

Herwartz, Helmut; Neumann, Michael H.

**Working Paper**

## A robust bootstrap approach to the Hausman test in stationary panel data models

Economics Working Paper, No. 2007-29

**Provided in Cooperation with:**

Christian-Albrechts-University of Kiel, Department of Economics

Suggested Citation: Herwartz, Helmut; Neumann, Michael H. (2007) : A robust bootstrap approach to the Hausman test in stationary panel data models, Economics Working Paper, No. 2007-29

This Version is available at:

<http://hdl.handle.net/10419/22045>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# **A robust bootstrap approach to the Hausman test in stationary panel data models**

by Helmut Herwartz and Michael H. Neumann

**C | A | U**

Christian-Albrechts-Universität Kiel

**Department of Economics**

*Economics Working Paper*

*No 2007-29*



# A robust bootstrap approach to the Hausman test in stationary panel data models

Helmut Herwartz\*      Michael H. Neumann†

December 5, 2007

## Abstract

In panel data econometrics the Hausman test is of central importance to select an efficient estimator of the models' slope parameters. When testing the null hypothesis of no correlation between unobserved heterogeneity and observable explanatory variables by means of the Hausman test model disturbances are typically assumed to be independent and identically distributed over the time and the cross section dimension. The test statistic lacks pivotalness in case the iid assumption is violated. GLS based variants of the test statistic are suitable to overcome the impact of nuisance parameters on the asymptotic distribution of the Hausman statistic. Such test statistics, however, also build upon strong homogeneity restrictions that might not be met by empirical data. We propose a bootstrap approach to specification testing in panel data models which is robust under cross sectional or time heteroskedasticity and inhomogeneous patterns of serial correlation. A Monte Carlo study shows that in small samples the bootstrap approach outperforms inference based on critical values that are taken from a  $\chi^2$ -distribution.

Keywords: Hausman test, random effects model, wild bootstrap, heteroskedasticity.  
JEL Classification: C12, C33.

---

\*Institut für Statistik und Ökonometrie, Christian-Albrechts-Universität zu Kiel, Ohlshausenstr. 40, D-24098 Kiel, E-mail: [Herwartz@stat-econ.uni-kiel.de](mailto:Herwartz@stat-econ.uni-kiel.de) (corresponding author)

†Institut für Stochastik, Friedrich-Schiller-Universität Jena, Ernst-Abbe-Platz 2, D-07743 Jena, E-mail: [mneumann@mathematik.uni-jena.de](mailto:mneumann@mathematik.uni-jena.de)

# 1 Introduction

The use of cross country or regional panel data is recently becoming more and more popular in macro- and spatial econometrics. Typical fields where panel data models are employed cover, for instance, tests of the purchasing power parity, models of growth, international or interregional trade, or empirical approaches to urban crime or environmental economics (Baltagi and Kao, 2000; Anselin, Florax and Rey, 2004). A core issue in panel specification testing is the selection of an efficient estimator in presence of unobserved heterogeneity. The Hausman test has become a prominent means of inference against correlation between individual effects and observable explanatory variables (Hausman, 1978).

Typical assumptions underlying traditional panel data models are absence of serial error correlation and homoskedasticity over both the time and cross section dimension. Model disturbances are easily justified to stem from an iid distribution in case of microeconomic studies where often a set of anonymous households or firms enters the analysis. Even for such widespread applications of panel models, however, (neglected) dynamic features might show up in autocorrelated error terms. In applied macroeconometrics serial error correlation is likely to emerge whenever an economy or region only partially absorbs a shock within the unit of time used as the sampling frequency. The presumption of time invariant error variances may also be criticized. In econometrics of financial data time dependent variances have attracted a huge theoretical and empirical interest (Bollerslev, Chou and Kroner, 1992). Similarly, with respect to macroeconometrics, shifts in the variations of disturbances may occur as a consequence of (fiscal or monetary) policy changes, central bank interventions or regime switches. Cross sectional patterns of second order heterogeneity are also more the rule than an exception. For instance, one may intuitively expect that larger economies (or urbanized regions) are likely to respond to exogenous shocks at a different scale in comparison with smaller economies (or rural regions).

Occasionally panel data models are formalized with some pattern of serial correlation (Lillard and Willis, 1979; Baltagi, 2001, Chapter 5). Then, correlation is typically specified parsimoniously with some first order autoregressive parameter. Over all members of the cross section a first order autocorrelation scheme might fail to provide a uniformly accurate approximation of the true underlying pattern of error dynamics. Moreover, it is likely that the autocorrelation parameter, if it exists, is cross section specific. Obviously, when allowing serially correlated disturbances within a panel data framework potential directions of covariance misspecification are manifold.

