Title
Tests of Specification in Econometrics

Author
Ruud, Paul A.

Publication Date
1984

Peer reviewed

TESTS OF SPECIFICATION IN ECONOMETRICS

Paul A. Ruud

May 1984

Research Papers in Economics No. 84-3

TESTS OF SPECIFICATION IN ECONOMETRICS

Paul A. Ruud
University of California, Berkeley, 94720

Key Words & Phrases: *hypothesis tests, specification tests, Chow tests, Hausman tests, maximum likelihood, pre-test estimation, limited dependent variables, exogeneity*

ABSTRACT

This survey of recent developments in testing for misspecification of econometric models reviews procedures based on a method due to Hausman. Particular attention is given to alternative forms of the test, its relationship to classical test procedures, and its role in pre-test estimation.

1. INTRODUCTION

Hausman (1978) proposed a test for model misspecification based on looking for a statisitically significant difference between an estimator that is efficient under the null hypothesis and an estimator that is consistent under the alternative hypothesis. Since Hausman's paper, several theoretical issues have arisen regarding the relationship between Hausman's test and conventional, or classical, tests. Many authors have also applied the method proposed by Hausman to such specific problems as testing for exogeneity and testing distributional assumptions. This survey reviews these developments of the "Hausman specification test."

The presentation is organized into three sections. The first section is introductory, summarizing basic findings in the literature through the example of testing the specification of a normal censored regression (Tobit). The second section is theoretical and methodological, reviewing the theory behind the Hausman specification test, its relationship to classical tests and pre-test estimation, and examining in detail a special class of problems closely related to the Chow test for structural change. The third section re-evaluates three Hausman specification tests that have been proposed in the light of the second section: the test of the multinomial logit model by Hausman and McFadden (1983); a test of exogeneity in simultaneous equations models; and the test of the Tobit model by Nelson (1981).

2. A HAUSMAN SPECIFICATION TEST FOR TOBIT

The censored normal regression model postulates an underlying regression which one does not observe

$$y^* = x'\beta + \varepsilon \tag{2.1}$$

where x is a vector of k explanatory variables, β is a vector of k unknown population constants, and ε is a random error term. Instead, one observes the regressors x and the censored dependent variable

$$y = \mathbf{1}(y^* \geq 0)\, y^* \tag{2.2}$$

where $\mathbf{1}(\bullet)$ is the indicator function. Only sample points of the dependent variable that exceed the threshold of zero are actually recorded. For other points one obtains the values of the explanatory variables only. By convention, the observed dependent variable is set equal to the censoring point zero when the latent y^* is negative.

The OLS estimator for β that uses y in place of y^* is biased. The source of this bias is interpreted often as a specification error: when the errors are i.i.d. $N(0,\sigma^2)$ random variables, the expectation of y conditional on $x'\beta$ is

$$E(y \mid x'\beta) = x'\beta\, \Phi\left(\frac{x'\beta}{\sigma}\right) - \sigma\, \phi\left(\frac{x'\beta}{\sigma}\right) \tag{2.3}$$

where ϕ and Φ are the standard normal probability density and cumulative distribution functions respectively. Therefore, the maximum likelihood (ML) estimator, called tobit, is used in place of the OLS estimator. The log-density for y conditional on x is

$$L_0 = 1(y=0)\ln\phi\left(-\frac{x'\beta}{\sigma}\right) - 1(y>0)\left[\frac{1}{2}\ln(2\pi\sigma^2) + \frac{1}{2}\left(\frac{y-x'\beta}{\sigma}\right)^2\right] \quad (2.4)$$

Summing (2.4) over observations and maximizing this sample log-likelihood function with respect to the unknown parameters β and σ^2 yields the maximum likelihood estimators b_0 and s_0^2. Under regularity conditions, this MLE is consistent and approximately normal for large sample sizes. We denote the variance-covariance matrix of this distribution by $V_0(\beta,\sigma)$ (Amemiya (1973)).

Two closely related estimators are the probit estimator and the truncated regression estimator. The probit estimator is the MLE that uses the sign of y^* to estimate the ratio of the regression slopes to the standard deviation, $\frac{\beta}{\sigma} = \gamma$. The log-density in this case is

$$\begin{aligned} L_1 &= 1(y=0)\ln\Phi\left(-\frac{x'\beta}{\sigma}\right) - 1(y>0)\ln\Phi\left(\frac{x'\beta}{\sigma}\right) \\ &= 1(y=0)\ln\Phi(-x'\gamma) - 1(y>0)\ln\Phi(x'\gamma) \end{aligned} \quad (2.5)$$

and, as above, the corresponding MLE for γ, c_1, has an approximately normal distribution, $N(\gamma, V_1(\gamma))$. On the other hand, the truncated regression estimator is based on the strictly positive, continuous observations of y only. Both β and σ can be estimated by maximizing the log-liklihood obtained from the log-density

$$L_2 = -\ln\Phi\left(\frac{x'\beta}{\sigma}\right) - \frac{1}{2}\ln(2\pi\sigma^2) - \frac{1}{2}\left(\frac{y-x'\beta}{\sigma}\right)^2 \quad (2.6)$$

and, the MLE for (β,σ), (b_2,s_2), is approximately a $N((\beta,\sigma),V_2(\beta,\sigma))$ random variable.

None of these estimators is robust to violations of the assumption that the errors are normally distributed. Although there are no general theoretical results, one suspects that whenever this distribution is actually non-normal, these ML estimators are inconsistent. Goldberger (1984) and Arabmazar and Schmidt (1982) give specific illustrations of such inconsistency. Furthermore, the three estimators are unlikely to converge in probability to the same point in the parameter space because the estimators depend on different functions of the sample data. For example, the probit estimator characteristically sets the sample fraction of censored observations equal to the estimated expected fraction whereas the truncated regression estimator is not a function of this fraction.

One can construct a simple diagnostic test for misspecification of distribution by comparing one estimator with another, testing for statistically significant differences with the Hausman specification test. Three possible comparisons between pairs of estimators are:

$$\begin{aligned} d_0 &= (b_0', s_0)' - (b_2', s_2)' \\ d_1 &= c_0 - c_1 \\ d_2 &= c_2 - c_1 \end{aligned} \tag{2.7}$$

where $c_i = \frac{b_i}{s_i}$ (i=0,2). The tobit, probit, and truncated regression estimators are subscripted by 0, 1, and 2, respectively. Note that contrast between the tobit and truncated regression estimators has $k + 1$ elements but the comparisons with the probit estimator are reduced to the k elements identified by that estimator.

Under the hypothesis of normality, the censored regression estimator is efficient relative to the other two estimators. Thus, the first two contrasts, d_0 and d_1 , are examples of the estimator difference emphasized by Hausman (1978): efficient versus inefficient estimators. The third difference contrasts two estimators that do not have a relative efficiency ranking and are not consistent. Ruud (1982), however, notes that this method of testing requires neither a relatively efficient estimator nor a consistent estimator. Rather a useful test derives from estimators that diverge greatly under alternative models and whose difference has a small sampling variance.

Several useful points about Hausman's test which can be applied to the present testing problem appear in the literature. First, the covariance matrix for the efficient versus inefficient contrast is simple and convenient: this matrix is the difference of the two estimators' covariance matrices. Because these matrices are estimated routinely as part of the estimation of the parameters themselves, all of the required components of the Hausman specification test are available from standard computer software. The actual statistic proposed by Hausman (1978) is

$$m_1 = N d_0'[V_2(b_2, s_2) - V_0(b_0, s_0)]^{-1} d_0 \tag{2.8}$$

Hausman and Taylor (1982) amended this formula by replacing the matrix inverse with its generalized inverse, denoted by a superscript "-", to account for a possibly singular covariance matrix. The statistic m_1 has a central chi-square distribution under the null hypothesis of normality. One carries out the test by comparing the computed statistic with the critical values of the chi-square distribution of the appropriate degrees of freedom. We defer, for the moment, a discussion of the role the generalized inverse plays and the determination of the degrees of freedom.

The second efficient versus inefficient estimator contrast, d_1 , is slightly different. This contrast contains one less parameter difference than does d_0 . However, one may still calculate an statistic analogous to m_1 . Perhaps the simplest way to do this, is to reparameterize L_0 in terms of $\gamma = \frac{\beta}{\sigma}$ and σ . One can then obtain directly the ML estimators and their estimated covariance matrix, and drop the terms related to σ . Then, proceeding as before:

$$m_2 = N d_1'[V_1(c_1) - S_\gamma' V_0(c_0,s_0) S_\gamma]^{-1} d_1 \qquad (2.9)$$

where $S_\gamma' = \frac{\partial(\gamma',\sigma)}{\partial\gamma}$ is a selection matrix containing zeroes and ones that deletes the column of V_0 corresponding to covariances with s_0 .

After Hausman's original proposal, attempts to use such statistics as m_1 and m_2 revealed that the estimated variance-covariance matrix formed from the difference of two estimated covariance matrices, might fail to be positive definite. Definiteness is required to take the matrix inverse and variance-covariance matrices must be positive semi-definite. Under the null hypothesis and as the sample size goes to infinity, the efficient estimator has a smaller covariance matrix than the inefficient estimator. In finite samples, however, nothing forces the estimated covariance matrices to satisfy this matrix inequality. White (1982) proposes an alternative estimator of the covariance of d_0 which is guaranteed to be positive semi-definite. Unfortunately, his estimator increases computational costs. A more straightforward solution is to evaluate both estimates of the estimators' covariances at one of the estimators. For example, m_1 would be modified to be

$$m_1' = N d_0'[V_2(b_2,s_2) - V_0(b_2,s_2)]^{-1} d_0 \qquad (2.10)$$

$V_2 - V_0$ is guaranteed to be positive semi-definite when both matrices are evaluated at the same point in the parameter space and one uses the information matrix to estimate these matrices. For local alternatives to the null hypothesis, this covariance estimator is consistent for the true asymptotic covariance no matter which estimator is used. In addition, Newey (1983) and Ruud (1982) discuss consistent estimators that, in this case, are based only on the first derivatives of the log-likelihood function and offer computational savings. When these estimators are employed, the Hausman specification test calculation can be reduced to the sample size multiplied by an unadjusted R^2 statistic, as happens commonly in LM statistics (Engle (1981)).

Another failure of positive definiteness of the covariance matrix of the estimators' difference occurs when the true asymptotic covariance matrix is singular. In other words, there may be a linear relationship between the estimator contrasts which holds under the null hypothesis for all parameter values. As a result, the number of nonredundant contrasts, and the degrees of freedom of the asymptotically chi-squared test statistic, is actually less than the dimension of d. This singularity prompted Hausman and Taylor (1982) to insert the generalized inverse in place of the usual matrix inverse in the calculation of the m statistic. Generalized inverses are not unique, but Holly and Monfort (1982) show that all generalized inverses give the same test statistic. Given this, a convenient method for calculating the Hausman specification test when there is a singular covariance matrix is to drop the linearly dependent contrasts and contstruct the test statistic using a linearly independent subset of the estimators' differences.

