The impact of a Hausman pretest on the size of a hypothesis test: The panel data case

Patrik Guggenberger *
Department of Economics, UCSD, United States

A R T I C L E   I N F O
Article history:
Received 10 September 2008
Received in revised form 13 July 2009
Accepted 10 November 2009
Available online 17 November 2009

JEL classification:
C12
C23
C52

Keywords:
Fixed effects estimation
Hausman specification test
Panel data
Pretest
Random effects estimation
Size distortion

A B S T R A C T
The size properties of a two-stage test in a panel data model are investigated where in the first stage a Hausman (1978) specification test is used as a pretest of the random effects specification and in the second stage, a simple hypothesis about a component of the parameter vector is tested, using a t-statistic that is based on either the random effects or the fixed effects estimator depending on the outcome of the Hausman pretest. It is shown that the asymptotic size of the two-stage test equals 1 for empirically relevant specifications of the parameter space. The size distortion is caused mainly by the poor power properties of the pretest. Given these results, we recommend using a t-statistic based on the fixed effects estimator instead of the two-stage procedure.

1. Introduction

When deciding between inference based on the random effects or the fixed effects estimator in a panel data model, it is quite standard in applied work to first implement a Hausman (1978) pretest. If the Hausman pretest rejects the pretest null hypothesis that the random effects specification is correct, inference based on the fixed effects estimator is used in the second stage, otherwise inference based on the random effects estimator is used which has favorable power properties. For example, Blonigen (1997) justifies the use of random effects inference based on a Hausman pretest, while Hasting (2004) uses fixed effects inference as a result of the Hausman pretest rejecting the random effects specification. The Hausman pretest is a common tool, used in hundreds of applied papers and discussed in most leading textbooks in Econometrics, see e.g. Baltagi (2005, chapter 4.3), Greene (2008, chapter 9.5.4), Hsiao (2003, chapter 3.5), and Wooldridge (2002, chapter 10.7.3).

It is shown in this article that the asymptotic size of the resulting two-stage test equals 1 for empirically relevant specifications of the parameter space. An explicit formula for the asymptotic size of the two-stage test is derived. It shows that the asymptotic size depends on the degree of time variation in the regressors and also on the relative magnitude of the error variance to the variance of the individual specific effect. Our results explain how these two quantities impact the size of the two-stage test. The result that the two-stage test is size-distorted is related to the findings in Guggenberger (in press). In that article, it is shown that the corresponding two-stage test in the linear instrumental variables (IV) model has asymptotic size 1, where the Hausman pretest is used as a test of exogeneity of a regressor. As outlined in more detail below, the analysis of the panel data example is more complicated than the analysis of the IV example, because in the former case the asymptotic size depends on a higher dimensional nuisance parameter vector than in the latter case.

Based on the general theory developed in Andrews and Guggenberger (forthcoming-a), (AG forthcoming-a) from now on), we characterize sequences of nuisance parameters that lead to the highest null rejection probabilities of the two-stage test asymptotically. It is shown that under certain local deviations from the random effects specification, the Hausman pretest statistic converges to a noncentral chi-square distribution. The noncentrality parameter is small when the error variance is large relative to the variance of the individual specific effect or when the regressors are...
strongly positively correlated over time. In this situation, the Hausman pretest has low power against local deviations of the pretest null hypothesis and consequently, with high probability, inference based on the random effects estimator is performed in the second stage which leads to size distortion. However, it is also shown that the conditional size of the two-stage test, conditional on the Hausman pretest rejecting the pretest null hypothesis, exceeds the nominal size of the test. In Monte Carlo simulations, we document that the asymptotic size distortion is well reflected in finite samples.

Given the results in the article, if controlling the size of a testing procedure is an objective, the use of the two-stage procedure cannot be recommended. Its asymptotic size is severely distorted and the size distortion is well reflected in finite sample simulations. On the other hand, use of a t-statistic based on the fixed effects estimator has correct asymptotic size and performs well in finite samples. If the random effects specification is correct, inference based on the random effects estimator has correct size and has favorable power properties, but of course, leads to size distortion otherwise. Given the results in the article, the random effects specification should not be tested using a Hausman pretest.

It has been long known that pretests have an impact on the risk properties of estimators and the size properties of tests, see Giles and Giles (1993) and Judge and Bock (1978) for an excellent comprehensive survey of the pretest literature, and for more recent references Guggenberger (in press). If one were to consider the regression-based version of the Hausman test (see e.g. Wooldridge, 2002, p. 290), then the problem considered here could potentially be embedded into the previously studied problem of pretesting whether a regressor variable should be included in a linear regression model, see e.g. Andrews and Guggenberger (2009a) and Leeb and Pötscher (2005).

