$$

APPENDIX II

In this appendix we generalize the discussion based on Eqs. (1) and (2), so that more than one right hand variables may be tested for exogeneity. We find that the results obtained with a single potential endogenous right hand variable carry over entirely, *mutatis mutndis* to the vector case. To this end, consider

$$y_{\cdot t} = X_1\beta_{(1)} + X_2\beta_{(2)} + u, \quad X_2 = W_1\Gamma + V,$$

where now X_1 is $n \times k_1$, X_2 is $n \times k_2$, $k_1 + k_2 = k$. Putting

$$x_{(2)} = \mathrm{vec}(X_2), \quad v = \mathrm{vec}(V), \quad \gamma = \mathrm{vec}(\Gamma),$$

we can rewrite the entire system compactly as

$$y = X\beta + u, \quad x_{(2)} = (I_{k_2} \otimes W_1)\gamma + v,$$

and a single observation as

$$y_i - x_{i\cdot}^{(2)}\beta_{(2)} = x_{i\cdot}^{(1)}\beta_{(1)} + u_i, \qquad x_{i\cdot}^{(2)} = w_{i\cdot}^{(1)}\Gamma + v_{i\cdot}.$$

Under the standard assumptions (including joint normality of the errors) we may write the likelihood function in terms of the errors as

$$L^*(u, V; \Sigma) = (2\pi)^{-n(k_2+1)/2}|\Sigma|^{-(n/2)}e^{-\frac{1}{2}\mathrm{tr}\Sigma^{-1}S^*}, \quad S^* = \begin{bmatrix} u'u & u'V \\ V'u & V'V \end{bmatrix}.$$

Viewing the equation for a single observation as a transformation from $(u_i, v_{i\cdot})'$ to $(y_i, x_{i\cdot}^2)'$, we find that the Jacobian matrix is

$$J = \begin{bmatrix} 1 & -\beta_{(2)}' \\ 0 & I_{k_2} \end{bmatrix},$$

so that the Jacobian of the transformation is $|J| = 1$; thus, the (log) likelihood function of the observations is

$$L = -\frac{n(k_2+1)}{2}\ln(2\pi) + \frac{n}{2}\ln|\Sigma^{-1}| - \frac{1}{2}\mathrm{tr}\Sigma^{-1}S, \quad S = \begin{bmatrix} s_{11} & s_{1\cdot} \\ s_{\cdot 1} & S_{22} \end{bmatrix}$$

$$s_{11} = (y - X\beta)'(y - X\beta), \quad s_{1\cdot} = (s_{12}, s_{13}, \ldots, s_{1(k_2+1)})$$

$$s_{1j} = (y - X\beta)'(x_{\cdot k_1+j-1} - W_1\gamma_{\cdot j-1}), \quad j = 2, 3, \ldots, k_2 + 1, \quad s_{\cdot 1} = s'_{1\cdot},$$

$$S_{22} = (s_{ij}), \quad s_{ij} = (x_{\cdot k_1+i-1} - W_1\gamma_{\cdot i-1})'(x_{\cdot k_1+j-1} - W_1\gamma_{\cdot j-1}), \quad j = 2, 3, \ldots, k_2 + 1.$$

Differentiating with respect to Σ^{-1} yields

$$\frac{\partial L}{\partial \Sigma^{-1}} = \frac{n}{2}\Sigma - \frac{1}{2}S = 0, \quad \text{or} \quad \hat{\Sigma} = \frac{1}{n}S.$$

Inserting this in the loglikelihood function we find the **concentrated** (log) likelihood function

$$L(\beta, \gamma) = -\frac{n(k_2+1)}{2}[\ln(2\pi) + 1] + \frac{n}{2}\ln(n) - \frac{n}{2}\ln|S|.$$

As we have done in the discussion surrounding Eq. (8), we note that the LRT statistic is some function of the likelihood ratio (LR)

$$\lambda = \frac{\max_{H_0} L_n^*}{\max_{H_1} L_n^*}.$$

In this instance the null is that $\sigma_{\cdot 1} = 0$, where

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{1\cdot} \\ \sigma_{\cdot 1} & \Sigma_{22} \end{bmatrix}.$$

