Econometric Methods Final Project, May 2021

Tests To The Test: Comparing Two Cluster-Robust Hausman Tests Using Monte Carlo

Louis CHARLOT, Chloé MAS and Rachel PAYA
Paris School of Economics - M2 APE

Abstract

This article uses a Monte Carlo experiment to compare two types of cluster-robust Hausman test: one based on the auxiliary regression method following Woolridge (2010) and the other relying on the pairs cluster bootstrap procedure for the variance detailed in Cameron and Miller (2015). Our findings indicate that the first type of cluster-robust Hausman test is better to avoid type II errors while the second is better to avoid type I errors. Since the failure to correctly reject the null is the primary concern in the Hausman test, we conclude that the first test should be preferred.

1 Introduction

The structure of clustered data requires special attention when making statistical inference. Indeed, the within-cluster error correlation implies that the standard deviation of the estimated coefficients will be underestimated when using estimators relying on the assumption of uncorrelated error. Cluster-robust standard errors are thus needed to ensure an accurate estimation of the precision of the estimators when dealing with clustered data. One particularly widespread type of clustered data is panel data, which is the focus of this article.

The random-effect (RE) model and the fixed-effect (FE) model are two popular models for meta-analysis and panel data. If these two models may be similar, they make very different assumptions about the data. Following Wooldridge (2010)[1] we denote y_{it} and x_{it} the outcome and the explanatory variables respectively, with i being the cluster and t being time (respectively g and i in Cameron and Miller [2]). The error term of the model ϵ_{it} can be decomposed into two components: a transitory component u_{it}, and a permanent component c_i that captures the unobserved individual effects. We thus have $\epsilon_{it} = u_{it} + c_i$. As the FE model does not demand uncorrelated individual effects over time, the sufficient condition for consistency can be written as $E(u_{it}|\mathbf{x}_i, c_i) = 0$ (**FE.1**). By contrast, the RE model is less flexible as it does require that individual effects are uncorrelated over time. In this model, the sufficient condition is $Cov(c_i; x_{it}) = 0$ (**RE.1.b**) and $E(u_{it}|\mathbf{x}_i, c_i) = 0$ (**RE.1.a** $\Leftrightarrow$ **FE.1**). Selecting the model that best suits the data between the two is necessary to have an exact estimation. By testing the uncorrelated effects hypothesis, Hausman (1978)[3] proposes a decision rule for choosing between the two models based on differences between those. The test aims to evaluate the

consistency of a RE estimator compared to a FE estimator that is consistent but less efficient. Indeed, under H_0 the conditional mean of the disturbances given the regressors is zero and the two estimators converge toward the same limit. As both estimators are consistent but the RE estimator is more efficient, we must choose the RE model. Under the alternative hypothesis, random and fixed effect estimators have a statistically significant difference. Then, the RE estimator is not consistent but the FE is and we must conclude against the RE model.

Nonetheless, Woolridge (2010)[1] emphasises two limits that may alter the decision rule. First, the Hausman test assumes that the FE estimator is at least a consistent estimator, while it is not the case when the sufficient condition does not hold (ie. when $E(u_{it}|\mathbf{x}_i, c_i) \neq 0$). The second issue is related to the validity of an underlying assumption on RE estimator that needs to hold in the implementation of the Hausman test. Indeed, for the classic Hausman test, we assume that the conditional variances are constants and the conditional covariances are null (ie. $E(u_i u_i'|\mathbf{x}_i c_i) = \sigma_u^2 I_T$) and the unobserved individual effect is homoskedastic (ie. $E(c_i^2|\mathbf{x}_i) = \sigma_c^2$)(**RE.3**). These two assumptions allow to assume that the RE estimator is fully efficient under the null hypothesis. Many papers (Wooldridge (2010)[1], Cameron and Miller (2015)[2]) insist on how strong this assumption is and thus how likely it is to fail. Using cluster-robust standards errors instead of default standard errors into a cluster-robust Hausman test allows to relax the assumption.