In our example, we have given two test statistics, m_1 and m_2, with an apparent difference in degrees of freedom. In fact, d_0 contains one redundant contrast and both test statistics have the same degrees of freedom: the number of explanatory variables in the regression function, k. Therefore, m_1 and m_1' require generalized inverses of $V_2 - V_0$. It is natural to view the difference in variance estimators, $s_2 - s_0$, as the redundant term, to drop this term from d_0, and to rewrite m_1' analogously to m_2.

$$m_1'' = N d_0' S_\gamma [S_\gamma' (V_2(b_2,s_2) - V_0(b_0,s_0)) S_\gamma]^{-1} S_\gamma' d_0$$

Any other parameter difference also could be dropped without affecting the asymptotic behavior of the test.

No one has provided a simple way to compute the correct degrees of freedom or to decide which subset of contrasts is redundant. One must either resort to the computer and software for calculating generalized inverses or tackle each problem analytically. Hence, the convenience of the test procedure is diminished if the covariance of d does not have full rank. It should be noted, however, that relying on the computer to compute the correct degrees of freedom is correct if the likelihood function is regular and one relies on asymptotic distributions. Singularities will only occur in finite samples by chance.

Further, the statistical relationship among the contrasts in (2.7) remains unexplored. In the second half of this paper, we will show that this relationship and the degrees of freedom calculation are closely related. Two results are that all the contrasts lead to asymptotically equivalent test statistics and that the degrees of freedom calculation is simple in such problems as our example.

Several researchers have noted test statistics which are asymptotically equivalent to the Hausman specification tests that we have constructed above. White (1982), Riess (1983), and Ruud (1982) all observe that a score (Lagrange Multiplier or LM) test statistic can be used. This test has the advantage that only one estimator is calculated. One example, using the censored regression problem, is to test whether the first derivatives of the probit log-likelihood function when evaluated at the censored regression MLE, $\frac{\partial L_1(c_0)}{\partial \gamma}$, are significantly different from zero. If they are, c_0 and c_1 must also be significiantly different because $\frac{\partial L_1(c_1)}{\partial \gamma} = 0$. We defer the mathematical expression for this score test statistic, and other equivalent ones, to the next section. Denoting derivatives with subscripts, this statistic is

$$m_1^* = N L_{1,\gamma}(c_0)' V_1(c_0)^{-1} [V_1(c_0) - S_\gamma' V_0(c_0,s_0) S_\gamma]^{-1} V_1(c_0)^{-1} L_{1,\gamma}(c_0) \quad (2.11)$$

Given the Hausman specification test methodology, researchers have an apparent alternative to the classical testing method of specifying a parametric null hypothesis and testing for deviations from that hypothesis. The relationship between classical and specification tests has been investigated by Holly (1982a). In Holly's framework, a parametric hypothesis is stated for a subset of the unknown parameters in the likelihood function. The classical tests of the parametric hypothesis are compared with the Hausman specification test that contrasts estimators for parameters left unrestricted by the null hypothesis. Although Hausman (1978) did not restrict his anyalysis in this way, Holly provides an insightful view of the new testing method through this structure.

To add concreteness, we continue the censored regression problem by nesting the normal distribution within a larger parametric family of distributions. We suppose that even if it is not normal, the distribution function of ε is a member of the parametric family

$$Prob(\varepsilon \leq z) = G(z) = \Phi\left[\frac{\Theta(z)}{\sigma}\right] \quad (2.12)$$

where $\Theta(z)$ is an M-th order power series,

$$\Theta(z) = \sum_{m=0}^{M} \theta_m z^m \qquad (2.13)$$

such that $\Theta'(z) \geq 0$ for all real z , and normalized so that the median of ε is zero ($\theta_0 = 0$) and $\theta_1 = 1$. In general, our hope is to approximate closely the transformation of ε to normality with Θ. The normalizations are required to identify separately σ and the regression intercept.

With this nesting, we can use the classical test statistics, the likelihood ratio (LR), the Wald (W), and the score (LM) tests, to examine deviations in the unrestricted estimates of the θ's from the values obtained by restricting the error distribution to the normal: $\theta_j = 0$, j=2,...,m. For limited dependent variable models, Lee (1981) also has advanced this methodology using the Pearson family of distributions for nesting. Denoting the unrestricted log-likelihood by $L_3(\gamma,\sigma,\theta)$, the unrestricted MLE by $(c_3{}',s_3,t_3{}')$, its approximate covariance matrix by $V_3(\gamma,\sigma,\theta)$, and using our notation above, the classical Wald test of $(\theta_3,\ldots,\theta_M)' = \theta = 0$ is

$$W = N t_3{}'\left[S_\theta{}' V_3(c_3,s_3,t_3) S_\theta\right]^{-1} t_3 \qquad (2.14)$$

where $S_\theta{}' = \dfrac{\partial(\gamma',\sigma,\theta')}{\partial\theta}$ is a selection matrix for θ .

The test statistics of Hausman examine deviations in the estimates of the slope (and perhaps variance) parameters with and without the restrictions to normality. Denoting the restricted MLE by $(c_0,s_0,0)$, his test statistic takes the form

$$m = N d'[V(d)]^{-1} d \qquad \text{where} \quad d = \left[(c_3 - c_0)', s_3 - s_0\right]' ,$$
$$V(d) = S_{\gamma\sigma}{}' V_3(c_3,s_3,t_3) S_{\gamma\sigma} - V_0(c_0,s_0) \quad , \text{ and } \quad S_{\gamma\sigma}{}' = \frac{\partial(\gamma',\sigma,\theta)}{\partial(\gamma',\sigma)'} \qquad (2.15)$$

Under these circumstances, there is yet another form for the Hausman specification test which parallels the classical approach. This is the Wald form discussed by Hausman and Taylor (1981), and implicit in Holly (1982a):

$$m = (I_{\gamma,\sigma;\theta}{}' t_3)' \left[I_{\gamma,\sigma;\theta}{}' S_\theta{}' V_3(c_3,s_3,t_3) S_\theta I_{\gamma,\sigma;\theta} \right]^{-1} I_{\gamma,\sigma;\theta}{}' t_3$$

$$\text{where} \quad I_{\gamma,\sigma;\theta} = E\left[-\frac{\partial^2 L_3}{\partial\theta\partial(\gamma',\sigma)} \right] \tag{2.16}$$

This statistic measures deviations in $I_{\gamma,\sigma;\theta}{}' t_3$ from zero. It also is an estimate of the asymptotic bias in the slope estimators under the restrictions on the θ's to the normality hypothesis. Thus, one can interpret the Hausman specification test as a classical test of the null hypothesis that the restricted parameter estimates are consistent for the population parameters. Holly (1982a) and Hausman and Taylor (1982) interpret the fundamental difference between the testing methods as a difference in the parametric null hypothesis that is under consideration. The classical test has the null hypothesis H_0: $\theta = 0$ but the Hausman specification test has the null hypothesis H_1: $I_{\gamma,\sigma;\theta}{}'\theta = 0$

Furthermore, the set of values of θ that satisfy H_1 is a subset of the set of values that satisfy H_0, so that the Hausman specification test is testing a subset of the restrictions tested by the classical Wald test. Therefore, as Holly showed, there are alternatives to H_0 for which the Hausman specification test has no power. On the other hand, for local alternatives to H_1, the Hausman specification test the appropriate classical Wald test and is, therefore, the locally uniformly most powerful unbiased test. This result is due to Hausman and Taylor (1982).

Riess (1983) provides another way to express the same result. He observes that the likelihood ratio test is asymptotically equivalent to a contrast between all of the parameter values at the restricted and unrestricted estimators. This result can also be found in Silvey (1959). The classical and Hausman test procedures are seen to differ, then, on the basis of which subset of contrasts one chooses to examine.

Finally, Holly (1982a) observes that there is a case where the classical and specification test methods are equivalent. If the dimension of θ is less than or equal to the number of remaining parameters (γ,σ) and $I_{\gamma,\sigma;\theta}$ has full rank, then H_0 is true if and only if H_1 holds, and the tests are identical asymptotically. The equivalence is exact in finite samples when the estimators are linear, as in linear regression.

3. THE STATISTICAL THEORY OF SPECIFICATION TESTS

Specification tests often are based on comparing a criterion function that measures the goodness of fit of competing models. The likelihood ratio test, for example, compares the values of the maximized likelihood function for the restricted and the unrestricted models. If the values of the log-likelihood differ enough, there is evidence that the restriction fails to hold in reality. Further examples are the chi-squared goodness of fit tests and non-nested hypothesis tests.

In the likelihood framework that compares the restricted and unrestricted models, there are several ways to calculate statistics that measure the change in goodness of fit. These different measures not only yield the classical trinity of tests, the score (LM), the likelihood ratio (LR), and the Wald (W) tests, but also yield various Hausman tests. In addition to the original test statistic proposed by Hausman (1978), there are other asymptotically equivalent tests that take the form of score and Wald tests.

3.1 Specification Tests in the Likelihood Framework

Following Holly (1982a), consider a family of models with log-likelihood function $L(\alpha)$ for N observations where α is a k-dimensional vector of unknown parameters. We will assume that the regularity conditions given in the appendix of Holly (1982a) are satisfied so that standard asymptotic distribution results hold for the maximum likelihood estimator a :

$$\max_\alpha L(\alpha) = L(a) \tag{3.1.1}$$

The parameter vector, α , is partitioned into two subvectors (θ,γ) of dimension p and q , respectively, because one is interested in the hypothesis H_0: $\theta = \theta^0$ and the significance of H_0 in the estimation of γ . In our censored regression example, primary interest is in the slope coefficients and the assumption of normality is largely for convenience. Two estimators for γ are immediately available. These are the restricted and the unrestricted maximum likelihood estimators, c_0 and (t_1,c_1) , defined by

$$L_\gamma(\theta^0,c_0) = 0 \tag{3.1.2}$$

and

$$L_\gamma(t_1,c_1) = 0 \ , \quad L_\theta(t_1,c_1) = 0 \tag{3.1.3}$$

where subscript parameters denote partial differentiation.

Hausman suggested a specification test based on detecting statistically significant differences between the alternative estimators c_0 and c_1 . Hausman and Taylor (1981) demonstrated the asymptotic equivalence of this test statistic and the Wald test statistic for the hypothesis that the restricted estimator c_0 is consistent for γ and thereby established the asymptotic local power properties of the test. They and Holly (1982a) also showed that when $p \leq q$ Hausman's test is equivalent to the classical test statistics for H_0 , the LM, LR, and W statistics. When $p > q$, the Hausman test checks a broader null hypothesis than H_0 , that c_0 is consistent. Reiss (1983), Ruud (1982), and White (1982) all noted score test versions of Hausman's test.

