Journal of Econometrics 29 (1985) 213–227. North-Holland

ON THE PERFORMANCE OF TESTS BY WU AND BY HAUSMAN FOR DETECTING THE ORDINARY LEAST SQUARES BIAS PROBLEM *

Alice NAKAMURA and Masao NAKAMURA

The University of Alberta, Edmonton, Alb., Canada T6G 2R6

Received October 1983, final version received March 1985

We first consider the performance of the Wu (1973) – Hausman (1978) (W-H) specification error test as a test for the existence of ordinary least squares (OLS) bias. We discuss power properties of the test under alternative null hypotheses, one of which has not previously been considered. We next consider how the W-H test performs as an indicator of the extent (rather than the existence) of an OLS bias problem, since this usage of the test seems common in applied studies. Finally Monte Carlo methods are used to evaluate Wu's two-step estimation procedure involving the W-H test as a pretest.

1. Introduction

In this paper we consider the performance of a specification error test proposed by Wu (1973) and by Hausman (1978) for detecting the ordinary least squares (OLS) bias problem in a linear simultaneous equations model.[1] For convenience, we refer to this test as the Wu–Hausman test. The Wu–Hausman test is closely related to, though not identical to, a test Durbin (1954) presents.

We first consider the performance of the Wu–Hausman test as a test for the existence of an OLS bias problem. We then consider whether the Wu–

*This research was supported in part by Social Sciences and Humanities Council of Canada Research Grant 410-77-0339 and a Leave Fellowship. Earlier versions of this paper were presented at the Econometric Society Meetings in New York, December 1982 and at an Econometrics Workshop at the University of Chicago in January 1983, as well as at the World Congress of the Econometric Society in Aix-en-Provence, August 1980. We are particularly grateful to Professors T. Amemiya, J.J. Heckman, J. Kmenta, J. Thursby and K.F. Wallis for their comments and encouragement at various crucial points in the development of this paper. We also thank Professors R. Carter, R.W. Farebrother, T. Kariya, J. Ramsey, N.S. Revankar, H. Tsurumi and D.M. Wu as well as anonymous referees and an associate editor for comments on earlier versions of the paper. We are, of course, solely responsible for any remaining errors or misinterpretations.

[1] We are referring here to Wu's T_2 and to the alternative formulation of the test by Hausman (1978, p. 1259). In this paper we do not discuss other proposed uses of the Wu and Hausman tests. See Nakamura and Nakamura (1981) for a proof of the equality of the Wu and Hausman test statistics in a linear model of the sort adopted in this paper.

Hausman test or test statistic might be used to judge the extent, as opposed to the existence, of an OLS bias problem. The test is often used in applied settings where there are strong theoretical or other a priori reasons for believing an OLS bias problem does exist. When the Wu–Hausman test is applied in such situations to determine whether there is a 'significant' OLS bias problem, perhaps it is really the extent of the bias problem that is at issue. Or the practitioner may implicitly be trying to judge the seriousness of various consequences of an OLS bias problem. For instance, an OLS bias problem will result in a departure of the actual from the stated probability of a Type I error for the usual t-test for the coefficient of an included endogenous variable. Monte Carlo methods are used to explore the extent to which a two-step estimation procedure proposed by Wu (1973), that uses the Wu–Hausman test as a pretest, overcomes this testing problem.

2. Model, test statistic and null hypotheses

Consider the two-equation linear system given by the structural equations

$$y_1 = \alpha_1 y_2 + Z_1\alpha_2 + \lambda_1 u_1, \tag{1}$$

and

$$y_2 = \gamma_1 y_1 + Z_1\gamma_2 + Z_2\gamma_3 + \gamma_4 u_2, \tag{2a}$$

or equivalently by (1) and the reduced form equation

$$y_2 = Z_1\beta_1 + Z_2\beta_2 + \lambda_2 u_1 + \lambda_3 u_2. \tag{2b}$$

In these equations y_1 and y_2 are $(n \times 1)$ vectors of observations;[2] Z_1 and Z_2 are $(n \times K_1)$ and $(n \times K_2)$ matrices of observations on K_1 predetermined variables that are included in the structural equation for y_1 and K_2 predetermined variables that are excluded from this equation; the u's are $(n \times 1)$ vectors of random disturbance terms that are each normally distributed with mean zero and a variance of one; the α's, γ's and β's are vectors of unknown parameters; and λ_1, λ_2 and λ_3 are unknown scalar parameters. It is assumed that the structural equation for y_1 is identified: hence we must have $K_2 \geq 1$. No assumption is made, however, concerning the identification of the structural equation for y_2.[3]

[2] In the relevant papers of Wu (1973, 1974) and Hausman (1978) y_2 is an $(n \times G_2)$ matrix. For expositional convenience, in this paper we have set $G_2 = 1$ as Durbin (1954) also does in his original paper.

[3] In fact, in the papers of Wu (1973, 1974) and Hausman (1978) the equation for y_2 is a reduced form equation in the sense that only predetermined variables appear on the right-hand side, but no structural equation for y_2 is given or asserted to exist nor is it assumed that any underlying structural equation for y_2 is identified. In the spirit of their papers the predetermined variables in Z_1 and Z_2 can be viewed simply as instruments for y_2.

Suppose we are concerned that the OLS estimator of α_1 in (1) is biased. Such a problem might be motivation for testing

$$\mathrm{H}_0^*\colon B=0 \quad \text{versus} \quad \mathrm{H}_a^*\colon B \neq 0, \tag{3}$$

where B is the asymptotic bias of the OLS estimator of α_1 in (1). We can rewrite the reduced form disturbance term as $\lambda_2 u_1 + \lambda_3 u_2 = \xi v_2$, where v_2 is a standard normal variable and ξ^2 is the variance of the reduced form disturbance term. Then for our model B is given by

$$B = \operatorname{plim}(b_1 - \alpha_1) = \frac{\operatorname{cov}(\lambda_1 u_1, \xi v_2)}{\operatorname{plim}(1/n)(y_2' A_1 y_2)} = \frac{\lambda_1 \xi \operatorname{cov}(u_1, v_2)}{\operatorname{var}^*(Z_2\beta_2) + \xi^2}$$