We note that Σ_2 is $k_2 \times k_2$ and thus $\sigma_{\cdot 1}$ is $k_2 \times 1$. Consequently, the LR is given by

$$\lambda = \left(\frac{|\tilde{S}_{H_0}|}{|\hat{S}_{H_1}|}\right)^{(-n/2)} = \left(\frac{\tilde{\sigma}_{11}|\tilde{\Sigma}_{22}|}{\hat{\sigma}_{11}|\hat{\Sigma}_{22}|)}(1 - \hat{\sigma}_{1\cdot}(\hat{\sigma}_{11}|\hat{\Sigma}_{22}|)^{-1}\hat{\sigma}_{\cdot 1}\right)^{-(n/2)}$$

$$\sim \left((1 - \hat{\sigma}_{1\cdot}(\hat{\sigma}_{11}|\hat{\Sigma}_{22}|)^{-1}\hat{\sigma}_{\cdot 1}\right)^{(n/2)} \quad \text{or}$$

$$1 - \lambda^{(2/n)} \quad \sim \hat{\sigma}_{1\cdot}(\hat{\sigma}_{11}|\hat{\Sigma}_{22}|)^{-1}\hat{\sigma}_{\cdot 1}.$$

If we use the sample analogs of these entities, we find

$$\hat{\sigma}_{1\cdot}(\hat{\sigma}_{11}|\hat{\Sigma}_{22}|)^{-1}\hat{\sigma}_{\cdot 1} = \hat{u}'\hat{V}(\hat{u}'\hat{u}\hat{V}'\hat{V})^{-1}\hat{V}'\hat{u},$$

which suggests the test statistic $\hat{V}'\hat{u}/\sqrt{n}$, where

$$\hat{u} = y - X\hat{\beta}_{3SLS} = u - X(\hat{\beta}_{3SLS} - \beta),$$

$$\hat{V} = X_2 - W_1\hat{\Gamma}_{3SLS} = V - W_1(\hat{\Gamma}_{3SLS} - \Gamma).$$

We can use 3SLS (instead of ML) estimators because, as we pointed out earlier, the Jacobian of the transformation is unity, and in such cases ML and iterated 3SLS will coincide.

Consider now the statistic

$$\frac{1}{\sqrt{n}}\hat{V}'\hat{u} = \frac{1}{\sqrt{n}}[V - W_1(\hat{\Gamma}_{3SLS} - \Gamma)]'[u - X(\hat{\beta}_{3SLS} - \beta)],$$

which upon multiplication and taking (probability) limits reduces to

$$\frac{1}{\sqrt{n}}\hat{V}'\hat{u} \sim \frac{1}{\sqrt{n}}V'u - (0, \Sigma_{22})\sqrt{n}(\hat{\beta}_{3SLS} - \beta).$$

But the second term may be easily shown to be (asymptotically) uncorrelated with the first and, thus, the covariance matrix of the limiting distribution of the entity above is simply the sum of the covariance matrices of the limiting distributions of the two terms, i.e.

$$\frac{1}{\sqrt{n}}\hat{V}'\hat{u} \xrightarrow{\text{d}} N(0, C), \quad C = C_1 + C_2,$$

where, under the null, $C_1 = \sigma_{11}\Sigma_{22}$ and C_2 is the submatrix of the covariance matrix of the limiting distribution of $\hat{\beta}_{3SLS}$ corresponding to its subvector $\beta_{(2)}$. To determine the limiting distribution of the 3SLS estimator of β , write the system as

$$y = X\beta + u, \quad x_{(2)} = (I_{k_2} \otimes W_1)\gamma + v, \quad \gamma = \text{vec}(\Gamma), \quad v = \text{vec}(V).$$