Returning to the general principle of measuring goodness of fit, a logical measure of the difference in log-likelihood function for restricted and unrestricted estimation is a derivative with respect to a parameter value. Three derivatives are used to obtain estimators in equations (3.1.2) and (3.1.3), and various permutations of those estimators as arguments yield score measures of fit:

$$L_\theta(\theta^0,c_0) \quad , \quad L_\theta(\theta^0,c_1) \quad , \quad L_\theta(t_1,c_0) \quad , \qquad L_\gamma(\theta^0,c_1) \quad , \quad L_\gamma(t_1,c_0) \tag{3.1.4}$$

The first score is the statistic used to compute the usual LM test of H_0 . This score measures the potential for improvement in the log-likelihood function due to changes in the estimate of θ if the restriction $\theta = \theta^0$ is dropped. As we will show, the other two expressions in the derivative with respect to θ often yield score tests that are asymptotically equivalent under local alternatives. The remaining score statistics measure the potential for improvement in the log-likelihood function due to changes in the estimated γ if the restrictions of H_0 are dropped. In general, these latter scores do not yield statistics similar to the first group because estimates of γ need not change in response to relaxing H_0 . In fact, score tests based on L_γ are equivalent to the Hausman specification test statistic, as we will now show.

Let (θ^0,γ) be the true parameter values. We will consider the sequence of local alternative hypotheses to H_0

$$H_a\colon \theta = \theta^0_N = \theta^0 + N^{-1/2}\beta \tag{3.1.5}$$

All of the test statistics that we consider are consistent: for a fixed alternative model, every test will reject the alternative model with certainty as the sample size N

approaches infinity. This sequence of local alternatives differentiates the asymptotic behavior of the tests by examining their power where power is low, that is, sufficiently close to the null hypothesis.

Like Holly, we can derive the asymptotic equivalences

$$0 \overset{a}{=} N^{-1/2}L_{\gamma,N} - I_{\gamma\gamma}N^{1/2}(c_0 - \gamma) + I_{\gamma\theta}\beta \tag{3.1.6}$$

$$0 \overset{a}{=} N^{-1/2}L_{\theta,N} - I_{\theta\theta}N^{1/2}(t_1 - \theta_N^0) + I_{\theta\gamma}N^{1/2}(c_1 - \gamma) \tag{3.1.7}$$

$$0 \overset{a}{=} N^{-1/2}L_{\gamma,N} - I_{\gamma\theta}N^{1/2}(t_1 - \theta_N^0) + I_{\theta\theta}N^{1/2}(c_1 - \gamma) \tag{3.1.8}$$

$$E\left[-L_{\alpha\alpha}(\alpha)\right] = I_{\alpha\alpha} = \begin{bmatrix} I_{\theta\theta} & I_{\theta\gamma} \\ I_{\gamma\theta} & I_{\gamma\gamma} \end{bmatrix} \tag{3.1.9}$$

by first order Taylor expansions of (3.1.2)-(3.1.3) around the parameter values (θ_N^0,γ) . The equivalence sign $\overset{a}{=}$ denotes that the difference between the two sides of the equivalence converges in probability to zero. I stands for the information matrix and its subscripts denote submatrices.

Now consider the asymptotic distribution of the score L_γ evaluated at the restricted value for $\theta = \theta^0$ and the consistent estimator c_1 :

$$N^{-1/2}L_\gamma(\theta^0,c_1) \overset{a}{=} N^{-1/2}L_{\gamma,N} - I_{\gamma\gamma}N^{1/2}(c_1 - \gamma) + I_{\gamma\theta}\beta \tag{3.1.10}$$

Combining (3.1.6) and (3.1.10) shows that

$$N^{-1/2}L_\gamma(\theta^0,c_1) \overset{a}{=} - I_{\gamma\gamma}N^{1/2}(c_1 - c_0) \tag{3.1.11}$$

$d = c_1 - c_0$ is the statistic upon which the Hausman specification test is based. Because $I_{\gamma\gamma}$ is nonsingular, Hausman's statistic is asymptotically equivalent to a statistic based on the score $L_\gamma(\theta^0,c_1)$. This score, however does not require the calculation of c_0 .

An alternative, intuitive explanation of the equivalence of the score and Hausman statistics is to consider the one-step, linearized maximum likelihood (LML) estimator

$$c_0^* = c_1 + I_{\gamma\gamma}^{-1} L_\gamma(\theta^0, c_1) \qquad (3.1.12)$$

c_0^* converges in distribution to c_0. Therefore the difference $d^* = c_1 - c_0^*$ is asymptotically equivalent to d. But

$$d^* = - I_{\gamma\gamma}^{-1} L_\gamma(\theta^0, c_1) \qquad (3.1.13)$$

showing, in turn, the equivalence of the Hausman test to the score test. Breusch and Pagan (1981) use this argument for the classical LM and Wald tests.

If we combine (3.1.8) and (3.1.10), we obtain

$$N^{-1/2} L_\gamma(\theta^0, c_1) \overset{a}{=} - I_{\gamma\theta}(t_1 - \theta^0) \qquad (3.1.14)$$

This equivalance demonstrates the relationship between Wald classical tests and the specification tests considered by Holly (1982a) and Hausman and Taylor (1981). While the classical tests evaluate $H_0: \beta = 0$, the Hausman specification tests evaluate $H_1: I_{\gamma\theta}\beta = 0$. If and only if $rank(I_{\gamma\theta}) = p \leq q$, H_0 and H_1 are identical and the classical and Hausman tests are equivalent. Otherwise, $rank(I_{\gamma\theta}) < p$ and H_1 is a broader null hypothesis than H_0. H_1 is, indeed, an identity when $rank(I_{\gamma\theta}) = 0$, that is, when estimates of θ contain no information about estimates of γ because $I_{\gamma\theta} = 0$. Therefore, if $rank(I_{\gamma\theta}) < p$, the Hausman tests are not the same as the classical tests; the degrees of freedom of the Hausman tests are $rank(I_{\gamma\theta})$, which is strictly less than the degrees of freedom of the classical tests. However, the Hausman tests are locally uniformly asymptotically most powerful for H_1.

Equation (3.1.14) can also be interpreted as equating the score test of estimator consistency with the classical Wald test of H_1, as Hausman and Taylor (1981) observe. They argue that this is the hypothesis that the restricted estimator has no asymptotic bias, a specification hypothesis of frequent interest. Under H_1, the asymptotic distribution of c_0 is a normal with mean γ and variance matrix $I_{\gamma\gamma}^{-1}$ by (3.1.6). Not only is the mean correct, but the restricted estimator also remains efficient so that the Hausman-Taylor interpretation can be extended to treat H_1 as the hypothesis of estimator *efficiency*.[1]

---

1 Engle (1981, p. 58) seems to suggest that this is not so. He rightly points out that if estimator consistency is the only concern, then one should be satisfied with the unrestricted estimator. However, Engle also claims that if estimator efficiency is important then acceptance of H_1 should not convince one to use the restricted estimator.

The equivalence of the Hausman test and a classical Wald test can also be found in Lemma 2 of Silvey (1959) where it is shown that the classical LR test of H_0 is asymptotically equivalent to

$$N[(t_1 - \theta^0)',(c_1 - c_0)'] \; I_{\alpha\alpha} \; [(t_1 - \theta^0)',(c_1 - c_0)']' \tag{3.1.15}$$

This quadratic form is simply Hausman's test of specification applied to all of the parameters. Reiss (1983) also notes this form of equivalence.

The score version of a Hausman test has two asymptotically equivalent forms. Besides $L_\gamma(\theta^0,c_1)$, one could look at the score $L_\gamma(t_1,c_0)$, the score for γ evaluated at the restricted estimator for γ and the unrestricted estimator for θ . Using (3.1.8),

$$N^{-1/2}L_\gamma(t_1,c_0) \overset{a}{=} I_{\gamma\gamma}N^{1/2}(c_1 - c_0) \tag{3.1.16}$$

As a result, if c_0 is easier to compute than c_1 , a simpler calculation is available in this alternative form. Following the approach of Durbin (1970), we can replace t_1 in L_γ by t_1^*

$$\max_\theta \; L(\theta,c_0) = L(t_1^*,c_0) \tag{3.1.17}$$

the MLE for θ treating $\gamma = c_0$, or by its LML equivalent yielding[2]

$$N^{-1/2}L_\gamma(t_1^*,c_0) \overset{a}{=} -\,(I^{\gamma\gamma})^{-1}N^{1/2}(c_1 - c_0) \tag{3.1.18}$$

where $I^{\gamma\gamma}$ is the indicated submatrix of $I_{\alpha\alpha}^{-1}$.

Similar arguments establish that the classical LM test can be constructed with two scores other than the usual one: $L_\theta(\theta^0,c_1)$ and $L_\theta(t_1,c_0)$ if $rank(I_{\gamma\theta}) = p$. All three scores are based on the function L_θ . The Hausman specification and classical tests differ essentially, therefore, in which function one uses, L_γ or L_θ .

According to Holly (1982a) and Hausman and Taylor (1981), the choice between the two tests depends on the hypothesis one wishes to test. If one is interested only in the verity of H_0 , then the LM, LR, and W tests are the obvious candidates, being

2 I am indebted to Alberto Holly for pointing out this connection. See also Engle (1981, p. 56-57).

locally uniformly most powerful for alternatives to H_0 . If one is interested only in consistency and efficiency of the restricted estimator, then a form of the Hausman specification test is optimal for alternatives to that hypothesis.

3.2 Classical versus Specification Tests

Holly's comparison of the Hausman test and the classical test began a controversy over the testing methodologies themselves. On one hand, it is advanced that "Hausman's procedure seems to be more general than the classical procedures for it does not seem to require that the null hypothesis be given in parametric form."[3] On the other hand, the Hausman test has been criticized because it does not test the "relevant" null hypothesis, suggesting a weakness in the procedure.[4] In response, Hausman and Taylor interpret the Hausman test as the appropriate classical test for the hypothesis that the restricted estimators are consistent. Almost all commentators conclude that the applied researcher must take care to specify clearly the null hypothesis to be tested in order to choose between the two test methods.

This conclusion is valid but vague. One wonders, for example, how to determine the "relevant" null hypothesis, especially if it does not take parametric form. We will attempt to clarify the issues by arguing (1) that there has been an unconventional, and hence misleading, use of statistical terminology and, (2) that there has been an inappropriate application of the classical hypothesis testing framework. The concept of null hypothesis has been confused with regions of a parameter space in which a test statistic has no statistical power. In turn, the classical method has been reversed to determine from a test statistic what its implicit null hypothesis is, rather than using an alternative hypothesis to derive a powerful test statistic. Despite these seemingly semantic distinctions, some writers have embraced or rejected the Hausman test based upon grounds related to them.

The first step in unraveling the methodological discussion in this literature is to recognize that Holly and those that followed him defined the null hypothesis of a test statistic to be the regions of a parameter space in which the size of a test is equal to the probability of the test statistic falling inside the critical region of the test. These are regions where the test has no power as well as regions that satisfy the maintained hypothesis. Within the classical framework, these regions and the null hypothesis are identical by construction. The classical method requires that one specify parametrically the *alternatives* to

3 See Holly (1983, fn. 2)

4 See Holly (1983, fn. 6) and Engle (1981, p. 58).

a null hypothesis and "it is important to realize that ... choosing such a specification is largely equivalent to, and not necessarily easier than, choosing a test statistic" (Cox and Hinkley (1974, p. 81)). Therefore, to be precise, the tests may be used to test the *same* null hypothesis but generally differ in their power under alternative hypotheses. Furthermore, a classical test is "tied" to a specific alternative but this characteristic is not a self-evident strength or weakness.