Cameron and Miller (2015)[2] cite two methods to implement a cluster-robust Hausman test. The first method follows Wooldridge (2010)[1] and consists in using a pooled OLS regression that contains the mean over t, denoted $\bar{w}_i$, of a subcompnent of x_{it} that varies within cluster, denoted w_{it}, and performing a Wald test on the coefficient associated to $\bar{w}_i$, based on cluster-robust estimate of the variance matrix. In the second method, they propose to perform a Hausman test where the variance is estimated thanks to the pairs cluster bootstrap method. The authors note that there is no study that compares the two versions of the cluster-robust Hausman test. This article contributes to the literature by filling this gap and comparing these two types of cluster-robust Hausman test using a Monte Carlo simulation. Monte Carlo is a method of estimation that holds on repeated random sampling. The use of Monte Carlo simulations allows to estimate the uncertainty in the model and helps us in the comparisons of the two tests.

This comparison allows us to highlight that the first type of cluster-robust Hausman test is better to avoid type II errors while the second is better to avoid type I errors. Yet, both methods always correctly reject the null hypothesis when the number of clusters is large enough ($G = 100$), while the second method has a significantly higher probability of failing to correctly reject the null when there are few clusters ($G = 10$). In the comparison of random and fixed effects estimators, failing to reject the absence of correlation when there is correlation is the main issue so we conclude in favour of the *Method 1* Hausman test.

The article is structured as follows: section 2 introduces the two cluster-robust Hausman tests, section 3 explains the data generating process underlying the Monte Carlo study, section 4

presents the results and section 5 concludes on the implications of the findings.

2 Building the methods of cluster-robust Hausman test for Fixed and Random Effects

We will now present the two different methods to carry out these cluster-robust Hausman tests, before comparing them in the next sections.

2.1 The first method (Wooldridge, 2010)

The first method to compute a Robust Hausman test for Random Effects (RE) against Fixed Effects (FE) lies on the regression-based approach for the Hausman statistic developed by Wooldridge (2010, chapter 10.7.3.)[1] and inspired from Mundlak (1978)[4]. This alternative formulation of the Hausman statistic can be robustified to violations of Assumption {**RE.3.**} $\stackrel{\text{def}}{=} \{E(u_i u_i' \mid \mathbf{x}_i, c_i) = \sigma_u^2 I_T \text{ and } E(c_i^2 \mid \mathbf{x}_i) = \sigma_c^2\}$.

Following Wooldridge (2010)[1], we depart from the model:

$$y_{it} = x_{it}\beta + c_i + u_{it} = z_i\gamma + w_{it}\delta + c_i + u_{it}$$

with c_i being the unobserved individual effect, u_{it} the idiosyncratic error and x_{it} the observed explanatory variables that can vary across i *and* t (w_{it}) or i (z_i which includes also an intercept). We assume that there are no aggregate time effects (d_t) in this model, because they could lead to problems of singularity when calculating the asymptotic variance matrix of the difference between β_{FE} and β_{RE}: this would lead to complications when deriving the Hausman test (see Problem 10.17 of Wooldridge (2010) for more details on this point). We implicitly assume here the validity of assumption {**RE.1.a**} $\stackrel{\text{def}}{=}$ {FE.1.} $\stackrel{\text{def}}{=} \{E(u_{it} \mid \mathbf{x}_i, c_i) = 0\}$, otherwise both $\hat{\delta}_{FE}$ and $\hat{\delta}_{RE}$ are inconsistent.

