

Örebro University  
Örebro University School of Business  
Master's program "Applied Statistics"  
Advanced level thesis I  
Supervisor: Prof. Panagiotis Mantalos  
Examiner: Sune Karlsson  
Autumn 2014

# The Accuracy of the Hausman Test in Panel Data: a Monte Carlo Study

---

Teodora Sheytanova  
90/04/04

## Abstract

The accuracy of the Hausman test is an important issue in panel data analysis. A procedure for estimating the properties of the test, when dealing with specific data, is suggested and implemented. Based on simulation that mimics the original data, the size and power of Hausman test is obtained.

The procedure is applied for different methods of estimating the panel data model with random effects: Swamy and Arora (1972), Amemiya (1971) and Nerlove (1971). Also, three types of critical values of the Hausman statistics distribution are used, where possible: asymptotical and Bootstrap (based on simulation and bootstrapping) critical values as well as Monte Carlo (based on pure simulation) critical values for estimating the small sample properties of Hausman test.

The simulation mimics the original data as close as possible in order to make inferences specifically for the data at hand, but controls the correlation between one of the variables and the individual-specific component in the panel data model.

The results indicate that Hausman test over-rejects the null hypothesis if performed based on its asymptotical critical values, when Swamy and Arora and Amemiya methods are used for estimating the random effects model. The Nerlove method of estimation leads to extreme under-rejection of the null-hypothesis. The Bootstrap critical values are more appropriate. With the example data used, the chosen bootstrap procedure and the specific number of bootstrap samples, the bootstrap size tends to follow the upper limit of the confidence interval of the nominal size, although sometimes it passes the limit line and a slight over-rejection is observed. The simulations show that the use of Monte Carlo critical values leads to an actual size very close to the nominal.

The power of Hausman test proved to be considerably low at least when a constant term is used in the modelling.

Keywords: Hausman Test, Panel Data, Random Effects, Fixed Effects, Monte Carlo, Bootstrap.

## Contents

1. Introduction .....	4
1.1. Background .....	4
1.2. Statement of the problem and purpose of the research.....	4
1.3. Organisation of the thesis.....	5
2. Theoretical basis .....	5
2.1. Panel data .....	5
2.2. Estimation methods.....	6
2.2.1. Pooled model .....	7
2.2.2. Fixed effects model.....	8
2.2.3. Random effects model.....	8
2.3. Hausman test for endogeneity – general case .....	10
2.4. Hausman test for differentiating between fixed effects model and random effects model	11
2.5. Simulation and resampling methods .....	12
2.5.1. Monte Carlo simulations.....	12
2.5.2. Bootstrap.....	12
3. Methodology: Design of the Monte Carlo simulation .....	13
3.1. Defining a set of data generated processes (DGP). .....	13
3.2. Generating $r$ independent replications.....	16
3.3. Generating bootstrap samples in order to obtain Hausman test critical values.....	16
3.4. Obtaining Hausman test critical values through the use of a pure Monte Carlo simulation	19
3.5. Performing Hausman test in each replication .....	19
3.6. Estimating the size and power.....	19
3.7. Note .....	20
4. Implementation: Reporting the Monte Carlo experiment .....	21
4.1. Data.....	21
4.2. Technical details.....	21
4.3. The model of the data generated process.....	22
4.4. Obtaining Hausman statistics critical values trough bootstrapping .....	23
4.5. Obtaining Hausman statistics critical values through pure Monte Carlo simulation .....	24
4.6. Results.....	24
4.6.1. Size and power of Hausman test (Amemiya method for estimating random effects model)	25
4.6.2. Size and power of Hausman test (Nerlove method for estimating random effects model)	33

4.6.3. Size and power of Hausman test (Swamy and Arora method for estimating random effects model) .....	40
4.7. Discussion.....	43
4.8. Recommendations .....	44
5. Summary and Conclusion .....	45
5.1. Summary .....	45
5.2. Conclusion.....	45
6. Bibliography .....	46
Appendix .....	48
6.1. Fixed effects estimation methods.....	48
6.1.1. Within-group method .....	48
6.1.2. Least squares dummy variable (LSDV) method .....	49
6.2. Algorithm .....	49

# 1. Introduction

## 1.1. Background

When performing a statistical hypothesis test an issue that must be considered is the accuracy of the test. There are two properties that define the accuracy of a hypothesis test: its size and power. The size is the probability of rejecting the null hypothesis, when it is the correct one and in social sciences tests are usually run at significance level 5%, which guarantees that if the null hypothesis is correct and a number of tests are made based on different samples of the same population, in 95% of the cases the null hypothesis won't be rejected. The power represents the probability of correctly rejecting the null hypothesis. Values of the power of 80% or above are considered "good" when corresponding to size of 5% (Cohen, 1988).

An issue arises of whether setting the significance level to a particular value will actually result in a risk of making a 1<sup>st</sup> type error of the same value. Based on a particular sample it might turn out that the test has bigger (or smaller) size than what is set initially. It is reasonable to conduct a research for finding the true value of the size of the performed test as well as of its power.

The focus on this thesis is Hausman test, used for choosing between models in panel data studies. Hausman test examines the presence of endogeneity in the panel model. The use of panel data gives considerable advantages over only cross-sectional or time series data, but the specification of the model to be used is of great importance for obtaining consistent results. One of the tests used to determine an appropriate model is Hausman test, which specifies whether fixed or random effects panel model should be used. As one of the most used tests in panel data analysis, the study of its properties should represent a great interest.

However, not many publications have been made in this sense. Many research papers like Wong (1996), Herwartz and Neumann (2007), Bole and Rebec (2013) focus on bootstrapping Hausman test in order to improve its finite sample properties and find the true distribution of the test, but there are less research efforts done on the subject of Hausman test's properties in specific. Jeong & Yoon (2007) explore the effect of instrumental variables on the performance of Hausman test by simulating data.

This thesis illustrates the estimation of Hausman test's size and power procedure, which can be implemented for a particular real data study. By applying the methods used in this thesis, one can examine the accuracy of Hausman test in a particular case for a specific research and not rely on general studies based on data, which might not fit their individual instance.

## 1.2. Statement of the problem and purpose of the research

The thesis illustrates a procedure that can be used for obtaining the actual size and power of the Hausman test in a particular study, for a specific sample. The research has been performed on real data and the methodology, which has been used, can be applied for different panel data cases.

The problem consists of obtaining information about the accuracy of Hausman test by controlling the presence of endogeneity in a panel data model, based on data of the investment demand, market value and value of the stock of plant and equipment for 10 large companies in a period of 20 years. This data is often used to illustrate panel data estimation methods: Greene (2008), Grunfeld and Griliches (1960), Boot and de Wit (1960). The model analyses the Gross investment as a function of the Market values and the Value of the stock of plant and equipment.

Monte Carlo simulation is used to compute robust critical values by generating data under the null hypothesis.

The main purpose of the research is to find out how reliable Hausman test is and its accuracy when applied for the panel data model of investment demand. The goal of this thesis is estimating the actual size and power of Hausman test and by doing so deriving and illustrating a procedure that can be used as a step-by-step guide for estimating the size and power of the Hausman test for any panel data.

### 1.3. Organisation of the thesis

The thesis is organized as follows. The second sections outline the theoretical basis. It details about panel data in general, models and methods of estimation. Section 3 describes the methodology of the research. It is described in details, but can be applied for different panel data cases in general. In section 4 the specific parameters of the research and simulation are presented with detailed information about the used data and other information necessary for replicating the results. Conclusions are drawn in Section 5.

## 2. Theoretical basis

### 2.1. Panel data

Panel data in general is obtained through a longitudinal study, when the same entities (e.g. individuals, companies, countries) are observed in time. The values of the variables of interest are registered for several time periods or at several time points for each individual. Thus, the panel dataset consists of both time series and cross-sectional data. Practice shows that panel data has an extensive use in biological and social sciences (Frees, 2004). There are considerable advantages of using panel data as opposed to using only time series or only cross-sectional data. They are extensively addressed by Frees (2004). The additional information that the panel data provides allows for more accurate estimations. The panel data estimation methods require less assumptions and are often less problematic than simpler methods. They combine the values of using both cross-sectional data and time series data and add further benefit in terms of problem-solving.

One advantage of using panel data is the use of individual-specific components in the models. For example a linear regression model of  $k$  factors can be expressed in the following way:

$$y_{it} = \beta_0 + \alpha_i + \beta_1 x_{1,it} + \beta_2 x_{2,it} + \dots + \beta_k x_{k,it} + \varepsilon_{it} \quad i = 1, \dots, n; t = 1, \dots, T. \quad (1)$$

where  $\alpha_i$  is specific for each individual. A model such as the above allows for the managing of the heterogeneity across individuals. The inclusion of this parameter in the model can explain correlation between the observations in time, which is not caused by dynamic tendencies. The individual specific component can be fixed for each individual or it can be random (and be treated like a random variable). This defines the existence of two major panel data models called fixed effects model and random effects model.

The individual-specific component explains part of the heterogeneity in the data, which reduces the unexplained variability and thus the mean squared error. This makes the estimates obtained from panel data models that use individual-specific components more efficient than the ones from models that don't include such parameter.

Panel data can also deal with the problem of omitted variable bias if those variables are time-invariant. Let  $\alpha$  be a vector with length  $N = n \cdot t$  and elements  $\alpha_i$ . Because of the perfect collinearity between the time-invariant omitted variable(s) and  $\alpha$  in models like the fixed effects, one can consider that this variable(s) has/have been incorporated in the individual-specific component. Thus, it is possible to deal with bias in some cases.

Another advantage is in terms of time series analysis and is expressed in the fact that panel data doesn't require very long series. In the classical time series analysis some methods require series of at least 30 observations and that can be a drawback for two reasons: one is the availability of data for so many consecutive time periods and the second is that sometimes it is unreasonable to use the same model for describing data in a very long period of time. In panel data the model can be more easily inferred by making observations on the series for all the individuals. By finding what is common among the individuals, one can construct a model accurately without having to rely on very long series. The available data across individuals compensates for the shorter series.

A benefit of the panel data over cross section analysis is that a model can be constructed for evaluating the impact that some time-varying variables (the values of which also vary across individuals) have on some dependent variable. The additional data over time increases the precision of the estimations.

As seen above, there are certain benefits in using panel data analysis. However, it also presents some drawback in terms of gathering data. Panel data is often connected with the continued burdening of a permanent group of respondents. This can considerably increase the non-response levels. Problems also occur with the control and traceability of the sampled individuals. This is a price for the benefit of being able to effectively monitor net changes in time.

## 2.2. Estimation methods

There are several estimation methods in panel data. The most general and frequently used panel data models are discussed below: fixed effects model and random effects model. The pooled model that does not use panel information has also been described for comparison reasons. Statistical hypothesis testing must be done in order to determine the appropriate

model for the available data. The pooled model would give inconsistent estimates if used when the fixed effect should have been, but it must be used in case the fixed effect model is inappropriate. The random effects model is more efficient than the fixed effects, when it is correct, but inconsistent if used inappropriately. When used appropriately the random effects model gives the best linear unbiased estimates (BLUE). The fixed effects model gives consistent results for the estimates.

### 2.2.1. Pooled model

The pooled model does not differ from the common regression equation. It regards each observation as unrelated to the others ignoring panels and time. No panel information is used. A pooled model can be expressed as:

$$y_{it} = \beta_0 + \beta_1 x_{1,it} + \beta_2 x_{2,it} + \dots + \beta_k x_{k,it} + \varepsilon_{it}. \quad (2)$$

A pooled model is used under the assumption that the individuals behave in the same way, where there is homoscedasticity and no autocorrelation. Only then OLS can be used for obtaining efficient estimates from the model in equation (2). The assumptions for the pooled model are the same as for the simple regression model as described by Greene (2012):

- 1) The model is correct:

$$E(\varepsilon_{it}) = 0.$$

- 2) There is no perfect collinearity:

$$\text{rank}(\mathbf{X}) = \text{rank}(\mathbf{X}'\mathbf{X}) = K,$$

where  $\mathbf{X}$  is the factor matrix with  $k$  columns and  $N = nt$  rows.

- 3) Exogeneity:

$$E(\varepsilon_{it} | \mathbf{X}) = 0; \text{Cor}(\varepsilon_{it}, \mathbf{X}) = 0,$$

- 4) Homoscedasticity:

$$\text{Var}(\varepsilon_{it} | \mathbf{X}) = E(\varepsilon_{it}^2 | \mathbf{X}) = \sigma^2$$

- 5) No cross section or time series correlation:

$$\text{Cov}(\varepsilon_{it}, \varepsilon_{js} | \mathbf{X}) = E(\varepsilon_{it} \varepsilon_{js} | \mathbf{X}) = 0 \quad i \neq j; t \neq s$$

- 6) Normal distribution of the disturbances  $\varepsilon_{it}$ .

These assumptions are also valid for the panel data models. Under the assumptions the parameter estimates are unbiased and consistent. However in panel data studies it is likely to come across autocorrelation of the disturbances within individuals in which case the fifth assumption is not met. This would lead to biased estimates of the standard errors. They will be underestimated leading to over-estimated t-statistics. The error must be adjusted and one way to do so is by using clustered standard errors.

Since the pooled model is not so different from the simple linear regression model, it doesn't encompass all the benefits and advantages of panel data mentioned in the previous section. This model is more restrictive compared to fixed effects or random effects models.

However it should be used, when the fixed effect is not appropriate. If it is used when the fixed effects should have been, then the estimates of the pooled OLS will be inconsistent.

### 2.2.2. Fixed effects model

One of the advantages of using panel data as mention in Section 2.1. is that models like the fixed effects model can deal with the unobserved heterogeneity. The fixed effects model for  $k$  factors can be expressed in the following way:

$$y_{it} = \alpha_i + \beta_1 x_{1,it} + \beta_2 x_{2,it} + \dots + \beta_k x_{k,it} + \varepsilon_{it}. \quad (3)$$

There is no constant term in the fixed effects model. Instead of the constant term  $\beta_0$  in pooled model (2), now we have an individual-specific component  $\alpha_i$  that determines a unique intercept for each individual. However, the slopes (the  $\beta$  parameters) are the same for all individuals.

The assumptions that are valid for the fixed effects model are as follows:

- 1) The model is correct:

$$E(\varepsilon_{it}) = 0.$$

- 2) Full rank:

$$\text{rank}(\mathbf{X}) = \text{rank}(\mathbf{X}'\mathbf{X}) = K;$$

- 3) Exogeneity:

$$E(\varepsilon_{it} | \mathbf{x}_i, \alpha_i) = 0,$$

but there is no assumption that  $E(\alpha_i | \mathbf{x}_i) = E(\alpha_i) = 0$ ;

- 4) Homoscedasticity:

$$E(\varepsilon_{it}^2 | \mathbf{x}_i, \alpha_i) = \sigma_u^2;$$

- 5) No cross section or time series correlation:

$$\text{Cov}(\varepsilon_{it}, \varepsilon_{js} | \mathbf{X}) = E(\varepsilon_{it} \varepsilon_{js} | \mathbf{X}) = 0 \quad i \neq j; t \neq s$$

- 6) Normal distribution of the disturbances  $\varepsilon_{it}$ .

Two methods for computing the estimates of the fixed effects model are presented in the Appendix: within-groups method and least squares dummy variable method (LSDV).

### 2.2.3. Random effects model

In the random effects model the individual-specific component  $\alpha$  is not treated as a parameter and it is not being estimated. Instead, it is considered as a random variable with mean  $\mu$  and variance  $\sigma_\alpha^2$ . The random effects model can thus be written as:

$$y_{it} = \mu + \beta_1 x_{1,it} + \beta_2 x_{2,it} + \dots + \beta_k x_{k,it} + (\alpha_i - \mu) + \varepsilon_{it}, \quad (4)$$

where  $\mu$  is the average individual effect. Let  $u_{it} = (\alpha_i - \mu) + \varepsilon_{it}$  and (4) can be rewritten as:

$$y_{it} = \mu + \beta_1 x_{1,it} + \beta_2 x_{2,it} + \dots + \beta_k x_{k,it} + u_{it}, \quad (5)$$

The assumptions for the random effects model are as follows:

1) The model is correct:

$$E(u_{it}) = E((\alpha_i - \mu) + \varepsilon_{it}) = E(\alpha_i - \mu) + E(\varepsilon_{it}) = 0 + E(\varepsilon_{it}) = 0$$

2) Full rank:  $rank(\mathbf{X}) = rank(\mathbf{X}'\mathbf{X}) = K$ ;

3) Exogeneity:

$$E(u_{it} | \mathbf{x}_i, \alpha_i) = 0; E(\alpha_i - \mu | \mathbf{x}_i) = E(\alpha_i - \mu) = 0;$$

$$Cov(u_{it}, \mathbf{x}_{it}) = Cov(\alpha_i, \mathbf{x}_{it}) + Cov(\varepsilon_{it}, \mathbf{x}_{it}) = 0;$$

4) Homoscedasticity:

$$E(u_{it}^2 | \mathbf{x}_i, \alpha_i) = \sigma_u^2; E(\alpha_i^2 | \mathbf{x}_i) = \sigma_\alpha^2;$$

5) Normal distribution of the disturbances  $u_{it}$ .

Because of the more specific error term special attention must be paid on some of the conditions. The estimates of the random effects model are consistent only if assumptions 1) and 3) are satisfied. However, the individual-specific component  $\alpha$  might be correlated with the independent variables, which means that the compliance of the exogeneity condition must be verified. If it turns out that there is correlation between the error term  $u_{it}$  and the factors used in the model, then either pooled or fixed effects models must be used.

The OLS estimators of the random effects model are inefficient, because the condition of serial independence is not met.

$$Cov(u_{it}, u_{is}) = \sigma_\alpha^2 \neq 0$$

To avoid inefficiency GLS method of estimation must be used. Let  $\theta = 1 - \frac{\sigma_\varepsilon}{\sigma}$ , where  $(\sigma')^2 = T\sigma_\alpha^2 + \sigma_\varepsilon^2$ ; and  $\mu^* = (1 - \theta)\mu$ . Then the following differences are computed:

$$y_{it}^* = y_{it} - \theta \bar{y}_i; \quad x_{l,it}^* = x_{l,it} - \theta \bar{x}_{l,i}. \quad l = 1, \dots, k$$

Next, OLS method is applied to the equation:

$$y_{it}^* = \mu^* + \beta_1 x_{1,it}^* + \beta_2 x_{2,it}^* + \dots + \beta_k x_{k,it}^* + u_{it}^*, \quad (6)$$

where  $u_{it}^*$  meets the assumptions of the OLS method.

If  $\sigma_\alpha^2$  and  $\sigma_\varepsilon^2$  are known, the estimates of the random effects model can be computed from the following formula:

$$\hat{\beta}_l^{RE} = \frac{\sum \sum (x_{l,it}^* - \bar{x}_{l,i}^*) (y_{l,it}^* - \bar{y}_{l,i}^*)}{\sum \sum (x_{l,it}^* - \bar{x}_{l,i}^*)^2}. \quad (7)$$

However, the variances of the error term and the individual effects are unknown. If the error terms  $\varepsilon_{it}$  and the component  $\alpha_i$  are known (and hence  $u_{it}$ ), then one would use the following estimates for  $(\sigma')^2$ ,  $\sigma_\alpha^2$  and  $\sigma_\varepsilon^2$ :

$$(\sigma')^2 = \frac{T}{N} \sum \bar{u}_i^2; \quad (8)$$

$$\sigma_\varepsilon^2 = \frac{1}{N(T-1)} \sum \sum (u_{it} - \bar{u}_i)^2 = \frac{1}{N(T-1)} \sum \sum (\varepsilon_{it} - \bar{\varepsilon}_i)^2; \quad (9)$$

$$\sigma_{\alpha}^2 = \frac{1}{N-1} \sum (\alpha_i - \mu)^2 \quad (10)$$

Because  $\varepsilon_{it}$  and  $\alpha_i$  are not known, the above estimations cannot be computed. This is why various estimation methods have been suggested for obtaining  $\theta$ . Wallace and Hussain's method (1969) substitutes  $u_{it}$  with the error terms obtained from the pooled model:  $\hat{u}_{it}^P$  to obtain estimates for  $(\sigma')^2$  and  $\sigma_{\varepsilon}^2$ . Amemiya's method (Amemiya, 1971) suggests that the error terms from the within method  $\hat{\varepsilon}_{it}^W$  should be used instead of  $\varepsilon_{it}$  and  $\hat{u}_{it}^W = \hat{\varepsilon}_{it}^W + (\hat{\alpha}_i^W + \bar{\alpha}_i^W)$  should be used instead of  $u_{it}$ , where  $\hat{\alpha}_i^W$  is the estimated individual-specific component from the within method, in order to obtain estimates for  $(\sigma')^2$  and  $\sigma_{\varepsilon}^2$ . Nerlove's method (Nerlove, 1971) uses  $\hat{\varepsilon}_{it}^W$  and  $\hat{\alpha}_i^W$  instead of  $\varepsilon_{it}$  and  $\alpha_i$  in order to obtain estimates for  $\sigma_{\varepsilon}^2$  and  $\sigma_{\alpha}^2$ . The most frequently used is Swamy and Arora's method (Swamy and Arora, 1972) of random effects model estimation. It uses  $\hat{\varepsilon}_{it}^W$  instead of  $\varepsilon_{it}$  in order to estimate  $\sigma_{\varepsilon}^2$  and for estimating  $(\sigma')^2$  it uses the residuals from the "between" regression:  $\hat{u}_{it}^B$ . The between regression has the following form:

$$\bar{y}_i = \alpha_i + \beta_1 \bar{x}_{1,i} + \beta_2 \bar{x}_{2,i} + \dots + \beta_k \bar{x}_{k,i} + \bar{u}_i \quad (11)$$

The Maximum Likelihood method (ML) uses one of the above methods to estimate the random effects model and obtain estimates of the random effects error. Then, it uses them to compute a new  $\theta$ . This procedure is iterative.