$$= \lambda_1 \rho \left(\frac{1}{\xi}\right) \frac{\xi^2}{\operatorname{var}^*(Z_2\beta_2) + \xi^2} = \lambda_1 \rho \left(\frac{1}{\xi}\right)\left(1 - R^2_{y_2 - Z_1\beta_1 \cdot Z_2}\right), \tag{4}$$

where b_1 is the OLS estimator of α_1, $A_1 = I - Z_1(Z_1'Z_1)^{-1}Z_1'$, $\operatorname{var}^*(Z_2 B_2) = \operatorname{var}(Z_2\beta_2) - \beta_2'\operatorname{cov}(Z_2, Z_1)\operatorname{var}(Z_1)^{-1}\operatorname{cov}(Z_2, Z_1)\beta_2$, $\rho = \operatorname{cov}(u_1, v_2)$, and $R^2_{y_2 - Z_1\beta_2 \cdot Z_2}$ denotes the multiple coefficient of determination from the regression of $(y_2 - Z_1\beta_1)$ on Z_2.[4] Or, if concern about an OLS bias problem arises from the potential correlation between y_2 and $\lambda_1 u_1$, we might test

$$\mathrm{H}_0'\colon \rho = 0 \quad \text{versus} \quad \mathrm{H}_a\colon \rho \neq 0, \tag{5}$$

where ρ is given for our model by

$$\rho = \operatorname{corr}(\lambda_1 u_1, \xi v_2) = \frac{\operatorname{cov}(\lambda_1 u_1, \xi v_2)}{\sqrt{\operatorname{var}(\lambda_1 u_1)\operatorname{var}(\xi v_2)}}$$

$$= \frac{\lambda_1 \xi \operatorname{cov}(u_1, v_2)}{\sqrt{\lambda_1^2 \xi^2}} = \operatorname{cov}(u_1, v_2). \tag{6}$$

Following a similar line of reasoning we might also consider testing

$$\mathrm{H}_0\colon \delta = 0 \quad \text{versus} \quad \mathrm{H}_a\colon \delta \neq 0, \tag{7}$$

where

$$\delta = \operatorname{cov}(\lambda_1 u_1, \xi v_2) = \lambda_1 \xi \operatorname{cov}(u_1, v_2) = \lambda_1 \xi \rho. \tag{8}$$

[4] Notice that $R^2_{y_2 - Z_1\beta_1 \cdot Z_2}$ will be larger the smaller ξ^2 is compared with the variability of $Z_1\beta_1 + Z_2\beta_2$ in (2b), and the larger the proportion is of the variability of y_2 that is explained by the variables that are excluded from, as opposed to the variables that are included in, the equation of interest for y_1.

The null hypothesis that Wu (1973) considers is H_0. Durbin (1954) and Hausman (1978) do not explicitly state the null hypothesis being tested. Hausman and Taylor (1981, p. 13) propose the null hypothesis H_0^* stating: 'It appears in practice... that H_0 is frequently tested in situations where we can infer from the subsequent actions taken that the hypothesis H_0^* was intended...'

The Wu–Hausman test may be used to test (3), (5) or (7). For our model, the statistic for this test may be written as[5]

$$T_2 = c_1(Q^*/Q_2), \tag{9}$$

where $c_1 = (n - K_1 - 2G_2)/G_2$, $Q^* = (b_1 - b_2)'[(y_2'A_2y_2)^{-1} - (y_2'A_1y_2)^{-1}]^{-1}(b_1 - b_2)$, $Q_2 = Q_4 - Q^*$, $Q_4 = (y_1 - y_2b_1)'A_1(y_1 - y_2b_1)$, b_1 and A_1 are defined as above, $b_2 = (y_2'A_2y_2)^{-1}y_2'A_2y_1$ is the instrumental variables (IV) method estimator of α_1 in (1), and $A_2 = Z(Z'Z)^{-1}Z' - Z_1(Z_1'Z_1)^{-1}Z_1'$ where $Z = (Z_1, Z_2)$. The statistic T_2 is not the only statistic that could be used to test (3), (5) or (7). Wu (1973, 1974) proposes three other statistics (T_1, T_3, T_4) which all share the same numerator, Q^*. Hausman (1978) also proposes an IV form of his test statistic that is identical to a statistic presented by Durbin (1954). Depending on the estimator used for the variance of $\lambda_1 u_1$, Durbin's statistic and the IV form of Hausman's statistic are identical to either Wu's T_3 or T_4 statistic. The statistics of Revankar and Hartley (1973) and Revankar (1978) also share the numerator Q^*, and are identical to Wu's T_2 when the structural equation for y_1 is just-identified.[6] The following discussion is in terms of the T_2 test because Wu (1973, 1974) gives theoretical and Monte Carlo results indicating that T_2 is to be preferred to his other statistics. Nevertheless, because of the close relationship among these statistics, our results hold for the other tests mentioned above as well.

3. Power properties

Except for degenerate cases, λ_1^2, ξ^2 and $(1 - R^2_{y_2 - Z_1\beta_1 \cdot Z_2})$ must all be non-zero for the model given by (1) and (2a), or by (1) and (2b). Thus from (4), (6) and (8) we see that the null hypotheses H_0^*, H_0' and H_0 are all equivalent[7] in the sense that in any particular case they are all either true or false, depending on whether ρ is zero or non-zero. The same test statistic and procedure are used for the Wu–Hausman test of each of these three null

[5] See Kariya and Hodoshima (1980, p. 47, eq. 3.16).

[6] See Nakamura and Nakamura (1981, p. 1587) and Kariya and Hodoshima (1980, p. 47).

[7] Holly (1982) argues that in some cases H_0 and H_0^* do not imply each other. For the model used in this paper, which is also the model treated by Durbin (1954), Wu (1973) and Hausman (1978), however, H_0, H_0^* and H_0' imply each other except for degenerate cases. See Nakamura and Nakamura (1982) for a proof based on a maximum likelihood approach.

hypotheses. For any given set of parameter values for (1) and (2a) or (2b), the rejection rate must be the same for the tests of H_0^*, H_0' or H_0. However, the power function for a test is the relationship between the rejection rate for the test of a given null hypothesis and the size of the departure from this null hypothesis, so it matters if we describe the power function for the Wu–Hausman test in terms of B, δ or ρ. Clearly the graph of the (common) rejection rate for the Wu–Hausman tests of H_0^*, H_0 or H_0' versus, say, B will not be the same as the graph of this rejection rate versus ρ or δ. These differences can be explored by considering the determinants of B, δ and ρ and the distribution of T_2.