The actual null hypothesis may be "a prediction of theory likely to be true or nearly so" or simply "circumstances we wish to assume hold" for analytical or conceptual convenience. As such, a fundamental statistical problem is to decide whether the null hypothesis is true or not by looking for evidence of inconsistency with the null hypothesis. To do this, we do *not* have to specify an alternative hypothesis. "In many situations test criteria may have to be obtained from intuitive considerations" alone (Rao (1973, p. 445)). Such tests have been called "pure significance tests" by Cox and Hinkley because no structure has been introduced to define statistical power. Goldfeld and Quandt (1972) call these tests "nonconstructive" and Ramsey (1974) calls them "general."

Hence, we prefer to view the Hausman test primarily as a pure significance test, like Thursby (1982) and Hall (1983). It is an attractive test, both computationally and intuitively. But the fact that the null hypothesis need not be in parametric form does not imply that the test is "more general;" generally, the Hausman test simply does not specify the *alternative* hypothesis. Clearly, there is only one "relevant" null hypothesis: the model that we wish to maintain holds, as distinct from many alternative models. For example, White (1982) provides an ingenious test statistic which requires specification of the parametric null hypothesis only, and which he interprets as a Hausman specification test. Based on the identity that $E[L_{\gamma\gamma}(\theta^0,\gamma)] = E[-L_\gamma(\theta^0,\gamma)L_\gamma(\theta^0,\gamma)']$ White suggests comparing estimates of each expectation. The test for heteroskedasticity of unknown form in White (1980) is a particular application of this general specification test.[5]

The researcher, then, holds one hypothesis and can use either testing procedure. But if a parametric alternative is specified, the choice of test procedure is the classical one. The delineation of alternative models that actually hold enables us to calculate the power of a test and the classical theory establishes test statistics that have properties of optimal power. One can, of course, use a statistic derived from a parametric alternative hypothesis to perform a pure significance test. However in that case, there is no *a priori*

---

5 Hall (1983) investigates the misspecifications that will be detected by White's information matrix test. Chesher and Lancaster (1983) note an alternative formula for the test based on a simplification of the estimator for the covariance matrix.

basis for preferring a "classical" test to a Hausman test. Alternatively, it is interesting to interpret, as Holly does, the Hausman test as a classical test for the intuition this yields about its use, but Hausman's specification test problem does not motivate the classical structure.

The polar cases of specifying the alternative hypothesis and not are surely too extreme to capture the actual position of empirical researchers. They operate at various degrees between these extremes. If no additional formal approaches to inference are adopted, the researcher must simply choose according to convenience and to intuition about alternative models of concern.

One alternative structure that is particularly relevent to this discussion is the statistical discrimination among nonnested hypotheses (Cox (1961)). In this problem, the specified alternative model "need not be a hypothesis which the investigator would seriously maintain" according Davidson and MacKinnon (1981). Thursby (1982) argues that Hausman's specification test method is the common basis of a number of tests, including nonnested hypothesis tests. Hausman and Pesaran (1982) show the asymptotic equivalence of the J-test of Davidson and MacKinnon (1981) and a Hausman test for the linear model.

Another issue which has clouded the comparison of classical and Hausman tests is an implicit interest in estimation that goes beyond testing hypotheses. The discussion above is limited to the problem of testing a parametric model. If one uses the Hausman test as a criterion for choosing between two estimators, then one faces a different problem, one to which we now turn.

3.3 Pre-Test Estimation

Hausman and Taylor (1982) argue that the Hausman specification test is the appropriate test statistic if one is interested in the effect of a set of restrictions on estimation, as opposed to the restrictions themselves. This suggests that the specification test be used as a preliminary step in estimation that determines which estimator, the restricted or the unrestricted estimator, one chooses. It is natural to ask, then, whether the specification test pre-test estimator is better than the classical pre-test estimator.

One's intuition might suggest that the specification pre-test estimator has a smaller mean-squared error than the classical pre-test estimator. Holly (1982a) shows that the Hausman specification test has no power against local alternatives that leave the restricted estimator consistent. Indeed, these local alternatives also preserve the relative efficiency of the restricted estimator as (3.1.6) reveals. Therefore, one expects the classical pre-test to reject the restricted estimator too frequently and to be less efficient. Similarily, the specification pre-test estimator may trade bias for a smaller mean-squared error than the classical pre-test estimator under local alternatives that leave the restricted estimator biased. For the case of the linear model, Gourieroux and Trognon (1981) showed that neither estimator is uniformly better than the other. They use a four-variable model to show numerically where one estimator performs better than the other.

In this section, the conjectures are shown to be incorrect for the general, regular likelihood problem. Not only does the specification pre-test estimator fail to dominate the classical pre-test estimator based on mean squared error, but the former may be dominated by the latter even when the restricted estimator is consistent and relatively efficient despite a failure of $H_0: \beta = 0$.

In our notation the classical estimator is

$$(t_2,c_2) = \begin{cases} (t_0,c_0) & \text{if} \quad LR \leq \chi^2_{p;\alpha} \\ (t_1,c_1) & \text{if} \quad LR > \chi^2_{p;\alpha} \end{cases} \tag{3.3.1}$$

where LR is the likelihood ratio test statistic of $H_0: \theta = \theta^0$ (or an asymptotically equivalent test statistic) and $\chi^2_{p;\alpha}$ is the α critical value of a central chi-squared distribution with p degrees of freedom. Sen (1979) shows that the asymptotic bias and mean squared error of $N^{1/2}(c_2 - \gamma)$ are

$$\lim_{N \to \infty} E\left[N^{1/2}(c_2 - \gamma)\right] = -\Pi_{p+2}(\chi^2_{p;\alpha},\Delta_2)\lambda \tag{3.3.2}$$

$$\begin{aligned} \lim_{n \to \infty} E\left[N(c_2 - \gamma)(c_2 - \gamma)'\right] = I^{\gamma\gamma} &+ \Pi_{p+2}(\chi^2_{p;\alpha},\Delta_2)\left(I_{\gamma\gamma}^{-1} - I^{\gamma\gamma}\right) \\ &+ \lambda\lambda'\left(2\Pi_{p+2}(\chi^2_{p;\alpha},\Delta_2) - \Pi_{p+4}(\chi^2_{p;\alpha},\Delta_2)\right) \end{aligned} \tag{3.3.3}$$

where

$$\lambda = I_{\gamma\gamma}^{-1} I_{\gamma\theta}\beta \quad \text{and} \quad \Delta_2 = \beta'\left(I^{\theta\theta}\right)^{-1}\beta \tag{3.3.4}$$

and $\Pi_r(\chi,\Delta)$ is the distribution function of the non-central chi-squared with r degrees of freedom and non-centrality parameter Δ . The same argument for the case of the specification pre-test estimator

$$(t_3, c_3) = \begin{cases} (t_0, c_0) & \text{if} \quad m_N \leq \chi^2_{q;\alpha} \\ (t_1, c_1) & \text{if} \quad m_N > \chi^2_{q;\alpha} \end{cases} \tag{3.3.5}$$

where

$$m_N = N\,(c_1 - c_0)'\,(I^{\gamma\gamma} - I^{-1}_{\gamma\gamma})^{-1}\,(c_1 - c_0) \tag{3.3.6}$$

yield

$$\lim_{N \to \infty} E[N^{1/2}(c_3 - \gamma)] = -\,\Pi_{q+2}(\chi^2_{q;\alpha}, \Delta_3)\lambda \tag{3.3.7}$$

$$\begin{aligned} \lim_{n \to \infty} E[N(c_3 - \gamma)(c_3 - \gamma)'] &= I^{\gamma\gamma} + \Pi_{q+2}(\chi^2_{q;\alpha}, \Delta_3)(I^{-1}_{\gamma\gamma} - I^{\gamma\gamma}) \\ &+ \lambda\lambda'(2\Pi_{q+2}(\chi^2_{q;\alpha}, \Delta_3) - \Pi_{q+4}(\chi^2_{q;\alpha}, \Delta_3)) \end{aligned} \tag{3.3.8}$$

where

$$\Delta_3 = \lambda'(I^{\gamma\gamma} - I^{-1}_{\gamma\gamma})^{-1}\lambda \tag{3.3.9}$$

These expressions are analogous to Sen's. Sen also points out that these pre-test estimators are not asymptotically normal; their distribution functions converge to a mixture of multivariate normal distribution functions. Holly observes, in connection with power comparisons for LR and m_N, that $\Delta_3 \leq \Delta_2$.

These results do not yield a general basis for preferring one pre-test estimator to another. The estimators share several properties: as α increases, so do the biases and as the Δ's increase, the biases fall. No ranking by biases or mean squared errors emerges. However, under the hypothesis that $\beta = 0$, the biases disappear and the mean squared errors simplify to weighted averages of the covariance matrices of the competing estimators, c_0 and c_1:

$$\lim_{N\to\infty} V[N^{1/2}(c_2-\gamma)] = \lim_{N\to\infty} V[N^{1/2}(c_0-\gamma)]\,\Pi_{p+2}(\chi^2_{p;\alpha},0) + \lim_{N\to\infty} V[N^{1/2}(c_1-\gamma)]\,[1-\Pi_{p+2}(\chi^2_{p;\alpha},0)] \tag{3.3.10}$$

$$\lim_{N\to\infty} V[N^{1/2}(c_3-\gamma)] = \lim_{N\to\infty} V[N^{1/2}(c_0-\gamma)]\,\Pi_{q+2}(\chi^2_{p;\alpha},0) + \lim_{N\to\infty} V[N^{1/2}(c_1-\gamma)]\,[1-\Pi_{q+2}(\chi^2_{q;\alpha},0)] \tag{3.3.11}$$

Using an argument in Das Gupta and Perlman (1974), one can show

$$\Pi_{q+s}(\chi^2_{q;\alpha},0) < \Pi_{p+s}(\chi^2_{p;\alpha},0) < 1-\alpha \;\;,\;\; p > q > 0 \;,\; s > 0 \tag{3.3.12}$$

Therefore, under the null hypothesis H_0 , the classical pre-test estimator dominates the specification pre-test estimator with respect to covariances because greater weight is given to the covariance matrix of the relatively efficient estimator, c_0 . (Recall that the estimators are not asymptotically normal, however, so that relative efficiency is not the only yardstick for comparison of the pre-test estimators.)

Indeed, the relative efficiency ranking of c_2 to c_3 extends to directions in the parameter space in which H_1 holds but H_0 does not. By continuity, there exists a non-zero Δ_2 , and corresponding non-zero β such that

$$\lambda = I_{\gamma\gamma}^{-1} I_{\gamma\theta}\beta = 0 \;\;,\;\; \beta'\left(I^{\theta\theta}\right)^{-1}\beta = \Delta_2 \tag{3.3.13}$$

so that the asymptotic variance of c_2 is still smaller than that of c_3 , contrary to our intuition. However, Δ_2 can become so large that the covariance inequality is reversed, in agreement with intuition. Nevertheless, we have established that the specification pre-test estimator may be less efficient than the classical pre-test estimator even in directions that the specification test is the preferred test.