As documented further below, the specification tests proposed in Hausman (1978) are routinely used as pretests in applied work. Furthermore, variants of Hausman pretests are still being developed in the theoretical literature, see e.g. Hahn et al. (2009). Besides Guggenberger (in press), where the case of the linear IV model is studied and the present article, I am not aware of any other theoretical results in the literature regarding the negative impact of the Hausman pretest on the size properties of a two-stage test.

The remainder of the article is organized as follows. Section 2 describes the model, the objective, and defines the test statistics. In Section 3, the asymptotic size of a two-stage test in the panel model is derived where in the first stage a Hausman pretest describes the model, the objective, and defines the test statistics. Opened in the theoreticalliterature, see e.g. Baltagi (2005, eqn. 2.1).² Let \( Z_i = (1, X_i) \). By \( y_i, X_i, Z_i, u_i \) and \( v_i \) we denote the matrices (or vectors) with \( T \) rows given by \( y_{it}, X_{it}, Z_{it}, u_{it} \), and \( v_{it} \), respectively. The observed data are \( (y_i, X_i) \in \mathbb{R}^{T \times 2} \), \( i = 1, \ldots, N \). In vector form, the model reads

\[
y = \alpha_{it} + X \beta + u = Z(\alpha, \beta)^\prime u + v \in \mathbb{R}^{NT},
\]

where \( u' = (u_{11}, \ldots, u_{1T}, u_{21}, \ldots, u_{2T}, \ldots, u_{NT}, \ldots, u_{NT}) \) with observations stacked such that the slower index is over \( t \) and the faster index is over \( t \), and analogous notation is used for \( Y, X \), and \( Z \in \mathbb{R}^{T \times 2} \), where in the latter case, the row vectors \( Z_{it} \) are being stacked.

The data \( (X_i', \mu_i', v_i)^\prime \), \( i = 1, \ldots, N \) are assumed to be i.i.d. with distribution \( F \) and \( E(t, v_t) = t \). Assume \( E(Z_{it}X_{it}) = \varepsilon_{it} \), \( E(Z_{it}v_{it}) = 0 \), \( E(\mu_{it}v_{it}) = 0 \), and define \( \sigma_{it}^2 = E(Z_{it}v_{it}^2) \) and \( \sigma_{it}^2 = E(\mu_{it}v_{it}^2) \). To simplify notation, we do not index \( \sigma_{it}^2 \) and \( \sigma_{it}^2 \) by \( F \). Our asymptotic framework has \( N \to \infty \), but \( T \) fixed.

The object of interest is to test the null hypothesis

\[
H_0 : \beta = \beta_0 \quad \text{versus} \quad H_1 : \beta \neq \beta_0.
\]

One-sided alternatives could be studied using the same approach. One way to test (2) is to use a t-statistic \( \tilde{\beta}_NE \) based on the random effects estimator \( \hat{\beta}_NE \). Denote by \( \tilde{\sigma}^2_\mu \) and \( \tilde{\sigma}^2_\mu \) estimators for \( \sigma^2_\mu \) and \( \sigma^2_\mu \). Possible choices for \( \tilde{\sigma}^2_\mu \) and \( \tilde{\sigma}^2_\mu \), and for the estimator \( \tilde{\sigma}^2_\mu \) used below, are discussed in the Appendix. Define \( Z_{it} = I_N \otimes I_T \).

Then,

\[
\tilde{\beta}_NE = (X'RX)^{-1}X'Ry \quad \text{and}
\]

\[
\tilde{\beta}_NE = N^{1/2}\tilde{\beta}_NE - \beta_0 / (\tilde{V}_NE)^{1/2}, \quad \text{where}
\]

\[
\tilde{V}_NE = N\tilde{\sigma}^2_\mu (X'RX)^{-1}
\]

and

\[
R = Q_{0} + \sigma^2_\mu \tilde{\sigma}^2_\mu \tilde{\sigma}^2_\mu + \left( \frac{1}{NT} I_N I_N' \right),
\]

see Baltagi (2005, (2.30)). Alternatively, the test of (2) can be based on the fixed effects estimator \( \hat{\beta}_FE \)

\[
\tilde{\beta}_FE = (X'Q_{0}X)^{-1}X'Q_{0}y \quad \text{by using the t-statistic}
\]

