Increasing the Power of Specification Tests

Tiemen Woutersen and Jerry A. Hausman*
University of Arizona and MIT

November 18, 2018

Abstract. This paper shows how to increase the power of Hausman's (1978) specification test as well as the difference test in a large class of models. The idea is to impose the restrictions of the null and the alternative hypotheses when estimating the covariance matrix. If the null hypothesis is true then the proposed test has the same distribution as the existing ones in large samples. If the hypothesis is false then the proposed test statistic is larger with probability approaching one as the sample size increases in several important applications, including testing for endogeneity in the linear model.

Keywords: Specification test, Hausman test, Power of tests.
JEL Codes: C01, C14, C18, C41

1. Introduction

Specification tests are important in empirical economics and other empirical research. Such tests help the researcher to evaluate whether the estimate of a quantity of interest would change if the model was changed. The purpose of this paper is to derive tests that are more powerful than the existing ones in important applications. The main idea is to impose the restrictions of the null and alternative hypotheses when estimating the variation of the test statistic. In particular, we impose the null and alternative hypotheses when calculating the Hausman (1978) test and other tests that are based on the difference between two estimators. Durbin (1954) and Wu (1973) propose such tests for the linear

*Forthcoming in the Journal of Econometrics. We thank Donald Andrews, Miriam Arden, Xiaohong Chen, Antonio Galvao, Carter Hill, Whitney Newey, and Gary Solon for helpful discussions. We also thank seminar participants at MIT, New York University Abu Dhabi, Pennsylvania State University, and Yale University. The comments of an associate editor and two referees improved and shortened the paper and we are grateful for those. All remaining errors are our own. Comments are welcome: woutersen@email.arizona.edu.

model, while Hausman (1978) considers a more general framework. These specification tests are very popular in empirical economics and other empirical research.[1] A reason for the popularity of these specification tests is that one can make a judgement about whether the estimates of the quantity of interest differ in a scientifically significant way in the two (economic) models, as well as whether this difference is statistically significant. Some tests such as score tests only yield whether a difference is statistically significant and are less suitable for economic interpretation.

So many papers use the Hausman (1978) test that we cannot review them all, but Zapata et al. (2012) and Adkins et al. (2012) review applications of the Hausman (1978) test. Guggenberger (2010) considers using the Hausman test for pretesting, but this paper is concerned with (i) testing whether two estimands are different and (ii) testing whether an efficient estimator differs from a robust one. A related paper is Woutersen (2016), which shows how to improve the power of the Hansen (1982) and Sargan (1958) tests. We extend the techniques of that paper to the Hausman (1978) test. This paper is organized as follows: section 2 presents two examples, section 3 gives the theorem and section 4 concludes.

2. Examples

This section presents examples that show how imposing the restrictions of the null *and* alternative hypotheses can yield a more powerful test.

2.1. Testing for Endogeneity. We first consider the linear model with a binary regressor that is potentially endogenous and with a binary instrument. Suppose we observe a random sample $\{X_i, Y_i, Z_i\}$, $i = 1, ..., N$. Let

$$Y_i = \alpha + \beta X_i + \varepsilon_i. \tag{1}$$

Further assume that the variation of the regressor X_i is strictly positive, the variation of the error term ε is bounded, and the instrument Z_i is correlated with the regressor X_i. The null hypothesis in Durbin (1954), Wu (1973), and Hausman (1978) is that

[1] For example Hausman's (1978) article is cited 4184 times in the Web of Science Core Collection as of the time of this writing; see also Kim, Morse, and Zingales (2006) for an earlier review. Some examples of textbooks that review the Hausman (1978) test are Ruud (2000), Cameron and Trivedi (2005), Wooldridge. (2010), and Greene (2012); the test is also reviewed by any other graduate econometric textbook that we checked; Romano and Shaikh, and Wolf (2010) review tests in econometrics.

the error term ε is homoscedastic and normally distributed, conditional on X_i and Z_i, i.e. $\varepsilon_i|X_i, Z_i \sim N(0, \sigma^2)$. In this case the least squares estimator for β is efficient. The alternative hypothesis in these papers is that ε_i has expectation zero for all Z_i and is homoscedastic given Z_i. In that case, the two-stage least squares estimator is consistent for β. Hausman (1978) observes that one can test this null versus alternative hypothesis by testing whether γ equals zero in the following equation:

$$Y_i = \alpha + \beta X_i + \gamma v_i + u_i, \tag{2}$$

where v_i is the residual of regressing X_i on a constant and Z_i. If the null hypothesis holds, then $\gamma = 0$, so u_i and ε_i are identical. Thus, in that case we have that $u_i|X_i, Z_i \sim N(0, \sigma^2)$. Let $\hat{u}_i$, $i = 1, ..., N$, denote the least squares residuals of regressing Y_i on a constant, X_i, and v_i. Then the following test statistic has a T-distribution with $N-3$ degrees of freedom under the null hypothesis,

$$T_{Hausman} = \gamma_{LS} \frac{\sqrt{\sum_i v_i^2 (1 - \frac{\sum_i v_i^2}{\sum_i X_i^2})}}{\sqrt{\frac{1}{N-3} \sum_i \hat{u}_i^2}}, \tag{3}$$

where γ_{LS} denotes the least squares estimator of γ. An observation that has not been used before is that the residuals $\hat{u}_i$, $i = 1, ..., N$, may be correlated with the interaction term $X_i Z_i$ under the alternative. In particular, if X_i and Z_i are binary, as above, then we have that the average value of the residual $\hat{u}_i$ given $X_i = 0$ is zero, but not necessarily that the average value of $\hat{u}_i$ is zero for the sub population with $X_i = Z_i = 0$ or other sub populations. We can remove this correlation by regressing $\hat{u}_i$, $i = 1, ..., N$, on a constant, X_i, Z_i, and $X_i Z_i$, or, equivalently, by regressing Y_i on these variables. Let $\hat{\eta}_i$ denote the residual of this regression. Then the following test statistic has a T-distribution with $N - 4$ degrees of freedom under the null hypothesis[2],

$$T_{New} = \gamma_{LS} \frac{\sqrt{\sum_i v_i^2 (1 - \frac{\sum_i v_i^2}{\sum_i X_i^2})}}{\sqrt{\frac{1}{N-4} \sum_i \hat{\eta}_i^2}}. \tag{4}$$