We will first consider T_2. The probability that the absolute value of T_2 will exceed some critical value is an increasing function of ρ^2, ξ^2 and $R^2_{y_2-Z_1\beta_1 \cdot Z_2}$.[8] However, this probability does not depend directly on λ_1. This can be established as follows. From Wu (1973, eq. 2.10) we have $b_1 - b_2 = C(\lambda_1 u_1)$ where $C = (y_2'A_1y_2)^{-1}y_2'A_2 - (y_2'A_2y_2)^{-1}y_2'A_2$. Also from Wu (1973, p. 737) we have $Q_2 = \lambda^2 u_1'(H - C'D^{-1}C)u_1$, where $H = A_1 - A_1y_2(y_2'A_1y_2)^{-1}y_2'A_1$ and $D = (y_2'A_2y_2)^{-1} - (y_2'A_1y_2)^{-1}$. Thus we can rewrite T_2 as

$$T_2 = c_1 \frac{\lambda_1^2 u_1'C'\left[(y_2'A_2y_2)^{-1} - (y_2'A_1y_2)^{-1}\right]^{-1}Cu_1}{\lambda_1^2 u_1'(H - C'D^{-1}C)u_1}, \tag{10}$$

and the parameter λ_1^2 can be cancelled out of (10) just as λ_1 cancels out of the expression for ρ given in (6). Of course, the distribution of T_2 will still depend indirectly on λ_1 if the distribution of y_2 involves λ_1. Without conditions on the coefficients of the *structural* equation for y_2, however, λ_1 may take on any value for any given values of λ_2 and λ_3. In fact, both the distribution of T_2 and the rejection rate of the Wu–Hausman test are invariant to changes in the structural equation for y_2 that leave the reduced form equation for y_2 unaltered. This point can perhaps be clarified with an example.

We let $Z_1 = (z_{1,1}, z_{1,2})$, $Z_2 = (z_{2,1}, z_{2,2}, z_{2,3})$, and we set α_1, all the elements of α_2 in (1) and all the elements of β_1 and β_2 in (2b) equal to 1. Thus the model is

$$y_1 = y_2 + z_{1,1} + z_{1,2} + \lambda_1 u_1, \tag{11}$$

and

$$y_2 = z_{1,1} + z_{1,2} + z_{2,1} + z_{2,2} + z_{2,3} + \lambda_1 u_1 + \lambda_3 u_2. \tag{12}$$

By controlling the values of λ_1, λ_2 and λ_3 we can determine the values of ρ^2 and B, and the population value of $R^2_{y_2 \cdot z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2}, z_{2,3}}$ denoted hereafter simply by R^2. Because the coefficients of the reduced form equation for y_2 are

[8]See Nakamura and Nakamura (1984a) for a proof.

fixed in (12), by determining the value of R^2 we also determine the value of $R^2_{y_2 - Z_1\beta_1 \cdot Z_2}$. For this model we see that $\rho^2 = \lambda_2/(\lambda_2^2 + \lambda_3^2)$, $B = \lambda_1\lambda_2/(3 + \lambda_2^2 + \lambda_3^2)$, and $R^2 = 5/(5 + \lambda_2^2 + \lambda_3^2)$. Thus for any given non-zero values for ρ^2, R^2 and B we have $\lambda^2 = \sqrt{5\rho^2(1 - R^2)/R^2}$, $\lambda_3 = \sqrt{5(1 - \rho^2)(1 - R^2)/R^2}$ and $\lambda_1 = (B/\lambda_2)(3 + \lambda_2^2 + \lambda_3^2)$. When $\rho^2 = 0$ (and $\lambda_2 = \delta = B = 0$), λ_1 may take on any non-zero value and the model is recursive.

We first choose the values of λ_1, λ_2 and λ_3 such that $\rho^2 = 0.5$, $R^2 = 0.5$ and $B = 0.2$. Thus we have

$$y_1 = y_2 + z_{1,1} + z_{1,2} + 1.012u_1, \tag{13}$$

and

$$y_2 = z_{1,1} + z_{1,2} + z_{2,1} + z_{2,2} + z_{2,3} + 1.581u_1 + 1.581u_2. \tag{14}$$

Next we choose the values of the λ's so $\rho^2 = 0.5$, $R^2 = 0.5$ and $B = 2$. Thus:

$$y_1 = y_2 + z_{1,1} + z_{1,2} + 10.120u_1, \tag{15}$$

and

$$y_2 = z_{1,1} + z_{1,2} + z_{2,1} + z_{2,2} + z_{2,3} + 1.581u_1 + 1.581u_2. \tag{16}$$

The following structural equations for y_2 are consistent with the models given by (13) and (14), and by (15) and (16), respectively:

$$\begin{aligned} y_2 = {} & 0.610y_1 - 0.22z_{1,1} - 0.22z_{1,2} \\ & + 0.39z_{2,1} + 0.39z_{2,2} + 0.39z_{2,3} + 0.616u_2, \end{aligned} \tag{17}$$

and

$$\begin{aligned} y_2 = {} & 0.135y_1 - 0.73z_{1,1} - 0.73z_{1,2} \\ & + 0.865z_{2,1} + 0.865z_{2,2} + 0.865z_{2,3} + 1.367u_2. \end{aligned} \tag{18}$$

For (13) and (17) we have $\lambda_2 = 1.562\lambda_1$, and for (15) and (18) we have $\lambda_2 = 0.156\lambda_1$. We could never determine these relationships empirically, though, because (17) and (18) are underidentified. The variance of the equation disturbance term is 100 times as large for eq. (15) as for eq. (13), and the asymptotic OLS bias measured as a percentage of the true value of α_1 is 20% for the model given by (13) and (14) and 200% for the model given by (15) and (16). However, for both models the power of the Wu–Hausman test of H_0^*, H_0 or H_0' using a critical region of 0.05 is approximately 45% for $n = 20$, 91% for $n = 40$ and 100% for $n = 100$. (These results were derived by Monte Carlo experiments conducted in the manner described later in this section.) This is because the distribution of the test statistic T_2 is the same for both models. It should be recalled that the existence and identification of a structural equation for y_2 is not a requirement for application of the Wu–Hausman test as this test was originally presented by Wu and by Hausman.