Differences in degrees of freedom explain this paradoxical result. The specification test is more powerful than the classical test when the non-centrality parameters are equal ($\Delta_2 = \Delta_3$) and the specification test has fewer degrees of freedom. In this case, the specification pre-test estimator weights the inefficient unrestricted estimator more heavily than the classical pre-test estimator does. As the non-centrality parameters approach zero but remain equal, this relative preference for the inefficient estimator remains so that in the limit, under the null hypothesis, the specification pre-test estimator itself is less efficient than the classical pre-test estimator.

Degrees of freedom can have a substantial influence on efficiency, too. At the ten percent significance level, the weight $\Pi_{r+2}(\chi^2_{r;0.1},0)$ in equations (3.3.10) and (3.3.11) varies from approximately 0.67 to 0.90 for r equals two to infinity. This range narrows, of course, as the significance level is increased; at five percent the range of the weight is (.80,.95) ($r = 2, \ldots, \infty$). Table 1 shows that the efficiency loss of the specification pre-test estimator is greatest when the tests differ greatly in degrees of freedom and q is small.

TABLE 1 $\Pi_{df+2}(\chi^2_{df;\alpha},0)$

df	*Critical Value* α				
	0.10	0.08	0.06	0.04	0.02
2	0.6697	0.7179	0.7712	0.8312	0.9018
3	0.7173	0.7606	0.8048	0.8600	0.9200
4	0.7452	0.7855	0.8289	0.8764	0.9312
5	0.7638	0.8020	0.8428	0.8871	0.9368
6	0.7773	0.8139	0.8527	0.8948	0.9415
7	0.7876	0.8229	0.8603	0.9006	0.9450
8	0.7958	0.8301	0.8663	0.9052	0.9478
9	0.8026	0.8360	0.8712	0.9089	0.9501
10	0.8082	0.8409	0.8753	0.9120	0.9519
11	0.8130	0.8451	0.8788	0.9146	0.9535
12	0.8171	0.8487	0.8818	0.9169	0.9549
13	0.8207	0.8518	0.8679	0.9188	0.9560
14	0.8239	0.8546	0.8867	0.9206	0.9571
15	0.8268	0.8571	0.8888	0.9221	0.9580
16	0.8297	0.8594	0.8906	0.9235	0.9588
17	0.8317	0.8614	0.8923	0.9248	0.9596
18	0.8338	0.8632	0.8938	0.9259	0.9603

Our relative efficiency finding for pre-test estimation of γ does not appear to carry over to the θ parameter vector. There is no analogous ordering of the asymptotic covariances of t_2 and t_3 . This failure is due to the difference in the conditional variance of t_1 given the random variables in the test statistics LR and m . t_1 has a degenerate distribution conditioned on itself as t_1 appears in the Wald version of the classical test. Conditional on $c_1 - c_0$ in the specification test, however, t_1 has a non-zero covariance matrix that leads to additional terms in the asymptotic variance of t_2 . Under the null hypothesis, the asymptotic variance of t_2 is

$$\lim_{N\to\infty} V[N^{1/2}(t_2 - \theta)] = I^{\theta\theta}[1 - \Pi_{p+2}(\chi^2_{p;\alpha},0)] \tag{3.3.14}$$

as one would expect from equation (3.3.10) but

$$\begin{aligned}\lim_{N\to\infty} V[N^{1/2}(t_3 - \theta)] &= I^{\theta\theta}[1 - \Pi_{q+2}(\chi^2_{p;\alpha},0)] \\ +[I^{\theta\theta} - I^{\theta\gamma}(I^{\gamma\theta}I^{\theta\theta\,-1}I^{\theta\gamma})^{-1}I^{\gamma\theta}]&[\Pi_{q+2}(\chi^2_{q;\alpha},0) - (1-\alpha)]\end{aligned} \tag{3.3.15}$$

in which the second term is a negative semi-definite matrix by (3.3.12). This term attenuates the asymptotic covariance matrix of t_3 which would otherwise be larger than the covariance matrix of t_2 . This extra term disappears, naturally, when $I^{\theta\gamma}$ is nonsingular and the pre-test estimators are asymptotically identical.

The theoretical results of this section confirm many of the computational results for the special linear model in Gourieroux and Trognon (1984). When H_0 is approximately satisfied, the classical pre-test estimator dominates the specification pre-test estimator. When H_1 is satisfied and H_0 is very wrong (Δ_2 large), the reverse ranking occurs. However, we deemphasize the intuition that c_3 is preferable to c_2 when H_1 holds. When the two tests' degrees of freedom are markedly different, c_2 has a smaller mean squared error.

3.4 A Factorization Theorem

In this last section on the theory of specification tests, we present a theorem about likelihoods that have a special property: they can be factored into two likelihoods.[1] Each likelihood yields an estimator for parameters of interest when maximized individu-

1 The material in this section and the next is drawn largely from Ruud (1982). Recently, Vuong (1983) has introduced related ideas based on conditional likelihood functions.

ally. A trivial example is splitting a sample of iid observations into two sub-samples and estimating within each sub-sample separately. Such sample stratification has been used to form specification tests in many econometric models, the textbook Chow test being an obvious example. This is one example of a more general method of forming Hausman tests from likelihood factors. Further examples are the applications of the Hausman method that have been advanced in the literature and several are reviewed in the last section of this paper.

Several results emerge from viewing these tests as "Chow" tests. First, in cases where one appears to have different tests because the tests compare different estimators, tests are actually identical. As a result, the comparison of two inefficient estimators can lead to a test statistic that is equivalent to the comparison of an efficient estimator with an inefficient one. Second, a likelihood ratio version of the Hausman test can be given. Such statistics may have a computational advantage over the the form of Hausman because they do not require the estimation of a variance matrix or the computation of a quadratic form. Third, the number of degrees of freedom of the test statistics is made obvious. One does not have to resort to a computer or an analysis of the rank of a variance matrix.

Consider the simple panel data discussed by Maddala (1971) for which Hausman (1978) proposed a specification test. The disturbance term in a linear regression model consists of two components, an iid error and an individual effect that is carried through time:

$$y_{nt} = x_{nt}\beta + \mu_n + \varepsilon_{nt} \quad (n = 1,...,N; t = 1,...,T) \tag{3.4.1}$$

Assume that the μ_n and ε_{nt} are iid normal random variables with means zero and variances σ_μ^2 and σ_ϵ^2 respectively. Three estimators are commonly applied to this problem. The efficient estimator is the feasible Aitken estimator called generalized least squares (GLS). One inefficient estimator is the within-groups or fixed-effects estimator which treats the μ_n as nuisance parameters. Another inefficient estimator is the between-groups estimator which runs the regression using group means, where groups are cross-section units indexed by n .

Associated with the three estimators are three Hausman specification tests: (1) GLS versus within-groups; (2) GLS versus between-groups; and (3) within-groups versus between-groups. Hausman (1978) conjectured that the first test is more powerful than the third because the covariance matrix for the first contrast is smaller than that for the third contrast. However, as Hausman and Taylor (1981b) showed, all three specification tests are asymptotically equivalent. Their proof was based on the result of Maddala (1971) that the GLS estimator is a matrix weighted average of the two inefficient estimators.

The panel data model is an example of a result which actually occurs frequently in econometric estimators based on the likelihood principle and can be easily recognized and understood given the following proposition.

Factorization Theorem: Suppose a random sample u has a regular log-likelihood function $L_0(\gamma,u)$ such that

$$L_0(\gamma;u) = L_1(\gamma;v(u)) + L_2(\gamma;w(u)) \tag{3.4.2}$$

where L_1 and L_2 are also regular log-likelihood functions for $v(u)$ and $w(u)$ respectively. Also suppose that the parameter vector γ is identified in each log-likelihood function so that there are three well-defined maximum likelihood estimators c_0, c_1, and c_2

$$L_i(c_i) = \max_{\gamma \in \Gamma} L_i(\gamma) \quad , \quad i = 1,2,3 \tag{3.4.3}$$

where Γ is a compact subspace of $\mathbf{R}^q$. Then asymptotically, (1) c_1 and c_2 are distributed independently, (2) there is a weighting matrix A such that

$$N^{1/2}(c_0 - \gamma) \overset{a}{=} N^{1/2}[Ac_1 + (I - A)c_2 - \gamma] \tag{3.4.4}$$

and (3) all tests based on pairwise estimator comparisons are equivalent among these three estimators.

Proof: (1) By successive differentiation of (3.4.2)

$$L_{0,\gamma}(c) = L_{1,\gamma}(c) + L_{2,\gamma}(c) \tag{3.4.5}$$

$$L_{0,\gamma\gamma}(c) = L_{1,\gamma\gamma}(c) + L_{2,\gamma\gamma}(c) \tag{3.4.6}$$

Taking expectations, equation (3.4.6) implies that as N approaches infinity

$$V[N^{-1/2}L_{0,\gamma}(\gamma)] = V[N^{-1/2}L_{1,\gamma}(\gamma)] + V[N^{-1/2}L_{2,\gamma}(\gamma)]$$
$$\text{or} \quad I_0 = I_1 + I_2 \tag{3.4.7}$$

where I_i (i = 1,2,3) is the information matrix of log-likelihood i, whereas (3.4.5) implies

$$V[N^{-1/2}L_{0,\gamma}(\gamma)] = V[N^{-1/2}L_{1,\gamma}(\gamma)] + V[N^{-1/2}L_{2,\gamma}(\gamma)]$$
$$+ Cov[N^{-1/2}L_{1,\gamma}(\gamma),N^{-1/2}L_{2,\gamma}(\gamma)] + Cov[N^{-1/2}L_{2,\gamma}(\gamma),N^{-1/2}L_{1,\gamma}(\gamma)] \quad (3.4.8)$$

so that the covariance of $L_{1,\gamma}$ and $L_{2,\gamma}$ is zero asymptotically. c_1 and c_2 are therefore asymptotically independent, given their asymptotic normality.

(2) By the usual Taylor series expansions

$$N^{-1/2}L_{i,\gamma}(\gamma) \overset{a}{=} I_i^{-1}N^{1/2}(c_i - \gamma) \quad , \quad (i = 1,2,3) \quad (3.4.9)$$

Combining (3.4.9) and (3.4.7) yields (3.4.4) where

$$A = I_0^{-1}I_1 \quad (3.4.10)$$

so that c_0 is the matrix weighted average of the two orthogonal estimators, asymptotically.

(3) Evaluating (3.4.5) at any of the three estimators implies that the two scores corresponding to the other estimators are equal, except for a possible difference in sign. Therefore, the score versions of the Hausman test that use those two statistics are identical in finite sample. In other words, for any two pairwise comparisons, one can find identical score test statistics. Asymptotically, all of the tests are therefore equivalent.