Under the null hypothesis, the estimators for the variation of u_i, $\frac{1}{N-3} \sum_i \hat{u}_i^2$ and $\frac{1}{N-4} \sum_i \hat{\eta}_i^2$, are unbiased. However, $\frac{1}{N-4} \sum_i \hat{\eta}_i^2$ is smaller, with probability approaching one under the alternative hypothesis. Our simulations below illustrate this. For example, we find in our

[2]See the appendix for details.

simulations that imposing the orthogonality condition can increase the power of the test by up to 29 percentage points. The simulations control for size so that the Hausman test uses the critical value of the T_{N-3} distribution while the new test uses the critical value of the T_{N-4} distribution. Both numbers converge to the critical value of the standard normal distribution, and the critical values of the T_{N-3} and T_{N-4} are very similar even for moderate sample sizes[3] and, therefore, do not have a noticeable role in the simulation results.

Suppose we observe a random sample $\{X_i, Y_i, Z_i\}, i = 1, ..., N$, and that these random variables satisfy the following conditions,

$$\begin{aligned} &X_i, Z_i \in \{0,1\}, \\ &P(Z_i = 1) = P(Z_i = 0) = \frac{1}{2},\ P(X_i = 1|Z_i) = \frac{2}{5} + \frac{Z_i}{5}, \\ &Y_i = X_i + \varepsilon_i, \text{ where} \\ \varepsilon_i|X_i, Z_i &\sim N(\mu_i, 1), \text{ and} \\ &\mu_i = \delta \cdot \{\frac{1}{3} \cdot 1(Z_i = X_i = 0) - \frac{1}{2} \cdot 1(Z_i = 0,\ X_i = 1) \\ &\qquad - 1(Z_i = 1,\ X_i = 0) + \frac{2}{3} \cdot 1(Z_i = X_i = 1)\}. \end{aligned}$$

Thus, for $\delta = 0$, the conditional expectation of the error term is zero. In that case, the least squares estimator is consistent and efficient. The two-stage least squares estimator is the 'robust estimator' in Hausman's (1978) terminology and is consistent for any value of δ. Also, this data generating process can be written as a function of four unobserved types[4].

The simulations below show that the proposed test improves on the Hausman test.

Table 1: 0.05 Rejection Frequencies

N	δ	New Test	Hausman Test	Ratio
200	1	0.14162	0.08453	1.67538
400	1	0.18110	0.11473	1.57849
800	1	0.25621	0.17233	1.48674
200	2	0.34411	0.12828	2.6824914
400	2	0.43576	0.18928	2.3021978
800	2	0.58675	0.31749	1.8480897

Results based on 100,000 simulations.

[3] We report the critical values in the appendix and these critical values are the same up to 4 digits for our sample sizes.

[4] Angrist, Imbens and Rubin (1996) introduce a terminology for these four unobserved types.

In table 1, the size of the tests is 5%, i.e. the critical value is such that the probability of falsely rejecting the null hypothesis is 5%. In table 2, the size of the tests is reduced to 1%, and we see the same pattern as in table 1. That is, the endogeneity bias of the least squares estimator is detected more frequently.

Table 2: 0.01 Rejection Frequencies

N	δ	New Test	Hausman Test	Ratio
200	1	0.05218	0.02268	2.30070
400	1	0.07427	0.03369	2.20452
800	1	0.11870	0.06015	1.97339
200	2	0.20891	0.040800	5.12034
400	2	0.28870	0.071180	4.05591
800	2	0.43717	0.142980	3.05756

Results based on 100,000 simulations.

We summarize this subsection by observing that the Hausman test for endogeneity has the feature that, under the alternative hypothesis, the least squares residuals are correlated with the instrument and interaction term while the two stage least squares residuals are correlated with the regressor and interaction term. These simulations show that imposing the conditions that the residuals are uncorrelated with the regressor, instrument, and interaction term yields a more powerful test. More generally, the idea is to impose the restrictions of the null and the alternative hypothesis when estimating the covariance matrix. We first consider a panel data model and then present our theorem.

2.2. Panel Data. In panel data we observe an individual more than once. An advantage of panel data is that we can test whether the regressors are exogenous, just as in the last section. However, with panel data we can test the exogeneity assumption against the alternative that the regressor is exogenous after we condition on an individual effect[5]. The following example illustrates this issue and shows how to apply our method. Consider the linear panel data model. The first estimator is the least squares estimator, and this estimator is consistent under the null hypothesis that the error term is randomly distributed. The second estimator is the individual (or fixed) effect least squares estimator. This estimator allows for time invariant heterogeneity that can depend on the regressors. Thus, the individual effect least squares estimator is consistent under the alternative hypothesis

[5] A motivation why it may be important to condition on an individual effect is that these regressors could be chosen by the firm or the individual and, therefore, would depend on the individual effect; Wooldridge (2010) gives an overview of panel data econometrics.

that allows for such dependence.

In the last example, we used the fact that under the alternative hypothesis, the least squares residuals are correlated with the instrument and the interaction term while the two-stage least squares residuals are correlated with the regressor. Here, we use that the residuals of the least squares (or random effects) estimator are correlated with the regressors. We can thus impose the null hypothesis of no correlation and improve the power of the Hausman (1978) test.