As mentioned in the introduction, because FE is consistent while RE is inconsistent when c_i and x_{it} are correlated, unveiling whether they are correlated is of paramount importance for the choice between RE and FE . Therefore, a statistically significant difference between the two estimators $\hat{\delta}_{FE}$ and $\hat{\delta}_{RE}$ is interpreted as evidence against the assumption {**RE.1.b**} $\stackrel{\text{def}}{=}$ $\{E(c_i \mid \mathbf{x}_i) = E(c_i) = 0\}$ (c_i and x_{it} are not correlated). This constitutes the central idea behind the Classical Hausman Test, but also behind the regression-based approach to compute a Hausman statistic developed by Mundlak (1978)[4] and exposed by Wooldridge (2010). More precisely, the aim of the regression-based approach is to derive another form of the usual Hausman statistic:

$$H = (\hat{\delta}_{FE} - \hat{\delta}_{RE})' \left[\hat{V}(\hat{\delta}_{FE} - \hat{\delta}_{RE})\right]^{-1} (\hat{\delta}_{FE} - \hat{\delta}_{RE})$$

which follows a χ^2_M distribution (where $M = dim(w_{it})$) under the null hypothesis H_0: {RE.1.b}.

When the usual RE assumptions are valid (including {**RE.3.**}), we have that $V(\delta_{FE} - \delta_{RE}) = V(\delta_{FE}) - V(\delta_{RE})$. Note that this equality doesn't hold anymore when {**RE.3.**} is not valid, and leads to a matrix $V(\delta_{FE}) - V(\delta_{RE})$ which might not meet the positive-definite requirement of the Hausman test. However, following Wooldridge (2010, p.331), we begin by assuming that all the usual RE assumptions are valid in a first step, and then suppress assumption {**RE.3.**} in a second step. Therefore, we can write:

$$H = (\hat{\delta}_{FE} - \hat{\delta}_{RE})' \left[\hat{V}(\hat{\delta}_{FE}) - \hat{V}(\hat{\delta}_{RE})\right]^{-1} (\hat{\delta}_{FE} - \hat{\delta}_{RE})$$

For this expression to be truly defined, $V(\hat{\delta}_{FE}) - V(\hat{\delta}_{RE})$ is required to be positive definite. By noting $\ddot{\mathrm{w}}_{it} = w_{it} - \bar{w}_i$, $\breve{\mathrm{w}}_{it} = w_{it} - \lambda\bar{w}_i$ the quasi-time demeaned time-varying regressors, $\tilde{w}_{it}$ the population residuals from the pooled regression of $\breve{\mathrm{w}}_{it}$ on z_i, and $N = dim(y_i)$ we can write:

$$V(\delta_{FE}) = \frac{\sigma_u^2}{N}\left[E(\ddot{\mathrm{W}}_i'\ddot{\mathrm{W}}_i)\right]^{-1} \text{ and } V(\delta_{RE}) = \frac{\sigma_u^2}{N}\left[E(\tilde{W}_i'\tilde{W}_i)\right]^{-1}$$

Therefore, as requiring $V(\hat{\delta}_{FE}) - V(\hat{\delta}_{RE})$ to be positive definite is equivalent to requiring $[V(\hat{\delta}_{RE})]^{-1} - [V(\hat{\delta}_{FE})]^{-1}$ to be positive definite, we have that H is truly defined if $E(\tilde{W}_i'\tilde{W}_i) - E(\ddot{\mathrm{W}}_i'\ddot{\mathrm{W}}_i)$ is positive definite. It can be shown (see Wooldridge (2010, p.330) that $\tilde{w}_{it} = \ddot{\mathrm{w}}_{it} + (1-\lambda)(\bar{w}_i - \bar{w}_i^*)$, where $\bar{w}_i^*$ is the linear projection of $\bar{w}_i$ on z_i. This last result can not only be used to show that $E(\tilde{W}_i'\tilde{W}_i) - E(\ddot{\mathrm{W}}_i'\ddot{\mathrm{W}}_i)$ is positive definite (and thus H is truly defined) if $\bar{w}_i$ and z_i are not perfectly collinear, but it also shows that the Hausman test is a test of:

$$H_0 : E\left[(\bar{w}_i - \bar{w}_i^*)'c_i\right] = 0$$

We can notice that if z_i is composed only of the intercept (with no other time constant variables), we have $\bar{w}_i^* = 0$ and thus the null hypothesis in this case is $H_0 : \{E\left[\bar{w}_i'c_i\right] = 0\} \Leftrightarrow \{Cov\left(\bar{w}_i, c_i\right) = 0\}$. In the case where there are time-constant variables in z_i, we first suppress the correlation between $\bar{w}_i$ and z_i by computing $\bar{w}_i - \bar{w}_i^*$, and then test its correlation with c_i. This shows that time-constant controls z_i have a determinant role in the success of a RE analysis.