A drawback in the estimation of the random effects model is the possibility of obtaining negative estimate of the variance for the individual-specific component (Magazzini and Calzolari, 2010). This would usually be the case, if the assumption for homoscedasticity  $E(u_{it}^2 | \mathbf{x}_i, \alpha_i) = \sigma_u^2$  is not met. Only Nerlove's method (Nerlove, 1971) explicitly estimates  $\sigma_{\alpha}^2$  by squaring the variations of  $\hat{\alpha}_i^W$  around its mean and thus is the only method that guarantees a positive estimate for  $\sigma_{\alpha}^2$ . The other methods derive  $\sigma_{\alpha}^2$  from  $\sigma_{\varepsilon}^2$  and  $(\sigma')^2$ , which can lead to a negative estimate. In such a case  $\theta$  can be set to 1, which would transform the random effects model into a fixed effects model.

When the random effects model has been used appropriately its estimates are efficient.

### 2.3. Hausman test for endogeneity – general case

There is a group of tests named after Hausman, which are used in model selection and compare the estimators of the tested models. Hausman test can be used if under the null hypothesis one of the compared models gives consistent and efficient results and the other – consistent, but inefficient, and at the same time under the alternative hypothesis the first model has to give inconsistent results and the second – consistent.

The general form of Hausman test statistic is:

$$H = (\hat{\beta}^I - \hat{\beta}^{II})' [Var(\hat{\beta}^I) - Var(\hat{\beta}^{II})]^{-1} (\hat{\beta}^I - \hat{\beta}^{II}), \quad (12)$$

and under the null hypothesis it is  $\chi^2(k)$  distributed, where  $k$  is the number of parameters.

Hausman test is often used when choosing between OLS and 2SLS methods for estimating a linear regression. 2SLS method incorporates instrumental variables in the model and is used to deal with endogeneity.

Hausman test is also useful for in panel data, when comparing the estimates of the fixed and random effects models.

#### 2.4. Hausman test for differentiating between fixed effects model and random effects model

The choice of model in panel data must be based on information about the individual-specific components and the exogeneity of the independent variables. Three hypotheses tests are used for choosing the correct model. One of them is for testing whether fixed or random effects model is appropriate, by identifying the presence of endogeneity in the explanatory variables: Hausman test. This section focuses entirely on discussing Hausman test as it is the topic of this work.

Table 1: Properties of the random and fixed effects models estimators.

Model	Random effects model used	Fixed effects model used
<b>Correct hypothesis</b> <b><math>H_0: Cov(\alpha_i, x_{it}) = 0</math></b> <b>Exogeneity</b>	Consistent Efficient	Consistent Inefficient
<b><math>H_1: Cov(\alpha_i, x_{it}) \neq 0</math></b> <b>Endogeneity</b>	Inconsistent	Consistent Possibly Efficient

As already mentioned in section 2.2., when used appropriately the random effects model gives the best linear unbiased estimates (BLUE). They are consistent, efficient and unbiased. However if there is correlation between the error term of the random effects model and the independent variables, its estimates would be inconsistent and thus fixed effects model would be preferred over the random effects model. The individual-specific component  $\alpha$  might be correlated with the independent variables in the random effects model, if there are omitted variables, to which the fixed effect model is robust. The fixed effects model estimates are always consistent, but they are inefficient compared to the random effects model estimates. Those properties of the panel data models estimates directs the researcher to Hausman test. The formulization of Hausman test and the steps for its implementation are described below.

- 1) Defining the null and alternative hypotheses:

**$H_0$ :** The appropriate model is Random effects. There is no correlation between the error term and the independent variables in the panel data model.

$$Cov(\alpha_i, x_{it}) = 0$$

**$H_1$ :** The appropriate model is Fixed effects. The correlation between the error term and the independent variables in the panel data model is statistically significant.

$$Cov(\alpha_i, x_{it}) \neq 0$$

- 2) A probability of first type error is chosen. For example  $\alpha = 0.05$ .
- 3) Hausman statistic is calculated from the formula:

$$H = (\hat{\beta}^{RE} - \hat{\beta}^{FE})' [Var(\hat{\beta}^{RE}) - Var(\hat{\beta}^{FE})]^{-1} (\hat{\beta}^{RE} - \hat{\beta}^{FE}),$$

where  $\hat{\beta}^{RE}$  and  $\hat{\beta}^{FE}$  are the vectors of coefficient estimates for the random and fixed effects model respectively. This statistic is  $\chi^2(k)$  distributed under the null hypothesis. The degrees of freedom  $k$  equal the number of factors.

- 4) The statistic, computed above is compared with the critical values for the  $\chi^2$  distribution for  $k$  degrees of freedom. The null hypothesis is rejected if the Hausman statistic is bigger than it's critical value.

## 2.5. Simulation and resampling methods

### 2.5.1. Monte Carlo simulations

Monte Carlo simulation can be particularly useful in estimating the properties of statistical hypothesis tests. For estimating the size data must be generated through simulations in such a way that it would satisfy the null hypothesis. The hypothesis testing is conducted and it is noted whether the null hypothesis has been rejected (a wrong decision) or not (a correct decision). The same procedure is repeated in  $r$  replications and in each replication the generated data is unique. The size is obtained by calculating the proportion of the wrong decisions.

The estimation of the power is similar. The difference is in the conditions used to generate the data. This time it must satisfy the alternative hypothesis. A correct decision would be to reject the null hypothesis. The estimation of power is computed by calculating the proportion of the correct decisions.

It must be noted that in this thesis work real data is used and its main aim is to derive results that are true for the particular data. Thus, the use of Monte Carlo simulation has to be done in a way that alters the original data in a minimal way. To conduct Hausman test within each Monte Carlo replication the random effects and fixed effects models are computed, but by adjusting the dependent variable  $y$  in such a way that would produce error terms that satisfy either the null or alternative hypothesis (for estimating the power and size respectively).

Monte Carlo simulation is also used for computing critical values for the Hausman statistic in order to estimate the small sample properties of the test.

### 2.5.2. Bootstrap

The bootstrap resampling method is used in the analysis of size and power of the Hausman test for the estimation of critical values that come from the true distribution of the Hausman statistic, specific for the given data. According to Herwartz and Neumann (2007) "in small samples the bootstrap approach outperforms inference based on critical values that are taken from a  $\chi^2$ -distribution" for Hausman test. The empirical distribution of the statistic, obtained from data, which satisfies the null hypothesis, is used for pinpointing the critical values at some level of significance. The bootstrap resampling method for obtaining

empirical critical values gives consistent results for pivotal test statistic, which the Hausman's is as long as the test is applied on a general panel data model with i.i.d. errors.

Bootstrapping is used within the Monte Carlo simulations and its methodology is described in details in the next section.

### 3. Methodology: Design of the Monte Carlo simulation

Having specified the data, the dependent variable and the factors for the tested models one can estimate the fixed effects and random effects models. This must be done before the start of any simulation procedures. By following the estimation methods described in Section 2 the  $k$ -length vectors of the parameters  $\hat{\beta}^{FE}$  and  $\hat{\beta}^{RE}$  are obtained. Hausman test can then be used for choosing the appropriate model. But before that, its properties can be assessed by simulation procedures. The generated data in the simulations has to resemble the original data as much as possible and therefore before proceeding to the simulations one has to obtain as much information about the data as possible. It is important to obtain not only  $\hat{\beta}^{FE}$  and  $\hat{\beta}^{RE}$ , but also the standard deviations of  $\hat{\alpha}^{RE}$ , and  $\hat{\varepsilon}^{RE}$  as estimated from the random effects model.

#### 3.1. Defining a set of data generated processes (DGP).

For estimating the size of Hausman test, it is necessary to provide such conditions, for which the null hypothesis would be the correct one. Then the data must satisfy the assumptions for using the random effects model. For estimating the power the data applied in the model must satisfy the alternative hypothesis. Note that equations (3) and (4) depicting the fixed effects model and the random effects model respectively have the same general panel data model form:

$$y_{it} = \beta_1 x_{1,it} + \beta_2 x_{2,it} + \dots + \beta_k x_{k,it} + \alpha_i + \varepsilon_{it} \quad (13)$$

The difference in the two types of models is only in the estimation methods and in the way we look at the individual-specific component  $\alpha$ . The hypotheses of the Hausman test for the general panel data model are:

$$H_0: Cov(\alpha_i, \mathbf{x}_{it}) = 0 \quad \text{against} \quad H_1: Cov(\alpha_i, \mathbf{x}_{it}) \neq 0.$$

Therefore when estimating the size in the simulation procedure  $\alpha_i$  must be generated in a way that guarantees no correlation with any of the independent variables. When estimating the power,  $\alpha_i$  must be generated correlated with a chosen variable from the factors. Next, a new variable  $y_{it}^*$  is generated using the random and fixed effects estimates of the parameters through the formula:

$$y_{it}^* = \hat{\beta}_1^{RE,FE} x_{1it} + \hat{\beta}_2^{RE,FE} x_{2it} + \dots + \hat{\beta}_k^{RE,FE} x_{kit} + \alpha_i + \varepsilon_{it} \quad (14)$$

The most crucial part of the simulation is generating the individual-specific component  $\alpha_i$ . There are two conditions that must be satisfied by  $\alpha_i$ :

- $\alpha_i$  must have the same values for all points in time within the same individual.
- $\alpha_i$  must be correlated with one of the factors  $x_{j,it}$  with a correlation coefficient  $\rho$ .

Note that  $\rho$  can be chosen as 0 when estimating the size.

The correlation between  $x_{j,it}$  and  $\alpha_i$  can be ensured through the use of Cholesky decomposition. However, because  $\alpha_i$  should stay constant in time, the task of generating it to be correlated with a time-varying variable by a predetermined correlation coefficient  $\rho$  becomes practically impossible. One possible workaround used in this thesis is to correlate  $\alpha_i$  not with  $x_{j,it}$ , but with its mean through time:  $\bar{x}_{j,i}$ . by specifying the correlation coefficient between them,  $\rho$ . Then, the true correlation  $\rho^*$  between  $\alpha_i$  and  $x_{j,it}$  can be computed for informative purposes. The true correlation  $\rho^*$  will always be smaller than  $\rho$  (or equal in case of no correlation), but still its increase or decrease can be controlled.

Cholesky decomposition is a frequently used method for generating correlated random variables. The method in general starts from generating observations for one variable A from a standard normal distribution, and then another variable B is generated by the use of Cholesky decomposition, correlated with A.

However, the focus of this thesis falls on real data and thus variable A is actually not generated, but already existing – in this case it is represented by a vector of elements  $\bar{x}_{j,i}$ , which might not be normally distributed and its mean and variance would almost certainly differ from 0 and 1. Even so, the Cholesky decomposition method can still be used to generate  $\alpha_i$  correlated with  $\bar{x}_{j,i}$ . and it produces vector  $\mathbf{a}$  with the same variance as the variance of the vector with elements  $\bar{x}_{j,i}$ . The details behind the method are shown below.

Let  $\bar{\mathbf{x}}_j^{ind}$  be the vector of length  $N$  with  $\bar{x}_{j,i}$  as elements and variance  $d_1$ . A vector  $\mathbf{a}$  of length  $N = n \cdot T$  and elements that are constant across time (for all  $t$  within individuals) is generated from a normal distribution  $N(0, d_2)$ . Consider the matrix:

$$\mathbf{P} = [\bar{\mathbf{x}}_j^{ind} \quad \mathbf{a}]' \sim F(\mathbf{m}, \mathbf{D}), \quad \mathbf{D} = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix},$$

where  $F$  is any distribution,  $\mathbf{D}$  is the covariance matrix of for  $\mathbf{P}$ . The task is to transform matrix  $\mathbf{P}$  into matrix:

$$\mathbf{Q} = [\bar{\mathbf{x}}_j^{ind} \quad \mathbf{a}_2]' \sim F(\mathbf{M}, \mathbf{\Sigma}),$$

where  $\bar{\mathbf{x}}_j^{ind}$  and  $\mathbf{a}_2$  are correlated (and  $\mathbf{a}_2$  has the same variance as  $\mathbf{a}$ ), by using the affine transformation  $\mathbf{Q} = \mathbf{L}\mathbf{P}$ . Matrix  $\mathbf{Q}$  after the transformation will have the distribution  $F(\mathbf{L}\mathbf{m}, \mathbf{L}\mathbf{D}\mathbf{L}')$  with covariance matrix  $\mathbf{\Sigma}$  that can be represented as follows:

$$\mathbf{\Sigma} = \mathbf{L}\mathbf{D}\mathbf{L}' = \begin{bmatrix} \text{Var}(\bar{\mathbf{x}}_j^{ind}) & \rho \sqrt{\text{Var}(\bar{\mathbf{x}}_j^{ind})} \sqrt{\text{Var}(\mathbf{a}_2)} \\ \rho \sqrt{\text{Var}(\bar{\mathbf{x}}_j^{ind})} \sqrt{\text{Var}(\mathbf{a}_2)} & \text{Var}(\mathbf{a}_2) \end{bmatrix} = \begin{bmatrix} d_1 & \rho \sqrt{d_1 d_2} \\ \rho \sqrt{d_1 d_2} & d_2 \end{bmatrix},$$

where  $\rho$  is the correlation between  $\bar{\mathbf{x}}_j^{ind}$  and  $\mathbf{a}$ . Consider the product:

$$\mathbf{J} = \mathbf{R}^{\frac{1}{2}} \mathbf{D} \mathbf{R}^{\frac{1}{2}} = \begin{bmatrix} \frac{1}{2}(\sqrt{1+\rho} + \sqrt{1-\rho}) & \frac{1}{2}(\sqrt{1+\rho} - \sqrt{1-\rho}) \\ \frac{1}{2}(\sqrt{1+\rho} - \sqrt{1-\rho}) & \frac{1}{2}(\sqrt{1+\rho} + \sqrt{1-\rho}) \end{bmatrix} \cdot \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}.$$

$$\begin{bmatrix} \frac{1}{2}(\sqrt{1+\rho} + \sqrt{1-\rho}) & \frac{1}{2}(\sqrt{1+\rho} - \sqrt{1-\rho}) \\ \frac{1}{2}(\sqrt{1+\rho} - \sqrt{1-\rho}) & \frac{1}{2}(\sqrt{1+\rho} + \sqrt{1-\rho}) \end{bmatrix} =$$

$$\begin{bmatrix} \frac{d_1}{2}(\sqrt{1+\rho} + \sqrt{1-\rho}) & \frac{d_2}{2}(\sqrt{1+\rho} - \sqrt{1-\rho}) \\ \frac{d_1}{2}(\sqrt{1+\rho} - \sqrt{1-\rho}) & \frac{d_2}{2}(\sqrt{1+\rho} + \sqrt{1-\rho}) \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{2}(\sqrt{1+\rho} + \sqrt{1-\rho}) & \frac{1}{2}(\sqrt{1+\rho} - \sqrt{1-\rho}) \\ \frac{1}{2}(\sqrt{1+\rho} - \sqrt{1-\rho}) & \frac{1}{2}(\sqrt{1+\rho} + \sqrt{1-\rho}) \end{bmatrix} =$$

$$\begin{bmatrix} \frac{d_1}{4}(\sqrt{1+\rho} + \sqrt{1-\rho})^2 + \frac{d_2}{4}(\sqrt{1+\rho} - \sqrt{1-\rho})^2 & \frac{d_1}{4}(1+\rho - (1-\rho)) + \frac{d_2}{4}(1+\rho - (1-\rho)) \\ \frac{d_1}{4}(1+\rho - (1-\rho)) + \frac{d_2}{4}(1+\rho - (1-\rho)) & \frac{d_1}{4}(\sqrt{1+\rho} - \sqrt{1-\rho})^2 + \frac{d_2}{4}(\sqrt{1+\rho} + \sqrt{1-\rho})^2 \end{bmatrix} =$$

$$\begin{bmatrix} \frac{1 + \sqrt{(1+\rho)(1-\rho)}}{2}d_1 + \frac{1 - \sqrt{(1+\rho)(1-\rho)}}{2}d_2 & \frac{\rho}{2}d_1 + \frac{\rho}{2}d_2 \\ \frac{\rho}{2}d_1 + \frac{\rho}{2}d_2 & \frac{1 - \sqrt{(1+\rho)(1-\rho)}}{2}d_1 + \frac{1 + \sqrt{(1+\rho)(1-\rho)}}{2}d_2 \end{bmatrix}$$

If  $d_1 = d_2 = d$ , then  $J = \mathbf{R}^{\frac{1}{2}}\mathbf{D}\mathbf{R}^{\frac{1}{2}} = \boldsymbol{\Sigma} = \mathbf{L}\mathbf{D}\mathbf{L}' = \begin{bmatrix} d & \rho d \\ \rho d & d \end{bmatrix}$  and  $\mathbf{R} = \mathbf{L}\mathbf{L}'$ . It is straightforward to obtain the lower triangular matrix  $\mathbf{L}$  by using Cholesky decomposition of  $\mathbf{R}$ . Once  $\mathbf{L}$  is computed, it is multiplied by  $\mathbf{P}$  and matrix  $\mathbf{Q} = \mathbf{L}\mathbf{P}$  is obtained. Vectors  $\mathbf{a}_2$  and  $\bar{\mathbf{x}}_j^{ind}$  in matrix  $\mathbf{Q}$  have covariance matrix  $\boldsymbol{\Sigma} = d\mathbf{R}$  and correlation matrix  $\mathbf{R}$ . The generated elements  $a_{2i}$  are used as  $\alpha_i$ . The only requirement is that vector  $\mathbf{a}$  in the beginning is generated with the same variance as  $\bar{\mathbf{x}}_j^{ind}$ .

As already mentioned, it is important to design the DGP in a way that would result in data as similar to the original as possible. However, by using the Cholesky decomposition constrictions are already put on  $\alpha_i$  that are likely unrealistic for the individual-specific components, estimated from the original data. Vector  $\boldsymbol{\alpha}$  is generated to have the same variance as  $\bar{\mathbf{x}}_j^{ind}$ , whereas the true variance of the individual-specific component is different. This difference between the simulated and real data is important, because it affects the the estimation of the Hausman test's power and size. The main reason for this is that in the estimation of the random effects model the following parameter is used:

$$\theta = 1 - \frac{\sigma_\varepsilon}{\sqrt{T\sigma_\alpha^2 + \sigma_\varepsilon^2}}$$

It is important to strive obtaining the same estimate of  $\theta$  using the data from the DGP as from the original data. This can be achieved by simulating the disturbances  $\varepsilon_{it}$  with a specific variance. To check what variance consider that  $\sigma_{\alpha_1}^2 = \text{Var}(\hat{\alpha}_i^{RE})$ ,  $\sigma_{\varepsilon_1}^2 = \text{Var}(\hat{\varepsilon}_{it}^{RE})$ ,  $\sigma_{\alpha_2}^2 = \text{Var}(\alpha_i)$ ,  $\sigma_{\varepsilon_2}^2 = \text{Var}(\varepsilon_{it})$  and  $m = \frac{\sqrt{\sigma_{\alpha_2}^2}}{\sqrt{\sigma_{\alpha_1}^2}}$ . The problem is reduced to finding

$\sigma_{\varepsilon_2}^2$  in the following equation:

$$1 - \frac{\sigma_{\varepsilon 1}}{\sqrt{T\sigma_{\alpha 1}^2 + \sigma_{\varepsilon 1}^2}} = 1 - \frac{\sigma_{\varepsilon 2}}{\sqrt{T\sigma_{\alpha 2}^2 + \sigma_{\varepsilon 2}^2}}$$

The solution is:

$$\begin{aligned} \frac{\sigma_{\varepsilon 1}}{\sqrt{T\sigma_{\alpha 1}^2 + \sigma_{\varepsilon 1}^2}} &= \frac{\sigma_{\varepsilon 2}}{\sqrt{Tm^2\sigma_{\alpha 1}^2 + \sigma_{\varepsilon 2}^2}} \\ \frac{\sigma_{\varepsilon 1}^2}{T\sigma_{\alpha 1}^2 + \sigma_{\varepsilon 1}^2} &= \frac{\sigma_{\varepsilon 2}^2}{Tm^2\sigma_{\alpha 1}^2 + \sigma_{\varepsilon 2}^2} \\ \sigma_{\varepsilon 1}^2(Tm^2\sigma_{\alpha 1}^2 + \sigma_{\varepsilon 2}^2) &= \sigma_{\varepsilon 2}^2(T\sigma_{\alpha 1}^2 + \sigma_{\varepsilon 1}^2) \\ Tm^2\sigma_{\varepsilon 1}^2\sigma_{\alpha 1}^2 + \sigma_{\varepsilon 1}^2\sigma_{\varepsilon 2}^2 &= T\sigma_{\varepsilon 2}^2\sigma_{\alpha 1}^2 + \sigma_{\varepsilon 2}^2\sigma_{\varepsilon 1}^2 \\ m^2\sigma_{\varepsilon 1}^2 &= \sigma_{\varepsilon 2}^2 \\ \sigma_{\varepsilon 2} &= m\sigma_{\varepsilon 1} \end{aligned}$$

Therefore  $\varepsilon_{it}$  must be generated from a normal distribution with mean 0 and standard deviation  $m\sqrt{Var(\hat{\varepsilon}_{it}^{RE})}$ .