We now specify the data generating process for the panel data example. Suppose we observe $\{X_{it}, Y_{it}\}, i = 1, ..., N$, $t = 1, 2$, and that these variables satisfy the following conditions,

$$Y_{it} = \alpha + \beta X_{it} + \varepsilon_{it}, \tag{5}$$

$$\varepsilon_{it}|X_{i1}, X_{i2} \sim N(\mu_i, \sigma^2),\ t = 1, 2 \text{ where}$$

$$\mu_i = \delta \cdot \{\frac{1}{4} + \frac{3}{2}X_{i1} - \frac{5}{4}X_{i2}\}.$$

Also, let the error term $\varepsilon_{it}|X_{i1}, X_{i2}$ be uncorrelated across time, $t = 1, 2$, and let it not depend on regressors or error terms of other individuals. Let $X_{i1} = X_{i2} = 0$ for $i = 1, 4, 7, ...$, let $X_{i1} = 0$ and $X_{i2} = 1$ for $i = 2, 5, 8, ...$, and let $X_{i1} = X_{i2} = 1$ for $i = 3, 6, 9, ...$etc. Note that for $\delta = 0$, the conditional expectation of the error term is zero. In that case, the least squares estimator with a common intercept is consistent and efficient, while the least squares estimator with individual effects is the 'robust estimator' in Hausman's (1978) terminology. One way to implement the Hausman test is to take the difference between these two estimators. Another way that is very popular in empirical work is to test whether the parameter γ equals zero in the following regression. These two versions of the Hausman test are equivalent, see for example Wooldridge (2010). Consider

$$Y_{it} = \alpha + \beta X_{it} + \gamma \bar{X}_i + u_{it},$$

where $\bar{X}_i = \frac{X_{i1}+X_{i2}}{2}$. Let $\hat{u}_{it}$ denote the residual from this regression. The efficiency of the estimator under the null implies that the errors are uncorrelated with all regressors. We impose this by regressing Y_{it} on a constant, X_{it}, X_{i1}, and X_{i2}. Let $\hat{\eta}_{it}$ denote the residual of this regression. We then construct an estimate of the variation using $\hat{u}_{it}$ and $\hat{\eta}_{it}$. The

Hausman test, based on $\hat{u}_{it}$, has an T-distribution with $NT-3$ degrees of freedom while the proposed test has an T-distribution with $NT-4$ degrees of freedom.

We use the following parameter values for our simulation, $\alpha = \beta = 0$, δ has the value $\frac{3}{4}$, or $\frac{3}{2}$ depending on the simulation. The simulations below show that the proposed test is more powerful than the Hausman test. The null hypothesis is violated in every case in the simulations shown here so that it is desirable to have high rejection frequencies. In the appendix we show that the tests have correct size and we show simulation results when the test is done at the 99% confidence level. For that case, the results are stronger than for the 95% confidence test shown below.

Table 3: 0.05 Rejection Frequencies; H_0: Common Intercept

N	δ	New Test	Hausman Test	Ratio
300	3/4	0.19660	0.13930	1.41134
600	3/4	0.33450	0.26040	1.28456
1200	3/4	0.57800	0.49210	1.17455
300	3/2	0.58150	0.27430	2.11994
600	3/2	0.85980	0.62170	1.38298
1200	3/2	0.99130	0.94460	1.04943

Results based on 10,000 simulations.

The simulations control for size. In particular, the Hausman test uses the critical value of the T_{N-3} distribution while the new test uses the critical value of the T_{N-4} distribution. Both numbers converge to the critical value of the standard normal distribution, and the critical values of the T_{N-3} and T_{N-4} are very similar even for moderate sample sizes[6] and, therefore, do not have a noticeable role in the simulation results.

The simulations in the last table have that the null hypothesis is a common intercept. An alternative null hypothesis is that every individual has her own intercept and that this intercept is random, i.e. independently distributed from all regressors. That model can be estimated using Generalized Least Squares (GLS), see Hausman (1978) or Wooldridge (2010) for details. GLS uses residuals to estimate the covariance matrix. Instead of using $\hat{u}_{it}$, defined above, we propose to use $\hat{\eta}_{it}$. The data generating process is the same as table 3, and the results are similar as well.

[6] We report the critical values in the appendix and these critical values are the same up to 5 digits for our sample sizes.

Table 4: 0.05 Rejection Frequencies; H_0: Random Effects

N	δ	New Test	Hausman Test	Ratio
300	3/4	0.19680	0.13460	1.46211
600	3/4	0.33500	0.25500	1.31372
1200	3/4	0.57830	0.48440	1.19384
300	3/2	0.58150	0.25610	2.27059
600	3/2	0.86030	0.60310	1.42646
1200	3/2	0.99120	0.93770	1.05705

Results based on 10,000 simulations.

3. Theorem

The last section gave examples where the power of the Hausman (1978) test could be improved. We now generalize these examples and state our theorem. In particular, we allow for the class of estimators that can be written in terms of an influence function as in Newey and McFadden (1994). The examples of the last section are in this class, and so are generalized method of moments estimators and many of the maximum likelihood estimators. Thus, we assume that we have two estimators, each with an influence functions. This assumption is nothing more than that the estimator can be approximated by an average.

Assumption 1

Let

$$\hat{\beta}_1 = \beta + G^{-1}\frac{\sum_i}{N} g_i(X_i, Y_i, Z_i) + o_p(1/\sqrt{N})$$
$$\hat{\beta}_2 = \beta + H^{-1}\frac{\sum_i}{N} h_i(X_i, Y_i, Z_i) + o_p(1/\sqrt{N}),$$

where the vectors $g_i(\cdot)$ and $h_i(\cdot)$ have the same dimension as β, N is the sample size, and G^{-1} and H^{-1} exist.

If a central limit theorem can be applied to the averages in the last assumption, then the estimators are asymptotically normally distributed, and that is what we assume in the next assumption.