Of course, even when the structural equation for y_2 is identified and the relationship between λ_2 and λ_1 can be empirically determined, this relationship is still model-specific. When the equation for y_2 is *not* derived from a structural model for y_1 and y_2, the distribution of T_2 may not involve λ_1 at all. One such situation might be when the suspected correlation between y_2 and u_1 in (1) is due to measurement errors in y_2.

We will now turn our attention to the determinants of B, δ and ρ. From (4) we see that the magnitude (or absolute value) of B, the parameter restricted under H_0^*, is an increasing function of the magnitudes of λ_1 and ρ, and a decreasing function of the magnitudes of ξ and of $R^2_{y_2 - Z_1\beta_1 \cdot Z_2}$ to the extent that these parameters can be varied independently.[9] From (8) we see that the magnitude of δ, the parameter restricted under H_0, is an increasing function of the magnitudes of λ_1, ξ and ρ, but does not depend on $R^2_{y_2 - Z_1\beta_1 \cdot Z_2}$. Finally from (6) we see that ρ, the parameter restricted under H_0', does not depend directly on λ_1, ξ or $R^2_{y_2 - Z_1\beta_1 \cdot Z_2}$.

Nuisance parameters are parameters that are assumed to be fixed in the sense that they take on the same values under both the null and alternative hypotheses. If we consider ρ, λ_1, ξ and $R^2_{y_2 - Z_1\beta_1 \cdot Z_2}$ *all* to be nuisance parameters we cannot consider a power function for the Wu–Hausman test because we cannot consider ranges of values for ρ, or for B and δ which are entirely determined by the nuisance parameters.

No one is suggesting that ρ be considered as a nuisance parameter, of course. Rather the spirit of most of the literature on the Wu–Hausman test is that λ_1, ξ and $R^2_{y_2 - Z_1\beta_1 \cdot Z_2}$, but *not* ρ, are to be treated as nuisance parameters (this parameterization has not been used by others). In this case, since B and δ are linear functions of ρ, although the power function of the Wu–Hausman test will have a different appearance depending on whether we describe it in terms of B, δ or ρ, these different representations of the power function will be uniquely related for any given model. However, these relationships will depend crucially on the values of the nuisance parameters. Also because B depends on λ_1, while ρ and the rejection rate for the Wu–Hausman test of H_0^*, H_0 or H_0' do not, the power of the Wu–Hausman test of H_0^* can be low for plausible models with large values of B. For such models, the Wu–Hausman test will perform poorly as an indicator of the existence of an OLS bias problem.

This point can be demonstrated by Monte Carlo methods. We use the model given by (11) and (12) in these experiments, except that we assume that the z's

[9] We can vary ξ while keeping $R^2_{y_2 - Z_1\beta_1 \cdot Z_2}$ fixed by varying $\text{var}^*(Z_2\beta_2)$. Similarly, $R^2_{y_2 - Z_1\beta_1 \cdot Z_2}$ can be varied while keeping ξ fixed by varying $\text{var}^*(Z_2\beta_2)$.

as well as the u's are identically distributed standard normal variables. For convenience we assume u_1 and u_2 are independently distributed.[10]

In table 1 we show rejection rates for the Wu–Hausman tests of H_0^* or H_0' or H_0, where all three of these null hypotheses will be rejected in exactly the same cases. The power for the test of H_0': $\rho = 0$ can be seen to rise as the values of ρ^2 rise, with the rate of this rise depending on the value of R^2.[11] From table 1 we also see that, when ρ^2 or R^2 are low, the power of the test of H_0^*: $B = 0$ can be quite low for fairly large values of B.

Suppose now that we do *not* consider λ_1, ξ and $R^2_{y_2 - Z_1\beta_1 \cdot Z_2}$ to be nuisance parameters when testing H_0^*. These parameters are not explicitly restricted under H_0^*, but neither is ρ. Moreover, ρ, λ_1, ξ and R^2 are all implicitly restricted under H_0^* since B depends on all of them. (Likewise ρ, λ_1 and ξ are implicitly restricted under H_0.) Viewing our results in this way, we see that it is possible to increase B without increasing the power of the Wu–Hausman test of H_0^*, as has been done in our experiments (see table 1). One desirable property for a statistical test to have is that, for a fixed sample size and size of test, the power of the test increases as the departure from the null hypothesis increases [see Rao (1973, p. 460)]. When parameters implicitly restricted under the null are not considered to be nuisance parameters, the Wu–Hausman test of H_0^* (or H_0) does not possess this property. This problem does not arise for the Wu–Hausman test of H_0': $\rho = 0$, but this null hypothesis has not been considered in the literature on the Wu–Hausman test.

4. Measuring the extent of a bias problem

The Wu–Hausman test was proposed as a test for the existence of an OLS bias problem.[12] Often, however, the existence of such a problem is known or

[10] In dealing with a related class of problems, Sawa (1969, p. 925) notes that 'without loss of generality we can assume the orthonormality of exogenous variables'. The variance and MSE of the IV estimator of α_1 are finite for the model given by (11) and (12), although this is not a requirement for application of the Wu–Hausman test. Normal random values for the z's and u's were generated using the polar method by the subroutine GGNPM in the IMSL library. We generated 200 sets of series for the z's and u's appearing in (11) and (12) for each combination of values for n, R^2, ρ^2 and B. In tables 2–4 results are shown for each pair of values of n and R^2 through the first value of ρ^2 for which the power of the test is 1 in table 1. In these experiments we are varying $R^2_{y_2 - Z_1\beta_1 \cdot Z_2}$ by varying R^2. The parameterization of our Monte Carlo experiments is motivated by Sawa's (1969) parameterization of the OLS and 2SLS estimators, and by our decompositions [given in Nakamura and Nakamura (1984a)] of the non-central parameters δ_1 and δ_2 of the doubly non-central F distribution which Kariya and Hodoshima (1980) show is the exact distribution of the Wu–Hausman statistic for the model adopted in this paper conditional on the values of the OLS estimates of β_2 and the variance of the disturbance term. When u_1 and u_2 are independently distributed, we have $\mathrm{var}(\lambda_1 u_1 + \lambda_3 u_2) = \lambda_2^2 + \lambda_3^2$.