\[
\tilde{\beta}_FE = N^{1/2}\tilde{\beta}_FE - \beta_0 / (\tilde{V}_FE)^{1/2}, \quad \text{where}
\]

\[
\tilde{V}_FE = N\tilde{\sigma}^2_\mu (X'Q_{0}X)^{-1}
\]

for an estimator \( \tilde{\sigma}^2_\mu \) of \( \sigma^2_\mu \) that may differ from \( \tilde{\sigma}^2_\mu \). Let \( \tilde{X}_t = T^{-1} \sum_{i=1}^T X_{it} \) be the time average of the regressor. Inference based on the t-statistic \( \tilde{\beta}_FE(\beta_0) \) is justified if \( \tilde{X} \) and \( \mu_i \) are uncorrelated but size-distorted otherwise. On the other hand, if \( \tilde{X} \) and \( \mu_i \) are uncorrelated, inference based on \( \tilde{\beta}_FE(\beta_0) \) provide power advantages over inference based on \( \tilde{\beta}_FE(\beta_0) \). Due to this trade-off between robustness and power, often in applied work, before testing (2), a Hausman (1978) pretest is undertaken. The pretest tests whether the pretest null hypothesis

\[
H_{P,0} : \text{Corr}(\mu_i, \tilde{X}_i) = 0
\]

is true. If the pretest rejects the pretest hypothesis, then, in the second stage, \( H_0 : \beta = \beta_0 \) is tested based on \( \tilde{\beta}_FE(\beta_0) \), the robust testing procedure when \( H_{P,0} \) is false. If the pretest does not reject (6), then in the second stage (2) is tested based on \( \tilde{\beta}_FE(\beta_0) \). Thus, denoting by

\[
H_N = N(\tilde{\beta}_FE - \hat{\beta}_FE)^2 \tilde{V}_FE - \tilde{V}_FE
\]

² Additional regressors \( \tilde{X}_t \) could be included into the model at the expense of more complicated notation. The asymptotic results of the article are identical if the intercept \( \alpha \) is not included into the model.
the Hausman statistic the resulting two-stage test statistic $T_N(\beta_0)$ is given by

$$T_N(\beta_0) = T_{RE}(\beta_0)H(H_T \leq \chi^2_{1,1-\epsilon}) + T_{FE}(\beta_0)H(H_T > \chi^2_{1,1-\epsilon}),$$

where $\epsilon$ denotes the nominal size of the pretest. The nominal size $\epsilon$ test rejects $H_0$ if

$$T_N(\beta_0) > z_{1-\epsilon/2}.\tag{9}$$

### 3. Asymptotic size of the test

The goal of this section is to explicitly calculate the asymptotic size of the test of $H_0: \beta = \beta_0$ in (9). By definition, the asymptotic size of the test equals

$$\text{AsySz}(\beta_0) = \lim_{N \to \infty} \sup_{\gamma \in \Gamma} P_{\beta_0,\gamma}(T_N(\beta_0) > z_{1-\epsilon/2}),$$

where $\gamma$ is the nuisance parameter vector described below, $\Gamma$ its parameter space, and $P_{\beta_0,\gamma}(\cdot)$ denotes probability when the true parameters are $(\beta_0, \gamma)$. The asymptotic size is an asymptotic approximation for the finite sample size.

Following AG (forthcoming-a), the parameter $\gamma$ is decomposed into three components $\gamma = (\gamma_1, \gamma_2, \gamma_3)$ according to the following criteria. The asymptotic distribution of the test statistic $T_N(\beta_0)$ is discontinuous in $\gamma_1$ in the sense that under sequences $\gamma_N = (\gamma_{N,1}, \gamma_{N,2}, \gamma_{N,3})$ that are such that $N^{1/2}/\gamma_{N,1} \to h_1$, the asymptotic distribution of $T_N(\beta_0)$ depends on $h_1$. The component $\gamma_2$ also affects the asymptotic distribution of $T_N(\beta_0)$, but unlike $\gamma_1$, it affects the distribution continuously, that is, under sequences that satisfy $\gamma_{N,2} \to h_2$, the asymptotic distribution of $T_N(\beta_0)$ only depends on $h_2$. Finally, $\gamma_3$ does not affect the limit distribution of $T_N(\beta_0)$. We will show below that for the test statistic $T_N(\beta_0)$ the appropriate decomposition is given by

$$\gamma_1 = \text{Corr}_{\tau}(\mu_i, \bar{X}_i), \quad \gamma_2 = (\gamma_{21}, \gamma_{22}), \quad \gamma_3 = (F, \alpha).$$