This theorem postulates a factorization of the likelihood which may be interpreted as a separation of the information about γ into two orthogonal parts. Conclusions (1) and (2) are consistent with this interpretation: the estimators derived from the orthogonal information are themselves orthogonal and the fully efficient estimator which uses all of the information is a combination of the orthogonal inefficient estimators. These facts about the within-groups, the between-groups, and the GLS estimators in a simple panel data model were proved by Maddala (1971). Result (2) also has been used in the panel case to construct an asymptotically efficient two-step estimator from the within-groups and the between-groups estimators. Even in cases where a useful two-step estimator does not arise, the third conclusion informs us that several alternative specification tests are asymptotically equivalent. A Hausman specification test does not require an efficient estimator. Because the information is split into only two orthogonal pieces, any pair of estimators contains all of the information about γ . Note, however, that more than two factors can arise easily. In those cases one contrast of estimators will not be equivalent to another. Note also, that the comparison of the two inefficient estimators has a computational advantage over the other two tests: an estimate of the covariance matrix of their difference is the sum of their estimated covariances which is always positive semi-definite. One need not re-compute the covariance matrix estimate of one estimator at the values for the other estimator to assure positive semi-definiteness.

Returning to the panel data model, we demonstrate the results of Maddala (1971) and Hausman and Taylor (1982) straightforwardly. Let

$$
\begin{aligned}
\eta_{nt} &= \mu_n + \varepsilon_{nt}, \\
\eta_{n\bullet} &= T^{-1} \sum_{t=1}^{T} \eta_{nt}, \\
\overset{*}{\eta}_{nt} &= \eta_{nt} - \eta_{n\bullet},
\end{aligned}
\tag{3.4.11}
$$

and note that $\eta_{n\bullet}$ is independently distributed from $\eta_{nt}{}^*$ and both are normally distributed. This independence is simple to verify directly. To understand why independence occurs, consider the variance of the $\eta_n = [\eta_{nt}]$ (t = 1,...,T) for a single group,

$$V(\eta_n) = \sigma_\varepsilon^2 I_T + \sigma_\mu^2 J_T \tag{3.4.12}$$

where I_T is the T by T identity matrix and J_T is a T by T matrix of ones. The orthogonal operators

$$P_T = T^{-1} J_T \quad \text{and} \quad Q_T = I_T - P_T \tag{3.4.13}$$

project η_n into orthogonal random variables because $P_T V(\eta_n) Q_T = 0$. Different groups are, of course, independent. Because of this independence, the likelihood for the sample can be written as

$$
\begin{aligned}
L_0(\beta,\sigma_\mu^2,\sigma_\varepsilon^2;\eta_{nt},n = 1,\ldots,N,t = 1,\ldots,T) = {}& L_1(\beta,\sigma_\mu^2,\sigma_\varepsilon^2;\eta_{n\bullet},n = 1,\ldots,N) \\
&+ L_2(\beta,\sigma_\mu^2,\sigma_\varepsilon^2;\overset{*}{\eta}_{nt},n = 1,\ldots,N,t = 1,\ldots,T)
\end{aligned}
\tag{3.4.14}
$$

The reader will recognize that the MLE of L_1 is the between-groups estimator and the MLE of L_2 is the within-groups estimator. Without any of the matrix algebra Maddala uses, we can conclude that the GLS estimator derived from maximizing L_0 , is asymptotically a matrix weighted average of the orthogonal between and within estimators. When σ_μ^2 and σ_ε^2 are known, these are linear estimators and so the result applies to finite samples. Finally, a specification test contrasting the GLS estimator with the within estimator or the between estimator is asymptotically equivalent to a consistency test comparing the within and the between estimators.

The likelihood function is not the only criterion function for which a decomposition leads to a specification test like Hausman's. Without giving the details, we note that the minimum distance criterion that is popular in simultaneous equations and nonlinear regression estimation can have a similar structure. Consider the minimum distance estimator

$$\varepsilon(c_0)' M_0 \varepsilon(c_0) = \min_{\gamma} \ \varepsilon(\gamma)' M_0 \varepsilon(\gamma) \tag{3.4.15}$$

that exploits the orthogonality condition

$$E[\varepsilon(\gamma)] = 0 \tag{3.4.16}$$

Hansen (1982) discusses the general conditions under which

$$\begin{gathered} N^{1/2}(c_0 - \gamma) = N[0,(\delta' M_0 \delta)^{-1} \delta' M_0 V(\varepsilon) M_0{}' \delta (\delta' M_0 \delta)^{-1}] \\ \textit{where} \ \ \delta = E\left[\frac{\partial \varepsilon(\gamma)'}{\partial \gamma}\right] \end{gathered} \tag{3.4.17}$$

and shows that the estimator is optimal among estimators based on the orthogonality condition if

$$(\delta' M_0 \delta)^{-1} \delta' M_0 V(\varepsilon) M_0{}' \delta (\delta' M_0 \delta)^{-1} = (\delta' M_0 \delta)^{-1} \tag{3.4.18}$$

Commonly, $M_0 = V(\varepsilon)^{-1}$. In addition, if the metric M_0 can be broken into M_1 and M_2 such that

$$M_0 = M_1 + M_2 \ \ , \ \ M_1 M_2 = 0 \tag{3.4.19}$$

(M_0 , M_1 , M_2 symmetric), then the minimum distance criterion function can be factored into two criterion functions that are analogous to the log-likelihood factors discussed above. For example, we might drop the assumptions of normal errors in the panel data model and employ the minimum distance estimator

$$b_0 : \ \min_{b} \ \ \eta(b)'[I_N \otimes (\sigma_\varepsilon I_T + \sigma_\mu J_T)]\eta(b) \tag{3.4.20}$$

which is simply a feasible GLS estimator based on initial, consistent estimates of σ_μ^2 and σ_ϵ^2 . $P_T V(\eta_n) Q_T = 0$ implies that

$$M_1 = \sigma_\varepsilon^2 (I_N \otimes Q_T)$$
$$M_2 = (T\sigma_\mu^2 + \sigma_\varepsilon^2)(I_N \otimes P_T) \tag{3.4.21}$$

satisfy the requirements. This decomposition of the minimum distance criterion function is, of course, behind the likelihood factorization above, and the various minimum distance estimators are closely related to their ML counterparts. Hausman specification tests can be constructed exactly as before outside of the likelihood framework of the normality assumption, and the estimators and test statistics retain the properties given by the factorization theorem.

To conclude this discussion of decomposable criterion functions, we note that the function itself is the basis of yet another test statistic that is asymptotically equivalent to the Hausman statistic. Within the special structure of the decomposable likelihood function, we can form a likelihood ratio statistic

$$LR = 2[L_1(c_1) + L_2(c_2) - L_0(c_0)]$$

that consists of the difference between the "unrestricted" log-likelihood function and the "restricted" log-likelihood function. This equivalence follows from the result (3.1.15) above. Although it requires the calculation of a third estimator, this test statistic is much simpler to calculate than the quadratic forms above, given the estimates of γ . This statistic is also non-negative, as a chi-square random variable should be, avoiding the inconvenience of failures of the estimated covariance matrix to be positive definite.

This likelihood ratio version of the Hausman specification test also makes the computation of degrees of freedom for the test statistics clear for such problems. As usual, one merely counts the number of restrictions that are dropped in the unconstrained criterion function. Several examples may illustrate the convenience afforded by this observation. In the panel data model, one might wonder about contrasting variance estimates as well as the regression slopes. Adding such contrasts will lead to a singularity because the unrestricted likelihood function estimates only an extra set of regression slopes. The number of variance parameters is still two, one for each factor of the likelihood. Furthermore, the degrees of freedom will be diminished by the number of time invariant regressors. This is because the corresponding slopes are not individually estimable by the within-groups estimator. In the censored regression example, the degrees of freedom of the tests equals the number of regressors because the Probit estimator does not include an estimate of the variance parameter. Here again, adding the variance contrast in a comparison of censored and truncated regression estimates would only introduce a singularity.

Finally, tying the specification test of Hausman to the likelihood ratio also clarifies the role of nuisance parameters that are not included in the estimator contrast. One can derive such tests simply by concentrating the nuisance parameters out of each of the likelihood functions first, and then proceeding as usual (Engle (1981)). All of the results that we have given under the restrictions that each factor of the likelihood identify the full parameter vector and that the full vector be compared for two estimators continue to hold when these restrictions are dropped. It is also clear by such a construction that the tests will be asymptotically similar under general conditions: the tests will not be influenced by the true value of the nuisance parameters.

4. APPLICATIONS OF THE HAUSMAN SPECIFICATION TEST

In this section, we review three applications of the Hausman specification test method. All of these applications fall into the structure of the factorization theorem, suggesting that, although it covers a special case, this theorem is of general interest.

4.1 The Multinomial Logit Model

Hausman and McFadden (1982) have suggested a diagnostic test of the multinomial logit (MNL) model based on the Hausman method. As a discrete choice model, the MNL specification is computationally convenient but suffers from the potential restrictiveness of the independence from irrelevent alternatives (IIA) property. According to this property, the ratio of the probabilities of choosing two alternatives is independent of the other alternatives in the choice set.

As a result of the IIA property, one can drop alternatives from the choice set and re-estimate the model consistently when the specification is correct. Hausman and McFadden use this result to form a test comparing the efficient maximum likelihood estimates to such inefficient estimates. It is not necessary that misspecification will lead to asymptotic divergence of the two estimators. If they are significantly different, however, one has evidence of a violation of the MNL specification.

Using the theorem on likelihood factorization, one can see the classical form of the test the Hausman-McFadden test. We also derive a score version of this test that requires the iterative calculation of only one estimator. Further, this test illustrates the similarity with Chow tests.

Using the notation in McFadden (1973), consider the conditional logit problem where one observes N random trials, without repetition, of choices from the finite alternative set $B = \{1,...,J\}$ by individuals with attributes x_{1n} and alternatives with attributes x_{2jn}, $n = 1,...,N, j = 1,...,J$. The selection probabilities satisfy

$$P_{in}(B) = P(i \mid z_n,B,\beta) = \frac{e^{z_{in}\beta}}{\sum_{j=1}^{J} e^{z_{jn}\beta}} \quad where \; z_{jn} = [x_{1n} \;\; x_{2jn}] \tag{4.1.1}$$

$P_{in}(B)$ is the probability that an arbitrary decision-maker faced with choice set B will choose alternative i , z_{jn} is a k-dimensional row vector, and β is a commensurate column vector of taste parameters. For convenience, we do not include alternative specific parameters. These must be treated as nuisance parameters, in the fashion explained previously.