Once, all the element on the right side of equation (14) are obtained,  $y_{it}^*$  is computed respectively for estimating the size or the power. From now on  $y_{it}^*$  will be used instead of  $y_{it}$  as dependent variable for estimating the fixed and random effects models. Since the value of the correlation  $\rho$  between  $x_{j,it}$  and  $\alpha_i$  can be controlled, this would be the way to fix the truthfulness of one of the hypotheses: null or alternative. If  $\rho$  is fixed to 0, one would now that the true hypothesis, when performing the Hausman test is the null, and by counting how many times on average the null hypothesis has been wrongfully rejected, the size of the test can be estimated. If  $\rho$  is set to be bigger than 0, then the true hypothesis will be the alternative and this gives the opportunity to estimate the power of the test.

### 3.2. Generating r independent replications

For the estimation of the properties of the Hausman test, one needs to perform a multiple replication process. This means that the data generated process is conducted r times, which defines r replications. In total r vectors of size N are generated for  $y_{it}^*$ . This would mean that Hausman test can be performed r times and the relative number of times a mistake has been made by the test can be counted.

### 3.3. Generating bootstrap samples in order to obtain Hausman test critical values

To check whether the Hausman test rejects or not the null hypothesis, it is necessary to compare the Hausman  $\chi^2$ -statistic with the critical value in accordance. Instead of using the asymptotic critical values of the  $\chi^2$  distribution, here Bootstrap critical values will be computed and used instead. As long as the disturbances in the model, tested by Hausman test, are independent and identically distributed, the Hausman test statistic is pivotal, which ensures the accuracy of the Bootstrap critical values estimation. It is necessary to ensure that the bootstrap DGP is similar to the one used in the testing itself. Bootstrap critical

values are computed in each replication separately. To compute them, it is necessary to generate B bootstrap samples in each replication and the data, obtained in each sample, must resemble the data from the DGP, used in the testing. Also, it must satisfy the null hypothesis, so that if a statistic, obtained from the Hausman test, is bigger than the Bootstrap critical value, one would know that the null hypothesis is rejected. The bootstrap procedure is in accordance to the non-parametric re-sampling of the error components procedure suggested by Andersson and Karlsson (2001), although an alternative bootstrapping procedure may be used instead too.

- 1) First, before the bootstrap samples are generated, it is necessary to estimate the fixed effects model (true under the alternative hypothesis) based on the new dependent variable  $y_{it}^*$ . This is done for all the replications from 1 to r. The estimates  $\hat{\beta}_l^{FE,q}$  are obtained, where  $l = 1, \dots, k, q = 1, \dots, r$ . The errors of the model are also obtained from the formula:

$$\hat{\varepsilon}_{it}^{FE,q} = (y_{it}^* - \bar{y}_{i\cdot}^*) - (x_{1,it} - \bar{x}_{1,i\cdot}) \hat{\beta}_1^{FE,q} - \dots - (x_{k,it} - \bar{x}_{k,i\cdot}) \hat{\beta}_k^{FE,q} + \bar{\varepsilon}_i^{FE,q}.$$

Additionally, the estimate of the individual-specific component  $\hat{\alpha}_i^{FE,q}$  must be computed:

$$\hat{\alpha}_i^{FE,q} = \bar{y}_{i\cdot}^* - \bar{x}_{1,i\cdot} \hat{\beta}_1^{FE,q} - \dots - \bar{x}_{k,i\cdot} \hat{\beta}_k^{FE,q}.$$

Also, the random effects model (true under the null hypothesis) must be estimated and  $\hat{\beta}_l^{RE,q}$  obtained ( $l = 1, \dots, k, q = 1, \dots, r$ ).

- 2) After the error estimates  $\hat{\varepsilon}_{it}^{FE,q}$  are computed, they must be centered:

$$\hat{\eta}_{it}^q = (\hat{\varepsilon}_{it}^{FE,q} - \bar{\varepsilon}^{Fe,q}) \gamma,$$

where  $\gamma$  adjusts for the degrees of freedom for the LSDV estimator:

$$\gamma = \sqrt{\frac{nT}{n(T-1) - k}}.$$

The same is done for  $\hat{\alpha}_i^{FE,q}$ :

$$\hat{\omega}_i^q = (\hat{\alpha}_i^{FE,q} - \bar{\alpha}^{Fe,q}) \gamma.$$

- 3) Next, B bootstrap samples of the fixed effect errors  $\hat{\eta}_{it}^q$  are obtained. B bootstrap vectors  $\hat{\eta}_j^{*q}$  of size N are obtained for each replication ( $j = 1, \dots, B, q = 1, \dots, r$ ). The bootstrap sampling is carried on two stages: first sampling individuals with replacement and then on the second stage indices are sampled for the time periods within the selected individuals again with replacement. The centered errors  $\hat{\eta}_j^{*q}$  corresponding to the sampled indices are included in the bootstrap samples.

B bootstrap samples for  $\hat{\omega}_i^q$  are also obtained. B bootstrap vectors  $\hat{\omega}_j^{*q}$  of size N are obtained for each replication ( $j = 1, \dots, B, q = 1, \dots, r$ ). Indices for the individuals are sampled with replacement. The centered components  $\hat{\omega}_i^q$  corresponding to the sampled indices are included in the bootstrap samples. Because of the sampling, each vector  $\hat{\omega}_j^{*q}$  is uncorrelated with the factors in the model and the bootstrapped individual-specific component satisfies the null hypothesis of Hausman test.

- 4) Next B bootstrap vectors of the dependent variable are obtained using the following equation:

$$y_{j,it}^{Boot,q} = \hat{\beta}_1^{RE,q} x_{1,it} + \hat{\beta}_2^{RE,q} x_{2,it} + \dots + \hat{\beta}_k^{RE,q} x_{k,it} + \hat{\omega}_{j,i}^{*q} + \hat{\eta}_{j,it}^{*q},$$

The bootstrapping in the previous step guarantees that the true hypothesis of the Hausman test, applied on the data  $y_{j,it}^{Boot,q}$  and  $x_{l,it}$  ( $l = 1, \dots, k$ ) is the null. For computing  $y_{j,it}^{Boot,q}$  ( $j = 1, \dots, B, q = 1, \dots, r$ ) the coefficients obtained from the model of the null hypothesis (the more restricted one) are used with combination of the centered errors from the model of the alternative hypothesis (the less restricted one). This procedure of creating bootstrapped vectors of the dependent variable for obtaining critical values is suggested by Davidson and MacKinnon (1999). The restricted estimates of the coefficients must be used in order to moderate the randomness of the DGP in the bootstrap procedure. In this case, efficient estimates of the critical values will be obtained. However, there are no constraints for using the unrestricted errors.

- 5) The random effects model and the fixed effects model are estimated for each bootstrap sample i.e. for the data  $y_{j,it}^{Boot,q}, x_{l,it}$  ( $l = 1, \dots, k, j = 1, \dots, B, q = 1, \dots, r$ ). The following bootstrap coefficient estimates are obtained:  $\hat{\beta}_{j,l}^{RE,q}$  and  $\hat{\beta}_{j,l}^{FE,q}$ .
- 6) Next, the Bootstrap Hausman statistics is calculated for all bootstrap samples:

$$H_j^{Boot,q} = (\hat{\beta}_j^{RE,q} - \hat{\beta}_j^{FE,q})' [Var(\hat{\beta}_j^{RE,q}) - Var(\hat{\beta}_j^{FE,q})]^{-1} (\hat{\beta}_j^{RE,q} - \hat{\beta}_j^{FE,q}),$$

Where  $\hat{\beta}_j^{RE,q}$  and  $\hat{\beta}_j^{FE,q}$  are  $k$ -long vectors of the random and fixed effects estimates, obtained from each bootstrap sample within each replication of the Monte Carlo simulation. Thus, B Bootstrap Hausman statistics are obtained in each replication.

- 7) The statistics obtained from the same replication are sorted in ascending order and the Bootstrap critical values are estimated by choosing the  $\left[\frac{1-\alpha}{100}\right] \cdot (B+1)^{th}$  (rounded to the next integer) value of the Bootstrap Hausman statistic, where  $\alpha$  is the nominal significance level of the test. Bootstrap critical values can be obtained for different values of the nominal size. Davidson and MacKinnon (1998) suggest obtaining critical values for 215 values of the significance level: 0.001, 0.002, ..., 0.010 (10 values); 0.015, 0.020, ..., 0.985 (195 values); 0.990, 0.991, ..., 0.999 (10 values).

Through the bootstrap procedure, specific set of critical values (as many as 215) are obtained for each replication and later the Hausman test statistic, obtained from the original DGP (outside the bootstrap loop), will be compared with those computed Bootstrap critical values in order to count the number of rejections of the null hypothesis. The bootstrap procedure is part of the Monte Carlo simulation and the obtained statistics of each replication of the DGP must be compared with the Bootstrap critical values from the same replication.

### 3.4. Obtaining Hausman test critical values through the use of a pure Monte Carlo simulation

As already mentioned the bootstrap procedure is a reliable method for obtaining the empirical critical values of a test statistic. It gives consistent results, when applied on pivotal test statistics. The disadvantage of the bootstrap procedure is determined by the substantial computational work, which consumes considerable amount of time. Since the bootstrap procedure requires the estimation of the same coefficients for B number of samples within each of the r replications of the Monte Carlo simulation, the total number of loops, that must be made for obtaining the end results, is  $B \cdot r$ . It is often required to go through hundreds of thousands loops. Therefore one has to have time and access to powerful enough computation equipment for applying the procedure.

A simpler and much faster procedure can be used for deriving the small sample properties of the Hausman test. A pure Monte Carlo simulation can be done, separately from the simulation for obtaining Hausman statistics (the main simulation). The number of replications in the two experiments doesn't have to be the same. It is important however that the DGP in the Monte Carlo simulation for the critical values (with  $r^{MC}$  replications) mimics the DGP in the main simulation (with  $r$  replications), but focusing only on obtaining data that satisfies the null hypothesis. Then, from each replication of the simulation a Hausman statistic will be computed. The  $r^{MC}$  statistics are sorted in ascending order and  $\left[\frac{1-\alpha}{100}\right] \cdot r^{MCth}$  value of the Monte Carlo Hausman statistic is saved as critical value for the nominal size  $\alpha$ . It gives only one set of critical values to be used in all replication of the main Monte Carlo simulation, whereas the bootstrap procedure gives r sets of critical values and in each replication a unique set of values is used. This method is more basic and applies only for the small sample properties. Also, its use in practice is ambiguous, since it largely depends on the DGP.

### 3.5. Performing Hausman test in each replication

Until now, a total of r vectors of size N are generated for  $y_{it}^*$ . The random and fixed effects models were estimated in each replication, where  $y^*$  is the dependent variable, and  $x_l$  are the independent variables. The estimates  $\hat{\beta}_l^{FE,q}$  and  $\hat{\beta}_l^{RE,q}$  ( $l = 1, \dots, k; q = 1, \dots, r$ ) were obtained. All information is available for estimating the Hausman statistic. This must be done in each replication by following the formula:

$$H^q = (\hat{\beta}^{RE,q} - \hat{\beta}^{FE,q})' [Var(\hat{\beta}^{RE,q}) - Var(\hat{\beta}^{FE,q})]^{-1} (\hat{\beta}^{RE,q} - \hat{\beta}^{FE,q})$$

Then, the computed statistic is compared with the set of the Bootstrap critical value, obtained for the specific replication or with the set of Monte Carlo critical values, obtained from another simulation. The number of replications in which the null hypothesis was rejected must be counted.

### 3.6. Estimating the size and power

The size is estimated, when the DGP is set in a way that satisfies the null hypothesis, or when the correlation coefficient  $\rho$  between any of the factors and  $\alpha_i$  is fixed to 0. Then the true hypothesis is the null, so one can count in how many replications on average has the

null hypothesis been rejected. The average times of making a mistake by rejecting the null hypothesis out of those  $r$  replications gives the estimated probability of rejecting the null hypothesis, when it is correct, i.e. an estimation of the size of Hausman test.

On the contrary, when  $\rho$  is set to be bigger than 0, thus satisfying the alternative hypothesis, the average times of rejecting the null hypothesis will be an estimate of the probability to correctly reject the null hypothesis (equal to one minus the probability to not reject the null hypothesis, when the alternative is correct). Then, the number of times the null hypothesis has been rejected over the total number of replications  $r$  in the case when  $\rho > 0$ , gives the power of Hausman test.

Note that a set of 215 Bootstrap critical values are computed for each replication for different values for the nominal size. This means that for estimating the size 215 times the Hausman statistic  $H^q$  is compared with a critical value in each replication. The average number of rejections will give 215 estimates for the actual size corresponding to 215 values of the nominal size. If the actual empirical distribution of the Bootstrap Hausman statistic is close to the  $\chi^2$ -distribution, the estimated actual size by using Bootstrap critical values won't differ much from the estimated actual size when using asymptotic critical values. By comparing the actual size with the nominal, one can draw conclusions for the accuracy of the test. The computation of the actual size for so many points of the nominal size, gives the opportunity to represent the results graphically by plotting the nominal and actual size together. Thus, finding the correspondence between the two becomes easier. One can set a nominal size of for example 5% when performing Hausman test, but the actual significance level could be bigger, if there is over-rejection of the null hypothesis and the information for the risk must be available for the researcher. The graphical illustration of the actual size and power plotted against the nominal size conveys much more information than the table could possibly do in a way that is easy to understand and interpret (Davidson and MacKinnon, 1998).

Similarly 215 estimates of the power are obtained corresponding to 215 values of the nominal size. It is interesting to see the power of the test for different values of  $\rho$ , the correlation between a factor in the model and the individual specific component  $\alpha_i$ .

### 3.7. Note

As already mentioned in Section 2.2.3., a serious problem in the estimation of the random effects model concerns the possibility of obtaining negative estimate of the variance. In such a case the estimation of the parameters is impossible. Different softwares react differently to negative estimates of the variance. In Stata, when a negative variance estimate is obtained, the variance is set to 0. This would actually mean that the random effects model is transformed to pooled model. R uses the same procedure if Amemiya method of estimation is used. However, it will stop the execution of the code if negative estimate of the variance is computed in Swamy and Arora method and an error message will be shown. There is a high risk of interruption of the simulation procedure if Swamy and Arora method is applied to the suggested methodology. Therefore, this methodology would work better on Nerlove's method for estimation of Random effects model. Another option is to use the Amemiya method, but knowing that pooled model will be used instead of random effects in case of

negative variance estimate. If negative variance estimate is obtained only in the bootstrapping, then Swamy and Arora method can still be used, but without using Bootstrap critical values.

## 4. Implementation: Reporting the Monte Carlo experiment

### 4.1. Data

The data that is used for implementing the methodology is often used in textbooks and panel data examples to illustrate the use of estimation methods in panel data: Greene (2008), Grunfeld and Griliches (1960), Boot and deWitt (1960). It is also known as Grunfeld data and consists of three variables: Gross investment (I), Market value (F) and Value of the stock of plant and equipment (C), information of which is obtained annually for 10 large companies: General Motors, Chrysler, General Electric, Westinghouse, U.S. Steel, Atlantic Refining, Diamond Match, Goodyear, Union Oil and IBM, for 20 years: from 1935 to 1954.<sup>1</sup> The definition of the variables is shown in the table below as taken from “The Grunfeld Data at 50” by Kleiber and Zeileis (2010):

*Table 2: Variables definition in Grunfeld data*

Gross Investment	<b>I</b>	Additions to plant and equipment plus maintenance and repairs in millions of dollars deflated by the implicit price deflator of producers' durable equipment (base 1947).
Market Value	<b>F</b>	The price of common shares at December 31 (or, for WH, IBM and CH, the average price of December 31 and January 31 of the following year) times the number of common shares outstanding plus price of preferred shares at December 31 (or average price of December 31 and January 31 of the following year) times number of preferred shares plus total book value of debt at December 31 in millions of dollars deflated by the implicit GNP price deflator (base 1947).
Value of the stock of plant and equipment	<b>C</b>	The accumulated sum of net additions to plant and equipment deflated by the implicit price deflator for producers' durable equipment (base 1947) minus depreciation allowance deflated by depreciation expense deflator (10 years moving average of wholesale price index of metals and metal products, base 1947).

### 4.2. Technical details

R software environment is used for the implementation part. To simulate the work, package ‘*plm*’ must be installed.

The bootstrap procedure has been applied by obtaining 299 bootstrap samples and 1000 Monte Carlo replications. This requires the estimation of 299 000 vectors of parameters for the fixed and random effects models. Because of the large number of estimations, there is a high chance of obtaining negative estimate of the variance of the individual-specific component in the random effects model in at least one bootstrap loop. When applying the methodology by comparing the estimates from the fixed effects model and the random effects model, estimated by Swamy and Arora method, if a negative variance estimate is

<sup>1</sup> The data is available on: <http://web.pdx.edu/~crkl/ec510/data/ifc10.txt>

obtained the execution of the R program stops and the procedure cannot be completed. If however there is no negative variance, obtained in the pure Monte Carlo simulation (when not using Bootstrap critical values) it is still possible to implement the procedure for estimating Hausman test's size and power in small samples by using Monte Carlo critical values. If Amemiya method is used and a negative estimate is computed, then R automatically sets the value of the variance for the individual-specific components to 0. This transforms the random effects model into pooled model. Nerlove method of estimation guarantees positive estimate of the variance. This is why the methodology has been applied three times using those three different methods for estimating the random effects model. The properties of the Hausman test based on the Swamy and Arora method of estimation for the random effects model are analyzed when using only asymptotic critical values as well as Monte Carlo critical values for estimating the small sample properties. The full procedure is implemented when using Amemiya and Nerlove methods.

#### 4.3. The model of the data generated process

The first step in the general case (no matter which method of estimation for the random effects model is used) is to estimate the fixed effects and random effects models. The parameter estimates are used later in the simulation. The following estimates are obtained:

- for the fixed effects model:  $\hat{\beta}_1^{FE} = 0.108, \hat{\beta}_2^{FE} = 0.312$ ;
- for the random effects model using Swamy and Arora method:  $\hat{\beta}_0^{RE,swar} = -57.449, \hat{\beta}_1^{RE,swar} = 0.108, \hat{\beta}_2^{FE,swar} = 0.310$ ;
- for the random effects model using Amemiya method:  $\hat{\beta}_0^{RE,am} = -57.392, \hat{\beta}_1^{RE,am} = 0.108, \hat{\beta}_2^{FE,am} = 0.309$ ;
- for the random effects model using Nerlove method:  $\hat{\beta}_0^{RE,ner} = -57.507, \hat{\beta}_1^{RE,ner} = 0.108, \hat{\beta}_2^{FE,ner} = 0.310$ .