Note that $\gamma_{21} \in [0, 1]$. The component $\gamma_1$ measures the degree of failure of the pretest hypothesis (6). $\gamma_2$ measures the expected time variation in the regressor, and finally $\gamma_3$ is a function of the distribution of the individual specific effect and the error term $v_{it}$. The parameter space $\Gamma$ of $\gamma$ is given as

$$\Gamma = \{ (\gamma_1, \gamma_2, \gamma_3): \gamma_1 \in [-1, 1], \gamma_2 \in [0, 1], \gamma_3 \in (-\infty, \infty) \},$$

where $\gamma_1 \in \Gamma_1(\gamma_1, \gamma_2)$ for some $0 < k_1 < k_2 < 0 < k_3 < k_4$, and

$$\Gamma_1(\gamma_1, \gamma_2) = \{ (F, \alpha) : \alpha < R; E_{F}(X_{it}^2 = E_{F} \mu_i = E_{F}v_{it} = 0, E_{F}v_{it}^2 = \sigma^2, E_{F} \mu_i^2 = \sigma^2 \mu^2 \}$$

and $\text{Corr}_{\tau}(\mu_i, \bar{X}_i)$.

$$\gamma_2 = (\gamma_{21}, \gamma_{22})$$

and

$$\gamma_3 = (F, \alpha).$$

The result holds by the Liptarounov CLT for independent, mean zero, $L^{2+\delta}$-bounded random variables using the moment restrictions in (13) and some calculations. Under any sequence $\{\gamma_N\}$ with $|h_1| < \infty$, the following convergence results hold jointly.

$$T_{RE}(\beta_0) \to d \{ \xi_{RE,H} \}, \quad \xi_{RE,H} \sim \frac{\psi_{\mu,\beta_2} - \beta_2 \psi_{\mu,\beta_3}}{(1 - h_2^2)^{1/2}},$$

and

$$T_{FE}(\beta_0) \to d \{ \xi_{FE,H} \}, \quad \xi_{FE,H} \sim \frac{\psi_{\mu,\beta_1} - \beta_1 \psi_{\mu,\beta_2}}{(1 - h_1^2)^{1/2}}.$$
under \( \{y_{N,h}\} \), where \( h \) is the distribution of the random variable
\[
\xi_h = \{\xi_{RE,h} \mid (\xi_{RE,h} \leq x_{1,1-\varepsilon}) + \{\xi_{RE,h} \mid (\xi_{RE,h} > x_{1,1-\varepsilon})\}. \tag{21}
\]

The theorem that follows is a special case of Theorem 1(a) in AG (forthcoming-a). It provides a formula for \( \text{AsySz} \) that — in contrast to the one in (10) — can be used for explicit calculation. It shows that the “worst case” sequence of nuisance parameters, that is, the sequence that yields the highest asymptotic null rejection probability, is of the type \( \{y_{N,h} : N \geq 1\} \).

**Theorem 1 (AG (forthcoming-a)).** For the test of \( H_0 : \beta = \beta_0 \) in (9) with parameter space given by (13) we have \( \text{AsySz}(\beta_0) = \sup_{h \in [1,10]} \{1 - J_h(21 - \varepsilon/2)\} \), where \( H \) and \( J_h \) are defined in (15) and (20), respectively.

This theorem holds for both the Hausman pretest and the current example. In the current case, the situation is more complex and two separate parameters impact the asymptotic size of the two-stage test through \( \gamma_2 \).

**Fig. 1.** Asy max rej prob over \( h \) of sym test as function of \( h_{21} \) and \( h_{22} \).

In these cases, the pretest has poor power properties and the two-stage test frequently uses inference based on \( T_{RE}(\beta_0) \) in the second stage. But the test based on \( T_{RE}(\beta_0) \) tends to reject frequently not only for moderate failures but also for larger failures (6) which leads to size distortion of the two-stage test. The parameter \( h_{21} \) is close to 1 when there is little time variation in the regressor, i.e. in the extreme case where \( X_0 = X_0 \) for all \( s = 1, \ldots, T, h_{21} = 1 \). In the case where \( X_0, t = 1, \ldots, T \) are i.i.d., \( h_{21} = T^{-1/2} \); for example, \( h_{21} = .71 \) and \( h_{21} = .58 \) when \( T = 2 \) or 3, respectively. Note that if \( \xi_1 = .71 \), the simulations for Fig. 1 show that \( \text{AsySz}(\beta_0) \) is about 30% if \( \xi_2 \) is small. So, even in the case where the regressor \( X_0 \) is uncorrelated for different time indices \( t \), the two-stage test is extremely size-distorted.