Let $W_1'W_1 = R_1R_1'$, where R_1 is a **non-singular** matrix; such a matrix exists because W_1 if of full (column) rank. Transform the system by premultiplying the first equation be $R_1^{-1}W_1'$ and the second by $I_{k_2} \otimes R_1^{-1}W_1'$

to obtain

$$\begin{bmatrix} R_1^{-1}W_1' & 0 \\ 0 & I_{k_2}\otimes R_1^{-1}W_1' \end{bmatrix}\begin{pmatrix} y \\ x_{(2)} \end{pmatrix} = \begin{bmatrix} R_1^{-1}W_1'X & 0 \\ 0 & I_{k_2}\otimes R_1' \end{bmatrix}\begin{pmatrix} \beta \\ \gamma \end{pmatrix}$$

$$+\begin{bmatrix} R_1^{-1}W_1' & 0 \\ 0 & I_{k_2}\otimes R_1^{-1}W_1' \end{bmatrix}\begin{pmatrix} u \\ v \end{pmatrix}$$

$$\sqrt{n}\left[\begin{pmatrix} \beta \\ \gamma \end{pmatrix}_{3SLS} - \begin{pmatrix} \beta \\ \gamma \end{pmatrix}\right] \xrightarrow{\mathrm{d}} N(0,\Phi),$$

where

$$\Phi = \operatorname*{plim}_{n\to\infty}\left(\frac{1}{n}\begin{bmatrix} \hat\sigma^{11}X'P_{w_1}X & \hat\sigma^{1\cdot}\otimes X'W_1 \\ \hat\sigma^{\cdot 1}\otimes W_1'X & \hat\Sigma^{22}\otimes W_1'W_1 \end{bmatrix}\right)^{-1}$$

$$\sqrt{n}(\hat\beta_{3SLS}-\beta) \xrightarrow{\mathrm{d}} N(0,\Phi_{11}), \quad \Phi_{11}^{-1} = \operatorname*{plim}_{n\to\infty}\frac{1}{n}(\hat\sigma^{11}-\hat\sigma^{1\cdot}(\hat\Sigma^{22})^{-1}\hat\sigma^{\cdot 1})X'P_{w_1}X.$$

¿From Proposition 2.31 in Dhrymes (2000), p. 43, we note that

$$(\hat\sigma^{11}-\hat\sigma^{1\cdot}(\hat\Sigma^{22})^{-1}\hat\sigma^{\cdot 1})^{-1} = \sigma_{11}, \quad \text{and thus} \quad \Phi_{11} = \sigma_{11}(X'P_{w_1}X)^{-1}.$$

Moreover,

$$\sqrt{n}(\hat\beta_{(2)3SLS}-\beta_{(2)}) \xrightarrow{\mathrm{d}} N(0,\Psi_{22}),$$

$$\Psi_{22} = \sigma_{11}\operatorname*{plim}_{n\to\infty}\left(\frac{X_2'(I_{k_2}-P_{w_1})X_2}{n}\right)^{-1}.$$

It follows, therefore, that **under the null**,

$$\frac{1}{\sqrt{n}}\hat V'\hat u \xrightarrow{\mathrm{d}} N(0,C), \quad C = C_1 + C_2, \quad C_1 = \sigma_{11}\Sigma_{22}$$

$$C_2 = \sigma_{11}\Sigma_{22}\hat B_{22}\Sigma_{22}, \quad \hat B_{22}^{-1} = \operatorname*{plim}_{n\to\infty}\frac{1}{n}X_2'(I-P_{w_1})X_2.$$

Notice that when $k_2 = 1$, the result above reduces to that obtained earlier in Eq. (26).

To test the null hypothesis $H_0 : \sigma_{\cdot 1} = 0$, we may thus use the statistic

$$t_m^2 = \frac{\hat{u}'\hat{V}}{\sqrt{n}} \left[\hat{\sigma}_{11}\hat{\Sigma}_{22}(I_{k_2} + \hat{B}_{22})\hat{\Sigma}_{22}\right]^{-1} \frac{\hat{V}'\hat{u}}{\sqrt{n}} \xrightarrow{\mathrm{d}} \chi^2_{k_2},$$

which is the multivariate generalization of the statistic given in Eq. (27).