Those estimates should be used in the simulations in order to generate data close to the original, but since they are close to each other, they can be generalized as:  $\hat{\beta}_1^{FE,RE} = 0.11$  and  $\hat{\beta}_2^{FE,RE} = 0.31$ . Those estimates are used for simulating the dependent variable in each replication of the Monte Carlo simulation:

$$I_{it}^* = \hat{\beta}_1^{FE,RE} F_{it} + \hat{\beta}_2^{FE,RE} C_{it} + \alpha_i + \varepsilon_{it}$$

To specify the parameters of the simulation it is important to obtain not only  $\hat{\beta}^{FE}$  and  $\hat{\beta}^{RE}$ , but also the standard deviations of  $\hat{\alpha}^{RE}$  and  $\hat{\varepsilon}^{RE}$  as estimated from the random effects model:

- for the random effects model using Swamy and Arora method:  $\sigma_{\alpha,RE} = \sqrt{Var(\hat{\alpha}_i^{RE})} = 82.89$  and  $\sigma_{\varepsilon,RE} = \sqrt{Var(\hat{\varepsilon}_{it}^{RE})} = 53.81$ ;
- for the random effects model using Amemiya method:  $\sigma_{\alpha,RE} = \sqrt{Var(\hat{\alpha}_i^{RE})} = 79.19$  and  $\sigma_{\varepsilon,RE} = \sqrt{Var(\hat{\varepsilon}_{it}^{RE})} = 53.81$ ;

- for the random effects model using Nerlvoe method:  $\sigma_{\alpha,RE} = \sqrt{Var(\hat{\alpha}_i^{RE})} = 84.44$  and  $\sigma_{\varepsilon,RE} = \sqrt{Var(\hat{\varepsilon}_{it}^{RE})} = 52.17$ ;

In the simulations  $\alpha_i$  is generated to be correlated with regressor C: Value of the stock of plant and equipment. It is necessary to generate  $\alpha_i$  with the same variance as  $\bar{C}_i$ . (they repeat across time within individuals) and  $\varepsilon_{it}$  must be generated with variance that keeps the same ratio between the standard deviations of  $\alpha_i$  and  $\varepsilon_{it}$  as in the original panel data model. The reason for this was explain in Section 3.1.1. The standard deviation of  $\bar{C}_i$  is 191.19. To keep the ration between  $\sqrt{Var(\alpha_i)}$  and  $\sqrt{Var(\varepsilon_{it})}$  ( $\varepsilon_{it}$  and  $\alpha_i$  are generated) the same as the ration between  $\sigma_{\alpha,RE}$  and  $\sigma_{\varepsilon,RE}$ ,  $\varepsilon_{it}$  must be simulated with standard deviation  $m \cdot \sigma_{\varepsilon,RE}$ . The coefficient  $m = \frac{\sigma_{\bar{C}_i}}{\sqrt{\sigma_{\alpha,RE}^2}}$  is:

- for the random effects model using Swamy and Arora method:  $m = 2.31$  and  $\varepsilon_{it}$  must be generated with standard deviation 124.12;
- for the random effects model using Amemiya method:  $m = 2.41$  and  $\varepsilon_{it}$  must be generated with standard deviation 129.90;
- for the random effects model using Nerlvoe method:  $m = 2.26$  and  $\varepsilon_{it}$  must be generated with standard deviation 118.13;

Having specified the parameters of the simulation, it can be implemented according to the methodology. 1000 replications are used in the Monte Carlo simulations to obtain estimates of the power and size of Hausman test. A seed is set to 1234. The power is estimated for 4 different values of the correlation coefficient  $\rho$  between  $\alpha_i$  and  $\bar{C}_i$ : 0.2, 0.4, 0.5 and 0.7. After the simulation the average correlation coefficient  $\bar{\rho}^*$  between  $\alpha_i$  and  $C_{it}$  across the replications can be computed in order interpret the results based on the real correlation coefficient  $\rho^*$  between  $\alpha_i$  and  $C_{it}$ .

The detailed algorithm of the Monte Carlo study including the computation of Hausman statistic's critical values is presented in a schematic way in the Appendix.

#### 4.4. Obtaining Hausman statistics critical values trough bootstrapping

As seen from the algorithm in the Appendix, the bootstrapping procedure is implemented inside the main Monte Carlo function for estimating the size and power of Hausman test. This must be done, because the bootstrapping is based on the estimates, obtained in the Monte Carlo simulation. For each replication specific Bootstrap critical values are obtained that are not valid for other replications.

The Bootstrap critical values are obtained according to the procedure described in Section 3.1.3. The number of bootstrap samples is  $B = 299$  and they are resampled in each separate replication. For 1000 replication a total of 299 000 loops must be executed. In each replication 299 bootstrap Hausman statistics are obtained. They are distributed under the null hypothesis, which means that taking the  $\left[\frac{1-\alpha}{100}\right] \cdot (B + 1)^{th}$  (rounded to the next integer)

value of the ordered Bootstrap statistics would give the critical value for significance level  $\alpha$ . If the size  $\alpha$  is specified with accuracy to three decimal places and  $B$  is less than a 1000 (in this case it is 299),  $\left\lceil \frac{1-\alpha}{100} \right\rceil \cdot (B + 1)$  won't be an integer and it must be since it represents an index. The significance level  $\alpha$  is indeed specified with accuracy of three digits after the decimal point: 0.001, 0.002,...,0.010 (10 values); 0.015, 0.020,...,0.985 (195 values); 0.990,0.991,...,0.999 (10 values). This is done in order to ensure the possibility for comparison between using Monte Carlo, Bootstrap and asymptotic critical values. That's why  $\left\lceil \frac{1-\alpha}{100} \right\rceil \cdot (B + 1)$  must be rounded to the next integer. Thus 215 critical values corresponding to 215 values for the size are obtained in each replication.

#### 4.5. Obtaining Hausman statistics critical values through pure Monte Carlo simulation

A pure Monte Carlo simulation separately from the main one that estimates the power and size is conducted in order to obtain critical values used for analyzing the small sample properties of Hausman test. Those values are referred to as Monte Carlo critical values as opposed to Bootstrap critical values. While the number of replications in the main Monte Carlo simulation is 1 000, this one doesn't include any time-consuming elements like bootstrapping and can be implemented using more replications. It has been executed with 10 000 replications. The DGP in the Monte Carlo simulation for the critical values follows the same idea as in the main one with the difference that data is generated following the null hypothesis. The seed also doesn't have to be the same. It is set to 6290. The dependent variable in each replication is in the same way generated according to the formula:

$$I_{it}^* = \hat{\beta}_1^{FE,RE} F_{it} + \hat{\beta}_2^{FE,RE} C_{it} + \alpha_i^{null} + \varepsilon_{it}$$

The fixed and random effects are estimated using  $I^*$  as dependent, and  $F$  and  $C$  as independent variables. Next Hausman test is performed and its statistic saved. The statistics obtained from all replications are sorted and similar to the bootstrap procedure the critical values can be obtained by extracting the  $\left\lceil \frac{1-\alpha}{100} \right\rceil \cdot r^{MC^{th}}$  statistic ( $r^{MC} = 1\ 000$ ), where  $\alpha$  is the significance level. This method would provide the critical values since the Hausman statistics are distributed under the null hypothesis. This Monte Carlo simulation produces only one set of critical values that can be used in all replications of the main simulation for estimating the size and power of Hausman test, but the use of those critical values in practice is ambiguous since they rely exclusively on the data generated process.

#### 4.6. Results

It is interesting to compare the size and power of Hausman test when applied on different methods for estimating the random effects model: Swamy and Arora method, Amemiya method and Nerlove method. It was not possible to compute bootstrap critical values for the Swamy and Arora method of estimation due to obtaining negative variance estimates of the individual-specific components, based on the bootstrap samples. The inferences about the power and size of Hausman test, comparing the estimates of the within method (fixed effects model) and the Swamy-Arora method (random effects model) are based only on asymptotic and Monte Carlo critical values (and not Bootstrap critical values). There was no

problem obtaining critical values when the random effects model is estimated with Amemiya method. If a negative estimate of the variance is obtained in the estimation of the random effects model with Amemiya, R replaces the model with pooled. Nerlove method of estimation guarantees positive estimates of the variance and can always estimate the random effects model. No negative variance estimates were obtained in either of the methods outside the bootstrapping procedure. This was an issue only in obtaining the Bootstrap critical values.

The next sub-sections present and interpret the results of the Hausman test properties estimation. First, the properties of the test are examined when Amemiya method is applied. Next, inferences are made about the properties under Nerlove method and finally, when using Swamy-Arora method.

#### 4.6.1. Size and power of Hausman test (Amemiya method for estimating random effects model)

The estimates of the within method for estimating the fixed effects model, based on the original data: Gross investment (I) as dependent variable and Market value (F), Value of the stock of plant and equipment (C) – as regressors, are:  $\hat{\beta}_1^{FE} = 0.108$ ,  $\hat{\beta}_2^{FE} = 0.312$ . Using Amemiya method for estimating the parameters of the random effects model, based on the same data, gives the results:  $\hat{\beta}_0^{RE,am} = -57.392$ ,  $\hat{\beta}_1^{RE,am} = 0.108$ ,  $\hat{\beta}_2^{RE,am} = 0.309$ .

By performing Hausman test on  $\hat{\beta}^{FE}$  and  $\hat{\beta}^{RE,am}$  one can see that the null hypothesis of no correlation between the individual-specific component and the factors is rejected for level 0.10 (fixed effects model should be used), but there is not enough evidence to support the rejection of H0 for significance level 0.01 and 0.05 and therefore according to the test random effects model should be used.

- 1) Defining the null and alternative hypotheses:

**H<sub>0</sub>**: The appropriate model is Random effects. There is no correlation between the error term and the independent variables in the panel data model.

$$Cov(\alpha_i, x_{it}) = 0$$

**H<sub>1</sub>**: The appropriate model is Fixed effects. The correlation between the error term and the independent variables in the panel data model is statistically significant.

$$Cov(\alpha_i, x_{it}) \neq 0$$

- 2) A probability of first type error is chosen. The test can be performed on  $\alpha = 0.01$ , 0.05 and 0.10.
- 3) Hausman statistic is calculated from the formula:

$$H = (\hat{\beta}^{RE} - \hat{\beta}^{FE})' [Var(\hat{\beta}^{RE}) - Var(\hat{\beta}^{FE})]^{-1} (\hat{\beta}^{RE} - \hat{\beta}^{FE}) = 5.109$$

- 4) The p-value of the statistic based on the  $\chi^2$  distribution with 2 degrees of freedom is 0.078.

Table 3: Hausman test: Within vs. Amemiya methods of estimating the panel data model

Hausman Test (Within vs. Amemiya method)	
$\chi^2$ - statistic	5.1088
p-value for $\chi^2$ (2)	0.07774

- 5) The null hypothesis can not be rejected for levels of significance 0.01 and 0.05, since  $0.07774 > 0.05$ .

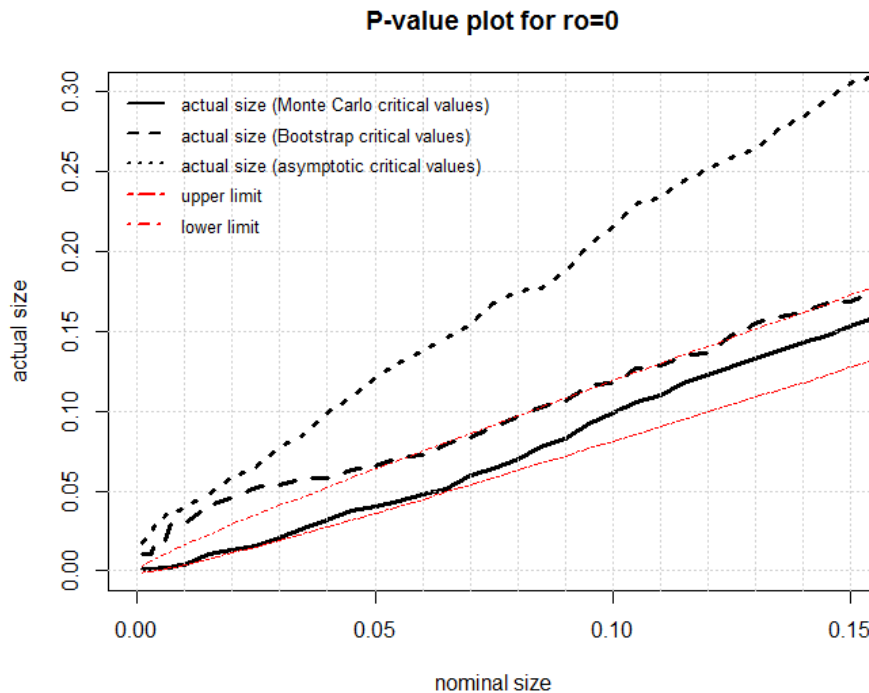
It is interesting to see whether Hausman test has the tendency to over- or under-reject the null hypothesis. From Table 4 as well as from Figure 1 one can see that when using asymptotic critical values, the actual size is much bigger than the nominal, which means that the actual p-values of the Hausman statistics are also bigger than the nominal and performing the Hausman test for a fixed significance level, there is over-rejection of the null hypothesis. For example when setting a significance level of 0.010, one actually works under the risk of making an error of type I as big as 0.040. This can substantially change the inference of a panel data study. The nominal p-value 0.078 corresponds to an actual probability around 0.175. Then, the rejection of the null hypothesis for significance level 0.10, in the test performed above, wouldn't be correct. In reality one shouldn't reject the null, since asymptotic critical values are used and the p-value was obtained based on the  $\chi^2(2)$  distribution.

Table 4: Actual size and power of Hausman test, which compares the estimates of the within method (fixed effects) vs. Amemiya method (random effects) of estimating the panel data model<sup>2</sup>

Size and power of Hausman test (Amemiya method for estimating random effects model)											
	$\rho$	$\rho^*$	Using asymptotic critical values			Using Bootstrap critical values			Using Monte Carlo critical values		
			0.010	0.050	0.100	0.010	0.050	0.100	0.010	0.050	0.100
<b>Nominal Size</b>	-	-	<b>0.010</b>	<b>0.050</b>	<b>0.100</b>	<b>0.010</b>	<b>0.050</b>	<b>0.100</b>	<b>0.010</b>	<b>0.050</b>	<b>0.100</b>
<b>Actual Size</b>	0	0	0.040	0.121	0.215	0.029	0.066	0.118	0.004	0.040	0.099
<b>Power</b>	0.20	0.125	0.061	0.131	0.241	0.035	0.081	0.124	0.010	0.061	0.115
<b>Power</b>	0.40	0.259	0.099	0.240	0.382	0.059	0.129	0.207	0.011	0.099	0.202
<b>Power</b>	0.50	0.327	0.127	0.325	0.484	0.070	0.166	0.263	0.013	0.125	0.275
<b>Power</b>	0.70	0.461	0.199	0.541	0.764	0.082	0.270	0.430	0.013	0.199	0.473

<sup>2</sup>  $\rho$  is the correlation coefficient between  $\alpha_i$  and  $\bar{C}_i$ , and  $\rho^*$  - the correlation coefficient between  $\alpha_i$  and  $C_{it}$ .

Figure 1: Zoomed in p-value plot for the Hausman statistic (Within vs. Amemiya):  $\rho = 0$ , 1000 replications. Critical values: Monte Carlo, Bootstrap and asymptotic.



Instead of asymptotic critical values one can use Bootstrap critical values or Monte Carlo critical values for deriving the small sample properties of Hausman test. Figure 1 shows that the bootstrap critical values also over-reject the null hypothesis for nominal size below 0.05, but the actual size doesn't differ significantly from the nominal for level 0.05 and 0.10. Also the over-rejection is smaller than what is observed with the asymptotic critical values. For significance level 0.01, the bootstrap critical values ensure 36.7% closer values of the actual size to the nominal than when using asymptotic critical values. For significance level above 0.05, the actual size is within the confidence interval of the nominal. Also, for this specific case with the used data, the problem of over-rejection for values of the nominal size below 0.10 didn't affect the outcome of the Hausman test, since the null hypothesis wasn't rejected. It wouldn't have been rejected even if bootstrap critical values were used, since the risk of wrongfully rejecting the null with bootstrap critical values is smaller. The over-rejection of  $H_0$  plays more crucial role for nominal size 0.10, but then, the bootstrap critical values ensure actual size within the confidence interval of the nominal. Therefore for this data, the bootstrap critical values ensure better size of Hausman test. The Monte Carlo small sample critical values provide actual size, which doesn't differ significantly from the nominal size regardless of the level of significance. Figure 2 shows that for generally the bootstrap and Monte Carlo critical values lead to the desired size levels without over-rejection of the null hypothesis. The actual size tends to follow the upper bound of the confidence interval.

Figure 2: P-value plot for the Hausman statistic (Within vs. Amemiya):  $\rho=0$ , 1000 replications. Critical values: Monte Carlo, Bootstrap and asymptotic.

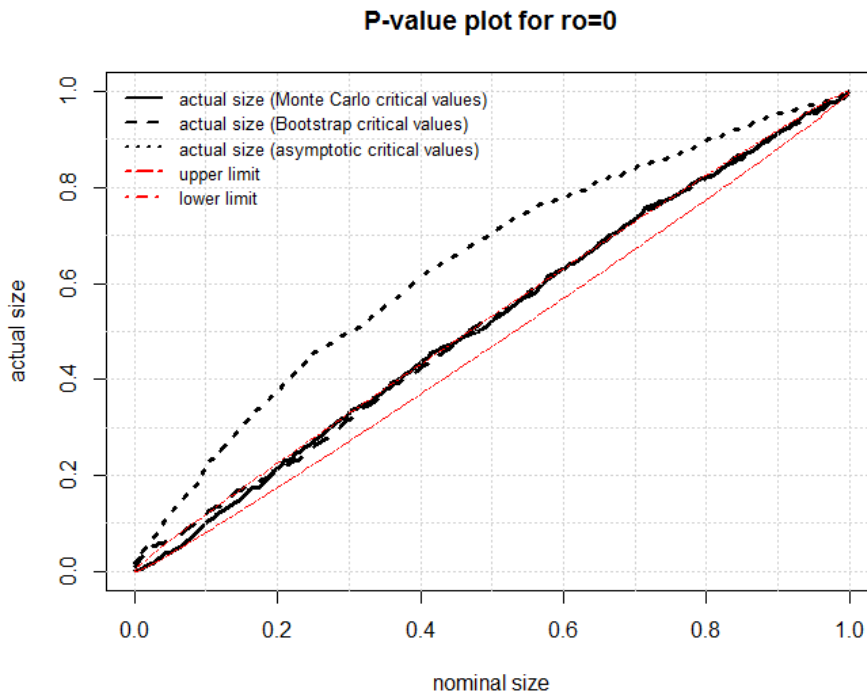


Table 4 as well as Figures 3-6 illustrate how the power of Hausman test changes for different values of the nominal size and correlation between the individual-specific component and the regressor C.

Figure 3: Size-power plot with nominal size for the Hausman statistic (Within vs. Amemiya):  $\rho=0.200$ ,  $\rho^* = 0.125$ , 1000 replications. Critical values: Monte Carlo, Bootstrap and asymptotic.

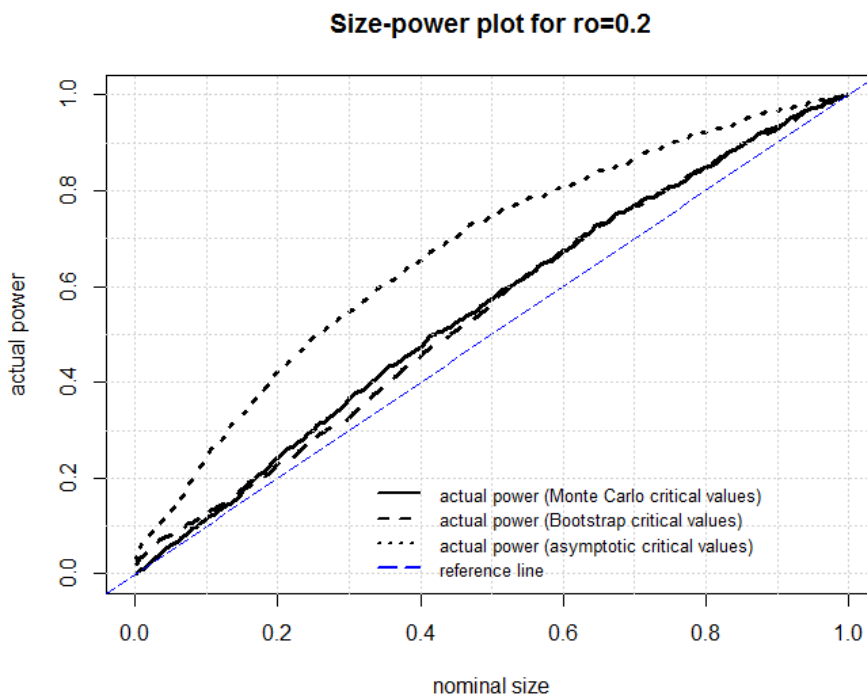


Figure 4: Size-power plot with nominal size for the Hausman statistic (Within vs. Amemiya):  $\rho=0.400$ ,  $\rho^* = 0.259$ , 1000 replications. Critical values: Monte Carlo, Bootstrap and asymptotic.

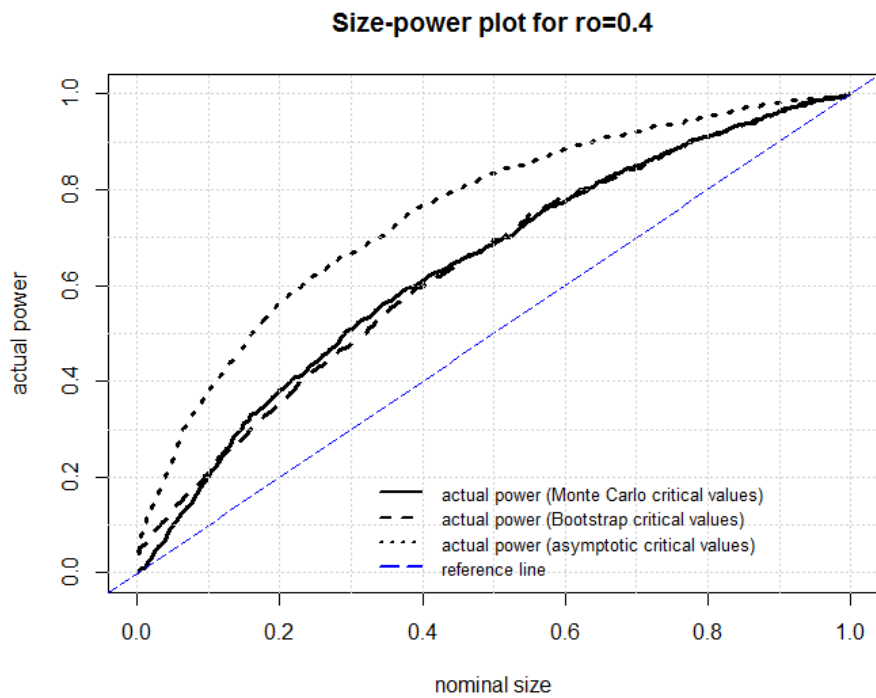


Figure 5: Size-power plot with nominal size for the Hausman statistic (Within vs. Amemiya):  $\rho=0.500$ ,  $\rho^* = 0.327$ , 1000 replications. Critical values: Monte Carlo, Bootstrap and asymptotic.