[11] By fixing the values of the coefficients of the reduced form equation for y_2 we also fix the relationship between R^2 and $R^2_{y_2 - Z_1\beta_1 \cdot Z_2}$. Thus we can vary $R^2_{y_2 - Z_1\beta_1 \cdot Z_2}$ by varying R^2.

[12] See Wu (1973, 1974) and Hausman (1978). See also Durbin (1954). Other possible uses of the test are suggested in these papers as well.

Table 1

Percentage rejection rates (power results) for the Wu–Hausman test of H_0^*: $B = 0$, H_0: $\delta = 0$ or H_0': $\rho = 0$, using a two-tailed critical region of 0.05.[a]

R^2, B, ρ^2	$n = 20$	$n = 40$	$n = 100$	$n = 250$	$n = 500$
0.2, 0, 0	6.0	5.0	5.5	4.5	4.5
0.2, 0.2, 0.1	6.5	9.0	14.0	45.0	74.5
0.2, 0.8, 0.1	6.5	8.5	19.5	49.5	72.5
0.2, 0.2, 0.3	8.0	19.0	57.5	95.5	100.0
0.2, 0.8, 0.3	13.0	23.5	60.0	93.0	100.0
0.2, 0.2, 0.5	13.0	40.5	86.5	100.0	
0.2, 0.8, 0.5	13.0	35.0	86.0	100.0	
0.2, 0.2, 0.7	28.5	71.5	99.5	100.0	
0.2, 0.8, 0.7	26.0	67.0	99.0		
0.2, 0.2, 0.9	57.0	94.5	100.0		
0.2, 0.8, 0.9	53.5	94.5	100.0		
0.5, 0, 0	5.5	5.5	5.5	5.0	4.5
0.5, 0.2, 0.1	7.5	20.5	45.5	89.0	99.5
0.5, 0.8, 0.1	8.0	17.5	47.0	89.0	99.5
0.5, 0.2, 0.3	23.5	57.5	96.5	100.0	100.0
0.5, 0.8, 0.3	27.5	54.0	99.5	100.0	100.0
0.5, 0.2, 0.5	45.0	91.5	100.0		
0.5, 0.8, 0.5	48.0	86.5	100.0		
0.5, 0.2, 0.7	70.5	98.0			
0.5, 0.8, 0.7	69.0	97.0			
0.5, 0.2, 0.9	90.5	100.0			
0.5, 0.8, 0.9	94.5	100.0			
0.8, 0, 0	3.0	4.5	5.0	4.5	5.5
0.8, 0.2, 0.1	19.0	34.0	72.5	100.0	100.0
0.8, 0.8, 0.1	16.0	29.5	73.0	100.0	100.0
0.8, 0.2, 0.3	42.0	81.0	99.5		
0.8, 0.8, 0.3	42.0	87.0	99.5		
0.8, 0.2, 0.5	67.5	99.5	100.0		
0.8, 0.8, 0.5	77.0	99.0	100.0		
0.8, 0.2, 0.7	90.5	100.0			
0.8, 0.8, 0.7	95.0	100.0			
0.8, 0.2, 0.9	98.5				
0.8, 0.8, 0.9	99.0				

[a]The Wu–Hausman test will reject H_0, H_0^* and H_0' in exactly the same cases. The number of repetitions in each case is 200. The length of the series is denoted by n.

assumed a priori.[13] In such cases, it is the extent, not the existence, of an OLS bias problem that is really in question.[14] In this section we consider the extent to which the Wu–Hausman test is able to pick out cases where the sample bias ($\hat{B}$) is small. For any given model with a given value of B, there is a sampling distribution of values for $\hat{B}$. However, if $\hat{B}=0$ (or is nearly zero) for some particular data sample, then we have a perfect (or nearly perfect) estimate of α_1 no matter what the value of B is. It is the bias of the estimate, not the estimator, that is of ultimate concern in many empirical studies. The Wu–Hausman test (or test statistic)[15] could be used to pick out cases where $\hat{B}$ is small if there were high positive correlations between $\hat{B}$ and T_2 for samples drawn from populations with any given values of ρ, λ_1, ξ and $R^2_{y_2 - Z_1\beta_1 \cdot Z_2}$.

Standard practice in developing a new test is to investigate the population properties of the test, including finite sample population properties. The property being considered in this section is a case selection, not a population, property.[16] Wu (1973) and Hausman (1978) make no claims concerning the case selection properties of their test, but it is of interest to see if the test could be used to pick out cases with small values of $\hat{B}$ for two reasons. First, the form of the respective formulas suggests there might be a positive correlation between $\hat{B}$ and T_2. Second, a similar case selection property has been established for another specification error test.[17] However, the correlations between the absolute values of T_2 and $\hat{B}$ in table 2 are all small, and are often negative for smaller values of n.

In table 3 we show the sample means for $\hat{B}$ for given values of n, ρ^2, R^2 and B for the cases where the Wu–Hausman (W-H) test accepts H_0^* (or H_0 or H_0')

[13] In fact, bias is implied by the assumed models in many applied settings. See, for example, Boulier and Rosenzweig (1984, pp. 719, 727), Eichengreen (1984, pp. 1002, 1005), and Nakamura and Nakamura (1984b).

[14] Arnold Zellner suggested in discussion with us that $(b_1 - b_2)$ might be directly used as a point estimator of $B = \text{plim}(b_1) - \alpha_1$, since b_2 is a consistent estimator of α_1 in (1) under H_0^*, H_0 and H_0' as well as under H_a^*, H_a and H_a'. It is only under H_0^*, H_0 and H_0', however, that a computationally tractable and consistent estimator of $\text{var}(b_1 - b_2)$ has been derived [see Durbin (1954, p. 29) and Hausman (1978)].