The asymptotic distribution of common panel specification test statistics derived under an iid assumption will depend on nuisance parameters if model disturbances are actually heteroskedastic over time, serially correlated or lack homogeneity over the cross section. On the one hand neglecting such forms of heterogeneity may invalidate conclusions obtained under an unrealistic modelling framework. On the other hand deriving first order asymptotic approximations is often cumbersome in presence of nuisance parameters. Under such circumstances bootstrap approaches are in widespread use to obtain robust critical values for a particular test statistic.

It is the purpose of this paper to contribute a robust approach to test for correlation between unobserved random effects and explanatory variables in panel data models by means of the Hausman test. It retains its validity in panels with finite time series dimension, under cross sectional heteroskedasticity, and (possibly time varying or cross section specific) serial error correlation. The proposed method exploits a convenient feature of the wild bootstrap

which copes with heteroskedasticity of model disturbances (Wu, 1986; Liu, 1988; Mammen, 1993) and cross sectional error correlation (Herwartz and Neumann, 2005).

The remainder of the paper is organized as follows: The panel model and the test statistic of interest are given in the next section. Then, Section 3 provides a bootstrap approach to generate critical values for the Hausman statistic. A simulation study, given in Section 4, illustrates the finite sample performance of the resampling scheme. Moreover, the effects of nuisance parameters on inference by means of critical values taken from the  $\chi^2$ -distribution will be addressed. Section 5 provides a few empirical examples which have attracted some interest in panel data econometrics. Conclusions are drawn in Section 6. An Appendix provides the proofs of the asymptotic results and gives further details on the empirically investigated data sets.

## 2 The model and the test statistic

### 2.1 A panel model with generalized covariance structure

Consider the common panel data model with random individual effects by observation

$$y_{it} = x'_{it}\boldsymbol{\beta} + \nu + u_{it}, \quad u_{it} = \alpha_i + e_{it} \quad (i = 1, \dots, N; t = 1, \dots, T). \quad (1)$$

Defining  $Y_i = (y_{i1}, \dots, y_{iT})'$ ,  $X_i = (x_{i1}, \dots, x_{iT})'$  and  $e_i = (e_{i1}, \dots, e_{iT})'$  we can rewrite (1) in matrix notation as

$$Y_i = X_i\boldsymbol{\beta} + \nu\mathbb{1}_T + u_i, \quad u_i = \alpha_i\mathbb{1}_T + e_i \quad (i = 1, \dots, N) \quad (2)$$

or, with  $Y = (Y_1', \dots, Y_N')'$ ,  $X = (X_1', \dots, X_N')'$ ,  $u = (u_1', \dots, u_N')'$ ,  $\alpha = (\alpha_1\mathbb{1}_T', \dots, \alpha_N\mathbb{1}_T')'$  and  $e = (e_1', \dots, e_N')'$ ,

$$Y = X\boldsymbol{\beta} + \nu\mathbb{1}_{NT} + u, \quad u = \alpha + e. \quad (3)$$

In (1)  $x_{it}$  is a  $K \times 1$  random vector of explanatory variables. Accordingly,  $\boldsymbol{\beta}$  is a  $K$  dimensional parameter vector,  $\nu$  denotes an intercept term and  $\mathbb{1}_R$  is an  $R$ -dimensional vector consisting of 1's. By assumption, the random individual effects  $\alpha_i \sim (0, \sigma_{\alpha_i}^2)$  are independent from the disturbances  $e_{it}$ . Note that in case  $\sigma_{\alpha_1}^2 = \dots = \sigma_{\alpha_N}^2 = 0$  the pooled regression is obtained as a special case of (1). With respect to the covariance of the mean zero innovations  $e_{it}$  we allow a pattern of serially correlated but cross sectionally uncorrelated error terms. Then, with  $E[e_i] = 0_T$  the latter scenario is formalized as

$$E[e_i e_j'] = \delta_{ij} \Sigma_{e_i}, \quad (4)$$

where  $\Sigma_{e_i}$  is a positive definite matrix of dimension  $T \times T$  and  $\delta_{ij}$  is the Kronecker delta.