This new formulation of the null hypothesis H_0 of the Hausman test leads to the regression-based approach of Wooldridge (2010), which lies on the RE assumption by Mundlak (1978) that c_i can be written as $c_i = \psi + \bar{w}_i\xi + a_i$, with $E(a_i) = 0$ and $Corr(a_i, (\mathbf{w}_i, z_i)) = 0$. We can then write y_{it} as:

$$y_{it} = x_{it}\beta + \bar{w}_i\xi + a_i + u_{it}$$

where ψ can be included in x_{it} (as z_i includes an intercept and is part of x_{it}). The estimation of ξ by pooled OLS is very interesting: as we have that $\hat{\xi} = \hat{\delta}_{Between} - \hat{\delta}_{FE}$, the estimation of ξ explicitly compares the Between and FE estimates of δ. It is even more interesting when we know that Hausman and Taylor (1981)[5] have shown that comparing the Between and FE estimates of δ is the same as comparing $\hat{\delta}_{FE}$ and $\hat{\delta}_{RE}$, which is the central idea of the Hausman test for RE against FE. Lying on this, Mundlak (1978)[4] suggested a new equivalent formulation of H_0: he proposed testing $H_0 : \{\xi = 0\}$ to determine if the heterogeneity c_i was

correlated with the time averages $\bar{w}_i$.

This method can be robustified to violations of assumption {**RE.3.**}. More precisely, to implement this method in a context where assumption {**RE.3.**} doesn't hold, we used a pooled OLS with a fully robust variance matrix estimator to test for $H_0 : \{\xi = 0\}$. Indeed, as mentioned by Wooldridge (2010, p.333)[1], it is feasible in this case to obtain a Wald statistic that is fully robust to violations of assumption {**RE.3.**}, which was the initial objective of this approach.

In our implementation, we used *vcovHC* function of *sandwich*[6] R package for heteroskedasticity-consistent estimation of the covariance matrix of the coefficient estimates in the regression model. We also took some inspiration from the formalism of the source functions of *plm*[7] R package, even though we decided to code our functions ourselves as *plm* functions implement other forms of robust Hausman estimation.

2.2 The second method based on bootstrap

The second method for the robust Hausman test proposed by Cameron and Miller (2015)[2] is more straightforward as it simply consists in estimating the variance V of the difference between the fixed-effect estimator and the random-effect estimator $(\hat{\delta}_{FE} - \hat{\delta}_{RE})$ using pairs cluster bootstrap, and plugging this $\hat{V}$ in the Hausman test statistic formula.

Boostrapping is a non-parametrical approach to statistical inference. It relies on sampling with replacement, which consists in drawing observations one by one randomly from a dataset and returning them to the dataset after each draw so that the resampled dataset can contain several times the same observation. The number of boostrap iterations determines the number of new samples produced, from which the sampling distribution of a statistic can be obtained. Brought to prominence by Efron (1979)[8], the bootstrapping procedure has the advantage of being simple to apply for statistics with sampling distributions that are difficult to derive even asymptotically.