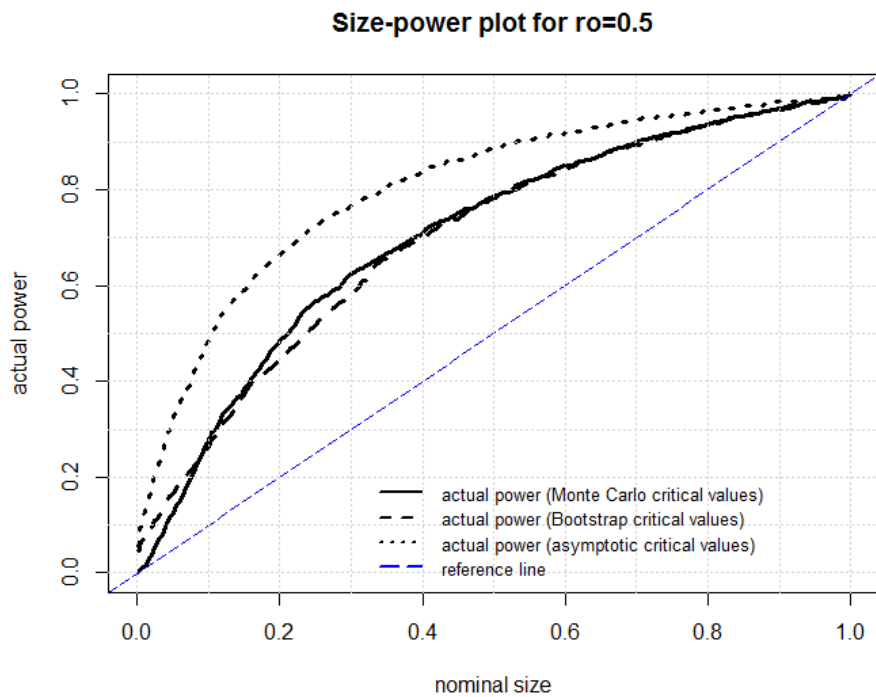
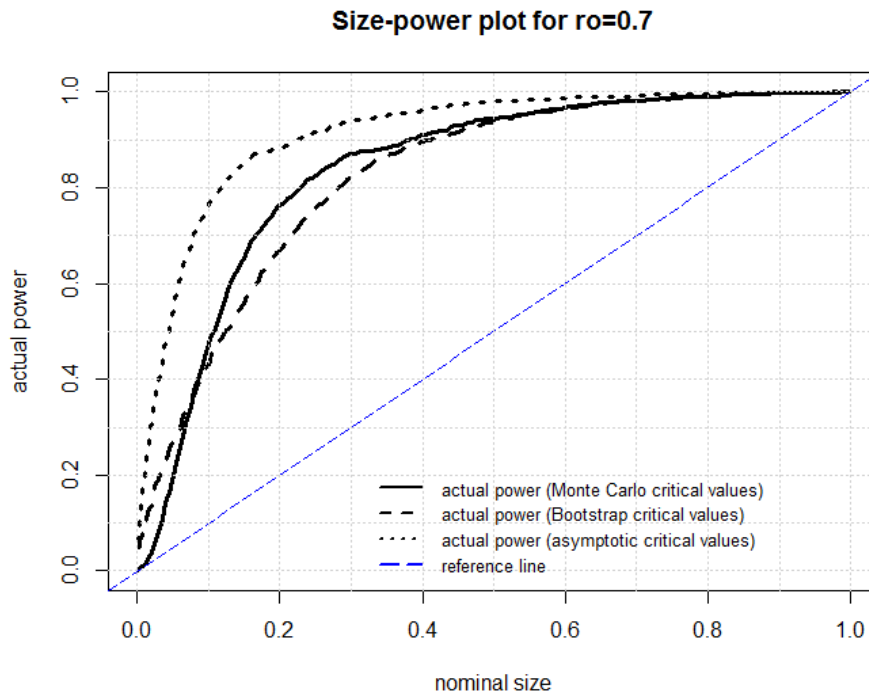


Figure 6: Size-power plot with nominal size for the Hausman statistic (Within vs. Amemiya):  $\rho=0.700$ ,  $\rho^* = 0.461$ , 1000 replications. Critical values: Monte Carlo, Bootstrap and asymptotic.



As expected the bigger the correlation coefficient, the better is the power. One can notice that for nominal size 0.05 the power of Hausman test does not reach the considered as “good” values of 0.8 (Cohen, 1988) for any of the used critical values even for  $\rho = 0.7$  – the correlation coefficient between  $\alpha_i$  and  $\bar{C}_{i.}$ , corresponding to  $\rho^* = 0.461$  - the correlation coefficient between  $\alpha_i$  and  $C_{it}$ . The Hausman test has small power unless there is very high correlation between the individual-specific component and the regressors. The power is very low for nominal size 0.01, especially for Monte Carlo critical values. The reason for this may be behind the fact that coefficient  $\rho$  is used instead of  $\rho^*$  - the real correlation coefficient between  $\alpha_i$  and  $\bar{C}_{i.}$ , which is smaller than the used coefficient  $\rho$ . However, even for  $\rho^* = 0.461$  the power for nominal size 0.01 and Monte Carlo critical values is only 0.013 – a value considerably small for a power, indicating high risk of not rejecting the null, when the alternative is correct. The Monte Carlo and Bootstrap power lines are similar in all graphs, but they cannot be compared since the pure Monte Carlo procedure is used only for analyzing small sample properties. Also, the comparison of the power measurements between Bootstrap and asymptotic critical values cannot be done unless they are plotted against the actual size. This is illustrated in Figures 7-10. They show that none of the critical values considerably predispose to better power of the test. No power is lost by using bootstrap critical values. Of course the power increases with the size, since by increasing the type I error, the type II error decreases, and the power is one minus the type II error. The size-adjusted power is very small even for the highest obtained correlation coefficient  $\rho^* = 0.461$ . For actual size of 0.05, the power is on bigger than 0.2 and for size 0.10 – the power is around 0.5. Perhaps, with grater correlation between the regressors and the individual-specific component, the power approaches more appealing levels.

Figure 7: Size-power plot with actual size for the Hausman statistic (Within vs. Amemiya):  $\rho=0.200$ ,  $\rho^* = 0.125$ , 1000 replications. Critical values: Bootstrap and asymptotic.

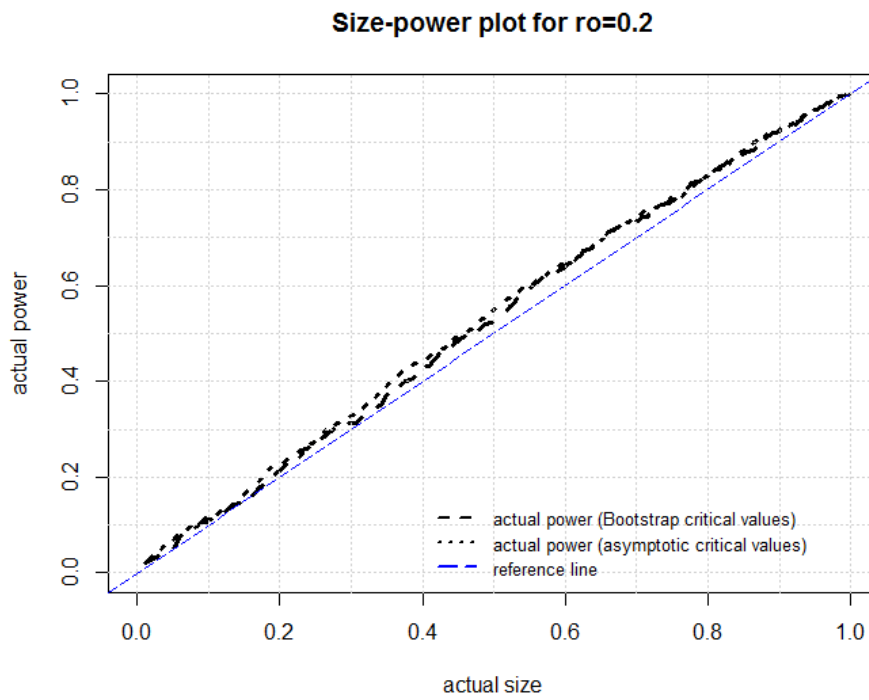


Figure 8: Size-power plot with actual size for the Hausman statistic (Within vs. Amemiya):  $\rho=0.400$ ,  $\rho^* = 0.259$ , 1000 replications. Critical values: Bootstrap and asymptotic.

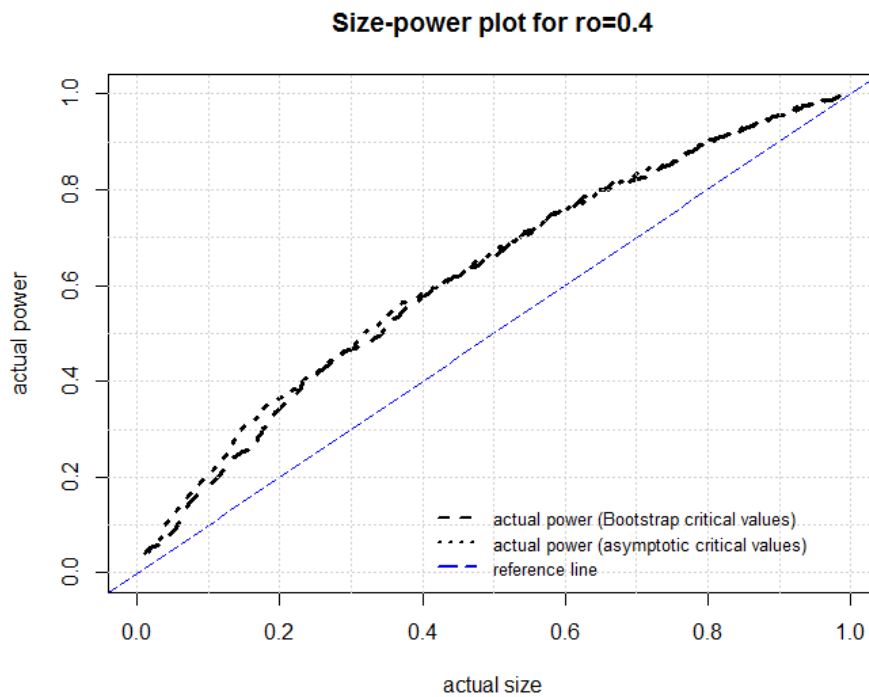


Figure 9: Size-power plot with actual size for the Hausman statistic (Within vs. Amemiya):  $\rho=0.500$ ,  $\rho^* = 0.327$ , 1000 replications. Critical values: Bootstrap and asymptotic.

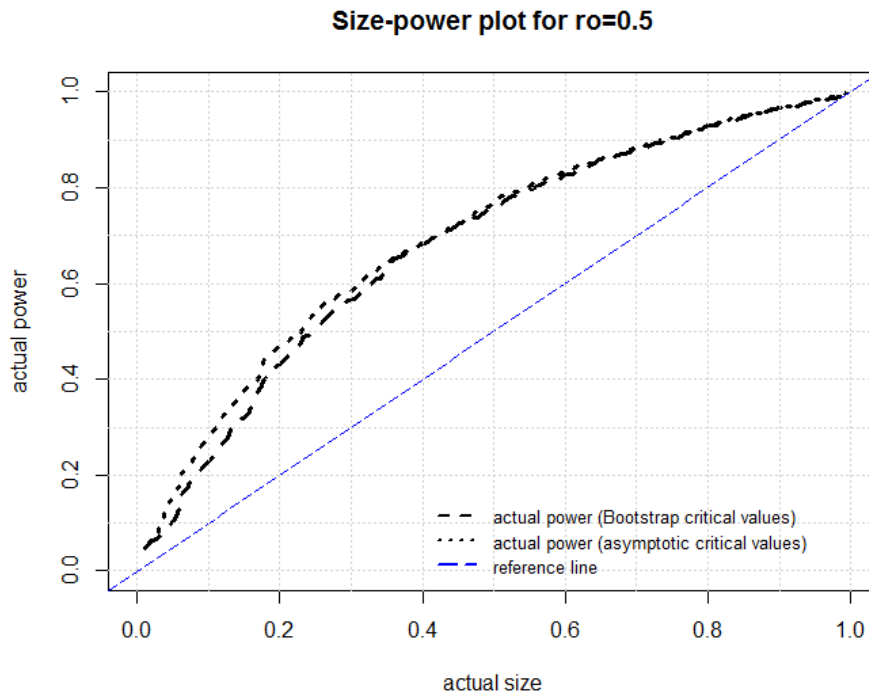
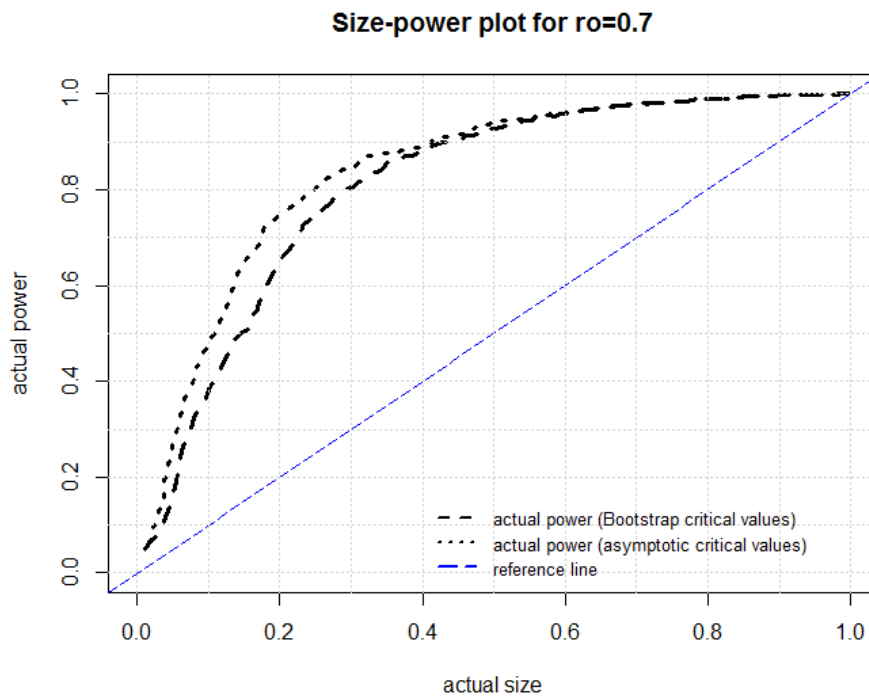


Figure 10: Size-power plot with actual size for the Hausman statistic (Within vs. Amemiya):  $\rho=0.700$ ,  $\rho^* = 0.461$ , 1000 replications. Critical values: Bootstrap and asymptotic.



#### 4.6.2. Size and power of Hausman test (Nerlove method for estimating random effects model)

Similarly to what was done in the previous section, when comparing the within estimates:  $\hat{\beta}_1^{FE} = 0.108$ ,  $\hat{\beta}_2^{FE} = 0.312$ , with the Nerlove estimates from the random effects model:  $\hat{\beta}_0^{RE,nerlove} = -57.507$ ,  $\hat{\beta}_1^{RE,nerlove} = 0.108$ ,  $\hat{\beta}_2^{FE,nerlove} = 0.310$ , Hausman test is performed:

*Table 5: Hausman test: Within vs. Nerlove methods of estimating the panel data model*

Hausman Test (Within vs. Nerlove method)	
$\chi^2$ - statistic	1.3073
p-value for $\chi^2$ (2)	0.5201

The null hypothesis cannot be rejected for all chosen levels of significance, since  $0.5201 > 0.10$ . It is interesting to see then whether Hausman test under-rejects the null hypothesis.

Table 6 shows that with Nerlove method the Hausman test would almost never reject the null hypothesis for small values of the nominal size when constant term is used in the models. And this applies even if the alternative hypothesis is actually true, which explains the 0 values of the power for nominal size 0.01 and 0.05. Even for size 0.10, the power remains smaller than the nominal size. In none of the 1000 replications of the Monte Carlo experiment was the null hypothesis rejected for significance levels 0.01 and 0.05. The actual size gets its first value above 0 for nominal size 0.08. This behavior can be observed in Figure 11, which illustrates how Hausman test, based on asymptotic critical values under-rejects the null hypothesis. The under-rejection starts getting smaller for nominal size bigger than 0.10 and this can be seen from Figure 12. The bootstrap critical values on the other hand lead to over-rejection of the null hypothesis or at least for nominal size equal or bigger than 0.05. However the over-rejection is not too big and it tends to follow the upper limit of the confidence interval of the nominal size. The small sample Monte Carlo critical values are accurate for estimating small sample properties. However, it must be noted that they also follow the upper limit of the nominal size. For nominal size 0.01, 0.05 and 0.10 the Monte Carlo critical values lead to actual size within the confidence interval of the nominal.

Concerning the test above, it seems that the null hypothesis wouldn't have been rejected for all levels 0.01, 0.05 and 0.10 even if the actual size were used.

Table 6: Actual size and power of Hausman test, which compares the estimates of the within method (fixed effects) vs. Nerlove method (random effects) of estimating the panel data model

Size and power of Hausman test (Nerlove method for estimating random effects model)											
	$\rho$	$\rho^*$	Using asymptotic critical values			Using Bootstrap critical values			Using Monte Carlo critical values		
Nominal Size	-	-	0.010	0.050	0.100	0.010	0.050	0.100	0.010	0.050	0.100
Actual Size	0	0	0.000	0.000	0.008	0.011	0.071	0.128	0.009	0.057	0.119
Power	0.20	0.125	0.000	0.000	0.008	0.013	0.079	0.140	0.013	0.070	0.137
Power	0.40	0.259	0.000	0.000	0.021	0.030	0.134	0.242	0.027	0.133	0.234
Power	0.50	0.327	0.000	0.000	0.032	0.051	0.200	0.347	0.038	0.183	0.344
Power	0.70	0.461	0.000	0.000	0.092	0.127	0.471	0.682	0.115	0.468	0.665

Figure 11: Zoomed in p-value plot for the Hausman statistic (Within vs. Nerlove):  $\rho = 0$ , 1000 replications. Critical values: Monte Carlo, Bootstrap and asymptotic.