As formalized in (4) covariance features of error terms are more general in comparison with traditional approaches to modelling panel data. According to (4) model disturbances may stem from cross sectionally heterogenous distributions and show time specific second order features. With regard to serial correlation the general specification covers e.g. the first order autoregressive model put forth by Lillard and Willis (1979). The latter formalizes time dependence of  $e_{it}$  in the usual manner, i.e.

$$e_{it} = \rho e_{it-1} + \epsilon_{it}, \quad \epsilon_{it} \sim \text{iid}(0, \sigma_\epsilon^2). \quad (5)$$

The autocorrelation structure assumed in Lillard and Willis (1979) is, however, very restrictive owing to the postulates of an exponentially decaying autocorrelation function on the one hand and cross sectional homogeneity on the other hand. As alternative specifications one may regard error terms following a higher order autoregression or some moving average structure. For a brief review of alternative parametric suggestions and their treatment for feasible GLS estimation in the context of panel data the reader may consult Baltagi (2001, Chapter 5.2). In any case, the more general error distribution complicates (E)GLS estimation of the models' slope parameters and, more importantly, introduces a source of potential misspecification of the model.

Given the likelihood of cross section specific covariance features it appears more natural to allow general unspecified patterns of second order features in macroeconomic or spatial panel data models as formalized in (1) and (4). Analyzing the latter specification this section will introduce generalized estimators one of which is efficient if individual effects and explanatory variables are uncorrelated. As it is typical in panel data modelling the latter feature is subjected to specification testing by means of a (generalized) Hausman statistic which is also provided below. Since misspecification might be seen as a crucial issue in this vein of econometric modelling we will also discuss the distributional features of the generalized Hausman statistic under misspecification of the covariance pattern. Before introducing generalized estimators and test statistics we make the following assumptions:

- (A1)** (i)  $\alpha_1, \dots, \alpha_N, e_1, \dots, e_N$  are conditionally on  $X = (X'_1, \dots, X'_n)'$  independent,  
(ii)  $E(e_i | X) = 0_T, \quad E(\alpha_i | X) = 0,$   
(iii) there exist positive constants  $C_1$  and  $C_2$  such that  
 $\text{var}(\alpha_i | X) = \sigma_{\alpha_i}^2, \quad 0 < C_1 \leq \sigma_{\alpha_i}^2 \leq C_2 < \infty,$   
 $\text{Cov}(e_i | X) = \Sigma_{e_i}, \quad C_1 I_T \preceq \Sigma_{e_i} \preceq C_2 I_T \quad (A \preceq B \text{ if } B - A \text{ is positive semidefinite}),$   
(iv) the random variables  $(e_{it}^2)_{i,t}$  and  $(\alpha_i^2)_i$  are conditionally on  $X$  uniformly integrable, that is,

$$\begin{aligned} \sup_{i \in \mathbb{N}} \max_{1 \leq t \leq T} E(e_{it}^2 I(|e_{it}| > c) | X) &\xrightarrow{c \rightarrow \infty} 0, \\ \sup_{i \in \mathbb{N}} E(\alpha_i^2 I(|\alpha_i| > c) | X) &\xrightarrow{c \rightarrow \infty} 0, \end{aligned}$$

- (v) the random variables  $(\|X'_i X_i\|)_{i \in \mathbb{N}}$  are uniformly integrable, that is,

$$\sup_{i \in \mathbb{N}} E[\|X'_i X_i\| I(\|X'_i X_i\| > c)] \xrightarrow{c \rightarrow \infty} 0.$$

## 2.2 Generalized estimators

Denote  $\bar{X} = (\mathbf{1}_{NT} X)$  and  $\Sigma := \text{Cov}(u) = \text{Diag}[\Sigma_1, \dots, \Sigma_N]$ , where  $\Sigma_i = \Sigma_{e_i} + \sigma_{\alpha_i}^2 \mathbf{1}_T \mathbf{1}'_T$ . An efficient estimator of  $\beta$  is given as

$$\begin{aligned} \hat{\beta}_{GLS} &= \left( X' \Sigma^{-1} X - \frac{X' \Sigma^{-1} \mathbf{1}_{NT} \mathbf{1}'_{NT} \Sigma^{-1} X}{\mathbf{1}'_{NT} \Sigma^{-1} \mathbf{1}_{NT}} \right)^{-1} \left( X' \Sigma^{-1} Y - \frac{X' \Sigma^{-1} \mathbf{1}_{NT} \mathbf{1}'_{NT} \Sigma^{-1} Y}{\mathbf{1}'_{NT} \Sigma^{-1} \mathbf{1}_{NT}} \right) \\ &= A_N^{-1} \sum_{i=1}^N a_{N,i} Y_i, \end{aligned} \tag{6}$$

where

$$A_N = \frac{1}{N} \left( X' \Sigma^{-1} X - \frac{X' \Sigma^{-1} \mathbb{1}_{NT} \mathbb{1}'_{NT} \Sigma^{-1} X}{\mathbb{1}'_{NT} \Sigma^{-1} \mathbb{1}_{NT}} \right),$$

$$a_{N,i} = \frac{1}{N} \left( X'_i \Sigma_i^{-1} - \frac{1}{\mathbb{1}'_{NT} \Sigma^{-1} \mathbb{1}_{NT}} X' \Sigma^{-1} \mathbb{1}_{NT} \mathbb{1}'_T \Sigma_i^{-1} \right).$$