To implement this robust Hausman test, we start by estimating the coefficients using the fixed-effect and the random-effect models using the *plm* function in the R package of the same name, obtaining $\hat{\delta}_{0;FE}$ and $\hat{\delta}_{0;RE}$. Then we use a bootstrap, which consists in repeating B times the following resampling procedure and estimation. The G clusters are resampled with replacement G times to obtain a new sample of the same dimension as the original one and with the same numbers of clusters but in which the original clusters can appear more than once. The clusters from the new sample are thus renamed so that they each have a unique identifier code. The fixed-effect and random-effect models coefficients are then estimated from each sample, again using the *plm* function, and the estimates of each model are stored in a matrix of dimension B×k, where k is the number of coefficients. The random-effect model estimates a coefficient for the time-constant variables while the fixed-effect model only identifies coefficients of time-varying variables, so the time-constant intercept of the random-effect model must be removed to compare the fixed-effect and the random-effect models coefficients. Once

we have the bootsrapped coefficients, we can compute the pairs cluster bootstrap variance. As we relax **RE3** assumption, we cannot use the formula from Wooldridge (2010, p.331)[1] $\hat{V}(\hat{\delta}_{FE}-\hat{\delta}_{RE}) = \hat{V}(\hat{\delta}_{FE})-\hat{V}(\hat{\delta}_{RE})$. For each value of b in $(1, B)$, we must compute the difference in coefficients. We note $\hat{\delta}_{Diff;b} = \hat{\delta}_{FE;b}-\hat{\delta}_{RE;b}$ this difference and subtract its mean value $\bar{\hat{\delta}}_{Diff}$. That allows to compute the pairs cluster bootstrapped variance of $\hat{\delta}_{Diff}$ using the formula in Cameron and Miller (2015, p.328):

$$\hat{V}_{clust;boot}(\hat{\delta}_{Diff}) = \frac{1}{B-1}\sum_{b=1}^{B}\left(\hat{\delta}_{Diff;b} - \bar{\hat{\delta}}_{Diff}\right)\left(\hat{\delta}_{Diff;b} - \bar{\hat{\delta}}_{Diff}\right)'$$

Where $\bar{\hat{\delta}}_{Diff} = \frac{1}{B}\sum_{b=1}^{B}\hat{\delta}_{Diff;b}$. To do so we first compute the 1×k vectors $\bar{\hat{\delta}}_{FE} - \bar{\hat{\delta}}_{RE}$, which we subtract from each row of the corresponding matrix of bootstrapped coefficients. Once both matrices have been 'demeaned', for each of them we multiply each row by its transpose and store the results in a list of B matrices of dimension k×k. We sum all the components within each list and divide the result by (B−1) to obtain $\hat{V}_{clust;boot}(\hat{\delta}_{FE} - \hat{\delta}_{RE})$. We can then readily compute the associated Hausman test statistic:

$$T = (\hat{\delta}_{0;FE} - \hat{\delta}_{0;RE})'(\hat{V}_{clust;boot}(\hat{\delta}_{FE} - \hat{\delta}_{RE}))^{-1}(\hat{\delta}_{0;FE} - \hat{\delta}_{0;RE})$$

which follows a Chi-square distribution with k degrees of freedom. Note that, from Cameron and Miller (2015) [2] we know that we do not need to use bootstrapped estimated coefficients outside of the variance in the Hausman test as the issue only comes from the variance.

3 Data generating process and Monte Carlo simulation

In this part, we will implement a Monte Carlo experiment to compare the performances of *Method 1* (the regression-based robust Hausman test described in Section 2.1) and *Method 2* (the bootstrap Hausman test described in Section 2.2). To do so, we must choose an appropriate data generating process, which is detailed below.