For given values of \( \xi_1, \xi_2, \xi_3, \) and \( \xi_4 \), the asymptotic size of the two-stage test is non-increasing in the pretest nominal size \( \varepsilon_p \). As an extreme case, when \( \varepsilon_p = 1 \), the test rejects with probability 1 and in the second stage, fixed effects inference is used. As a result, the asymptotic size equals the nominal size \( \varepsilon \) in this case. Therefore, the “optimal choice” of the pretest nominal size from a size perspective is the pathological choice \( \varepsilon_p = 1 \). So, in order to alleviate the asymptotic size distortion of the two-stage test, one might consider implementing the test with a larger pretest nominal size \( \varepsilon_p \) than 1% or 5% that are typically used in applied work. There are several criticisms for that approach.

First, while the asymptotic size of the two-stage test decreases in \( \varepsilon_p \) for a given specification of \( \Gamma \), it is still distorted, and considerably so, when \( \xi_1 \) is close to 1 and \( \xi_2 \) is close to 0. Since these interval endpoints are unknown in practice, the asymptotic size might, therefore, be severely distorted even if one picks a large pretest nominal size \( \varepsilon_p \). Expanded on Fig. 1, Table 1 lists \( f(h_{21}, h_{22}) \) for \( \varepsilon = .05 \) and \( \varepsilon_p = .05, .25, .5, \) and .75 for a grid of values for \( h_{21} \) and \( h_{22} \). The table makes it clear that for given \( h_{21} \) and \( h_{22} \) the maximal asymptotic null rejection probability decreases with increasing \( \varepsilon_p \) but that for any pretest level considered here, there are choices of \( h_{21} \) and \( h_{22} \) for which the asymptotic null rejection probability of the test by far exceeds the nominal size \( \varepsilon = .05 \) of the test.

Second, the rationale for using the Hausman pretest is to improve on the power properties of a simple one-stage test that always uses fixed effects inference. As \( \varepsilon_p \) is increased in the two-stage test, the probability that fixed effects inference is being used

---

4 For each \( h \), the results are based on \( R = 30,000 \) random draws from the distribution of \( h \). When calculating the sup in (22), I consider \( h \) values in \([-2000,2000]\) using a grid with stepsize .01 on [0,1], stepsize .1 on [1,2], stepsize 1 on [1,10], stepsize 10 on [10,100], and stepsize 50 on [100,1000] and the analogous grid for negative \( h \) values.

5 Values of \( h_{22} \) close to zero are possible if \( X_t \) is negatively correlated across different values for \( t \), a case which is probably of lesser importance in applied work.
in the second stage increases and, therefore, the power advantage of the two-stage test diminishes, while the asymptotic size is still distorted. Even if one could size-correct the two-stage test (by appropriately increasing the critical value in the second stage) it is unclear how the size-corrected two-stage test would compare in terms of power relative to the simple one-stage test that always uses fixed effects inference.

Third, while the asymptotic size of the two-stage test decreases as \( \varepsilon \) increases, the probability of rejecting in the second stage conditional on not rejecting in the first stage seems unaffected by the value of \( \varepsilon \). That is, whenever the two-stage procedure uses random effects inference in the second stage in an attempt to improve the power properties of the two-stage test, there is an extreme danger of incorrectly rejecting the null.

To document this phenomenon, Table 2 reports asymptotic conditional rejection probabilities of the two-stage test, conditional on the Hausman pretest rejecting the pretest null hypothesis, \( R - C = R = P(\hat{\beta}_h > z_{1-\varepsilon/2}\mid \hat{\varepsilon}_T = 1) \), and conditional on the pretest not rejecting the pretest null hypothesis, \( R - C = NR = P(\hat{\beta}_h > z_{1-\varepsilon/2}\mid \hat{\varepsilon}_T = 0) \), when \( \varepsilon = .05, .05, .5 \), and \( h_1 = 15 \) for a grid of \( h_1 \) and \( h_2 \) values. Table 2 also reports asymptotic rejection probabilities of the Hausman pretest \( P(\hat{\varepsilon}_T \mid h_1, h_2 > \chi_{1,1-\varepsilon}^2) \). The results in Table 2 are based on \( R = 3,000,000 \) simulation repetitions.