In the special case of  $\Sigma_{e_i} = \sigma_e^2 I_T$  and  $\sigma_{\alpha_i}^2 = \sigma_\alpha^2$  we have that  $\Sigma^{-1} = \frac{1}{\sigma_e^2} \text{Diag}[I_T - \phi_T \mathbb{1}_T \mathbb{1}'_T, \dots, I_T - \phi_T \mathbb{1}_T \mathbb{1}'_T]$ , where  $\phi_T = \sigma_\alpha^2 / (\sigma_e^2 + T \sigma_\alpha^2)$ . Therefore, we obtain that  $X' \Sigma^{-1} X = \frac{1}{\sigma_e^2} (X' X - X' \text{Diag}[P_T, \dots, P_T] X + (1 - T \phi_T) X' \text{Diag}[P_T, \dots, P_T] X)$ ,  $\frac{1}{N} X' \Sigma^{-1} \mathbb{1}_{NT} \mathbb{1}'_{NT} \Sigma^{-1} X = \frac{1}{\sigma_e^2} (T - T^2 \phi_T) X' P_{NT} X$  and  $\mathbb{1}'_{NT} \Sigma^{-1} \mathbb{1}_{NT} = \frac{1}{\sigma_e^2} N (T - T^2 \phi_T)$ . This implies

$$\begin{aligned} X' \Sigma^{-1} X - \frac{X' \Sigma^{-1} \mathbb{1}_{NT} \mathbb{1}'_{NT} \Sigma^{-1} X}{\mathbb{1}'_{NT} \Sigma^{-1} \mathbb{1}_{NT}} \\ &= \frac{1}{\sigma_e^2} \left\{ X' (I_{NT} - \text{Diag}[P_T, \dots, P_T]) X + (1 - T \phi_T) X' \text{Diag}[P_T, \dots, P_T] X \right. \\ &\quad \left. - N (T - T^2) X' P_{NT} X \right\} \\ &= \frac{1}{\sigma_e^2} \left\{ \sum_{i,t} \tilde{x}_{it} \tilde{x}'_{it} + T (1 - T \phi_T) \sum_i \tilde{x}_i \tilde{x}'_i \right\}, \end{aligned}$$

where  $\bar{x}_i = \frac{1}{T} \sum_{t=1}^T x_{it}$ ,  $\tilde{x}_{it} = x_{it} - \bar{x}_i$ ,  $\bar{x} = \frac{1}{N} \sum_{i=1}^N \bar{x}_i$  and  $\tilde{x}_i = \bar{x}_i - \bar{x}$ . Analogously we can see that

$$\left( X' \Sigma^{-1} - \frac{X' \Sigma^{-1} \mathbb{1}_{NT} \mathbb{1}'_{NT} \Sigma^{-1}}{\mathbb{1}'_{NT} \Sigma^{-1} \mathbb{1}_{NT}} \right) Y = \frac{1}{\sigma_e^2} \left\{ \sum_{i,t} \tilde{x}_{it} \tilde{y}_{it} + T (1 - T \phi_T) \sum_i \tilde{x}_i \tilde{y}_i \right\},$$

which leads, because of  $T(1 - T \phi_T) = T \frac{\sigma_e^2}{\sigma_e^2 + T \sigma_\alpha^2}$ , to the common random effects estimator

$$\hat{\beta}_{GLS} = \left( \sum_{i,t} \tilde{x}_{it} \tilde{x}'_{it} + T \frac{\sigma_e^2}{\sigma_e^2 + T \sigma_\alpha^2} \sum_i \tilde{x}_i \tilde{x}'_i \right)^{-1} \left( \sum_{i,t} \tilde{x}_{it} \tilde{y}_{it} + T \frac{\sigma_e^2}{\sigma_e^2 + T \sigma_\alpha^2} \sum_i \tilde{x}_i \tilde{y}_i \right);$$

see also Baltagi (2001, Chapter 2).