Assumption 2

Let

$$\sqrt{N}(\hat{\beta}_1 - \hat{\beta}_2 - \frac{c}{\sqrt{N}}) \underset{d}{\to} N(0, \Omega),$$

for some vector constant c and positive definite Ω. Also, let there exist a consistent estimator for Ω,

$$\widehat{\Omega} = \Omega + o_p(1).$$

Assumption 2 does not require $\hat{\beta}_1$ or $\hat{\beta}_2$ to be efficient, i.e. we allow for our test to be based on a difference. We use the information set W, which consists of data and knowledge of the model that the researcher has. In example 1, W consists of the regressor and the instrument. In example 2, W consists of the regressors in the first and second period. The contribution of the individual to the influence function have expectation zero in many models. For example, the contributions to the score function in likelihood models generally have mean zero. Also, in many applications, the contributions to the generalized method of moment influence function have mean zero, see for example Newey and McFadden (1994). De Jong and Woutersen (2011) give examples of estimating functions in a dynamic model, and these also have this property. The properties (i) $E\{G^{-1}g_i(X_{i,}Y_i, Z_i)|W\} = 0$ for all i and (ii) $E\{H^{-1}h_i(X_{i,}Y_i, Z_i)|W\} = 0$ for all i imply that $E\{G^{-1}g_i(X_{i,}Y_i, Z_i) - H^{-1}h_i(X_{i,}Y_i, Z_i)|W\} = 0$ for all i. The last condition then implies that the $\text{plim}\{\hat{\beta}_1 - \hat{\beta}_2\} = 0$. Of course, the reverse, $E\{G^{-1}g_i(X_{i,}Y_i, Z_i) - H^{-1}h_i(X_{i,}Y_i, Z_i)|W\} \neq 0$ for some i does not imply $\text{plim}\{\hat{\beta}_1 - \hat{\beta}_2\} \neq 0$, so we also use c to state the hypotheses.

Hypotheses

H_0 : Let $c = 0$ and $E\{G^{-1}g_i(X_{i,}Y_i, Z_i) - H^{-1}h_i(X_{i,}Y_i, Z_i)|W\} = 0$ for all i.

H_1 : Let $c \neq 0$ or $E\{G^{-1}g_i(X_{i,}Y_i, Z_i) - H^{-1}h_i(X_{i,}Y_i, Z_i)|W\} \neq 0$ for some i.

The hypothesis H_1 states that a conditional expectation is nonzero and we can use this to reduce the variation of the Hausman test under H_1, just as in the examples. Define $\delta_i = G^{-1}g_i(X_{i,}Y_i, Z_i) - H^{-1}h_i(X_{i,}Y_i, Z_i)$ so that $\hat{\beta}_1 - \hat{\beta}_2 = \frac{\sum_i}{N}\delta_i + o_p(1/\sqrt{N})$. Also define the conditional expectation $\phi_i = E(\delta_i|W)$. Note that $var\{\frac{\sum_i}{N}(\delta_i - \phi_i)\} = var\{\frac{\sum_i}{N}\delta_i\} - var\{\frac{\sum_i}{N}\phi_i\}$, i.e. the conditional expectation reduces the variation. Hypothesis H_1 implies that the conditional expectation is nonzero so that assuming that $var\{\frac{\sum_i}{N}\phi_i\}$ is positive definite is a mild assumption. The case that the conditional expectation absorbs all the variation is not realistic so that we also assume that $var\{\frac{\sum_i}{N}(\delta_i - \phi_i)\}$ is positive definite.

A simple way to approximate the conditional expectation $\phi_i = E(\delta_i|W)$ is by using a projection. In particular, we can use regressors from the information set W to construct the matrix S and the projection matrix P_S. Also, let $P_{S,ii} = S_i(S'S)^{-1}S_i'$, $\hat{\delta}_i = \hat{G}^{-1}g_i(X_{i,}Y_i, Z_i) - \hat{H}^{-1}h_i(X_{i,}Y_i, Z_i)$ where $\hat{G}$ and $\hat{H}$ are estimators for G and H. Instead of subtracting a conditional expectation from $\hat{\Omega}$ we can subtract the projection. That is, define $\tilde{\Lambda} = \hat{\Omega} - \frac{1}{N}\sum_i \hat{\delta}_i' P_{S,ii}\hat{\delta}_i$. This is very similar to using a residual matrix $I - P_S$.

An alternative to using the Hausman test in example 1 is to use the Wald test to test whether $\gamma = \lambda = 0$ in the following model,

$$Y_i = \alpha + \beta X_i + \gamma v_i + \lambda X_i Z_i + \varepsilon_i$$

where $E(\varepsilon_i|X_i, Z_i) = 0$. This Wald test has a χ^2-distribution with two degrees of freedom under the null, while the Hausman test in example 1 has a χ^2-distribution with one degree of freedom (after squaring the test statistic). Of course the Hausman test has the benefit that one can make a judgement about whether the estimates of the quantity of interest differ in a scientifically significant way in the two economic models as well as whether this difference is statistically significant while the Wald test may be less suitable for economic interpretation. Usually an empirical researcher would like to have a test that is powerful against many violations of the model. However, there is only so much information in a data set, so the empirical researcher may have to choose between focusing on β and using the Hausman test (and allowing λ to be nonzero) or using the Wald test and having power against both $\gamma \neq 0$ and $\lambda \neq 0$. The proposed test has more power than the Hausman test against the probability limits of estimators for β to be different and also has some power against $\lambda \neq 0$. This can be an additional advantage for empirical researchers who are choosing between the proposed test and the Wald test. The theorem that follows states that T_{New} and $T_{Hausman}$ have the same asymptotic distribution under H_0 but that, under the conditions of the theorem, T_{New} is more powerful against violations of H_0.

Theorem

Let assumption 1 and 2 hold. Let $\Lambda = \Omega - var(\frac{1}{\sqrt{N}}\sum_i \phi_i)$ where $\phi_i = E\{G^{-1}g_i(X_{i,}Y_i, Z_i) - H^{-1}h_i(X_{i,}Y_i, Z_i)|W\}$ be positive definite and let and $\hat{\Lambda} = \Lambda + o_p(1)$.