Remark AII.1 The asymptotic equivalence between the 2SLS and 3SLS estimators of β, in the context of this model, in part explains why in the body of the paper we had dealt essentially with the "2SLS" estimator only.

APPENDIX III

In this appendix we report the results of a limited Monte Carlo study as follows: The results of the Monte Carlo study are reported in the tables below. They involve 10,000 replications with samples of size 50, 75, 100 and 2000 and for correlations between the structural error u and the error in the "reduced form" or instrumental equation v, of the order 0, .1, .3, .5, .7 and .9 The variance of the errors was set to $\sigma_{11} \approx 1$, $\sigma_{22} \approx 1$ and σ_{12} was set to the level appropriate so as to obtain the desired correlation.

Hausman 1, refers to the nave application of the test, i.e. the quadratic form is given by

$$n(\hat{\beta}_{OLS}-\hat{\beta}_{2SLS})'[\hat{\sigma}_{11(2SLS)}(X'P_{w_1}X/n)^{-1}-\hat{\sigma}_{11(OLS)}(X'X/n)^{-1}]^{-1}(\hat{\beta}_{OLS}-\hat{\beta}_{2SLS}),$$

which, under the null, in this case is claimed to be chi-squared with 5 degrees of freedom.

Hausman 2 refers to the test that recognizes that under the null the two estimators of σ_{11} will converge to the same entity and thus uses as the test statistic

$$n(\hat{\beta}_{OLS}-\hat{\beta}_{2SLS})'[\hat{\sigma}_{11(OLS)}((X'P_{w_1}X/n)^{-1}-(X'X/n)^{-1}]_g(\hat{\beta}_{OLS}-\hat{\beta}_{2SLS}),$$

while Hausman 3 refers to the test that uses the test statistic

$$n(\hat{\beta}_{OLS}-\hat{\beta}_{2SLS})'[\hat{\sigma}_{11(2SLS)}((X'P_{w_1}X/n)^{-1}-(X'X/n)^{-1})]_g(\hat{\beta}_{OLS}-\hat{\beta}_{2SLS}),$$

where A_g denotes the **generalized inverse of** A, see Dhrymes (2000).

The test labeled Eq. (26) refers to the test that uses residuals from the 2SLS estimation of parameters in the first equation and OLS residuals from the second equation.

The test labeled Eq. (36) uses OLS residuals **for both** equations.

In all the tables below the model is

$$\text{Equation 1}: y_i = .5 - .6x_{i1} + .8x_{i2} - .4x_{i3} + .2x_{i4} + .3x_{i5} + u_i$$

$$\text{Equation 2}: x_{i5} = .3 - .8z_{i1} - .7z_{i2} + .6z_{i3} - .4z_{i4} + .2z_{i5} + .6x_{i1} - .3x_2 + 1.0x_{i3} - .6x_{i4} + v_i.$$

TABLE 1

Correlation between u and v = 0.

	50 Observations		75 Observations		100 Observations		200 Observations	
Test Type	Size	NCP	Size	NCP	Size	NCP	Size	NCP
Hausman 1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Hausman 2	0.049	0.000	0.052	0.000	0.049	0.000	0.045	0.000
Hausman 3	0.060	0.000	0.058	0.000	0.055	0.000	0.047	0.000
Test Eq. 16	0.060	0.000	0.058	0.000	0.055	0.000	0.047	0.000
Test Eq. 26	0.065	0.000	0.063	0.000	0.058	0.000	0.049	0.000

NCP stands for non-centrality parameter.

TABLE 2

Correlation between u and v = 0.1.