If  $E(\alpha_i | X_i)$  does not necessarily vanish, then  $\hat{\beta}_{GLS}$  is in general not a consistent estimator of  $\beta$ . In this case we can augment the design matrix  $X$  with the  $NT \times N$ -matrix  $W = \text{Diag}[\mathbb{1}_T, \dots, \mathbb{1}_T]$  and obtain from (3) the equation

$$Y = \bar{X} \begin{pmatrix} \delta \\ \beta \end{pmatrix} + \bar{u}, \quad (7)$$

where  $\bar{X} = (W \ X)$ ,  $\delta = (\nu + E(\alpha_1 | X_1), \dots, \nu + E(\alpha_N | X_N))'$  and  $\bar{u} = u - W(E(\alpha_1 | X_1), \dots, E(\alpha_N | X_N))'$ . Then, an efficient estimator of  $\beta$  is the (generalized) fixed effect or least squares dummy variable estimator (LSDV)

$$\begin{aligned} \hat{\beta}_{FE} &= (X' \Sigma^{-1} X - X' \Sigma^{-1} W (W' \Sigma^{-1} W)^{-1} W' \Sigma^{-1} X)^{-1} \\ &\quad (X' \Sigma^{-1} - X' \Sigma^{-1} W (W' \Sigma^{-1} W)^{-1} W' \Sigma^{-1}) Y. \end{aligned}$$

$\hat{\beta}_{FE}$  is the (unique) best linear unbiased estimator (BLUE) of  $\beta$  in model (7), that is, it can be written in the form  $LY$ , where unbiasedness of  $\hat{\beta}_{FE}$  requires that  $LX = I_K$  and

$LW = 0_{K \times N}$  while optimality means that  $\text{Cov}(\widehat{\beta}_{FE}) = L\Sigma L'$  is minimal under these side conditions. However, since all matrices  $L$  satisfying these side conditions fulfill  $L\Sigma L' = L\text{Diag}[\Sigma_{e_1}, \dots, \Sigma_{e_N}]L'$  it follows that  $\widehat{\beta}_{FE}$  is equal to the BLUE in model (7) with  $\text{Cov}(\bar{u}) = \Sigma_e := \text{Diag}[\Sigma_{e_1}, \dots, \Sigma_{e_N}]$ . Therefore,  $\widehat{\beta}_{FE}$  can also be written as

$$\begin{aligned}\widehat{\beta}_{FE} &= (X'\Sigma_e^{-1}X - X'\Sigma_e^{-1}W(W'\Sigma_e^{-1}W)^{-1}W'\Sigma_e^{-1}X)^{-1} \\ &\quad (X'\Sigma_e^{-1} - X'\Sigma_e^{-1}W(W'\Sigma_e^{-1}W)^{-1}W'\Sigma_e^{-1})Y \\ &= B_N^{-1} \sum_{i=1}^N b_{N,i}Y_i,\end{aligned}\tag{8}$$

where

$$\begin{aligned}B_N &= \frac{1}{N}X'\Sigma_e^{-1}X - X'\Sigma_e^{-1}W(W'\Sigma_e^{-1}W)^{-1}W'\Sigma_e^{-1}X \\ &= \frac{1}{N} \sum_{i=1}^N \left( X_i'\Sigma_{e_i}^{-1}X_i - \frac{1}{\mathbb{1}'_T\Sigma_{e_i}^{-1}\mathbb{1}_T} X_i'\Sigma_{e_i}^{-1}\mathbb{1}_T\mathbb{1}'_T\Sigma_{e_i}^{-1}X_i \right), \\ b_{N,i} &= \frac{1}{N} \left( X_i'\Sigma_{e_i}^{-1} - \frac{1}{\mathbb{1}'_T\Sigma_{e_i}^{-1}\mathbb{1}_T} X_i'\Sigma_{e_i}^{-1}\mathbb{1}_T\mathbb{1}'_T\Sigma_{e_i}^{-1} \right).\end{aligned}$$

In the special case of  $\Sigma_{e_i} = \sigma_e^2 I_T$  this estimator simplifies to the standard LSDV estimator

$$\begin{aligned}\widehat{\beta}_{FE} &= (X'(I_{NT} - W(W'W)^{-1}W')X)^{-1} X'(I_{NT} - W(W'W)^{-1}W')Y \\ &= \left( \sum_{i,t} \tilde{x}_{it}\tilde{x}'_{it} \right)^{-1} \sum_{i,t} \tilde{x}_{it}\tilde{y}_{it};\end{aligned}$$

see Baltagi (2001, Chapter 2).