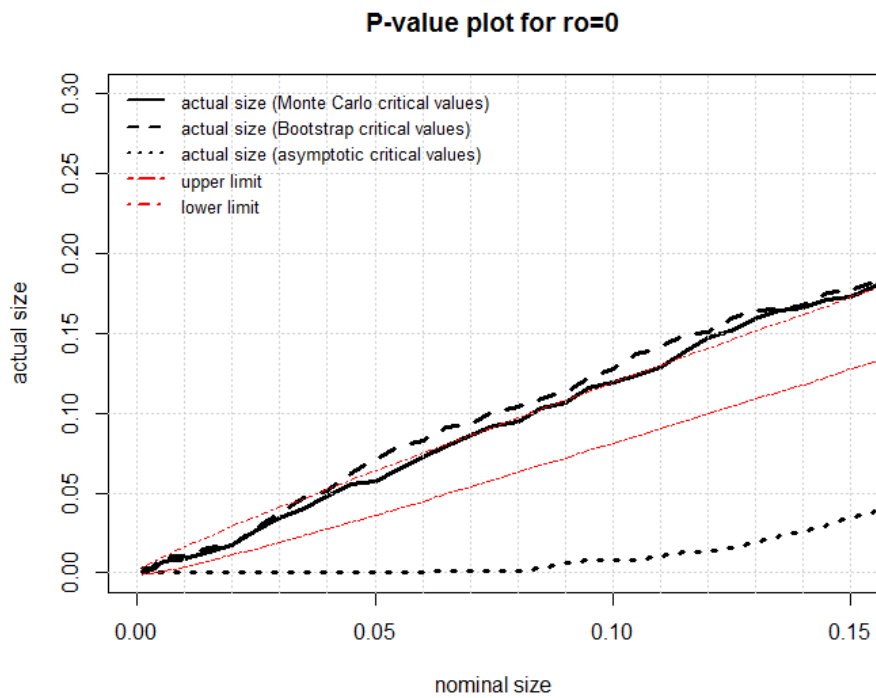
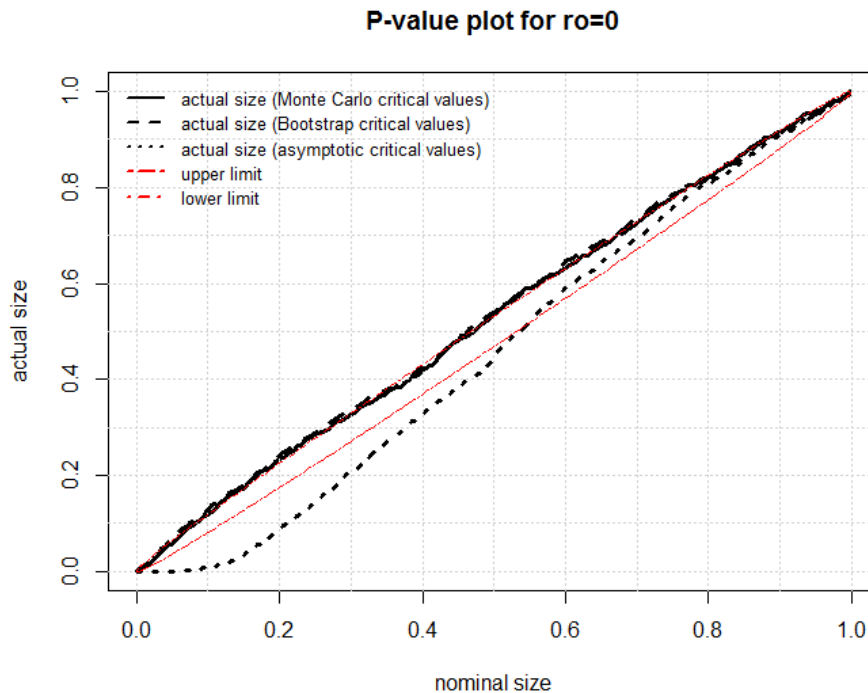


Figure 12: P-value plot for the Hausman statistic (Within vs. Nerlove):  $\rho=0$ , 1000 replications. Critical values: Monte Carlo, Bootstrap and asymptotic.



As in the previous section Table 6 as well as Figures 13-16 map the power with the nominal size for different correlation coefficients and Figures 14-20 plot the power and the actual size. The asymptotic critical values produce extremely poor power of the Hausman test (even equal to 0), corresponding to the nominal size. The power starts increasing after nominal size of 0.10. The null hypothesis wasn't rejected in any of the 1000 replications for nominal size 0.01 and 0.05. The size is relatively small when using Bootstrap and Monte Carlo critical values too, although they are not as extreme as in the case with asymptotic critical values and it increases more rapidly with the increase of the nominal size. Especially small is the power for nominal size 0.01. From the size-adjusted power figures, one can see that none of the critical values guarantees better power of Hausman test, which means that no power is lost by using bootstrap critical values over asymptotic. The power in general is quite low for small values of the size. For actual size of 0.100 naturally, the biggest power estimate was obtained for correlation coefficient 0.70 between  $\alpha_i$  and  $\bar{C}_i$ . and corresponding correlation coefficient of 0.461 between  $\alpha_i$  and  $C_{it}$ . Then the power is around 0.6. This power values are bigger than the ones obtained in the previous section, when Amemiya method is used instead of Nerlove.

Figure 13: Size-power plot with nominal size for the Hausman statistic (Within vs. Nerlove):  $\rho=0.200$ ,  $\rho^* = 0.125$ , 1000 replications. Critical values: Monte Carlo, Bootstrap and asymptotic.

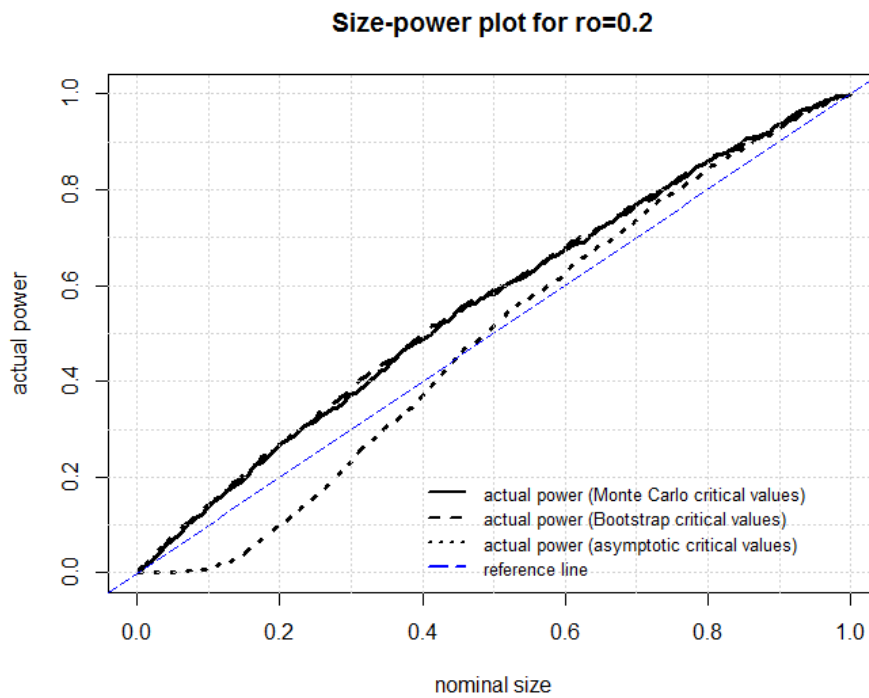


Figure 14: Size-power plot with nominal size for the Hausman statistic (Within vs. Nerlove):  $\rho=0.400$ ,  $\rho^* = 0.259$ , 1000 replications. Critical values: Monte Carlo, Bootstrap and asymptotic.

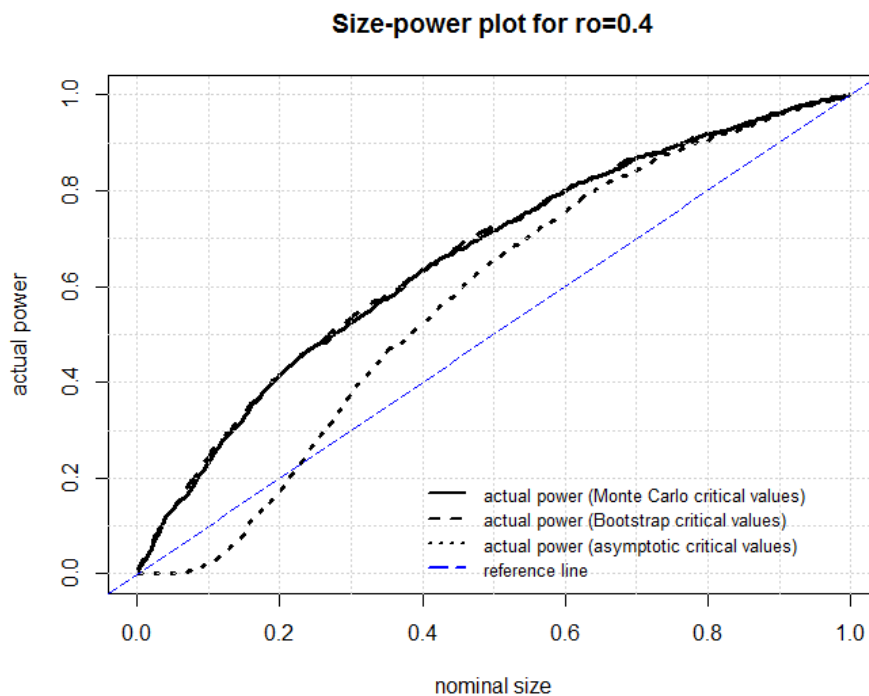


Figure 15: Size-power plot with nominal size for the Hausman statistic (Within vs. Nerlove):  $\rho=0.500$ ,  $\rho^* = 0.327$ , 1000 replications. Critical values: Monte Carlo, Bootstrap and asymptotic.

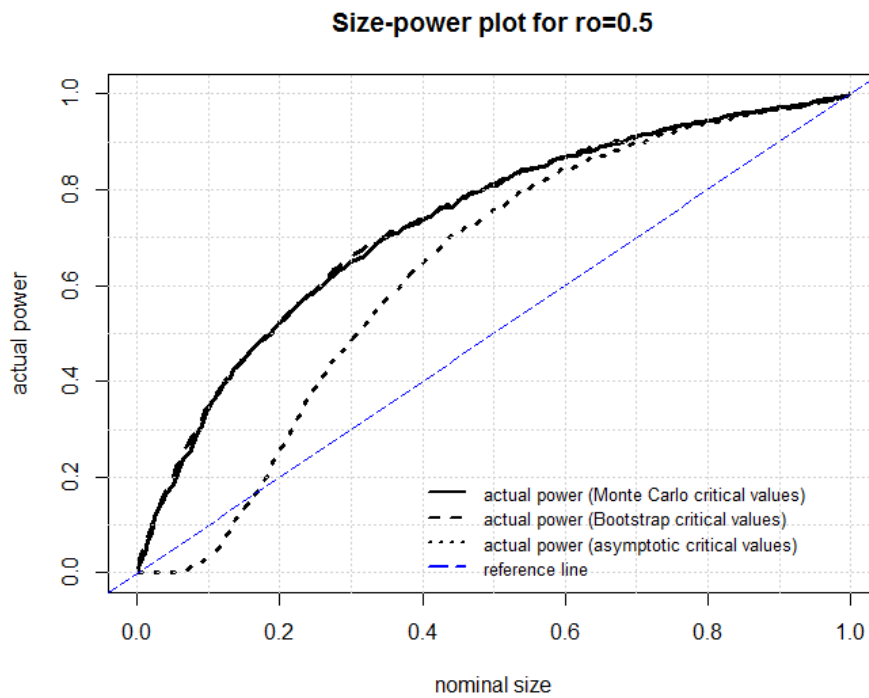


Figure 16: Size-power plot with nominal size for the Hausman statistic (Within vs. Nerlove):  $\rho=0.700$ ,  $\rho^* = 0.461$ , 1000 replications. Critical values: Monte Carlo, Bootstrap and asymptotic.

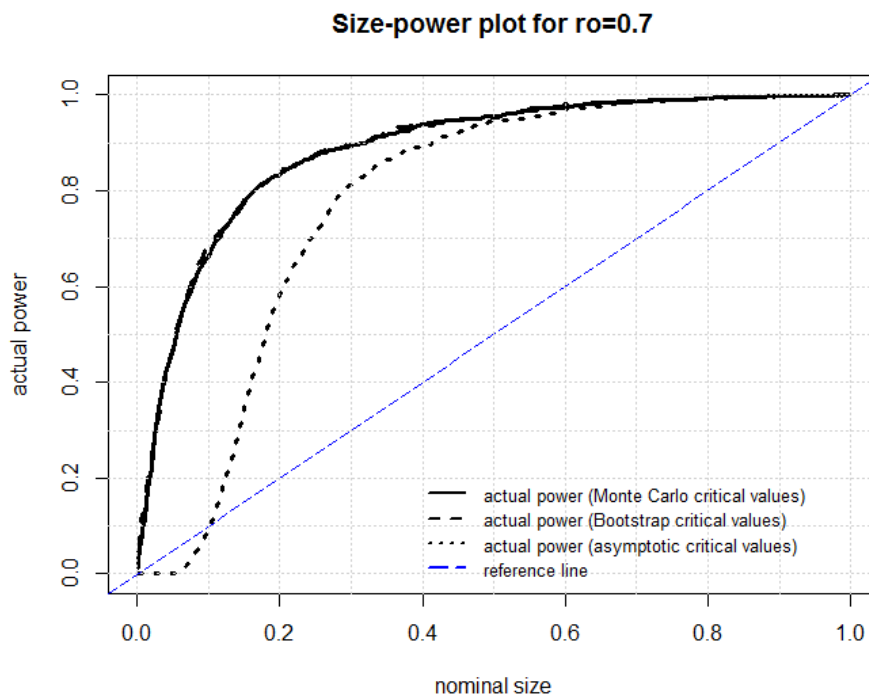


Figure 17: Size-power plot with actual size for the Hausman statistic (Within vs. Nerlove):  $\rho=0.200$ ,  $\rho^* = 0.125$ , 1000 replications. Critical values: Bootstrap and asymptotic.

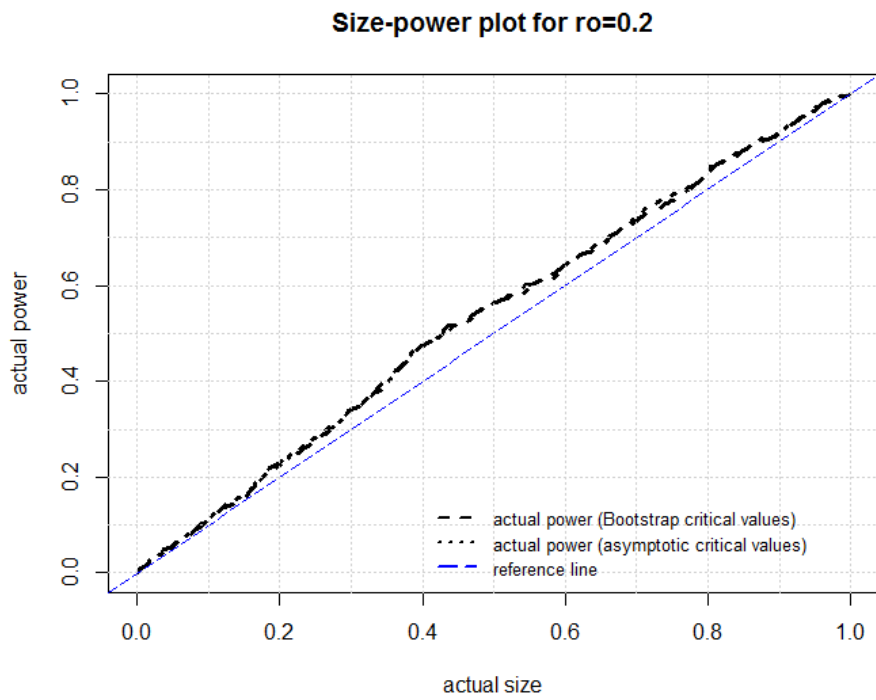


Figure 18: Size-power plot with actual size for the Hausman statistic (Within vs. Nerlove):  $\rho=0.400$ ,  $\rho^* = 0.259$ , 1000 replications. Critical values: Bootstrap and asymptotic.

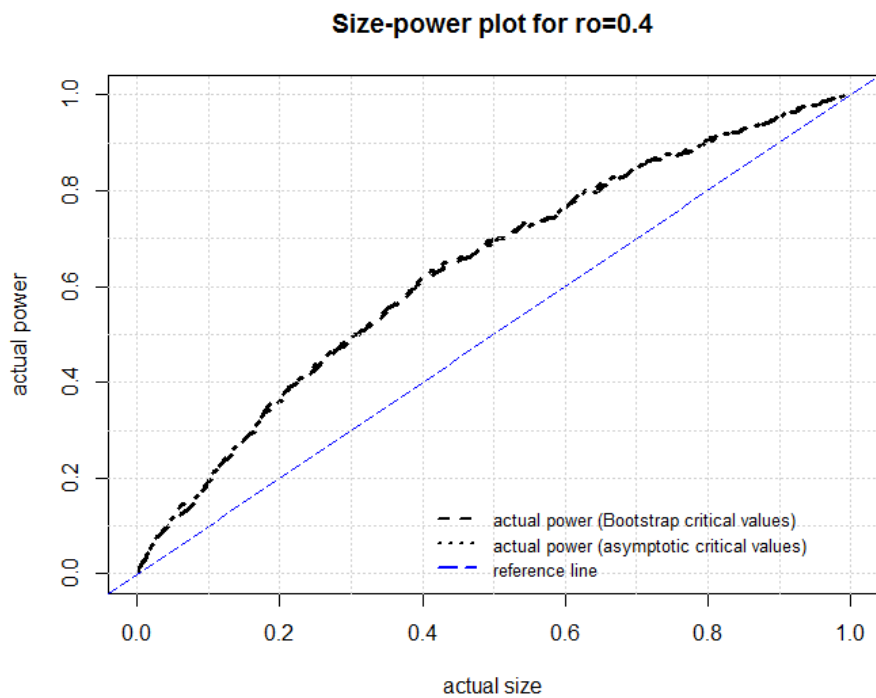


Figure 19: Size-power plot with actual size for the Hausman statistic (Within vs. Nerlove):  $\rho=0.500$ ,  $\rho^* = 0.327$ , 1000 replications. Critical values: Bootstrap and asymptotic.

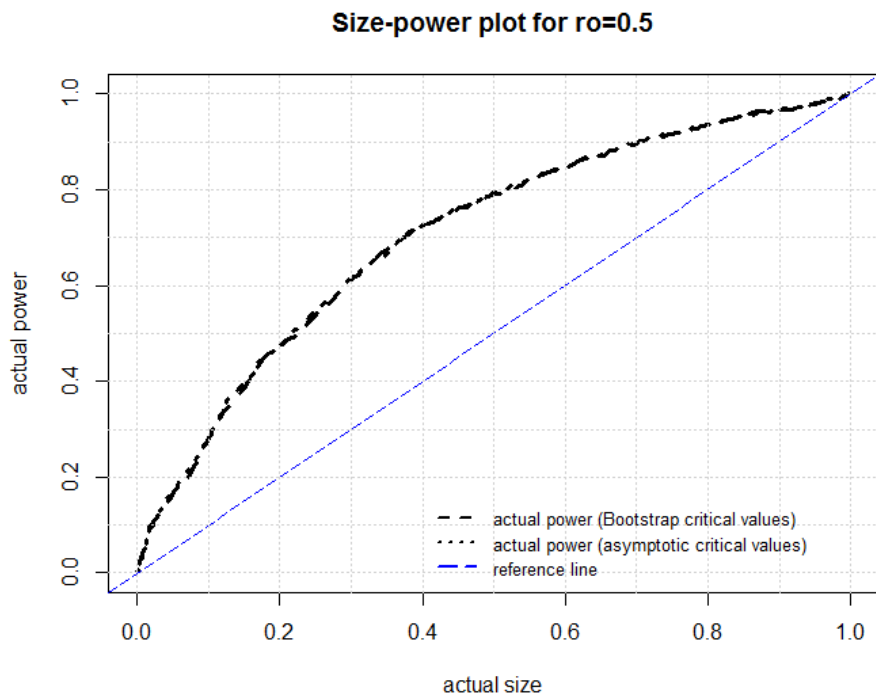
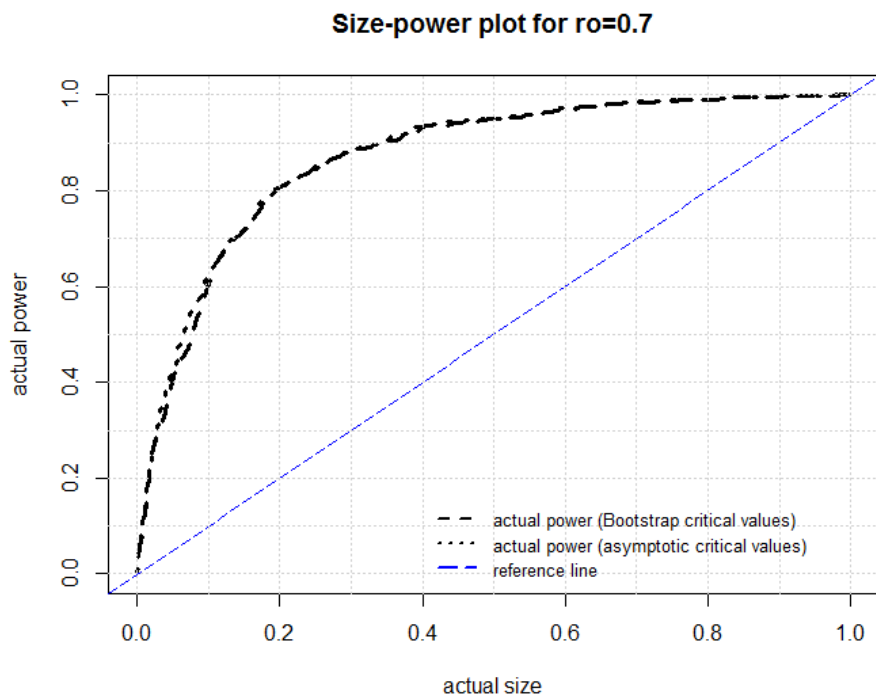


Figure 20: Size-power plot with actual size for the Hausman statistic (Within vs. Nerlove):  $\rho=0.700$ ,  $\rho^* = 0.461$ , 1000 replications. Critical values: Bootstrap and asymptotic.