(i) If H_0 is true then (a) $T_{Hausman} = N \cdot (\hat{\beta}_1 - \hat{\beta}_2)' \hat{\Omega}^{-1} (\hat{\beta}_1 - \hat{\beta}_2) \underset{d}{\rightarrow} \chi^2$-distribution with $\dim(\beta_1)$ degrees of freedom and (b) $T_{New} = N \cdot (\hat{\beta}_1 - \hat{\beta}_2)' \hat{\Lambda}^{-1} (\hat{\beta}_1 - \hat{\beta}_2) \underset{d}{\rightarrow} \chi^2$-distribution with $\dim(\beta_1)$ degrees of freedom.

(ii) If H_1 is true and $var(\frac{1}{\sqrt{N}} \sum_i \phi_i)$ is positive definite then T_{New} is more powerful then $T_{Hausman}$ in the sense that $T_{New} > T_{Hausman}$ with probability approaching one.

(iii) If H_1 is true and $\hat{\Lambda} = \hat{\varphi} \hat{\Omega}$ where $\hat{\varphi} = \varphi + o_p(1)$ and $0 < \varphi < 1$ then (a) $T_{Hausman} \underset{d}{\rightarrow}$ noncentral χ^2-distribution with $\dim(\beta_1)$ degrees of freedom and noncentrality parameter $c' \Omega^{-1} c$ (b) $T_{New} = \frac{T_{Hausman}}{\varphi} + o_p(1)$.

(iv) Let $\tilde{\Lambda} = \hat{\Omega} - \frac{1}{N} \sum_i \hat{\delta}_i' P_{S,ii} \hat{\delta}_i$ where $\hat{\delta}_i = \hat{G}^{-1} g_i(X_i, Y_i, Z_i) - \hat{H}^{-1} h_i(X_i, Y_i, Z_i)$, $\hat{G} = G + o_p(1)$, $\hat{H} = H + o_p(1)$, and $\text{plim}\{\frac{1}{N} \sum_i \hat{\delta}_i' P_{S,ii} \hat{\delta}_i\}$ is positive definite. Let $T^*_{New} = (\hat{\beta}_1 - \hat{\beta}_2)' \tilde{\Lambda}^{-1} (\hat{\beta}_1 - \hat{\beta}_2)$. Then $T^*_{New} \underset{d}{\rightarrow} \chi^2$-distribution with $\dim(\beta_1)$ degrees of freedom under H_0. If H_1 is true then T^*_{New} is more powerful then $T_{Hausman}$ in the sense that $T^*_{New} > T_{Hausman}$ with probability approaching one.

Proof: See appendix.

Adjusting for the degrees of freedom does not effect the results in the theorem. However, we suggest to make such a correction if the number of regressors that is used in P_S is large. That would yield $\tilde{\Lambda} = \frac{N - K_1}{N - K_2} \hat{\Omega} - \frac{1}{N - K_2} \sum_i \hat{\delta}_i' P_{S,ii} \hat{\delta}_i$ where K_1 is the number of parameters used to construct $\hat{\Omega}$ while K_2 is the total number of parameters (including regressors) used for $\hat{\Omega}$ and P_S (our examples make such corrections).

The theorem is stated in terms of covariance matrices and conditional expectations and allows for unobservables to be dependent. In particular, the information set W is the same for all individuals so that, conditional on W, the conditional expectation of one individual does not contain information about the conditional expectation of another individual. An application of the idea of this paper to an information set that is predetermined is beyond the scope of the paper. Our line of attack would be to use the covariance estimator by Newey and West (1987) to estimate $var(\frac{1}{\sqrt{N}} \sum_i \phi_i)$.

Further, the motivation to use local asymptotics in assumption 1 is to ensure that the Hausman test and the proposed test have the same distribution under H_0, and, therefore, have the same critical values. In example 2 we have that μ_i is different for individuals

with $X_{i1} = X_{i2} = 1$ and $X_{i1} = X_{i2} = 0$. In local asymptotics this difference in μ_i has to be proportional to $\frac{1}{\sqrt{N}}$. This means that the reduction in the variation that we use in the theorem cannot be this difference but, rather, the difference between the individuals who have the same values of the regressor versus those who have $X_{i1} = 0$ and $X_{i2} = 1$. Such a difference may be plausible in applications and goes against the common intercept and random effects assumptions so that a test is warranted.

Also, Hahn et al. (2011) consider the linear model with endogeneity and assume that the applied researcher has a weak set of instruments that is valid and a strong set of instruments that is invalid. They then use the local asymptotics that was introduced by Staiger and Stock (1997) and also regress the residuals on the instruments. Thus, our theorem can also be viewed as a generalization of that approach.

Finally, the theorem above also applies to subset inference. The Hausman (1978) test, as well as the difference test, is a convenient way to do subset inference (i.e. test the parameters that you care about), and the new test has this advantage as well.

4. Conclusion

This paper shows how to increase the power of Hausman's (1978) specification test, as well as the difference test, in a large class of models. For example, consider the case where a researcher compares her results that assume that a regressor in the linear model is exogenous to another study that allows for endogeneity and uses instruments. We propose to estimate the covariance matrix based on residuals that are orthogonal to these exogenous regressors, instruments, and interaction terms. This differs from using least squares residuals that are only orthogonal to the exogenous regressors. More generally, the idea is to impose the restrictions of the null and the alternative hypotheses when estimating the covariance matrix. If the null hypothesis is true, then the proposed test has the same distribution as the existing one in large samples. If the hypothesis is false, then the proposed test statistic is larger with probability approaching one as the sample size increases in several important applications, including testing for endogeneity in the linear model. As the Hausman (1978) test is very popular in empirical work, we expect the current results to be useful as well.