## 2.3 The Hausman statistic

As mentioned OLS or GLS estimation of the slope parameters in (1) will be biased if the individual effects  $\alpha_i$  are correlated with (some of) the explanatory variables  $x_{it,1}, \dots, x_{it,K}$ . On the other hand, under assumption (A1,iii) the GLS estimator is to be preferred over the OLS or LSDV estimator since it exploits the underlying error covariance structure efficiently. Moreover, estimation of  $N$  fixed effects is avoided such that model evaluation does not suffer from problems arising from incidental parameters. Summarizing the latter arguments a test for correlation between individual effects and explanatory variables is essential for the selection of an efficient estimator that is to be applied for the model in (1). The Hausman statistic (Hausman, 1978) has become a prominent tool to test the null hypothesis that individual effects are uncorrelated with the variables in  $x_{it}$  against the alternative of correlation, i.e.

$$H_0 : E(\alpha_i|X_i) = 0 \quad \text{vs.} \quad H_1 : E(\alpha_i|X_i) \neq 0 \quad \text{for at least one } i.\tag{9}$$

For this paper we allow the error terms  $e_i$  to have some general covariance pattern as formalized in (4). Accordingly, we consider a GLS based modification of the Hausman statistic. It follows from least squares theory that  $\text{Cov}(\widehat{\beta}_{GLS} | X) = \frac{1}{N}A_N^{-1}$ ,  $\text{Cov}(\widehat{\beta}_{FE} | X) = \frac{1}{N}B_N^{-1}$  and, since  $\widehat{\beta}_{GLS}$  is efficient under (A1), we have also that  $\text{Cov}(\sqrt{N}(\widehat{\beta}_{FE} - \widehat{\beta}_{GLS}) | X) = B_N^{-1} - A_N^{-1}$  (Hausman, 1978). Moreover, it follows from assumption (A2) below that



$B_N^{-1} - A_N^{-1}$  is a positive definite matrix if  $N$  is sufficiently large. For simplicity of presentation we assume that this holds true even for all  $N$ . Now the Hausman statistic is defined as

$$\mathcal{H}_N = N(\widehat{\boldsymbol{\beta}}_{FE} - \widehat{\boldsymbol{\beta}}_{GLS})' (B_N^{-1} - A_N^{-1})^{-1} (\widehat{\boldsymbol{\beta}}_{FE} - \widehat{\boldsymbol{\beta}}_{GLS}). \quad (10)$$

Note that

$$\sqrt{N}(\widehat{\boldsymbol{\beta}}_{FE} - \widehat{\boldsymbol{\beta}}_{GLS}) = \sum_{i=1}^N C_{N,i} u_i, \quad (11)$$

where  $C_{N,i} = \sqrt{N}(B_N^{-1}b_{N,i} - A_N^{-1}a_{N,i})$ . Accordingly, the Hausman statistic allows a representation as a quadratic form in the underlying model disturbances, i.e.

$$\mathcal{H}_N = \left\| (B_N^{-1} - A_N^{-1})^{-1/2} \sum_{i=1}^N C_{N,i} u_i \right\|^2. \quad (12)$$

In the case that the covariance parameters are unknown but respective estimators are available we can estimate the matrices  $A_N$  and  $B_N$  by  $\widehat{A}_N$  and  $\widehat{B}_N$ . In this case we consider the statistic with estimated covariance parameters

$$\widehat{\mathcal{H}}_N = N(\widehat{\boldsymbol{\beta}}_{FE} - \widehat{\boldsymbol{\beta}}_{GLS})' (\widehat{B}_N^{-1} - \widehat{A}_N^{-1})^{-1} (\widehat{\boldsymbol{\beta}}_{FE} - \widehat{\boldsymbol{\beta}}_{GLS}), \quad (13)$$

where  $\widehat{\boldsymbol{\beta}}_{FE}$  and  $\widehat{\boldsymbol{\beta}}_{GLS}$  are estimated LSDV and GLS (ELSDV and EGLS) estimators. The corresponding quadratic form representation of  $\widehat{\mathcal{H}}_N$  is analogous to (12).

To derive the asymptotic properties of the Hausman statistic we make the following assumptions:

**(A2)** It holds that  $A_N \xrightarrow{P} A$  and  $B_N \xrightarrow{P} B$ , as  $N \rightarrow \infty$ , where  $B$  and  $A - B$  are positive definite matrices.