In the top panel, \( \varepsilon = .05 \). Even though \( h_1 = 15 \) is quite large, \( P(\hat{\varepsilon}_T > \chi_{1,1-\varepsilon}^2) \) is quite small, especially when \( h_2 \) is close to one and/or \( h_2 \) is close to 0. This is consistent with the local power result of the Hausman pretest described in the previous paragraph because when \( h_2 = 1 \) is close to one and/or \( h_2 \) is close to 0 then the noncentrality parameter \( h_2^2(\hat{h}_T)^{-1} \) is close to 0. For example, when \( h_2 = .75 \) and \( h_2 = .1 \), then the Hausman pretest rejects in only 16.8% of the cases. However, in cases where the Hausman pretest does not reject—despite the fact that \( h_1 = 15 \)—the rejection probability in the second stage can be very high. This is because then in the second stage inference is based on \( T_{H0}(\hat{\beta}_0) \) which takes on relatively large values when \( h_1 \) is nonzero. For example, in the case \( h_2 = .75 \) and \( h_2 = 2 \), conditional on the Hausman pretest rejecting (which happens in 49.5% of the cases), the test rejects with probability 59.1% in the second stage. Perhaps more surprisingly, size distortion of the two-stage test is also caused by the two-stage test rejecting at high frequency conditional on the Hausman

<table>
<thead>
<tr>
<th>( \varepsilon )</th>
<th>( R - C - R )</th>
<th>( R - C - NR )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( h_1 )</td>
<td>( \varepsilon = .05 )</td>
<td>( \varepsilon = .05 )</td>
</tr>
<tr>
<td>.75</td>
<td>26.0</td>
<td>3.0</td>
</tr>
<tr>
<td>.8</td>
<td>31.0</td>
<td>3.5</td>
</tr>
<tr>
<td>.85</td>
<td>37.5</td>
<td>3.9</td>
</tr>
<tr>
<td>.9</td>
<td>46.6</td>
<td>5.1</td>
</tr>
<tr>
<td>.95</td>
<td>60.8</td>
<td>8.9</td>
</tr>
<tr>
<td>( \varepsilon = .5 )</td>
<td>8.2</td>
<td>3.9</td>
</tr>
<tr>
<td>.8</td>
<td>8.7</td>
<td>3.8</td>
</tr>
<tr>
<td>.85</td>
<td>9.2</td>
<td>3.8</td>
</tr>
<tr>
<td>.9</td>
<td>9.6</td>
<td>3.8</td>
</tr>
<tr>
<td>.95</td>
<td>9.9</td>
<td>3.8</td>
</tr>
</tbody>
</table>

Note that if we evaluate the limiting distribution \( \hat{\varepsilon}_T \) of \( H_0 \) at \( h_1 = 1 \) and \( h_2 = 0 \) the result is \( \hat{\varepsilon}_T^2 \) which is the squared limiting distribution of \( T_{H0}(\hat{\beta}_0) \), see (18).
that lead to a positive definite matrix $C_{p,q}$. Note that different values for $(p, q, \sigma^2)$ translate into values for $\gamma_1$, $\gamma_2$, and $\gamma_{22}$ through the relation $\gamma_1 = q(T + p(T^2 - T)^{-1/2})$, $\gamma_2 = (T^{-1} + p(1 - T^{-1}))^{1/2}$, and $\gamma_{22} = (T/\sigma^2)^{1/2}$. For example, when $T = 2$ the choices for $\sigma^2$ translate into the values $7.07, 4.47, 3.16, 1.41,$ and $6.3$ for $\gamma_{22}$. When $T = 10$ and $p = q = .9$, we have $\gamma_1 = .94$.

Finally, we choose nominal sizes $\varepsilon = \varepsilon_p = .05$. Small choices for the pretest nominal size $\varepsilon_p$ such as $5\%$ are common in applied work when applying Hausman pretests, both in panel data and linear IV applications. E.g. Deschênes (2006, p. 1757) states “A Hausman–Wu test does not reject the null hypothesis of exogeneity ($p$ value = 0.06). . . Given these results... we treat volume as exogenous hereafter.” and Bedard and Deschênes (2006, p. 189) state “... the Hausman (1978) test, testing the null hypothesis that the difference between TSLS and OLS coefficients is due only to sampling error, is rejected at the 5-percenter level.” Also see Owusu-Gyapong (1986). However, many times the pretest nominal size $\varepsilon_p$ is not even reported in applied articles. E.g. Blonigen (1997, p. 453) states “A Hausman test indicated that the random effects model estimates are consistent for these data, and thus I report only the more efficient random effects model estimates.” and Banerjee and Iyer (2005, p. 1205) state “A Hausman test does not reject the null hypothesis that the OLS and IV coefficients are equal.” As a further indication that small values of $\varepsilon_p$ are common in applied work, consider Bradford (2003, p. 1757) that states that the Hausman statistic “which is distributed as a chi-square with two degrees of freedom under the null is calculated at 1.46. This fails to reject the null at any reasonable level of significance. Consequently, these two variables are treated as exogenous regressors hereafter.” Note that the $p$ value in this case is .48. Thus, choices of $\varepsilon_p$ of that magnitude are considered unreasonable.