Acknowledgments

We thank Donald Andrews, Miriam Arden, Xiaohong Chen, Antonio Galvao, Carter Hill, Whitney Newey, and Gary Solon for helpful discussions. We also thank seminar participants at MIT, New York University Abu Dhabi, Pennsylvania State University, and Yale University. The comments of an associate editor and two referees improved and shortened the paper and we are grateful for those. All remaining errors are our own.

References

[1] Abadir, K. M. and J. R. Magnus (2005): *Matrix Algebra, Econometric Exercises 1*, Cambridge University Press, Cambridge.

[2] Adkins, L.C., R. C. Campbell, V. Chmelarova, R. C. Hill (2012): "The Hausman Test, and Some Alternatives, with Heteroskedastic Data," in *Advances in Econometrics: Essays in Honor of Jerry Hausman*, edited by B. H. Baltagi, R. C. Hill, W. K. Newey, H. L. White, Volume 29, Emerald Group Publishing, UK, 515-546.

[3] Angrist, J.D., G.W. Imbens and D.B. Rubin (1996): "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, 91: 444-455.

[4] Cameron, A. C. and P. K. Trivedi (2005): *Microeconometrics, Methods and Applications*, Cambridge University Press, Cambridge.

[5] de Jong, R. M., and T. Woutersen (2011): "Dynamic Time Series Binary Choice," 2011, Econometric Theory 27, 673-702.

[6] Durbin, James (1954). "Errors in variables," *Review of the International Statistical Institute,* 22: 23–32.

[7] Engle, R. F., (1984): "Wald, likelihood ratio, and Lagrange multiplier tests in econometrics," in: Z. Griliches & M. D. Intriligator (ed.), Handbook of Econometrics, edition 1, volume 2, chapter 13: 775-826. North-Holland, Amsterdam.

[8] Greene, W. (2012): *Econometric Analyses*, 7th edition, Prentice Hall, New York.

[9] Guggenberger, P. (2010): "The Impact of a Hausman Pretest on the Asymptotic Size of a Hypothesis Test", *Econometric Theory* 26:369-382.

[10] Hansen, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica,* 50: 1029-1054.

[11] Hahn, J., J. C. Ham, and H. R. Moon (2011): "The Hausman test and weak instruments", *Journal of Econometrics*, 160:289–299

[12] Hausman, J.A., (1978): "Specification Tests in Econometrics", *Econometrica*, 46, 1251-1272.

[13] Hausman, J. A. and T. Woutersen (2014a): "Estimating a Semi-Parametric Duration Model without Specifying Heterogeneity," *Journal of Econometrics*, 178: 114–131.

[14] Hausman, J. A. and T. Woutersen (2014b): "Estimating the Derivative Function and Counterfactuals in Duration Models with Heterogeneity," *Econometric Reviews*, 33: 472-496.

[15] Kim, E. H. , A. Morse, and L. Zingales (2006): "What Has Mattered to Economics since 1970," *Journal of Economic Perspectives*, 20(4): 189-202.

[16] Lehmann, E. L. and J. P. Romano (2005): *Testing Statistical Hypotheses*, Third Edition, Springer-Verlag, New York.

[17] Newey, W. K., and D. McFadden (1994): "Large Sample Estimation and Hypothesis Testing," in the *Handbook of Econometrics,* Vol. 4, ed. by R. F. Engle and D. MacFadden. North-Holland, Amsterdam.

[18] Newey, W. K. and K. D. West (1987): "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica,* 55: 703–708.

[19] Romano, J. P., A. M. Shaikh, and M. Wolf (2010): "Hypothesis Testing in Econometrics," *Annual Review of Economics,* 75-104.

[20] Ruud, Paul A. (2000): *An Introduction to Classical Econometric Theory.* New York: Oxford University Press. pp. 578–585.

[21] Sargan, J.D. (1958): "The estimation of economic relationships using instrumental variables," *Econometrica,* 26:393-415.

[22] Staiger, D., and Stock, J. (1997): "Instrumental Variables Regression with Weak Instruments," *Econometrica*, 65:557-586.

[23] Wooldridge, J. M. (2010): *Econometric Analysis of Cross Section and Panel Data,* 2nd edition, MIT Press, Cambridge MA.

[24] Woutersen, T. (2016): "Increasing the Power of Moment-based Tests," manuscript.

[25] Wu, De-Min (July 1973): "Alternative Tests of Independence between Stochastic Regressors and Disturbances". *Econometrica* 41:733–750.

[26] Zapata, H. O., and C. M. Caminita (2012): "The Diffusion of Hausman's Econometric Ideas," in *Advances in Econometrics: Essays in Honor of Jerry Hausman*, edited by B. H. Baltagi, R. C. Hill, W. K. Newey, H. L. White, Volume 29, Emerald Group Publishing, UK, 1-29.

Appendix 1: T-distribution in Example 1

Let ε be a vector with length N with ε_i, $i = 1, ..., N$, as its elements. Let Q be an N by 3 matrix with $\{1, X_i, Z_i\}$ as its rows, $i = 1, ..., N$. Let R be an N by 4 matrix with $\{1, X_i, Z_i, X_i Z_i\}$ as its rows, $i = 1, ..., N$.

Lemma A1: Let $Y_i = \alpha + \beta X_i + \varepsilon_i$ for $i = 1, ..., N$. Let $\varepsilon | X_i, Z_i, i = 1, ..., N \sim N(0, \sigma^2 I_N)$. Let the matrices $Q'Q$ and $R'R$ have full rank. Then $T_{Hausman} = \gamma_{LS} \frac{\sqrt{\sum_i v_i^2 (1 - \frac{\sum_i v_i^2}{\sum_i X_i^2})}}{\sqrt{\frac{1}{N-3} \sum_i \hat{u}_i^2}}$ and $T_{New} = \gamma_{LS} \frac{\sqrt{\sum_i v_i^2 (1 - \frac{\sum_i v_i^2}{\sum_i X_i^2})}}{\sqrt{\frac{1}{N-4} \sum_i \hat{\eta}_i^2}}$ have a T-distributions with $N-3$ and $N-4$ degrees of freedom respectively.