**Proposition 1.** *Suppose that (A1) and (A2) are fulfilled. Then, as  $N \rightarrow \infty$ ,*

$$\mathcal{H}_N \xrightarrow{d} \chi^2(K). \quad (14)$$

Furthermore, if  $\widehat{A}_N$  and  $\widehat{B}_N$  are consistent estimators of  $A_N$  and  $B_N$ , that is,  $\|\widehat{A}_N - A_N\| \xrightarrow{P} 0$  and  $\|\widehat{B}_N - B_N\| \xrightarrow{P} 0$ , and if  $\sqrt{N}(\widehat{\boldsymbol{\beta}}_{FE} - \widehat{\boldsymbol{\beta}}_{GLS}) \xrightarrow{d} \mathcal{N}(0_K, B^{-1} - A^{-1})$ , then

$$\widehat{\mathcal{H}}_N \xrightarrow{d} \chi^2(K). \quad (15)$$

The asymptotic results in (14) and (15) are both derived for the case of a finite time dimension  $T$ . Owing to consistency of  $\widehat{\boldsymbol{\beta}}_{GLS}$  and  $\widehat{\boldsymbol{\beta}}_{FE}$  their difference vanishes under (A1) and (A2) as  $T \rightarrow \infty$ . For the case of an underlying iid covariance structure Ahn and Moon (2001) show that as  $T \rightarrow \infty$   $\text{Cov}[\widehat{\boldsymbol{\beta}}_{FE} - \widehat{\boldsymbol{\beta}}_{GLS}]$  converges sufficiently fast to ensure a nondegenerate limit distribution of the standard Hausman statistic.

As argued before any (cross sectionally homogeneous) a-priori formalization of panel covariance features is likely subjected to misspecification. Therefore we also consider the realistic case where the presumed covariance pattern differs from the unknown covariance

structure. We will still assume that the true covariances are given by (A1,iii), that is,  $\Sigma$  is the covariance matrix of  $u$  as above. Let now  $\tilde{\Sigma}$  denote the covariance specification which is actually used for constructing the (feasible) LSDV and GLS estimators and the test statistic. Denote by  $\tilde{A}_N$ ,  $\tilde{a}_{N,i}$ ,  $\tilde{B}_N$ ,  $\tilde{b}_{N,i}$  and  $\tilde{C}_{N,i}$  the analogues of  $A_N$ ,  $a_{N,i}$ ,  $B_N$ ,  $b_{N,i}$  and  $C_{N,i}$ , respectively, where for each term the true covariance matrix  $\Sigma$  is replaced by the presumed (false) covariance matrix  $\tilde{\Sigma}$ . In this case, the difference between the two panel estimators writes as  $\sqrt{N}(\hat{\beta}_{FE} - \hat{\beta}_{GLS}) = \sum_{i=1}^N \tilde{C}_{N,i} u_i$ , which leads to the test statistic

$$\begin{aligned} \tilde{\mathcal{H}}_N &= N \left( \hat{\beta}_{FE} - \hat{\beta}_{GLS} \right)' \left( \tilde{B}_N^{-1} - \tilde{A}_N^{-1} \right)^{-1} \left( \hat{\beta}_{FE} - \hat{\beta}_{GLS} \right) \\ &= \left\| \left( \tilde{B}_N^{-1} - \tilde{A}_N^{-1} \right)^{-1/2} \sum_{i=1}^N \tilde{C}_{N,i} u_i \right\|^2. \end{aligned} \quad (16)$$

Analogously to the correctly specified case, the matrix  $\tilde{B}_N^{-1} - \tilde{A}_N^{-1}$  is necessarily positive semidefinite since it is equal to  $\text{Cov}(\sqrt{N}(\hat{\beta}_{FE} - \hat{\beta}_{GLS}) | X)$  if  $\text{Cov}(u)$  were equal to  $\tilde{\Sigma}$  rather than  $\Sigma$ . Regularity of  $\tilde{B}_N^{-1} - \tilde{A}_N^{-1}$  follows from assumption (A3) below, for  $N$  sufficiently large. For simplicity of presentation, we assume again that this holds true for all  $N$ .

Finally, if estimates of the covariances are involved, we obtain the statistic

$$\hat{\tilde{\mathcal{H}}}_N = \left\| \left( \hat{\tilde{B}}_N^{-1} - \hat{\tilde{A}}_N^{-1} \right)^{-1/2} \sum_{i=1}^N \hat{\tilde{C}}_{N,i} u_i \right\|^2, \quad (17)$$

where  $\hat{\tilde{A}}_N$ ,  $\hat{\tilde{B}}_N$  and  $\hat{\tilde{C}}_{N,i}$  are the analogues of  $\tilde{A}_N$ ,  $\tilde{B}_N$  and  $\tilde{C}_{N,i}$ , respectively, with the presumed (false) covariances  $\tilde{\Sigma}$  replaced by their estimates  $\hat{\tilde{\Sigma}}$ .