The log-likelihood of the sample is

$$L_0(\beta) = \sum_{n=1}^{N} \left(\sum_{j=1}^{J} S_{jn} z_{jn} \right) \beta - \ln \left(\sum_{j=1}^{J} e^{z_{jn}\beta} \right) \tag{4.1.2}$$

$$where \;\; S_{jn} = \begin{Bmatrix} 1 & \text{if the } n^{th} \text{ individual chooses } j \\ 0 & \text{otherwise} \end{Bmatrix}$$

McFadden (1973) discusses the asymptotic properties of the maximum likelihood estimator, b_0 , which is the solution to the normal equations

$$\begin{aligned} L_{0,\beta}(b_0) &= \sum_{n,j} \left[S_{jn} z_{jn} - z_{\bullet n}(b_0) \right]' = 0 \\ where \;\; z_{\bullet n}(b_0) &= \sum_{j} P_{jn}(B) z_{jn} \end{aligned} \tag{4.1.3}$$

The information matrix is

$$I = \sum_{n,j} P_{jn}(B)[z_{jn} - z_{\bullet n}(\beta)]\,[z_{jn} - z_{\bullet n}(\beta)]' \tag{4.1.4}$$

Now suppose, for example, alternative J is dropped from the choice set, individuals who chose J are dropped from the sample, and the MNL model is re-estimated to yield the comparison estimator, b_1, corresponding to the choice set $C = \{1,...,J-1\}$. The normal equation for b_1 is the same as (4.1.3) with the summation of j running from 1 to $J-1$. The score for b_0 evaluated at b_1 is, using (4.1.1) and (4.1.3)

$$L_{0,\beta}(b_1) = \sum_{n} [S_{Jn} - P_{Jn}(b_1,B)]\,[z_{Jn} - z_{\bullet n}(b_1)]' \tag{4.1.5}$$

After calculating one estimator, b_1, one computes an estimate of the covariance matrix of this score and forms the usual quadratic form to carry out the score test.

Equation (4.1.5) is the difference between the sample mean of $z_{Jn} - z_{\bullet n}$ and a prediction of this mean based on b_1 which did not use alternative J information. Therefore one can view this procedure as an out-of-sample test of the model's predictive power. Such out-of-sample comparisons of predictions versus realizations have been used informally as a model check in a variety of econometric problems. Intuitively, then, this specification test of the MNL model is a sensible test for any choice model. We therefore expect the test to perform well against a broad class of alternative specifications.

Further insight about the nature of the test can be gained by viewing this test of specification as an example of a generalized Chow test. One can construct a LR version of the Hausman-McFadden test, the appropriate treatment of nuisance parameters is clear, and the classical flavor of this test is apparent.

First, we must determine how to factor the likelihood function. Implicitly, Hausman and McFadden have supplied a factor of the complete likelihood in constructing their test: that is, the likelihood of a particular choice from the subset of alternatives C given that the chosen alternative is in C. The second factor is immediately clear; it is the likelihood of the observed choice between two alternatives: alternative J and choosing an alternative from C. Thus, one can form a likelihood ratio statistic that is asymptotically equivalent to the Hausman-McFadden statistic (although this cannot be done with standard MNL computer software).[2]

Hausman and McFadden allow the presence of attributes in z that do not vary within the restricted choice set C. The coefficients of such variables are not identified in the likelihood for C and must therefore be treated as nuisance parameters. In this case, then, the available degrees of freedom for the Hausman-McFadden test is the number of estimable parameters in the restricted choice set likelihood and one can concentrate these nuisance parameters out of the problem without any loss of generality.

The implicit factorization that Hausman and McFadden employ breaks the original decision problem into a sequential two-step decision, first choosing between J and C and second, if C is chosen, choosing from C. The classical interpretation of this test is, then, as a particular parametric test of a single-step decision against the alternative of a two-step decision in which the parameters used to choose between J and C differ from those used to choose among the alternatives in C. This suggests that the subset of alternatives chosen to construct a specification test in this fashion should reflect alternative sequential decisions thought to be likely. There seems to be no other *a priori* guideline on how to choose the set C and Hausman and McFadden suggest that several be tried in actual application.

4.1.2 A Limited Information Test of Exogeneity

The motivation given by Hausman (1978) for his test of specification was to attack the problem of testing the exogeneity of explanatory variables in linear regression models. In the cases considered, the estimator contrast was an instrumental variables estimator versus the efficient estimator under the exogeneity hypothesis. Recently, several researchers have used conventional and Hausman methods to find tests of exogeneity within a limited information framework for simultaneous systems of equations (Wu (1973,1974), Hwang (1980), Engle (1981), Nakamura and Nakamura (1981)). Holly and Sargan (1982) give a systematic treatment of the exogeneity test problem, based on a likelihood framework. They show the equivalence of the Hausman, Wu, and classical tests. Spencer and Berk (1981) provide a closely related test within the instrumental variables (IV) framework which we review here. In addition, we note the more general treatment of such test problems by Newey (1983), who derives optimal instrumented score tests.

Consider the i^{th} structural equation in a system of simultaneous equations:

---

2 Small and Hsiao (1982) propose another LR statistic based on a random division of the sample. This test, a more conventional Chow test, has the same distribution under the null hypothesis as these tests, but does not under local alternatives. Small and Hsiao report computational stability of their LR statistic in finite samples compared to the Hausman-McFadden statistic. Presumably, our LR equivalent will share this stability. Vuong (1983) has also suggested a competing LR statistic.

$$y_i = Y_i\beta_i + X_{1i}\gamma_{1i} + X_{2i}\gamma_{2i} + u_i \tag{4.2.2}$$

where y_i is a T by one vector of observations on the chosen endogenous variable, Y_i is the matrix of remaining endogenous variables in the equation, X_{1i} is a T by K_1 matrix of variables known to be orthogonal to u_i, X_{2i} is a T by K_2 matrix of variables suspected to be correlated with u_i, and u_i is a random disturbance with zero expectation and a scalar covariance matrix. Spencer and Berk propose a test of the hypothesis that X_{2i} is orthogonal to u_i.

The Spencer-Berk procedure compares the two stage least squares (2SLS) estimator to another IV estimator that does not include X_{2i} in the list of instruments. The 2SLS estimator is efficient relative to other IV estimators that use instruments which are linear combinations of the exogenous variables in the system. Thus, the statistical comparison of these two estimators follows the method outlined by Hausman. Rewriting equation (4.2.1) as

$$y_i = Z_i\alpha + u_i \tag{4.2.2}$$

where $Z_i = [Y_i\ X_{1i}\ X_{2i}]$ and $\alpha' = [\beta_i'\ \gamma_{1i}'\ \gamma_{2i}']$, the 2SLS estimator can be written as the minimum distance estimator

$$a_0 : \min_a u(a)'X(X'X)^{-1}X'u(a) \tag{4.2.3}$$

where X is the matrix of all exogenous variables in the system. The alternative estimator that remains consistent despite correlation between X_{2i} and u_i is

$$a_1 : \min_a u(a)'X^*(X^{*\prime}X^*)^{-1}X^{*\prime}u(a) \tag{4.2.4}$$

where X^* is the matrix of exogenous variables after removing the submatrix X_{2i}. Spencer and Berk then form the test statistic

$$d'V(a_1,a_0)^{-1}d \tag{4.2.5}$$

where $d = a_1 - a_0$ and $V(a_1,a_0) = V_1(a_1) - V_0(a_0)$ is a consistent estimate of the asymptotic variance of d. They claimed that this statistic is asymptotically a chi square with degrees of freedom $\dim(\alpha)$ under the null hypothesis.[3]

3 This claim was corrected in Spencer and Berk (1982).

Using the well-known matrix identity

$$X(X'X)^{-1}X' = X^*(X^{*\prime}X^*)^{-1}X^{*\prime} + W(W'W)^{-1}W'$$
$$\text{where} \quad W = [I - X^*(X^{*\prime}X^*)^{-1}X^{*\prime}]X_{2i} \qquad (4.2.6)$$

one sees that the Spencer-Berk method falls within the family of factorized criterion function tests, where the third criterion function is

$$u(a)'W(W'W)^{-1}W'u(a) \qquad (4.2.7)$$

The implicit third estimator contains only K_2 estimable the parameters, given the dimensions of W, so that the test can really only have K_2 degrees of freedom. The Spencer-Berk d, therfore, has a singular asymptotic distribution. One can either amend (4.2.5) to use a generalized inverse or simply reduce d to the competing estimates of γ_{2i}.

The method of factorization used in this problem is, of course, exactly analogous to the method described in the panel data problem. The error term is projected into the space spanned by a subset of explanatory variables and the orthogonal space to create two orthogonal estimators. These may be compared with the relatively efficient estimator. If one adds the assumption of normality of the u, there is an immediate likelihood ratio test based on this same factorization, as Holly and Sargan (1982) conjectured.

4.1.3 Censored Normal Regression

Finally, let us return to the introductory example of testing for misspecification of the distribution in the censored regression model (Tobit).[4] In the example, we constructed several Hausman specification tests based on a likelihood factorization into Probit and truncated regression likelihoods. Nelson (1981) has proposed a test of the Tobit model based on Hausman's method which we will compare with the test of our example.[5] Both tests can be viewed as tests based on estimators that are both inconsistent under the alternative hypothesis, even though Nelson constructed a comparison of efficient

4 See also Vuong (1983).

5 Amemiya (1982) suggests an alternative comparison of the Tobit estimator with the robust estimator by Powell (1981). Lin and Schmidt (1984) propose a specification test based on a classical score test for a two-stage alternative model.

and and inefficient, but consistent, estimators of raw data sample moments. Implicitly Nelson's test is also a comparison of alternative estimates of the regression slope parameters.

In the simplest case, the Nelson test amounts to comparing Probit and Tobit MLE's for the standardized mean. In the extension to full regression models, Nelson's test is a score test which compares the Tobit MLE with an unusual, inefficient, binary data estimator, and not Probit. We will argue that the relative efficiency of the Probit estimator may make it a better choice for comparison. In addition, such an ammended test has the interpretation of comparing an out of sample predictor for the dependent variable with its realization, as in our multinomial logit example.

Nelson proposes a test of the normality assumption in the Tobit model by comparing $N^{-1}X'y$ with its efficient estimator

$$d = N^{-1}\{X'y - [X'P(c_0)Xc_0 + s_0X'p(c_0)]\}$$
$$\text{where } P(c_0) = diag[\Phi(x_n{}'c_0)] \text{ and } p(c_0) = [\phi(x_n{}'c_0)] \tag{4.3.1}$$

based on the Tobit MLE. But by concentrating out σ from the the Tobit log-likelihood function, this difference can be rewritten as

$$N^{-1}s_0\Big\{X'[1(y>0) - P(c_0)]Xc_0 - X'[1(y>0) - P(c_0)]\,[I - P(c_0)]^{-1}p(c_0)\Big\} \tag{4.3.2}$$

where $1(y>0)$ is a diagonal matrix with $1(y_n>0)$ in the diagonal positions. The expression in the brace brackets implicitly gives the estimator with which the Tobit estimator is being compared in this test when the test is viewed as a score test:

$$c: \; X'[1(y>0) - P(c)]\,[I - P(c)]^{-1}p(c) \tag{4.3.3}$$

Note that this estimator uses only the sign information about the dependent variable and is therefore similar to the Probit estimator which solves the normal equations

$$L_{1,\gamma}(c_1) = \sum_{n=1}^{N} \frac{p_n(c_1)}{P_n(c_1)[1 - P_n(c_1)]}\,[1(y_n>0) - P_n(c_1)]\,x_n = 0 \tag{4.3.4}$$

but these are not identical estimators.