	50 Observations		75 Observations		100 Observations		200 Observations	
Test Type	Power	NCP	Power	NCP	Power	NCP	Power	NCP
Hausman 1	0.000	0.26	0.000	0.32	0.001	0.39	0.004	0.70
Hausman 2	0.073	0.26	0.087	0.32	0.109	0.39	0.188	0.70
Hausman 3	0.087	0.26	0.097	0.33	0.117	0.40	0.194	0.71
Test Eq. 26	0.087	0.15	0.097	0.23	0.117	0.31	0.194	0.62
Test Eq. 36	0.093	0.14	0.102	0.22	0.121	0.30	0.196	0.61

NCP stands for non-centrality parameter.

TABLE 3

Correlation between u and v = 0.3.

	50 Observations		75 Observations		100 Observations		200 Observations	
Test Type	Power	NCP	Power	NCP	Power	NCP	Power	NCP
Hausman 1	0.000	2.26	0.010	2.80	0.040	3.41	0.349	6.03
Hausman 2	0.307	2.26	0.488	2.80	0.626	3.41	0.916	6.03
Hausman 3	0.342	2.38	0.512	2.95	0.642	3.60	0.919	6.37
Test Eq. 26	0.342	1.36	0.512	2.07	0.642	2.78	0.919	5.60
Test Eq. 36	0.358	1.28	0.522	1.99	0.650	2.70	0.922	5.53

NCP stands for non-centrality parameter.

TABLE 4

Correlation between u and v = 0.5.

	50 Observations		75 Observations		100 Observations		200 Observations	
Test Type	Power	NCP	Power	NCP	Power	NCP	Power	NCP
Hausman 1	0.013	6.03	0.211	7.32	0.534	8.93	0.989	15.50
Hausman 2	0.764	6.03	0.931	7.32	0.982	8.93	1.000	15.50
Hausman 3	0.790	6.81	0.938	8.33	0.984	10.20	1.000	17.82
Test Eq. 26	0.790	3.78	0.938	5.75	0.984	7.71	1.000	15.58
Test Eq. 36	0.800	3.55	0.941	5.53	0.984	7.50	1.000	15.38

NCP stands for non-centrality parameter.

TABLE 5

Correlation between u and v = 0.7.

	50 Observations		75 Observations		100 Observations		200 Observations	
Test Type	Power	NCP	Power	NCP	Power	NCP	Power	NCP
Hausman 1	0.232	11.11	0.904	13.23	0.995	15.86	1.000	27.31
Hausman 2	0.989	11.11	1.000	13.23	1.000	15.86	1.000	27.31
Hausman 3	0.991	13.96	1.000	16.86	1.000	20.35	1.000	35.39
Test Eq. 26	0.991	7.42	1.000	11.26	1.000	15.12	1.000	30.51
Test Eq. 36	0.991	6.96	1.000	10.83	1.000	14.71	1.000	30.13

NCP stands for non-centrality parameter.

TABLE 6

Correlation between u and v = 0.9.

	50 Observations		75 Observations		100 Observations		200 Observations	
Test Type	Power	NCP	Size	NCP	Power	NCP	Power	NCP
Hausman 1	0.973	16.71	1.000	19.59	0.995	23.20	1.000	39.67
Hausman 2	1.000	16.71	1.000	19.59	1.000	23.20	1.000	39.67
Hausman 3	1.000	24.53	1.000	29.13	1.000	34.64	1.000	59.60
Test Eq. 26	1.000	12.25	1.000	18.63	1.000	25.00	1.000	50.43
Test Eq. 36	1.000	11.48	1.000	17.93	1.000	24.34	1.000	49.77

NCP stands for non-centrality parameter.

In Table 1 the first row lists only zeros; this is not because the actual statistic is precisely zero, or the test did not function as anticipated; rather it is zero because Hausman 1gets very few results "right". Table 1 shows that the Hausman 1, 2 tests do equally well but the test based on Eq. (36) is slightly closer to the true size. In terms of the other tables that entail correlated errors (u and v), Hausman 3 does generally better than Hausman

2, but the test based on Eq. (36), exhibits marginally higher power. As a generalization, the power of all tests is rather low for a correlation of .1, but it increases with the magnitude of the correlation, as well as the size of the sample.