Table 3 reports a subset of the simulation results when $N = 100$, $T = 2$, 5, and $\sigma^2 = .1$ and 5. In Table 3(a), we list null rejection probabilities of the two-stage test, in Table 3(b) rejection probabilities of the Hausman pretest, and, for additional information, in Table 3(c) conditional rejection probabilities in the second stage conditional on the pretest not rejecting in the first stage. The simulations are based on $R = 30,000$ repetitions. Table 3(a) reveals that the two-stage test may severely overreject. At the nominal size of 5\%, the null rejection rates reach values close to 90\% for the parameter combinations considered here. When $q = 0$ (not shown in Table), i.e. when $\mu_i$ and the regressors $X_i$ are independent, the null rejection probabilities are relatively close to the nominal size of the test and do not exceed 10\% over all the parameter combinations considered here. In this case, inference based on both $T_{RE}(\beta_0)$ and $T_{RE}(\beta_0)$ is justified. However, when $q$ is nonzero and thus $\mu_i$ and the regressors $X_i$ are correlated, only inference based on $T_{RE}(\beta_0)$ is justified but inference based on $T_{RE}(\beta_0)$ is size-distorted. The simulations show that overrejection increases in $p$ and, most of the times, also in $\sigma^2$. Picking $p$ close to 1 and/or large enough values for $\sigma^2$, the null rejection probability of the two-stage test can be made arbitrarily close to 100\%. On the other hand, if $p$ and/or $\sigma^2$ are relatively small, the finite sample null rejection probabilities of the two-stage test are close to the nominal size of the test.

As Table 3(b) and (c) document, the main reason why the two-stage test overrejects in cases with large $p$ and/or $\sigma^2$ is because the Hausman pretest does not reject frequently enough (as documented in Table 3(b)) in cases where $q \neq 0$, but in these cases the resulting inference based on $T_{RE}(\beta_0)$ in the second stage leads to very frequent rejections because $T_{RE}(\beta_0)$ is not well approximated by a $\mathcal{N}(0, 1)$ distribution when $q \neq 0$. For example, when $N = 100$, $T = 2$, $p = q = .6$ and $\sigma^2 = 5$ the Hausman pretest only rejects in 45.2\% of the cases. The two-stage test rejects in 51.8\% of the cases for this constellation because in the second stage, inference based on $T_{RE}(\beta_0)$ leads to a 90\% rejection probability conditional on the pretest not rejecting. This is consistent with the theoretical results in (23) that imply that the Hausman pretest has little power when $p$ and/or $\sigma^2$ is small because then the noncentrality parameter of the chi-square distribution is small.

For parameter constellations where the two-stage test has good finite sample null rejection probabilities, the two-stage test virtually coincides with a one-stage test that always uses fixed effects inference. For example, when $N = 100$, $T = 5$, $p \in \{.6, .9\}$ and $\sigma^2 = .1$, Table 3(a) and (b) show that the null rejection probabilities are close to 5\% but that in these cases the Hausman pretest rejects in 100\% of the cases. In summary, when the two-stage test does not overreject it is close to a simple one-stage test based on $T_{RE}(\beta_0)$—if the test aims at improving power by using random effects inference in the second stage, a prize is paid in terms of severe size distortion.

The additional simulations reveal that there is no uniform effect of $N$ or $T$ on the rejection probabilities of the two-stage test even though a larger $T$ often reduces the overrejection.

---

### Table 3

<table>
<thead>
<tr>
<th>$\sigma^2$</th>
<th>$T = 2$</th>
<th>$T = 5$</th>
<th>$\sigma^2 = .1$</th>
<th>$T = 2$</th>
<th>$T = 5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p/q$</td>
<td>.9</td>
<td>.6</td>
<td>.3</td>
<td>.9</td>
<td>.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.9</td>
<td>87.7</td>
<td>85.7</td>
<td>37.6</td>
<td>58.1</td>
<td>79.7</td>
</tr>
<tr>
<td>.9</td>
<td>77.4</td>
<td>81.2</td>
<td>37.3</td>
<td>27.7</td>
<td>61.7</td>
</tr>
<tr>
<td>.6</td>
<td>51.8</td>
<td>31.9</td>
<td>–</td>
<td>7.3</td>
<td>29.6</td>
</tr>
<tr>
<td>.3</td>
<td>24.4</td>
<td>.7</td>
<td>27.8</td>
<td>5.0</td>
<td>12.6</td>
</tr>
</tbody>
</table>