3.1 Data Generating process

The choice of the underlying data generating process (DGP) is important as it has to respect some particular assumptions and relax others. Indeed, as previously mentioned, with these two Robust Hausman methods, we can relax assumption {**RE.3.**}. Therefore, our data doesn't need to respect the very particular form required by this assumption. On the contrary, our DGP must respect assumption {**RE.1.a**}. To respect it, we give to the different u_{it} independent normal distributions $N(0, 1)$. In addition, we decide to take only time-varying variables x_{it}, so

that $x_{it} = w_{it}$. To summarize these first points, we can write our DGP as follows:

$$y_{it} = x_{it}\delta + c_i + u_{it}$$

Given these first steps, we can now focus on the most important point, which is the respect (or not) of assumption $\{\textbf{RE.1.b}\} \overset{\text{def}}{=} \{c_i \text{ and } x_{it} \text{ are not correlated}\}$. Indeed, we want to implement two different DGPs: a first one (*DGP 1*) that respects this assumption and should therefore lead to Hausman tests that don't reject the RE validity, and a second one (*DGP 2*) that doesn't respect this assumption and should therefore lead to a Hausman tests that reject the RE validity. Therefore, for *DGP 1*, we will take $c_i = c_i^{Normal}$, with c_i^{Normal} that follow independent normal distributions $N(0, 1)$ for each i. On the contrary, for *DGP 2*, we implement a correlation between c_i and the x_{it} by taking $c_i = \phi \cdot c_i^{Normal} + \kappa \cdot \bar{w}_i$ (with $\kappa \neq 0$ and $\bar{w}_i$ being the mean on i of time-varying x_{it} as mentioned in Section 2.1). In our simulations, we have chosen $\phi = 0.01$ and $\kappa = 0.99$ to implement a strong correlation, and $\delta = 5$.

3.2 Monte Carlo simulations

As we are interested in the distribution of the results with the two Robust Hausman statistics, we carry out a Monte Carlo analysis. More precisely, rather than comparing the Hausman statistics and the associated p-values of the two methods for a single draw of *DGP 1* and *DGP 2*, we run $num = 50$ Monte Carlo iterations to get the distribution of the results for 50 different draws of *DGP 1* and *DGP 2*. We then not only take the means of the Hausman statistics and associated p-values, but we also look at the frequency of rejection of the null hypothesis $H_0 : \{RE.1.b.\}$ at the 5% and 1% level. This last point allows to understand with which frequency we falsely reject or accept H_0, which is very important for a Hausman test. Indeed, a false acceptance would make us choose a RE estimator while there is correlation, and lead us to inconsistent estimates. On the contrary, a false rejection would make us choose a FE analysis, instead of the more efficient RE analysis.

4 Results and comparisons

In order to properly compare the two tests, we must compare both the probability that the tests reject the true null hypothesis, that is the probability of committing a type I error, which is 1 minus the size of the test, and the probability that the tests fail to reject the null when it is false, which is 1 minus the power of the test and defined as the probability of a type II error. With the *Method 1* and over 50 Monte Carlo iterations, the cluster-robust Hausman test rejects the true null at the 5% significance level 18% of the time when the number of clusters equals 10 and 14% of the time when there are 100 clusters (with the total number of observations set to 100 and 1000 respectively). At the 1% significance level, it makes a type I error 12% of the time and 4% of the time with 10 and 100 clusters respectively. On the other hand, the Hausman test of *Method 2* never rejects the true null when the number of clusters is set to

10 and rejects it only 8% of the time at the 5% significance level and 2% of the time at the 1% significance level when there are 100 clusters. These results highlight that the Hausman test of *Method 2* is better to avoid falsely rejecting the null, even if on average both tests do not reject the null when it is true. Regarding the probability of a failure to reject the null when it is false, the *Method 1* Hausman test seems to perform better than the test of *Method 2*. Indeed, with 10 clusters it rejects the false null on average and 84% of the time at the 5% significance level (76% of the time at the 1% level) while the *Method 2* Hausman test fails to reject it on average even at the 10% significance level and only rejects it 58% of the time at the 5% significance level. With 100 clusters, however, both tests correctly reject the null 100% of the time. To conclude on which test to use, it must be emphasised that the failure to correctly reject the null hypothesis is much more problematic in the situation under analysis than the incorrect rejection of the true null as the first case would lead to an inconsistent estimation while the second case would simply mean that the estimator used (FE rather than RE) is less efficient. Hence, the cluster-robust Hausman test of *Method 1* outperforms the *Method 2* test as the latter under rejects the false null when the number of clusters is small.