4.6.3. Size and power of Hausman test (Swamy and Arora method for estimating random effects model)

In the Swamy and Arora method for estimating the random effects model, the Hausman test compares the following estimates:  $\hat{\beta}_1^{FE} = 0.108$ ,  $\hat{\beta}_2^{FE} = 0.312$  (for the within method) and  $\hat{\beta}_0^{RE,swar} = -57.449$ ,  $\hat{\beta}_1^{RE,swar} = 0.108$ ,  $\hat{\beta}_2^{FE,nerlove} = 0.310$ , Hausman test is performed:

Table 7: Hausman test: Within vs. Swamy-Arora methods of estimating the panel data model

Hausman Test (Within vs. Swamy-Arora method)	
$\chi^2$ - statistic	2.308
p-value for $\chi^2$ (2)	0.3154

The null hypothesis cannot be rejected for any of the chosen levels of significance, since  $0.3154 > 0.10$ . According to the test, fixed effects model should be used.

From Table 8 one can see that the asymptotic critical values when Swamy-Arora method of estimation is used also lead to over-rejection of the null hypothesis, similar to the Amemiya method. Figure 22 shows that the asymptotic critical values stop over-rejecting for values around 0.35 and above. The p-value from Table 7 is 0.3154 (close to 0.35) and thus the actual p-value is very similar to the one obtained. In this particular case the asymptotic critical values wouldn't affect negatively the outcome of the result.

The small sample Monte Carlo critical values produce very accurate actual size. After nominal size of 0.5 the Monte Carlo critical values lead to an actual size that follows the upper bound of the confidence interval of the nominal size, whereas the asymptotic actual size follow the lower limit. However, in practice, the interest in the size is limited on a level of up to 0.10 unless the p-value is regarded. Also, the use of Monte Carlo critical values in practice is restricted and dubious due to the dependency of the Monte Carlo results on the DGP.

Table 8: Actual size and power of Hausman test, which compares the estimates of the within method (fixed effects) vs. Swamy-Arora method (random effects) of estimating the panel data model

Size and power of Hausman test (Swamy-Arora method for estimating random effects model)								
	$\rho$	$\rho^*$	Using asymptotic critical values			Using Monte Carlo critical values		
<b>Nominal Size</b>	-	-	<b>0.010</b>	<b>0.050</b>	<b>0.100</b>	<b>0.010</b>	<b>0.050</b>	<b>0.100</b>
<b>Actual Size</b>	0	0	0.087	0.139	0.194	0.004	0.047	0.097
<b>Power</b>	0.20	0.125	0.123	0.185	0.229	0.012	0.056	0.140
<b>Power</b>	0.40	0.259	0.163	0.252	0.325	0.015	0.096	0.185
<b>Power</b>	0.50	0.327	0.205	0.308	0.390	0.015	0.101	0.222
<b>Power</b>	0.70	0.461	0.317	0.466	0.563	0.028	0.174	0.335

Figure 21: Zoomed in  $p$ -value plot for the Hausman statistic (Within vs. Swamy-Arora):  $\rho = 0$ , 1000 replications. Critical values: Monte Carlo and asymptotic.

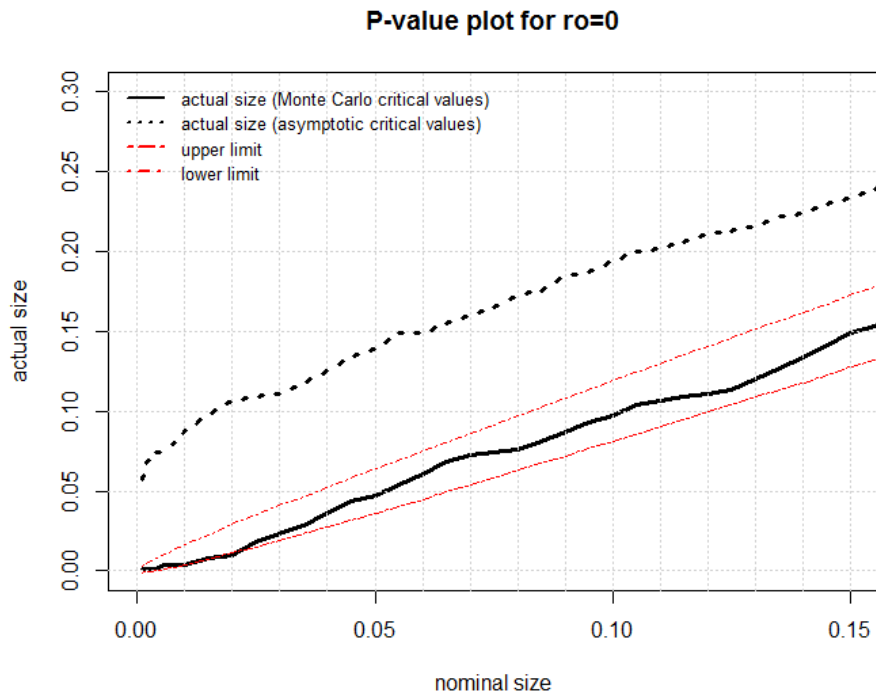
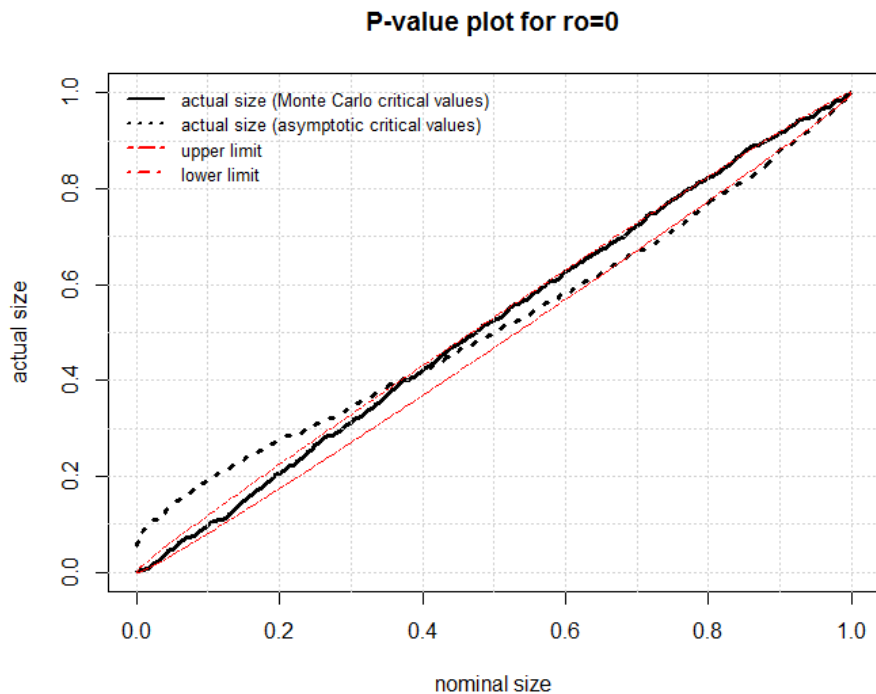


Figure 22:  $P$ -value plot for the Hausman statistic (Within vs. Swamy-Arora):  $\rho=0$ , 1000 replications. Critical values: Monte Carlo and asymptotic.



Since the difference between the actual size estimates when using asymptotic and Monte Carlo critical values is big only for smaller values of the nominal size, the power also differs mainly for those values in Table 8 and in Figures 23-26. For bigger values of the nominal size, the power doesn't differ so much. As in the previous used methods, the power is generally

low, especially when using Monte Carlo critical values, where it is only 0.028 for nominal size 0.01 and correlation coefficient 0.70 between  $\alpha_i$  and  $\bar{C}_i$ . corresponding to correlation coefficient of 0.461 between  $\alpha_i$  and  $C_{it}$ .

Figure 23: Size-power plot with nominal size for the Hausman statistic (Within vs. Swamy-Arora):  $\rho=0.200$ ,  $\rho^* = 0.125$ , 1000 replications. Critical values: Monte Carlo and asymptotic.

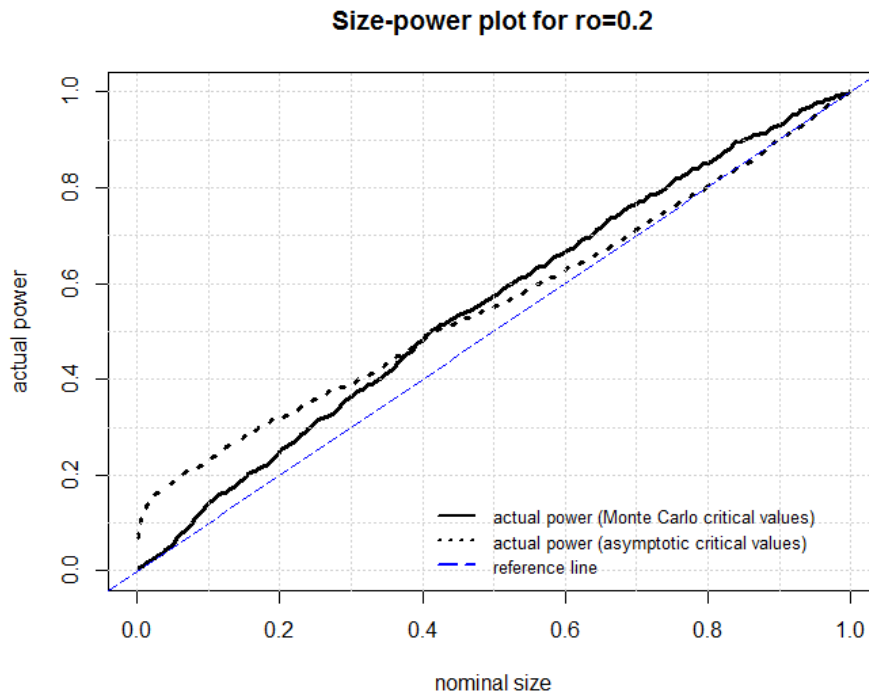


Figure 24: Size-power plot with nominal size for the Hausman statistic (Within vs. Swamy-Arora):  $\rho=0.400$ ,  $\rho^* = 0.259$ , 1000 replications. Critical values: Monte Carlo and asymptotic.

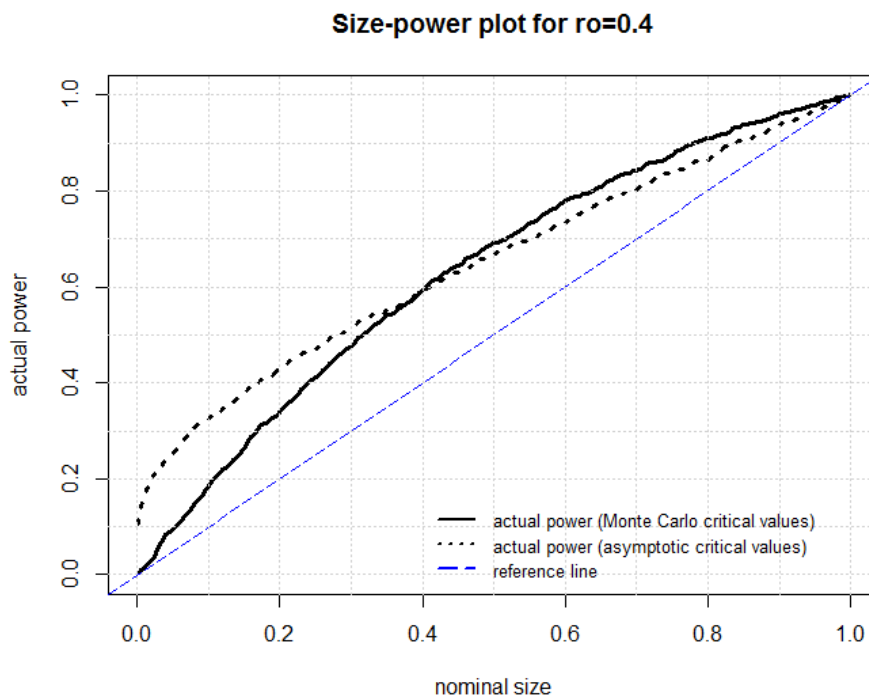


Figure 25: Size-power plot with nominal size for the Hausman statistic (Within vs. Swamy-Arora):  $\rho=0.500$ ,  $\rho^*=0.327$ , 1000 replications. Critical values: Monte Carlo and asymptotic.

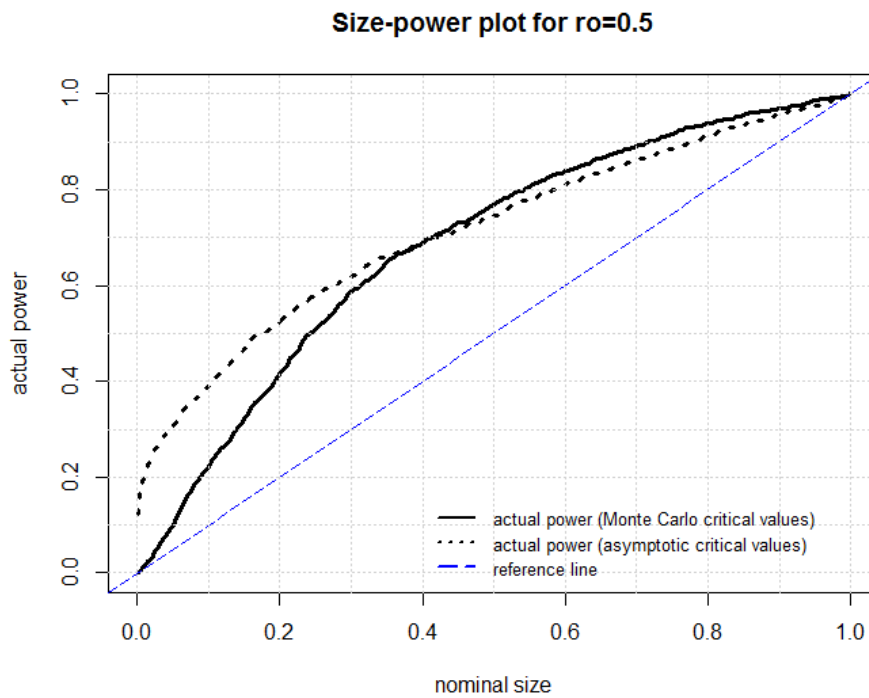
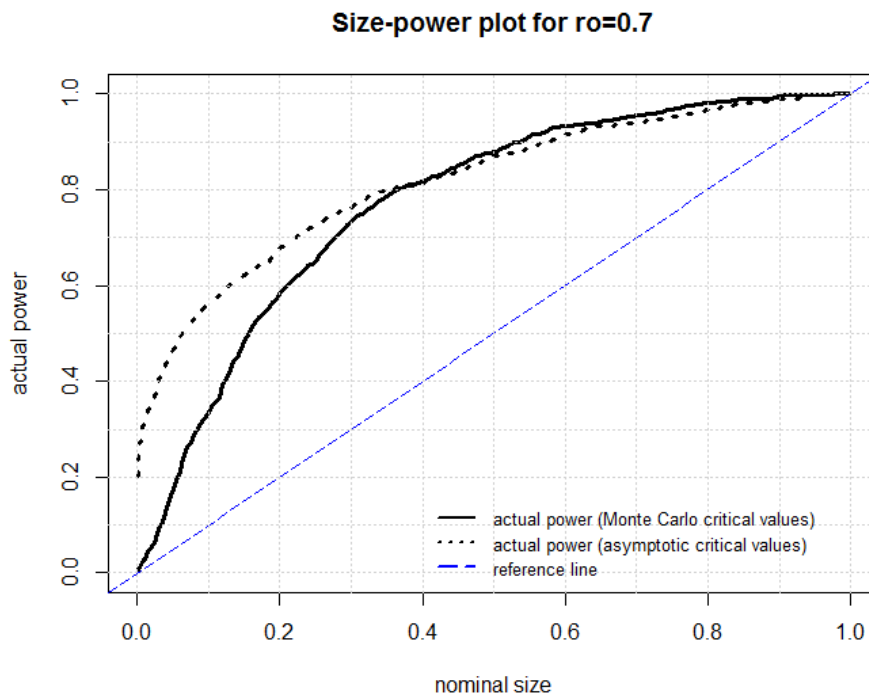


Figure 26: Size-power plot with nominal size for the Hausman statistic (Within vs. Swamy-Arora):  $\rho=0.700$ ,  $\rho^*=0.461$ , 1000 replications. Critical values: Monte Carlo and asymptotic.



#### 4.7. Discussion

As expected there was a difference between the estimated properties of the Hausman test, when different methods for estimating the random effects model were used. However, in all

of the methods, the small sample Monte Carlo critical values seem to give actual size close to the nominal. The Bootstrap critical values lead to slight over-rejection, but are still much more accurate than asymptotic critical values. Since the procedure was applied by using only one of the suggested by Andersson and Karlsson (2001) methods of bootstrapping, one must consider the possibility of obtaining critical values by implementing any of the other suggested methods. Since the individual-specific component is being bootstrapped, its variance changes in the bootstrap samples, which affects the estimation of the random effects model during the bootstrapping. Therefore, better approximation of the bootstrap actual size to the nominal would be achieved, when more individuals (firms) are included in the model and more variability is achieved during the sampling. Only 10 companies are used in the example of this thesis, which affects the bootstrapping performance. An increase of the number of bootstrap samples should also improve the bootstrapping procedure. Previous work done on bootstrapping for obtaining the true empirical distribution of Hausman test has proved the usefulness of the method, specifically because the Hausman statistic is pivotal (does not depend on unknown parameters). And the bootstrap indeed improves the properties of the Hausman test compared to using asymptotic critical values, when applied for the Amemiya and Nerlove methods of estimation. Due to the specific way the package 'plm' in R is programmed however, bootstrapping was impossible for obtaining critical values when using Swamy and Arora method of estimation. Nevertheless, the small sample Monte Carlo critical values provide an opportunity for estimating the properties of Hausman test within the simulation. The computation of the Monte Carlo critical values is quick and straightforward.

Worth mentioning in this section is the fact there was no exact way to control the precise correlation between  $\alpha_i$  and  $C_{it}$ . Instead only the correlation between  $\alpha_i$  and  $\bar{C}_i$  was controlled and the first one could only be observed. This indirect way of manipulating the correlation probably means that in reality the same results that were obtained for the power and size would correspond to slightly more different correlation coefficients. However, it was still possible to observe the increase in the power, when the correlation is increased and the exact value of this correlation is not so consequential.

#### 4.8. Recommendations

A subsequent work can answer the question whether there would be any difference in the results and how big would it be, if the individual-specific components were correlated with  $F_i$  instead of with  $C_i$ .

Also, different software could be used for obtaining bootstrap critical values, when Swamy and Arora method is used.

Another guideline for future work in the area would be the estimation of the size and power when different time periods and number of individuals are used.

## 5. Summary and Conclusion

### 5.1. Summary

A procedure for obtaining the actual size and power of the Hausman test was suggested and illustrated based on data, used for implementing the investment function.

Three types of critical values were used for the Hypothesis testing: asymptotical, small sample Monte Carlo (based on pure simulation) and Bootstrap (based on simulation and bootstrapping) critical values. The conditions tested in the Hausman test were controlled in a Monte Carlo simulation of 1000 replications in order to manipulate the correct outcome of the test and count how many times a wrong or correct output was returned. By counting the average number of the false rejections of the null hypothesis the size of Hausman test was estimated and contrariwise the power was estimated based on the average number of correct rejections.

The simulation mimics the original data as close as possible in order to make inferences specifically for the data at hand, but controls the correlation between one of the variables and the individual-specific component in the panel data model.

### 5.2. Conclusion

The suggested procedure works as expected with no serious deviations from previous work done in the area in terms of identifying the size and power of Hausman test.

Most noticeably, using asymptotic critical values when performing Hausman test on the available data would lead to over-rejection of the null hypothesis when Swamy and Arora and Amemiya methods of estimating the random effects model are used. Hausman test would tend indicating that the fixed effects model should be used more than it has to and it would produce a bigger risk of type I error than specified. When Nerlove method is used for estimating the random effects, the asymptotic critical values lead to extreme under-rejection of the null hypothesis with very close to 0 values of the size and even power.

For small samples Monte Carlo critical values produce actual size similar to the nominal.

A slight over-rejection of the null hypothesis was observed, when bootstrap critical values were used. Nonetheless, the bootstrap actual size follows the upper limit of the nominal size and perhaps by increasing the number of bootstrap samples, the procedure would provide more accurate results, provided the small number of companies. A noticeable problem of the bootstrapping was the high possibility of obtaining negative estimate of the variance of the individual-specific component in at least one run of the 299 000 loops of the bootstrap cyclical procedure. Problem, which cannot be sidestepped when using Swamy and Arora method of estimation of the random effects model in R.

The estimates of the power of Hausman test show high risk of type II error for small value of the size. The power of Hausman test is considerably low, at least when a constant term is used in the modelling. The correlation between the individual-specific components and the regressors should be very high in order to achieve adequate power levels.