[15] The Wu–Hausman test of H_0^* (or H_0 or H_0') is a consistent test. [See Nakamura and Nakamura (1984a) for a proof of the consistency of this test that does not assume local misspecification alternatives.] Thus for large samples, H_0^* will always be rejected when B is non-zero no matter how small B may be.

[16] Notice that case selection properties are irrelevant for classical tests. If, for instance, we are testing the null hypothesis that α_1 equals some specified value in (1), we only care about what this test tells us about the *population* value of α_1. On the other hand, in the Wu–Hausman test of H_0^*: $B=0$, we care about the value of $\hat{B}$ because $\hat{B}$ is the distance between the OLS estimate of α_1 and the true value of this parameter.

[17] From results presented in Nakamura, Nakamura and Orcutt (1976), it can be seen that when the sample autocorrelation coefficient for the error term of a simple regression is close to zero, there is no 'autocorrelation problem' in terms of the usual testing problems resulting from the autocorrelation of the error term, regardless of the population value of the autocorrelation coefficient.

Table 2

Simple correlations between the absolute sample values of the Wu–Hausman test statistic and the absolute values of $\hat{B}$, the deviations of the OLS estimates of α_1 in (11) from the true value of $\alpha_1(=1)$.[a]

R^2, B, ρ^2	$n=20$	$n=40$	$n=100$	$n=250$	$n=500$
0.2, 0, 0	−0.07	−0.07	0.05	0.10	0.06
0.2, 0.2, 0.1	0.04	−0.03	0.04	−0.01	0.04
0.2, 0.8, 0.1	0.06	0.03	0.05	0.09	0.13
0.2, 0.2, 0.3	−0.05	−0.06	0.02	−0.01	
0.2, 0.8, 0.3	0.01	−0.02	−0.05	0.02	
0.2, 0.2, 0.5	−0.00	−0.03	0.14		
0.2, 0.8, 0.5	−0.15	−0.11	0.16		
0.2, 0.2, 0.7	−0.06	−0.14	−0.05		
0.2, 0.8, 0.7	−0.20	−0.04	−0.08		
0.2, 0.2, 0.9	−0.42	−0.18			
0.2, 0.8, 0.9	−0.29	−0.39			
0.5, 0, 0	−0.07	0.09	−0.05	−0.01	−0.05
0.5, 0.2, 0.1	0.08	−0.00	0.00	−0.02	0.05
0.5, 0.8, 0.1	−0.04	0.02	−0.02	0.07	0.02
0.5, 0.2, 0.3	0.08	0.06	−0.01		
0.5, 0.8, 0.3	−0.19	−0.01	0.03		
0.5, 0.2, 0.5	0.01	−0.03			
0.5, 0.8, 0.5	−0.12	−0.06			
0.5, 0.2, 0.7	−0.10	−0.20			
0.5, 0.8, 0.7	−0.11	−0.19			
0.5, 0.2, 0.9	−0.22	—[b]			
0.5, 0.8, 0.9	−0.27				
0.8, 0, 0	−0.10	0.03	−0.03	−0.03	0.00
0.8, 0.2, 0.1	−0.02	−0.12	0.01		
0.8, 0.8, 0.1	−0.01	−0.02	0.00		
0.8, 0.2, 0.3	0.01	−0.00	0.01		
0.8, 0.8, 0.3	0.18	0.16	−0.09		
0.8, 0.2, 0.5	0.00	−0.08			
0.8, 0.8, 0.5	−0.06	−0.01			
0.8, 0.2, 0.7	−0.05				
0.8, 0.8, 0.7	−0.09				
0.8, 0.2, 0.9	−0.12				
0.8, 0.8, 0.9	−0.16				

[a] See footnote to table 1.

[b] Results are only shown through the first pair of values of R^2 and ρ^2 (for the given value of n) for which the power of the test is found to be one in table 1.

and the cases where H_0^* (or H_0 or H_0') is rejected using a two-tailed critical region of 0.05. The sample means for $\hat{B}$ for the cases where H_0^* is accepted rise and fall with the true values of B in the same way as for the cases where H_0^* is rejected. For any given non-zero values of ρ^2, R^2 and B, what we would like to have found is

$$\mathrm{E}(\hat{B} \mid \text{W-H tests accepts } H_0^*) \simeq 0, \tag{19}$$

Table 3

Mean sample biases for OLS estimator of α_1 in (11) for cases where H_0^*: $B=0$ was accepted and cases where H_0^*: $B=0$ was rejected.[a]