(a) Finite sample null rejection prob's (in %) of two-stage test

(b) Finite sample rejection prob's (in %) of Hausman pretest

(c) Conditional rejection prob's (in %) conditional on pretest not rejecting

---

*The remaining results are available upon request. When calculating the estimators $\hat{\sigma^2_E}, \hat{\sigma^2_R},$ and $\hat{\sigma^2}$ in (28), the degrees of freedom correction $K = 2$ is used. In finite samples, the Hausman statistic in (7) may take on negative values. To prevent this, I use $\hat{\sigma^2}$ when calculating the random and fixed effects estimator, see Wooldridge (2002, p. 290). A “—” in Table 3 indicates that for this constellation the matrix $C_{p,q}$ is not positive definite, and a “NA” stands for “not applicable” because the conditional event never occurs.*
In contrast to the size-distorted two-stage procedure, the simple one-stage test that always uses the test statistic $T_{FE}$ ($\beta_0$) has very reliable null rejection probabilities. Over all the parameter combinations in (25), the null rejection probabilities of the one-stage test never exceeded 5.6%! On the other hand, the two-stage test rejected at rates up to 90% for these constellations. Therefore, if controlling the size is an objective, use of the two-stage procedure can not be recommended.

5. Conclusion

The article shows that if one uses a Hausman pretest to decide between random and fixed effects inference in a panel data context, the size of the resulting two-stage test is distorted. The size distortion is documented in finite samples through Monte Carlo simulation and is also theoretically proven asymptotically. The advice then is to refrain from the pretest. Consonant with this finding, Baltagi (2005, chapter 2.3.1) points out that “unfortunately … nonrejection of the Hausman pretest has been interpreted as an adoption of the random effects model.” Baltagi (2005) goes one step further and states that “a rejection of the Hausman pretest should not be interpreted automatically as an adoption of the fixed effects model!”. In fact, Chamberlain (1984) points out that the fixed effects model imposes testable restrictions on the parameters of the reduced form model that should be tested. Finally, Hausman and Taylor (1981) bridges the gap between the fixed and random effects specifications allowing for a model where some of the regressors are correlated with the individual specific effects. See Baltagi et al. (2003) for the finite sample performance of a Hausman and Taylor (1981) type pretest estimator.

Acknowledgement

The author would like to thank a co-editor, an associate editor, two referees, Iván Fernández-Val, Yixiao Sun, and Jeffrey Wooldridge for excellent comments, the NSF for support under grant number SES-0748922, the Alfred P. Sloan Foundation for a 2009 fellowship, and seminar and conference participants at various places.

Appendix. Possible choices for $\delta_0^2$, $\delta_\mu^2$, and $\delta_{\gamma}^2$

Following Wooldridge (2002, p. 260 and 271), one way to define the variance estimators is as follows. Let

$$\tilde{\beta}_{OLS} = (Z'Z)^{-1} Z'y$$

be the pooled OLS estimator of $(\alpha, \beta)'$ and

$$\tilde{u}_a = y_a - Z_a(\tilde{\beta}_{OLS}, \tilde{\alpha}_{OLS})'$$

the residuals from pooled OLS regression estimating $\mu_i = \mu + \tilde{u}_a$. Then, let

$$\delta_a^2 = (NT - K)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{u}_a^2,$$

$$\delta_\mu^2 = (NT(1 - 1)/2 - K)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{u}_\mu \tilde{u}_a,$$

$$\delta_{\gamma}^2 = \delta_a^2 - \delta_\mu^2,$$

for $K = 0$ or 2 depending on whether a degrees-of-freedom correction is desired. We can also estimate $\delta_\mu^2$ based on the fixed effects estimator $\hat{\beta}_{FE}$ in (4). Let $\tilde{y}_i = T^{-1} \sum_{t=1}^{T} \tilde{y}_i$ and let

$$\tilde{\tilde{u}}_a = (y_a - \tilde{y}_i) - (X_a - \tilde{X}_i) \hat{\beta}_{FE}$$

be the fixed effects residuals estimating $\tilde{u}_a$ and define the estimator

$$\delta_{\gamma}^2 = (N(T - 1) - K)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{\tilde{u}}_a^2.$$