## 6. Bibliography

- Amemiya, T. (1971). The Estimation of Variances in Variance Components Model. *International Economic Review* , 12, 1-13.
- Andersson, M.K., Karlsson, S. (2001). Bootstrapping Error Component Models. *Computational Statistics* , 16, 221-231.
- Bole, V., Rebec, P. (2013). Bootstrapping the Hausman Test in Panel Data Models. *Communications in Statistics - Simulation and Computation* , 42 (3), 650-670.
- Boot, J.C.G.; de Wit, G.M. (1960). Investment Demand: An Empirical Contribution to the Aggregation Problem. *International Economic Review* , 27-28.
- Christian Kleiber & Achim Zeileis. (2010). The Grunfeld Data at 50. *German Economic Review, Verein für Socialpolitik* , 11, 404-417.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. New York: Lawrence Erlbaum Associates, Publishers.
- Davidson, R., & MacKinnon, J. (1999). Bootstrap Testing in Nonlinear Models. *International Economic Review* , 40(2), 487-508.
- Davidson, Russell & MacKinnon, James G. (1998). Graphical Methods for Investigating the Size and Power of Hypothesis Tests. *The Manchester School of Economic & Social Studies* , 66(1), 1-26.
- Frees, E. W. (2004). *Longitudinal and Panel Data: Analysis and Applications in the Social Sciences*. Cambridge: Cambridge University Press.
- Greene, W. (2008). *Econometric Analysis* (6th ed.). Upper Saddle River: NJ: Prentice Hall.
- Greene, W. H. (2012). *Econometric Analysis: International Edition* (7th ed.). Boston: London: Pearson.
- Grunfeld, Y., & Griliches, Z. (1960, February). Is Aggregation Necessarily Bad? *Review of Economics and Statistics* , 1-13.
- Hausman, J. (1978). Specification Tests in Econometrics. *Econometrica* , 46, 1251-1271.
- Hausman, J., & Taylor, W. (1981). Panel Data and Unobservable Individual Effects. *Econometrica* , 49, 1377-1398.
- Herwartz, H., & Neumann, M. H. (2007). *A robust bootstrap approach to the Hausman test in stationary panel data models*. Kiel: Christian-Albrechts-University of Kiel, Department of Economics.
- Jeong J., Yoon B. (2007). *The Effect of Pseudo-exogenous Instrumental Variables on Hausman Test*. Munich: University Library of Munich.
- Magazzini Laura, G. Calzolari. (2010). *Negative Variance Estimates in Panel Data Models*. Verona: University of Verona.
- Nerlove, M. (1971). Further Evidence on the Estimation of Dynamic Economic Relations from a Time Series of Cross Sections. *Econometrica* , 39(2), 359-82.
- Swamy, P.A.V.B. and S. S. Arora. (1972). The Exact Finite Sample Properties of Estimators of Coefficients in the Error Components Regression Models. *Econometrica* , 40, 253-260.

Wallace, T., & Hussain, A. (1969). The use of error components in combining cross section with. *Econometrica* , 37, 55-72.

Wong, K. (1996, November). Bootstrapping Hausman's exogeneity test. *Economics Letters* , 53 (2), pp. 139-143.

## Appendix

### 6.1. Fixed effects estimation methods

#### 6.1.1. Within-group method

As mentioned the fixed effects model is:

$$y_{it} = \alpha_i + \beta_1 x_{1,it} + \beta_2 x_{2,it} + \dots + \beta_k x_{k,it} + \varepsilon_{it}. \quad (15)$$

For the within-group method, first, one has to compute the means of all observed variables within individuals across time:

$$\bar{y}_{i\cdot} = \frac{1}{T} \sum_{t=1}^T y_{it}; \quad \bar{x}_{l,i\cdot} = \frac{1}{T} \sum_{t=1}^T x_{l,it} \quad l = 1, \dots, k.$$

Mean values are obtained for each individual and used instead of  $y_{it}$  and  $x_{l,it}$  ( $l = 1, \dots, k$ ) in equation (15). The equation takes the following form:

$$\bar{y}_{i\cdot} = \alpha_i + \beta_1 \bar{x}_{1,i\cdot} + \beta_2 \bar{x}_{2,i\cdot} + \dots + \beta_k \bar{x}_{k,i\cdot} + \bar{\varepsilon}_{i\cdot}. \quad (16)$$

The term  $\bar{\varepsilon}_{i\cdot}$  is considered to be 0 by assumption. Also, since  $\alpha_i$  is time-invariant, its mean across time would stay as the original value for each individual.

In the next step of the computation, equation (16) is subtracted from (15). The individual-specific component disappears. Also if a constant term had been used, it would have also disappeared. That is why the fixed effects model doesn't have a constant term:

$$y_{it} - \bar{y}_{i\cdot} = \beta_1 (x_{1,it} - \bar{x}_{1,i\cdot}) + \beta_2 (x_{2,it} - \bar{x}_{2,i\cdot}) + \dots + \beta_k (x_{k,it} - \bar{x}_{k,i\cdot}) + \varepsilon_{it} - \bar{\varepsilon}_{i\cdot}. \quad (17)$$

Let  $\tilde{y}_{it} = y_{it} - \bar{y}_{i\cdot}$ ,  $\tilde{x}_{l,it} = x_{l,it} - \bar{x}_{l,i\cdot}$  for  $l = 1, \dots, k$  and  $\tilde{\varepsilon}_{it} = \varepsilon_{it} - \bar{\varepsilon}_{i\cdot}$ . Then the equation can be written as:

$$\tilde{y}_{it} = \beta_1 \tilde{x}_{1,it} + \beta_2 \tilde{x}_{2,it} + \dots + \beta_k \tilde{x}_{k,it} + \tilde{\varepsilon}_{it}. \quad (18)$$

The computation of the parameters and the individual-specific component follows the formulas:

$$\hat{\beta}_l^{FE} = \frac{\sum \sum \tilde{x}_{l,it} \tilde{y}_{it}}{\sum \sum \tilde{x}_{l,it}^2}; \quad \hat{\alpha}_i^{FE} = \bar{y}_{i\cdot} - \hat{\beta}_1^{FE} \bar{x}_{1,i\cdot} - \hat{\beta}_2^{FE} \bar{x}_{2,i\cdot} - \dots - \hat{\beta}_k^{FE} \bar{x}_{k,i\cdot}.$$

The estimates are consistent.

The deviations around the mean of the dependent variable are regressed on the deviations around the mean of the independent variables. Thus, one can control the heterogeneity in the data. The degrees of freedom in this case are  $n(T - 1) - k$  and not  $nT - k$  as if they would have been if OLS is used on equation (18) with no regard of the fact that panel data is used.

A drawback in the estimation of the fixed effects model is the loss of the intercept term as well as the possibility to include in the model time-invariant variables even if they vary across individuals. Also, due to the change in the dependent variable, its variance would be smaller. This can lead to bigger measurement error bias if such errors are present.

### 6.1.2. Least squares dummy variable (LSDV) method

Another method for estimating the fixed effect parameters that will give equivalent results is the LSDV method. The model is then estimated as:

$$y_{it} = \sum_{i=1}^n \alpha_i A_i + \beta_1 x_{1,it} + \beta_2 x_{2,it} + \dots + \beta_k x_{k,it} + \varepsilon_{it}, \quad (19)$$

where  $A_i$  represent dummy variables that equal one for the  $i^{th}$  individual throughout all observations and zero for the other individuals. Model (19) can be estimated by using OLS. The components  $\alpha_i$  are treated as parameters to the additional variables  $A_i$ . Any intercept term must be dropped in order to avoid perfect multicollinearity with the dummy variables. This method is harder applied in case of a large number of individuals, but it has the benefit of directly obtaining  $n(T - 1) - k$  degrees of freedom.

## 6.2. Algorithm

1. Load the panel data: **companies, years, I, F, C**.
2. Set  $N=n*t$  ( $N=10*20=200$ ).
3. Compute  $\bar{C}_i$ . (mean within companies, across time for **C**).
4. Obtain  $\sigma_{\bar{C}_i}$ . (standard deviation).
5. Estimate the fixed effects model using the data: **I** – dependent, and **F** and **C** – independent variables. Obtain  $\hat{\beta}_1^{FE}$  and  $\hat{\beta}_2^{FE}$ .
6. Estimate the random effects model using the data: **I** – dependent, and **F** and **C** – independent variables. Obtain  $\hat{\beta}_1^{RE}$  and  $\hat{\beta}_2^{RE}$ .
7. Save the estimated variance of the individual-specific component in the random effects model:  $\sigma_{\alpha,RE}^2 = Var(\hat{\alpha}_i^{RE})$ .
8. Save the estimated variance of the residuals in the random effects model:  $\sigma_{\varepsilon,RE}^2 = Var(\hat{\varepsilon}_{it}^{RE})$ .
9. Compute the ratio between  $\sigma_{\bar{C}_i}$  and  $\sigma_{\alpha,RE}$ :  $m = \frac{\sigma_{\bar{C}_i}}{\sqrt{\sigma_{\alpha,RE}^2}}$ .
10. Perform Hausman test, based on the estimates  $\hat{\beta}^{RE}$ ,  $\hat{\beta}^{FE}$ .

---

11. Set a function for the Monte Carlo critical values (input:  $\hat{\beta}_1^{FE,RE} = 0.11$ ,  $\hat{\beta}_2^{FE,RE} = 0.31$ , number of replications=10 000){
  - 11.1. Set seed = 6290
  - 11.2. For(k from 1 to 10 000){
    - 11.2.1. Generate individual-specific components  $\alpha_i^{null}$  with length N (should be constant within companies, across time)  $\sim N(0, \sigma_{\bar{C}_i})$ .
    - 11.2.2. Generate errors:  $\varepsilon_{it} \sim N(0, m * \sigma_{\varepsilon,RE})$ .
    - 11.2.3. Compute  $I_{it}^* = \hat{\beta}_1^{FE,RE} F_{it} + \hat{\beta}_2^{FE,RE} C_{it} + \alpha_i^{null} + \varepsilon_{it}$
    - 11.2.4. Estimate the fixed effects model using as data: **I\*** – dependent, and **F** and **C** – independent variables. Save estimates  $\hat{\beta}^{FE,k}$ .

- 11.2.5. Estimate the random effects model with using as data:  $I^*$  – dependent, and  $F$  and  $C$  – independent variables. Save estimates  $\hat{\beta}^{RE,k}$ .
- 11.2.6. Perform Hausman test, based on  $\hat{\beta}^{RE,k}$ ,  $\hat{\beta}^{FE,k}$ .
- 11.2.7. Save the obtained statistic as the  $k^{\text{th}}$  element of a vector  $\mathbf{h}$ .
- }
- 11.3. Sort the elements in vector  $\mathbf{h}$ .
- 11.4. Create a vector with chosen values for the nominal size: vector  $\mathbf{l}$  with 215 elements: 0.001, 0.002, ..., 0.010 (10 values); 0.015, 0.020, ..., 0.985 (195 values); 0.990, 0.991, ..., 0.999 (10 values).
- 11.5. Find the indices for the critical values within vector  $\mathbf{h}$  by using the formula:  $ind_p = (1 - l_p) \cdot 10\,000$ ,  $p=1, \dots, 215$ .
- 11.6. Save the Hausman statistic from vector  $\mathbf{h}$  with indices  $ind_p$ . In total 215 statistics are saved and they represent the Monte Carlo generated critical values. They are saved in vector  $\mathbf{h}^{MC}$  of length 215.
- }
- 
12. Set a function for estimating the size and power of Hausman test (input:  $\rho$ ,  $\hat{\beta}_1^{FE,RE} = 0.1$ ,  $\hat{\beta}_2^{FE,RE} = 0.3$ , number of replications=1 000){
- 12.1. Set seed = 1234
- 12.2. For(k from 1 to 1 000){
- 12.2.1. Generate  $a_i$  with length N (should be constant within companies, across time)  $\sim N(0, \sigma_{\bar{c}_i})$ .
- 12.2.2. Use Cholesky decomposition to transform  $a_i$  into  $\alpha_i$  correlated with  $\bar{c}_i$ . with correlation  $\rho$ .
- 12.2.3. Generate errors:  $\varepsilon_{it} \sim N(0, m * \sigma_{\varepsilon,RE})$ .
- 12.2.4. Save the true value of the correlation coefficient between  $\alpha$  and  $C$  as the  $k^{\text{th}}$  element of vector  $\rho^*$  with length 1 000 (the number of replications).
- 12.2.5. Compute  $I_{it}^* = \hat{\beta}_1^{FE,RE} F_{it} + \hat{\beta}_2^{FE,RE} C_{it} + \alpha_i + \varepsilon_{it}$
- 12.2.6. Estimate the fixed effects model using as data:  $I^*$  – dependent, and  $F$  and  $C$  – independent variables. Save estimates  $\hat{\beta}^{FE,k}$ .
- 12.2.7. Save the errors  $\hat{\varepsilon}_{it}^{FE,k}$  and individual-specific components  $\hat{\alpha}_i^{FE,k}$ , ( $k = 1, \dots, r$ ) from the fixed effects model estimated above.
- 12.2.8. Compute  $\gamma = \sqrt{\frac{nT}{n(T-1)-k}}$ .
- 12.2.9. Center the errors:  $\hat{\eta}_{it}^k = (\hat{\varepsilon}_{it}^{FE,k} - \bar{\varepsilon}^{Fe,k})\gamma$ .
- 12.2.10. Center the estimated individual-specific components:  $\hat{\omega}_i^k = (\hat{\alpha}_i^{FE,k} - \bar{\alpha}^{Fe,k})\gamma$

- 12.2.11. Estimate the random effects model with using as data:  $I^*$  – dependent, and  $F$  and  $C$  – independent variables. Save estimates  $\hat{\beta}^{RE,k}$ .
- 12.2.12. Perform Hausman test, based on  $\hat{\beta}^{RE,k}$ ,  $\hat{\beta}^{FE,k}$ .
- 12.2.13. Save the obtained statistic as the  $k^{\text{th}}$  element of a vector  $\mathbf{h}$ .
- 12.2.14. For(j in 1:299){
- 12.2.14.1. Sample indices for the bootstrap samples for  $\hat{\eta}_{it}^k$ : row indices for selecting companies and indices, indicating the time within the selected companies.
- 12.2.14.2. The centred errors  $\hat{\eta}_{it}^k$ , which correspond to the sampled indices are included in the  $j^{\text{th}}$  bootstrap sample. The sample is a vector of elements:  $\hat{\eta}_{j,it}^*$ .
- 12.2.14.3. Sample indices for the bootstrap samples for  $\hat{\omega}_i^k$ : row indices for selecting companies.
- 12.2.14.4. The centred components  $\hat{\omega}_i^k$ , which correspond to the sampled indices are included in the  $j^{\text{th}}$  bootstrap sample. The sample is a vector of elements:  $\hat{\omega}_{j,i}^k$ .
- 12.2.14.5. Compute  $I_{j,it}^{Boot,k} = \hat{\beta}_1^{RE,k} F_{it} + \hat{\beta}_2^{RE,k} C_{it} + \hat{\omega}_{j,i}^k + \hat{\eta}_{j,it}^*$ .
- 12.2.14.6. Estimate the fixed effects model using as data:  $I^{Boot,k}$  – dependent, and  $F$  and  $C$  – independent variables. Save estimates  $\hat{\beta}_j^{FE,q}$ .
- 12.2.14.7. Estimate the random effects model using as data:  $I^{Boot,k}$  – dependent, and  $F$  and  $C$  – independent variables. Save estimates  $\hat{\beta}_j^{RE,q}$ .
- 12.2.14.8. Perform Hausman test, based on  $\hat{\beta}_j^{FE,q}$  and  $\hat{\beta}_j^{RE,q}$ .
- 12.2.14.9. Save the obtained statistic as element  $j$  of a replication specific vector  $\mathbf{h}_k^B$  ( $k$  for each replication and  $j$  for each bootstrap sample).
- }
- 12.2.15. Sort the elements of vector  $\mathbf{h}_k^B$  (done separately for each replication).
- 12.2.16. Create a vector with chosen values for the nominal size: vector  $\mathbf{l}$  with 215 elements: 0.001, 0.002,...,0.010 (10 values); 0.015, 0.020,...,0.985 (195 values); 0.990,0.991,...,0.999 (10 values).
- 12.2.17. Find the indices for the critical values within vector  $\mathbf{h}_k^B$  by using the formula:  $ind_p = (1 - l_p) \cdot 299$  (rounded to the next integer);  $p=1, \dots, 215$ .
- 12.2.18. Save the Hausman statistic from vector  $\mathbf{h}_k^B$  with indices  $ind_p$ . In total 215 statistics are saved in each replication and they represent the Monte Carlo generated critical values for the specific replication. The critical values are saved in a matrix  $\mathbf{H}^B$  with dimensions [1 000, 215]

```

12.2.19. Save the asymptotic critical values of the  $\chi^2$ -distribution with  $df_1 = 1 - l_p$  and
         $df_2 = 2$ . The asymptotic critical values are saved in a vector  $\mathbf{h}^A$  of length 215.
12.2.20. For(i from 1 to 215){
    12.2.20.1. If( $h_k > h_i^{MC}$ ){
        12.2.20.1.1. The element [k,i] of a counter matrix  $\mathbf{C}^{MC}$  with dimensions
            1 000 x 215 takes the value 1. Otherwise 0.
    }
    12.2.20.2. If( $h_k > H_{k,i}^B$ ){
        12.2.20.2.1. The element [k,i] of a counter matrix  $\mathbf{C}^B$  with dimensions
            1 000 x 215 takes the value 1. Otherwise 0.
    }
    12.2.20.3. If( $h_k > h_i^A$ ){
        12.2.20.3.1. The element [k,i] of a counter matrix  $\mathbf{C}^A$  with dimensions
            1 000 x 215 takes the value 1. Otherwise 0.
    }
}
}

```

12.3. Calculate the mean correlation between  $\alpha$  and  $\mathbf{C}$  as  $\bar{\rho}^*$ .

12.4. Estimate the size/power when using Monte Carlo critical values by calculating the average number of rejections of the null hypothesis for each critical value (that correspond to a specific nominal size), or the column means of matrix  $\mathbf{C}^{MC}$ . Save the results in a vector of 215 elements:  $\mathbf{SP}^{MC}$

12.5. Prepare for output specifically elements with indices 10, 18 and 28, which represent the actual size/power corresponding to nominal size of 0.01, 0.05 and 0.10 respectively.

12.6. Estimate the size/power when using Bootstrap critical values by calculating the average number of rejections of the null hypothesis for each critical value (that correspond to a specific nominal size), or the column means of matrix  $\mathbf{C}^B$ . Save the results in a vector of 215 elements:  $\mathbf{SP}^B$

12.7. Prepare for output specifically elements with indices 10, 18 and 28, which represent the actual size/power corresponding to nominal size of 0.01, 0.05 and 0.10 respectively.

12.8. Estimate the size/power when using asymptotic critical values by calculating the average number of rejections of the null hypothesis for each critical value (that correspond to a specific nominal size), or the column means of matrix  $\mathbf{C}^A$ . Save the results in a vector of 215 elements:  $\mathbf{SP}^A$

12.9. Prepare for output specifically elements with indices 10, 18 and 28, which represent the actual size/power corresponding to nominal size of 0.01, 0.05 and 0.10 respectively.

12.10. Compute the upper and lower limits of the confidence intervals for the nominal size.

When checking whether the actual size over-rejects or under-rejects the null hypothesis consider those intervals.

The nominal size is:  $x_j = 0.001, 0.002, \dots, 0.010$  (10 values);  $0.015, 0.020, \dots, 0.985$  (195 values);  $0.990, 0.991, \dots, 0.999$  (10 values).

The lower limit is calculated as:  $x_j^L = x_j - 2\sqrt{x_j \cdot (1 - x_j)/1000}$  (1000 = number of replications).

The upper limit is calculated as:  $x_j^U = x_j + 2\sqrt{x_j \cdot (1 - x_j)/1000}$  (1000 = number of replications).

12.11. If( $\rho=0$ ) {

12.11.1. Plot the nominal size ( $x_j$ ) on the x-axis and the actual size ( $SP^{MC}$ ,  $SP^B$  and  $SP^A$ ) on the y-axis. Include the upper and lower limit in the graph.

12.11.2. Return as result specifically the 10<sup>th</sup>, 18<sup>th</sup> and 28<sup>th</sup> elements of matrices  $SP^{MC}$ ,  $SP^B$  and  $SP^A$ : the actual size corresponding to nominal of 0.01, 0.05 and 0.10 in each way of obtaining critical values.

12.12.} else {

12.12.1. Save the 215 long vectors for the actual size (when  $\rho=0$ ):  $SP^{MC0}$ ,  $SP^{B0}$  and  $SP^{A0}$ .

12.12.2. Plot the nominal size ( $x_j$ ) on the x-axis and the power ( $SP^{MC}$ ,  $SP^B$  and  $SP^A$ ) on the y-axis.

12.12.3. Plot separately the actual size on the x-axis ( $SP^{B0}$  and  $SP^{A0}$ ) and the power ( $SP^B$  and  $SP^A$ ) on the y-axis. Match  $SP^{B0}$  with  $SP^B$ , and  $SP^{A0}$  with  $SP^A$ .

12.12.4. Return as result specifically the 10<sup>th</sup>, 18<sup>th</sup> and 28<sup>th</sup> elements of matrices  $SP^{MC}$ ,  $SP^B$  and  $SP^A$ : the power corresponding to nominal size of 0.01, 0.05 and 0.10 in each way of obtaining critical values.

}

12.13. Return all results.

}

---

13. Run the functions with the desired  $\rho$ .