5 Conclusion

In this paper, we use Cameron and Miller (2015)[2] proposal to implement two versions of a cluster-robust Hausman test that allow to overcome the strict assumptions required by the traditional Hausman test: constant conditional variances and null conditional covariances. The first method follows Woolridge (2010)[1] and is based on an auxiliary regression while the second one, detailed by Cameron and Miller (2015)[2], relies on a pairs cluster bootstrap procedure to compute the variance of the difference between the FE and RE estimators. If these two methods have already been implemented, Cameron and Miller (2015)[2] pointed to the absence of comparison between them. We contribute to the literature by implementing these methods using the same data generating process and Monte Carlo simulations. Our main finding is that the Hausman test of *Method 1* under-performs relative to the test of *Method 2* when the null must not be rejected, but outperforms for the correct rejection of the null, especially when the number of clusters is low. We thus conclude in favour of the *Method 1* cluster-robust Hausman test as correctly rejecting the null is our main concern.

References

[1] Wooldridge, J. (2010). Econometric Analysis of Cross Section and Panel Data. Cambridge, Massachusetts; London, England: The MIT Press. Retrieved May 3, 2021, from http://www.jstor.org/stable/j.ctt5hhcfr.

[2] Cameron, A.C., and Miller, D.L. (2015) "Practitioner's Guide to Cluster- Robust Inference" The Journal of Human Resources, 50(2): 317-372. doi: 10.3368/jhr.50.2.317.

[3] Jerry A Hausman. 'Specification tests in econometrics'. In: *Econometrica: Journal of the econometric society* (1978), pp. 1251–1271.

[4] Mundlak, Y. (1978) "On the Pooling of Time Series and Cross Section Data" Econometrica, 46(1): 69-85. doi:10.2307/1913646.

[5] Hausman, Jerry A. Taylor, William E., (1981) "Panel data and unobservable individual effects" Journal of Econometrics,16(1): 155-155.

[6] Zeileis A, Köll S, Graham N (2020). "Various Versatile Variances: An Object-Oriented Implementation of Clustered Covariances in R." Journal of Statistical Software, 95(1), 1–36. doi: 10.18637/jss.v095.i01.

[7] Croissant Y, Millo G (2008). "Panel Data Econometrics in R: The plm Package." Journal of Statistical Software, 27(2), 1–43. doi: 10.18637/jss.v027.i02.

[8] Efron, B. (1979) "Bootstrap Methods: Another Look at the Jackknife." The Annals of Statistics, 7(1): 1-26. https://doi.org/10.1214/aos/1176344552.

Appendix

Results of the Monte Carlo Analysis

Table 1: Mean of p-values p_{Mean}, frequency of p-values inferior to 0.05 $freq(p < 0.05)$, frequency of p-values inferior to 0.01 $freq(p < 0.01)$ and mean of Cluster-Robust Hausman statistic H_{Mean} over 50 Monte Carlo iterations, with $B = 399$ Bootstrap iterations, with 10 time periods and $G = 10$ and 100 clusters for {RE.1.b.} assumption respected and not respected

Case	{RE.1.b.} and G=10	NOT {RE.1.b.} and G=10	{RE.1.b.} and G=100	NOT {RE.1.b.} and G=100
Method 1:				
p_{Mean}	0.3955996	0.04652651	0.4664159	1.732134e-10
$freq(p < 0.05)$	0.18	0.84	0.14	1
$freq(p < 0.01)$	0.12	0.76	0.04	1
H_{Mean}	2.753435	20.36881	1.453387	91.66104
Method 2:				
p_{Mean}	0.7279315	0.1065866	0.4930527	1.640855e-05
$freq(p < 0.05)$	0	0.58	0.08	1
$freq(p < 0.01)$	0	0.32	0.02	1
H_{Mean}	0.2181224	5.08248	1.086238	33.11566