Proof: Let $M_Q = I_N - Q(Q'Q)^{-1}Q'$ and $M_R = I_N - R(R'R)^{-1}R'$. Note that $\sum_i \hat{u}_i^2 = \frac{\varepsilon' M_Q \varepsilon}{\sigma^2}$ has a χ^2-distribution with $N-3$ degrees of freedom and $\sum_i \hat{\eta}_i^2 = \frac{\varepsilon' M_R \varepsilon}{\sigma^2}$ has a χ^2-distribution with $N-4$ degrees of freedom. The next step is to show that the vector $\hat{u} = M_Q \varepsilon$ and γ_{LS} are independently distributed. Note that $\hat{u}$ and γ_{LS} are jointly normally distributed, so we only have to show that every element of $\hat{u}$ is uncorrelated with γ_{LS}. Consider

$$\begin{pmatrix} \alpha_{LS} \\ \beta_{LS} \\ \gamma_{LS} \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \\ 0 \end{pmatrix} + (Q'Q)^{-1}Q'\varepsilon.$$

Next consider the expectation of $(Q'Q)^{-1}Q'\varepsilon$, a 3 by 1 matrix, times $\hat{u}'$, a 1 by N matrix, i.e.

$$\begin{aligned} E\{(Q'Q)^{-1}Q'\varepsilon \cdot \hat{u}'\} &= E\{(Q'Q)^{-1}Q'\varepsilon \cdot \varepsilon' M_Q\} \\ &= \sigma^2 E\{(Q'Q)^{-1}Q' M_Q\} \\ &= \sigma^2 E\{(Q'Q)^{-1}Q' - (Q'Q)^{-1}Q'\} = 0. \end{aligned}$$

The next step is to show that the vector $\hat{\eta} = M_R \varepsilon$ and γ_{LS} are independently distributed. Note that $\hat{\eta}$ and γ_{LS} are jointly normally distributed, so we only have to show that every element of $\hat{\eta}$ is uncorrelated with γ_{LS}. Consider

$$\begin{pmatrix} \alpha_{LS} \\ \beta_{LS} \\ \gamma_{LS} \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \\ 0 \end{pmatrix} + (Q'Q)^{-1}Q'\varepsilon.$$

Next consider the expectation of the 3 by 1 matrix $(Q'Q)^{-1}Q'\varepsilon$ times the 1 by N matrix $\hat{\eta}'$, i.e.

$$E\{(Q'Q)^{-1}Q'\varepsilon \cdot \hat{\eta}'\} = E\{(Q'Q)^{-1}Q'\varepsilon \cdot \varepsilon' M_R\}$$
$$= \sigma^2 E\{(Q'Q)^{-1}Q' M_R\} = 0$$

since $Q'M_R = 0$ because all the regressors in Q are also in R. This completes the proof.

Appendix 2: Simulations Example 1

In the simulations we use sample sizes of 200, 400 and 800. Thus, the relevant critical values are the critical values of the T-distribution with 196, 197, 396, 397, 796, and 797 degrees of freedom. Usually, empirical researchers use critical values that have three digits (e.g. 1.96). The critical values for $N-3$ and $N-4$ are the same if one uses three digits or four digits. We use 7 digits, and then one can observe a slight (although basically irrelevant) change. The following table has the number of degrees of freedom in its first column, the critical values for the two-sided test at the 95% significance level as its second column, and the critical values for the two-sided test at the 99% significance level as its third column.

Table A1: Critical Values T-distribution

Degrees of Freedom	95% confidence, 2 sided test	99% confidence, 2 sided test
196	1.972141	2.600887
197	1.972079	2.601016
396	1.965973	2.588301
397	1.965957	2.588270
796	1.962945	2.582020
797	1.962945	2.582012

Both the proposed test and the Hausman test have T-distributions. Therefore, the size of the test is exact. Our simulations confirm that. That is, at the 95% confidence level we reject the truth about 5% of the time (see table A2), and at the 99% confidence level we reject the truth about 1% of the time (see table A3).

Table A2: 0.05 Rejection Frequencies:

N	δ	New Test	Hausman Test	Ratio
200	0	0.04974	0.04989	0.99699
400	0	0.05019	0.05011	1.00160
800	0	0.04978	0.04981	0.99939

Results based on 100,000 simulations.

Table A3: 0.01 Rejection Frequencies:

N	δ	New Test	Hausman Test	Ratio
200	0	0.01017	0.01007	1.00993
400	0	0.01012	0.01010	1.00198
800	0	0.01007	0.01009	0.99801

Results based on 100,000 simulations.

In about 1 in 10,000 simulations for $N = 200$, we had that $X_i = Z_i$ for all i. In that case, the least squares and two-stage least squares estimators are equal, and we did not reject the null. We never had this issue with $N = 400$ or $N = 800$.

Appendix 3: Simulations Example 2

Arguments similar to the ones used in appendix 1 yield that the Hausman test in the simulations has a T-distribution with $NT - 3$ degrees of freedom under the null of a common intercept, while the new test has a T-distribution with $NT-4$ degrees of freedom under the null. In the simulations we use sample sizes of 300, 600 and 900 individuals, and we observe everybody twice. We thus use the critical values of the T-distribution with 596, 597, 1196, 1197, 2396, and 2397 degrees of freedom. As the simulations below show, these values yielded the correct size in simulations.