			H_0^*: $B=0$ accepted					H_0^*: $B=0$ rejected				
R^2,	B,	ρ^2	$n=20$	$n=40$	$n=100$	$n=250$	$n=500$	$n=20$	$n=40$	$n=100$	$n=250$	$n=500$
0.2,	0,	0	−0.00	−0.00	0.00	0.00	0.00	0.02	−0.01	−0.00	0.00	0.00
0.2,	0.2,	0.1	0.19	0.21	0.20	0.20	0.19	0.25	0.17	0.20	0.19	0.20
0.2,	0.8,	0.1	0.85	0.84	0.80	0.78	0.79	0.67	0.76	0.82	0.81	0.83
0.2,	0.2,	0.3	0.21	0.20	0.20	0.20		0.21	0.19	0.20	0.20	
0.2,	0.8,	0.3	0.82	0.81	0.80	0.76		0.84	0.81	0.80	0.81	
0.2,	0.2,	0.5	0.20	0.20	0.20			0.20	0.20	0.20		
0.2,	0.8,	0.5	0.82	0.80	0.77			0.73	0.77	0.79		
0.2,	0.2,	0.7	0.19	0.21	0.21			0.19	0.20	0.20		
0.2,	0.8,	0.7	0.80	0.80	0.80			0.76	0.80	0.80		
0.2,	0.2,	0.9	0.21	0.21				0.19	0.20			
0.2,	0.8,	0.9	0.83	0.87				0.78	0.79			
0.5,	0,	0	−0.00	−0.00	0.00	0.00	0.00	0.03	0.01	−0.00	0.00	0.00
0.5,	0.2,	0.1	0.19	0.20	0.21	0.20	0.20	0.21	0.21	0.20	0.20	0.19
0.5,	0.8,	0.1	0.81	0.78	0.83	0.74	1.14	0.55	0.71	0.85	0.80	0.82
0.5,	0.2,	0.3	0.19	0.20	0.21			0.24	0.21	0.20		
0.5,	0.8,	0.3	0.83	0.80	0.83			0.68	0.77	0.82		
0.5,	0.2,	0.5	0.20	0.21				0.20	0.20			
0.5,	0.8,	0.5	0.83	0.80				0.77	0.78			
0.5,	0.2,	0.7	0.20	0.25				0.20	0.19			
0.5,	0.8,	0.7	0.83	0.78				0.80	0.78			
0.5,	0.2,	0.9	0.23					0.19				
0.5,	0.8,	0.9	0.95					0.79				
0.8,	0,	0	−0.00	0.01	−0.00	0.00	0.00	−0.01	−0.02	−0.00	−0.00	−0.00
0.8,	0.2,	0.1	0.23	0.20	0.22			0.13	0.17	0.21		
0.8,	0.8,	0.1	0.80	0.69	0.77			1.19	0.68	0.83		
0.8,	0.2,	0.3	0.20	0.22	0.27			0.18	0.20	0.20		
0.8,	0.8,	0.3	0.80	0.78	0.96			0.76	0.85	0.81		
0.8,	0.2,	0.5	0.18	0.21				0.19	0.20			
0.8,	0.8,	0.5	0.79	0.37				0.82	0.80			
0.8,	0.2,	0.7	0.20					0.21				
0.8,	0.8,	0.7	1.06					0.81				
0.8,	0.2,	0.9	0.23					0.19				
0.8,	0.8,	0.9	1.13					0.78				

[a]See footnote a to table 1 and footnote b to table 2.

or, at least,

$$\mathrm{E}(\hat{B}|\text{W-H test accepts } \mathrm{H}_0^*) < \mathrm{E}(\hat{B}|\text{W-H test rejects } \mathrm{H}_0^*). \tag{20}$$

But what we have found instead is

$$\mathrm{E}(\hat{B}|\text{W-H test accepts } \mathrm{H}_0^*) \simeq \mathrm{E}(\hat{B}|\text{W-H test rejects } \mathrm{H}_0^*) \simeq \mathrm{E}(\hat{B}). \tag{21}$$

Table 4

Percentage rejection rates for the null hypothesis that α_1 in eq. (11) equals its true value (Type I errors) using OLS, IV and our MIXED method[a] to estimate (11) and using a two-tailed critical region of five percent for the test for α_1.[b]

			$n=20$			$n=40$			$n=100$			$n=250$			$n=500$		
R^2,	B,	ρ^2	OLS	IV	MIXED	OLS	IV	MIXED	OLS	IV	MIXED	OLS	IV	MIXED	OLS	IV	MIXED
0.2,	0,	0	4.0	1.5	4.5	5.5	1.0	5.5	7.5	4.5	10.0	6.0	4.5	8.5	6.0	4.0	8.5
0.2,	0.2,	0.1	20.0	3.0	18.5	45.0	3.0	43.0	87.5	5.0	74.5	100.0	5.5	33.3	100.0	3.0	26.5
0.2,	0.8,	0.1	28.0	4.0	27.0	49.5	3.0	45.0	85.0	6.5	69.5	100.0	4.5	50.5	100.0	4.5	27.3
0.2,	0.2,	0.3	64.0	14.0	60.0	92.5	13.0	74.0	100.0	8.5	42.5	100.0	7.0	10.0			
0.2,	0.8,	0.3	64.0	9.0	58.0	92.5	11.0	71.5	100.0	6.5	40.0	100.0	7.5	9.5			
0.2,	0.2,	0.5	88.5	20.0	76.5	100.0	16.0	60.5	100.0	9.5	16.5						
0.2,	0.8,	0.5	89.5	22.0	79.5	100.0	20.0	66.5	100.0	12.5	19.5						
0.2,	0.2,	0.7	99.5	29.0	75.0	100.0	20.0	34.0	100.0	10.0	10.5						
0.2,	0.8,	0.7	99.5	30.0	74.5	100.0	27.0	41.0	100.0	12.5	13.0						
0.2,	0.2,	0.9	100.0	38.3	57.5	100.0	27.5	29.0									
0.2,	0.8,	0.9	100.0	48.0	62.5	100.0	29.0	31.0									
0.5,	0,	0	5.0	2.0	4.0	4.5	2.5	5.0	6.5	5.0	8.0	7.0	5.5	7.5	6.0	4.0	7.5
0.5,	0.2,	0.1	14.5	3.0	13.0	35.5	5.0	29.0	74.5	5.5	45.0	99.0	5.5	13.5	100.0	2.5	2.5
0.5,	0.8,	0.1	19.5	3.0	19.5	33.5	5.0	29.5	77.5	5.0	42.5	98.0	5.0	13.5	100.0	4.0	4.0
0.5,	0.2,	0.3	48.5	5.0	35.5	84.0	6.5	35.5	99.5	3.5	6.0						
0.5,	0.8,	0.3	42.0	7.5	34.0	77.5	6.5	38.5	99.5	7.0	9.5						
0.5,	0.2,	0.5	77.0	10.0	46.5	97.5	9.0	14.0									
0.5,	0.8,	0.5	74.0	16.0	44.5	94.5	8.5	18.5									
0.5,	0.2,	0.7	90.0	16.5	31.5	99.5	9.0	9.5									
0.5,	0.8,	0.7	90.5	11.5	33.5	99.5	10.0	11.0									
0.5,	0.2,	0.9	97.5	16.5	21.5												
0.5,	0.8,	0.9	96.5	19.5	22.0												
0.8,	0,	0	6.5	5.0	6.5	7.5	5.0	8.0	6.5	4.5	7.5	6.5	5.0	6.5	7.0	4.0	7.5
0.8,	0.2,	0.1	10.5	6.5	11.0	15.5	5.5	12.5	46.0	7.0	15.5						
0.8,	0.8,	0.1	11.0	3.5	8.5	14.0	5.5	12.5	42.0	6.0	14.5						
0.8,	0.2,	0.3	20.0	5.5	17.5	46.5	7.5	14.0	88.5	4.0	4.0						
0.8,	0.8,	0.3	23.0	6.5	14.5	48.0	7.5	11.0	88.0	5.5	5.5						
0.8,	0.2,	0.5	34.0	7.5	16.0	71.0	6.5	7.0									
0.8,	0.8,	0.5	35.5	5.0	8.5	70.5	8.0	8.5									
0.8,	0.2,	0.7	56.0	8.5	10.5												
0.8,	0.8,	0.7	54.5	8.5	10.0												
0.8,	0.2,	0.9	56.5	7.0	7.0												
0.8,	0.8,	0.9	62.0	8.0	8.5												