Given that one can construct a variety of binary data estimators for γ , which should we compare with c_1 ? For guidance, note that the properties of a powerful test of this kind are (1) a large difference between estimators under the alternative model and (2) a relatively efficient alternative estimator to the MLE. Without specifying the alternative model parametrically, there is no information about the estimators' differences. However, Probit is the most efficient binary data estimator for the normal model and is in this sense the preferred contrast to Tobit. Our results on factorization indicate that a Probit-Tobit contrast is asymptotically equivalent to the contrast of Tobit and the truncated regression MLE and of Probit and the truncated regression MLE. All three tests are asymptotically chi square with degrees of freedom equal to the number of regressors.

We muster several observations in support of using this as a diagnostic test for censored regression. First, one version of the score test computes the truncated regression estimator, c_2 , and examines the Probit score

$$L_{1,\gamma}(c_2) = \sum_{n=1}^{N} \frac{p_n(c_2)}{P_n(c_2)[1 - P_n(c_2)]} [1(y_n > 0) - P_n(c_2)] \, x_n. \tag{4.3.5}$$

Once again, the interpretation of (4.3.5) is that one is comparing a prediction with a realization; that is, the predicted probability of a positive observation, $P_n(c_2)$, and whether the observation is positive, $1(y > 0)$. This is an intuitive verification of the usefulness of the test in detecting model misspecification. Second, Nelson's Monte Carlo experiment adds further support. Third, the contrast of the binary data with the continuous data is also consistent with the kinds of misspecification that have concerned econometricians (see Cragg (1971) and Lin and Schmidt (1984)). Finally, the test uses estimators available from computer software. Note that normality is inessential to our arguments for testing the censored regression model and the factorization is just as convenient for other distributions. This is not true for the logit problem discussed above, but such generalized Chow tests as these seem natural for general application to limited dependent variable models. These models typically yield likelihood functions with conditional and marginal factors.

ACKNOWLEDGEMENTS

The author wishes to thank the editor, Dale J. Poirier, and the associate editors, Alberto Holly and Peter Schmidt, for their comments on an earlier draft of this paper. Thomas Rothenberg and Jiro Hodoshima also provided valuable comments. Katherine Rockett provided research assistance. The author also acknowledges the research support of the National Science Foundation.

REFERENCES

Aitchison, J. & Silvey, S.D., (1958). Maximum Likelihood Estimation of Parameters Subject to Restraints. *Annals of Mathematical Statistics 29*, 813-828.

Amemiya, T., (1973). Regression Analysis When the Dependent Variable is Truncated Normal. *Econometrica 41(6)*, 997-1016.

Amemiya, T., (1982). Tobit Models: A Survery. *Criminal Justice Research Series 4.* Palo Alto: Rhodes Associates.

Arabmazar, A., & Schmidt, P., (1982). An Investigation of the Robustness of the Tobit Estimator to Non-normality. *Econometrica 50(4)*, 1055-1063.

Breusch, T.S. & Pagan, A.R., (1978). The Lagrange Multiplier Test and Its Application to Model Specification in Econometrics. *Review of Economic Studies 67(1)*, 239-253.

Burgete, J., Gallant, A.R., & Souza, G., (1982). On the Unification of the Asymptotic Theory of Nonlinear Econometric Models. *Econometric Reviews 1(2)*, 151-190.

Chesher, A., & Lancaster, T., (1983). The Covariance Matrix of the Information Matrix Test. Mimeograph, University of Florida.

Cox, D.R., (1961). Tests of Separate Families. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, Berkeley: Univ. of California Press.

Cox, D.R. & Hinkley, D.V., (1974). *Theoretical Statistics*. London: Chapman and Hall.

Cragg, J., (1971). Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods. *Econometrica 39(5)*, 829-845.

Das Gupta, S. & Perlman, M.D., (1974). Power of the Noncentral F-test. *JASA 69*, 174-180.

Durbin, J., (1970). Testing for Serial Correlation in Least-Squares Regression When Some of the Regressors are Lagged Dependent Variables. *Econometrica 38(3)*, 410-421.

Engle, R.F., (1979). A General Approach to the Construction of Model Diagnostics Based Upon the Lagrange Multiplier Principle. Warwick Economic Research Paper 156, University of Warwick, Coventry.

Engle, R.F., (1981). Wald, Likelihood Ratio, and Lagrange Multiplier Tests in Econometrics. *Handbook of Econometrics* (Griliches & Intriligator, Eds.), draft.

Goldberger, A.S., (1983). Abnormal Selection Bias. *Studies in Econometrics, Time Serieis and Multivariate Statistics*, New York: Academic Press.

Godfrey, L.F., (1978). Testing for Multiplicative Heteroskedasticity. *Journal of Econometrics 8*, 227-236.

Goldfeld, S.M. & Quandt, R.E., (1972). *Nonlinear Methods in Econometrics*, Amsterdam: North-Holland.

Gourieroux, C. & Trognon, A., (1984). Specification Pre-test Estimator. *Annals of Econometrics*, forthcoming.

Griliches, Z., Hall, B.H., & Hausman, J.A., (1978). Missing Data and Self-Selection in Large Panels. *Annales de l'Insee* 30-31, 137-176.

Hall, A.R., (1983). The Information Matrix Test for the Linear Model. Mimeograph, University of Warwick.

Hansen, L.P., (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica 50(4)*, 1029-1054.1.

Hausman, J.A., (1978). Specification Tests in Econometrics. *Econometrica 46(6)*, 1251-1271.

Hausman, J.A. & McFadden, D., (1983). Specification Tests for the Multinomial Logit Model. Mimeograph, Massachusetts Institute of Technology, Cambridge, MA.

Hausman, J.A. & Pesaran, H., (1982). The J-test as a Hausman Specification Test. Discussion Paper No. 945, Harvard Institute for Economics Research, Harvard University, Cambridge, MA.

Hausman, J.A. & Taylor, W.E., (1981). Comparing Specification Tests and Classical Tests. Working Paper, Massachusetts Institute of Technology, Cambridge, MA.

Hausman, J.A. & Taylor, W.E., (1981). Panel Data and Unobservable Individual Effects. *Econometrica 49(6)*, 1377-1398.

Hausman, J.A. & Taylor, W.E., (1982). A Generalized Specification Test. *Economics Letters 8*, 239-245.

Holly, Alberto, (1982a). A Remark on Hausman's Specification Test. *Econometrica 50(3)*, 749-760.

Holly, Alberto, (1982b). Tests de Specification. *Cahiers du Seminaire d'Econometrie 24*, 151-173.

Holly, Alberto, (1982c). Some Useful Equivalence Properties of Hausman's Test. Working Paper No. 8201, University of Lausanne.

Holly, Alberto, & Sargan, J.D., (1982). Testing for Exogeneity within a Limited Information Framework. Working Paper No. 8204, University of Lausanne.

Hsiao, C. & Small, K.A., (1982). Multinomial Logit Specification Tests. Mimeograph, Princeton University.

Hwang, H.-S., (1981). Test of Independence Between a Subset of Stochastic Regressors and Disturbances. *International Economic Review 21*, 749-760.

Judge, G.G. & Bock, M.E., (1978). *The Statistical Implications of Pre-test and Stein-rule Estimators in Econometrics.* Amsterdam: North Holland.

Lee, L.-F., (1981). A Specification Test for Normality Assumption for the Truncated and Censored Tobit Models. Discussion Paper No. 81-147, Department of Economics, University of Minnesota.

Lin, T.-S. & Schmidt, P., (1984). A Test of the Tobit Specification Against an Alternative Suggested by Cragg. *Review of Economics and Statistics 66(1)*, 174-176.

Maddala, G.S., (1971). The Use of Variance Components Models in Pooling Cross-Section and Time Series Data. *Econometrica* 39(2), 341-358.

McFadden, D., (1973). Conditional Logit Analysis of Qualitative Choice Behavior. *Frontiers in Econometrics* (Zarembka, Ed.). New York: Academic Press, 105-142.

Morey, M.J., (1982). The Statistical Implications of Preliminary Specification Error Testing. Mimeograph, Indiana University.

Nakamura, A. & Nakamura, M., (1981). On the relationships Among Several Specification Error Tests Presented by Durbin, Wu, and Hausman. *Econometrica 49(6)*, 1583-1588.

Nelson, Forrest D., (1981). A Test for Misspecification in the Censored Normal Model. *Econometrica 49(5)*, 1317-1330.

Newey, W.K., (1983). Maximum Likelihood Specification Testing and Instrumented Score Tests. Mimeograph, Princeton University.

Plosser, C.I., Schwert, G.W., & White, H., (1982). Differencing as a Test of Specification. *International Economic Review 23(3)*, 535-552.

Poirier, D.J. & Ruud, P.A., (1979). A Simple Lagrange Multiplier Test for Lognormal Regression. *Economics Letters 4*, 251-255.

Powell, J.L., (1981). Least Absolute Deviations Estimation for Censored and Truncated Regression Models. Technical Report 356, IMSSS, Stanford University.

Ramsey, J.B., (1969). Tests for Specification Errors in Classical Linear Least Squares Regression Analysis. *JRSS, Series B* 31(2), 350-371.

Ramsey, J.B., (1974). Classical Model Selection Through Specification Error Tests. *Frontiers in Econometrics* (P. Zarembka, Ed.), New York: Academic Press.

Rao, C.R., (1947). Large Sample Tests of Statistical Hypotheses Concerning Several Parameters with Applications to Problems of Estimation. *Proceedings of the Cambridge Philosophical Society 44*, 50-57.

Rao, C.R., (1973). *Linear Statistical Inference and Its Applications.* New York: Wiley.

Riess, P., (1983). Alternative Interpretations of Hausman's m Test. Mimeograph, Yale University.

Ruud, P.A., (1982). A Score Test of Consistency. Mimeograph, University of California, Berkeley.

Sen, P.K., (1979). Asymptotic Properties of Maximum Likelihood Estimators Based on Conditional Specification. *Annals of Statistics 7(5)*, 1019-1033.

Silvey, S.D., (1959). The Lagrangian Multiplier Test. *Annals of Mathematical Statistics 30(2)*, 389-407.

Spencer, D.E. & Berk, K.N., (1981). A Limited Information Specification Test. *Econometrica*, 49(4), 1079-1085.

Spencer, D.E. & Berk, K.N., (1982). Erratum. *Econometrica 50(4)*, 1087.

Thursby, Jerry G., (1982). The Common Structure of Regression-Based Specification Error Tests. Mimeograph, Department of Economics, Ohio State University.

Vuong, Q.H., (1983). Misspecification and Conditional Maximum-Likelihood Estimation. Mimeograph, California Institute of Technology.

White, Halbert, (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica 48(4)*, 817-838.

White, Halbert, (1982). Maximum Likelihood Estimation of Misspecified Models. *Econometrica 50(1)*, 1-26.

Wu, De-Min, (1973). Alternative Tests of Independence Between Stochastic Regressors and Disturbances. *Econometrica* 41(4), 733-750.

Wu, De-Min, (1974). Alternative Tests of Independence Between Stochastic Regressors and Disturbances: Finite Sample Results. *Econometrica 42(3)*, 529-546.