Table A4: Critical Values T-distribution

Degrees of Freedom	95% confidence, 2 sided test	99% confidence, 2 sided test
596	1.963952	2.584104
597	1.963946	2.584090
1196	1.961949	2.579946
1197	1.961948	2.579943
2396	1.960955	2.577883
2397	1.960954	2.577882

Table A5: 0.05 Rejection Frequencies; H_0: Common Intercept

N	δ	New Test	Hausman Test	Ratio
300	0	0.05150	0.05150	1
600	0	0.05060	0.05060	1
1200	0	0.05300	0.05280	1.00379

Results based on 10,000 simulations.

Table A6: 0.01 Rejection Frequencies; H_0: Common Intercept

N	δ	New Test	Hausman Test	Ratio
300	0	0.00990	0.00990	1
600	0	0.00960	0.00960	1
1200	0	0.00960	0.00960	1

Results based on 10,000 simulations.

Table A7: 0.05 Rejection Frequencies; H_0: Random Effects

N	δ	New Test	Hausman Test	Ratio
300	0	0.05130	0.05140	0.99805
600	0	0.05040	0.05030	1.00199
1200	0	0.05290	0.05280	1.00189

Results based on 10,000 simulations.

Table A8: 0.01 Rejection Frequencies; H_0: Random Effects

N	δ	New Test	Hausman Test	Ratio
300	0	0.00990	0.00970	1.02061
600	0	0.00940	0.00940	1
1200	0	0.00970	0.00970	1

Results based on 10,000 simulations.

Table A9: 0.01 Rejection Frequencies; H_0: Common Intercept

N	δ	New Test	Hausman Test	Ratio
300	3/4	0.066100	0.0373000	1.77211
600	3/4	0.14820	0.0905000	1.63757
1200	3/4	0.33690	0.23950	1.40668
300	3/2	0.33080	0.07230	4.57538
600	3/2	0.67770	0.29750	2.27798
1200	3/2	0.96000	0.77080	1.24545

Results based on 10,000 simulations.

Table A10: 0.01 Rejection Frequencies; H_0: Random Effects

N	δ	New Test	Hausman Test	Ratio
300	3/4	0.066400	0.035900	1.84958
600	3/4	0.14850	0.086700	1.71280
1200	3/4	0.33680	0.234000	1.43931
300	3/2	0.33170	0.0627000	5.29027
600	3/2	0.67770	0.27290	2.48332
1200	3/2	0.95990	0.74850	1.28243

Results based on 10,000 simulations.

Also, an alternative to running the augmented regression in the main text is to take the difference between the least squares estimator, β_{LS}, and the least squares estimator with individual effects, $\beta_{\text{Individual Effects LS}}$. These estimators and their variance estimates are

$$\beta_{LS} = \frac{\sum_i \sum_t (X_{it} - \bar{X}) Y_{it}}{\sum_i \sum_t (X_{it} - \bar{X})^2} \text{ and } \widehat{Var(\beta_{LS})} = \frac{\frac{1}{NT-2} \sum_i \sum_t e_{it}^2}{\{\sum_i \sum_t (X_{it} - \bar{X})^2\}^2}$$

where $e_{it} = Y_{it} - \bar{Y} - \beta_{LS}(X_{it} - \bar{X})$, and

$$\beta_{\text{Individual Effects LS}} = \frac{\sum_i \sum_t (X_{it} - \bar{X}_i) Y_{it}}{\sum_i \sum_t (X_{it} - \bar{X}_i)^2} \text{ and } \widehat{Var(\beta_{\text{Individual Effects LS}})} = \frac{\frac{1}{NT-N-1} \sum_i \sum_t u_{it}^2}{\{\sum_i \sum_t (X_{it} - \bar{X}_i)^2\}^2}$$

where $u_{it} = Y_{it} - \bar{Y}_i - \beta_{\text{Individual Effects LS}}(X_{it} - \bar{X}_i)$. In the simulation $T = 2$ and $N = 300$, 600, or 1200.

Appendix 4: Proof of Theorem

(i) If H_0 is true then $\phi_i = E\{G^{-1}g_i(X_i,Y_i,Z_i) - H^{-1}h_i(X_i,Y_i,Z_i)|W\} = 0$ for all i so that $\Lambda \equiv \Omega - var(\frac{1}{\sqrt{N}}\sum_i \phi_i) = \Omega$. The result then follows from the properties of the χ^2-distribution (see, e.g., Lehmann and Romano (2005)).

(ii) Let A and B denote invertible matrices. Abadir and Magnus (2005, exercise and solution 12.16) show that if $A-B$ is positive definite then $\{B^{-1}-A^{-1}\}$ is positive definite as well. Thus, if $\Omega - \Lambda$ is positive definite then $\{\Lambda^{-1} - \Omega^{-1}\}$ is positive definite as well. We have that

$$\begin{aligned} T_{New} - T_{Hausman} &= (\hat{\beta}_1 - \hat{\beta}_2)'\{\hat{\Lambda}^{-1} - \hat{\Omega}^{-1}\}(\hat{\beta}_1 - \hat{\beta}_2) \\ &= (\hat{\beta}_1 - \hat{\beta}_2)'\{\Lambda^{-1} - \Omega^{-1}\}(\hat{\beta}_1 - \hat{\beta}_2) + o_p(1) \\ &> 0 \text{ with probability approaching one under } H_1. \end{aligned}$$

(iii) Note that

$$\begin{aligned} T_{New} &= (\hat{\beta}_1 - \hat{\beta}_2)'\hat{\Lambda}^{-1}(\hat{\beta}_1 - \hat{\beta}_2) \\ &= (\hat{\beta}_1 - \hat{\beta}_2)'(\frac{1}{\hat{\varphi}})\hat{\Omega}^{-1}(\hat{\beta}_1 - \hat{\beta}_2) \\ &= \frac{T_{Hausman}}{\varphi} + o_p(1) \end{aligned}$$

and the result follows.

(iv) Note that $\hat{\Omega} - \hat{\Lambda}$ is positive definite with probability approaching one and the result follows from (ii).