[a] In the MIXED method, OLS is used to estimate (11) when H_0': $\rho = 0$ is accepted and IV is used to estimate (11) when H_0': $\rho = 0$ is rejected. The test of significance for ρ (the pretest) is carried out using the Wu–Hausman test with a two-tailed critical region of 0.05.

[b] See footnote a to table 1 and footnote b to table 2.

The results in tables 2 and 3 show that neither the Wu–Hausman test nor T_2 can be used to pick out estimated models where $\hat{B}$ is close to zero.

5. Use of the Wu–Hausman test as a pretest

One reason for concern about OLS bias is that it leads to Type I errors that are larger than the stated size for the usual test of significance for the coefficient of the included endogenous variable. In this section we consider whether this testing problem can be alleviated by basing the test of significance for the coefficient of the included endogenous variable on the OLS estimation results when H_0^* (or H_0 or H_0') is accepted by the Wu–Hausman test and on the IV estimation results otherwise.[18]

We use the term MIXED to indicate that OLS or IV estimation is used, respectively, depending on whether the Wu–Hausman test with a two-tailed five-percent critical region accepts or rejects H_0^*. In table 4 we compare the observed proportions (or probabilities) of Type I errors for the OLS, IV and MIXED method t-tests of H_0'': $\alpha_1 = 1$ versus H_a'': $\alpha_1 \neq 1$ using a two-tailed five-percent critical region, where one is the true value of α_1. For non-zero values of ρ, as n increases the probability of a Type I error tends toward hundred percent for OLS and toward the specified level of five percent for IV. The probabilities of a Type I error reported in table 4 for the MIXED method are always smaller than for OLS, but are still large. The advantage of IV over the MIXED method in this respect is large even for values of ρ quite close to zero, until n becomes large enough that IV is used virtually all the time in the MIXED method. For instance, when $n = 100$, $R^2 = 0.2$, and $\rho^2 = 0.1$ (hence $\rho = 0.316$), the probabilities of a Type I error are 5 percent for IV, 87.5 percent for OLS, and 74.5 percent for the MIXED method.

References

Boulier, B.L. and M.R. Rosenzweig, 1984, Schooling, search, and spouse selection: Testing economic theories of marriage and household behavior, Journal of Political Economy 92, 712–732.

Durbin, J., 1954, Errors in variables, Review of the International Statistical Institute 22, 23–32.

Eichengreen, B., 1984, Mortgage interest rates in the populist era, American Economic Review 74, 995–1015.

Hausman, J.A., 1978, Specification tests in econometrics, Econometrica 46, 1251–1271.

Hausman, J.A. and W.E. Taylor, 1981, Comparing specification tests and classical tests, Bell Laboratories economics discussion paper, presented at the Econometric Society Meeting, San Diego, June.

[18] Wu (1974, p. 546) suggests this particular mixed procedure. Of course, IV does not have any optimal finite sample properties [see, for instance, Zellner (1979)]. Unless the use of OLS is ruled out on *a priori* grounds, however, the problem of choosing between OLS and *some* alternative estimation method still remains. Use of the Wu–Hausman test as a basis for making this choice will result in the application of OLS in exactly the same cases no matter what the alternative to OLS is taken to be.

Holly, A., 1982, A remark on Hausman's specification test, Econometrica 50, 749–759.
Kariya, T. and J. Hodoshima, 1980, Finite sample properties of the tests for independence in structural systems and LRT, Economic Studies Quarterly 21, 45–56.
Nakamura, A. and M. Nakamura, 1981, On the relationships among several specification error tests presented by Durbin, Wu and Hausman, Econometrica 49, 1583–1588.
Nakamura, A. and M. Nakamura, 1982, A clarification concerning the consistency of certain specification error tests under alternative null hypotheses, Working paper (Faculty of Business, University of Alberta, Edmonton).
Nakamura, A. and M. Nakamura, 1984a, Power considerations of certain specification error tests, Working paper (Faculty of Business, University of Alberta, Edmonton).
Nakamura, A. and M. Nakamura, 1984b, Rational expectations and the firm's dividend behavior, Review of Economics and Statistics, forthcoming.
Nakamura, A., M. Nakamura and G.H. Orcutt, 1976, Testing for relationships between time series, Journal of the American Statistical Association 71, 214–222.
Rao, C.R., 1973, Linear statistical inference and its applications, 2nd ed. (Wiley, New York).
Revankar, N.S., 1978, Asymptotic relative efficiency analysis of certain tests of independence in structural systems, International Economic Review 19, 165–179.
Revankar, N.S. and M.J. Hartley, 1973, An independence test and conditional unbiased predictions in the context of simultaneous equation system, International Economic Review 14, 625–631.
Sawa, T., 1969, The exact sampling distribution of ordinary least squares and two-stage least squares estimator, Journal of the American Statistical Association 64, 923–927.
Wu, D., 1973, Alternative tests of independence between stochastic regressors and disturbances, Econometrica 41, 733–750.
Wu, D., 1974, Alternative tests of independence between stochastic regressors and disturbances: Finite sample results, Econometrica 42, 529–546.
Zellner, A., 1979, Statistical analysis of econometric models, Journal of the American Statistical Association 74, 628–643.