For our asymptotic considerations we assume additionally

- (A3)** It holds that  $\tilde{A}_N \xrightarrow{P} \tilde{A}$  and  $\tilde{B}_N \xrightarrow{P} \tilde{B}$ , as  $N \rightarrow \infty$ , where  $\tilde{B}$  and  $\tilde{A} - \tilde{B}$  are positive definite matrices. Furthermore,  $\sum_{i=1}^N \tilde{C}_{N,i} \Sigma_i \tilde{C}'_{N,i} \xrightarrow{P} D$ , where  $D$  is a non-vanishing matrix.

In contrast to the result in Proposition 1, the Hausman statistics  $\tilde{\mathcal{H}}_N$  and  $\hat{\tilde{\mathcal{H}}}_N$  converge now in distribution to a random variable which is a weighted sum of independent  $\chi^2(1)$  random variables.

**Proposition 2.** *Suppose that (A1), (A2) and (A3) are fulfilled. Then, as  $N \rightarrow \infty$ ,*

$$\tilde{\mathcal{H}}_N \xrightarrow{d} \sum_{i=1}^K \lambda_i Z_i^2,$$

where  $Z_1, \dots, Z_K$  are independent standard normal random variables and  $\lambda_1, \dots, \lambda_K$  are the eigenvalues of the matrix  $(\tilde{B}^{-1} - \tilde{A}^{-1})^{-1/2} D (\tilde{B}^{-1} - \tilde{A}^{-1})^{-1/2}$ .

Furthermore, if  $\widehat{\tilde{A}}_N$  and  $\widehat{\tilde{B}}_N$  are consistent estimators of  $\tilde{A}_N$  and  $\tilde{B}_N$ , that is,  $\|\widehat{\tilde{A}}_N - \tilde{A}_N\| \xrightarrow{P} 0$  and  $\|\widehat{\tilde{B}}_N - \tilde{B}_N\| \xrightarrow{P} 0$ , and if  $\sum_{i=1}^N \widehat{\tilde{C}}_{N,i} u_i \xrightarrow{d} \mathcal{N}(0_K, D)$ , then

$$\widehat{\mathcal{H}}_N \xrightarrow{d} \sum_{i=1}^K \lambda_i Z_i^2.$$

### 3 Bootstrapping the Hausman statistic

As stated in Proposition 2 misspecification of the covariance structure of the innovations entering a panel model is likely to imply that the (generalized) Hausman test statistic lacks pivotalness. In this case its asymptotic distribution might be difficult if not impossible to derive analytically. As a particularly important case of misspecification in macroeconomic models one may regard the imposition of cross sectionally homogeneous covariance features whenever the true error distribution varies over the cross section. Under such circumstances a bootstrap approach could be adopted to generate robust critical values for a respective test statistic promising valid inference under the null hypothesis. Owing to particular issues raised for the Hausman test, as cross and intra sectional heteroskedasticity, the so-called wild or external bootstrap (Wu, 1986) can be seen as a natural tool for determining critical values in our context. Addressing the case of heteroskedasticity of unknown form Liu (1988) and Mammen (1993) have established the wild bootstrap to approximate the distribution of studentized statistics and F-type tests in static linear regression models, respectively. Recently, Herwartz and Neumann (2005) have used the wild bootstrap to mimic a cross sectional correlation pattern observed in systems of error correction models. The latter idea can be adopted for the present case of the Hausman test to imitate covariance features characterizing model disturbances over the finite time dimension. For the general convenience of the wild bootstrap it is worthwhile to mention that its implementation does not require any a-priori parametric guess concerning the actual error covariance structure from which model disturbances are drawn.

According to the different versions of the Hausman statistic in (10), (13), (16) and (17) resampling the statistic may proceed under alternative degrees of knowledge of the underlying covariance structure. For the following exposition we assume that an analyst has access to the true covariance structure ( $\Sigma$ ) but will have to estimate a set of structural parameters describing the latter. As an example we take the case where error terms  $e_{it}$  follow a first order autoregression as given in (5). Resampling the remaining variants of the Hausman statistic will be completely analogous except for the (estimated) covariance matrix ( $\Sigma$ ,  $\tilde{\Sigma}$ ,  $\widehat{\Sigma}$ ) entering the initial statistic and its bootstrap counterparts. Resampling  $\widehat{\mathcal{H}}_N$  proceeds along the following steps:

1. Estimate consistently the structural variance and correlation parameters entering  $\Sigma$  as e.g.  $\sigma_\epsilon^2$ ,  $\sigma_\alpha^2$  and  $\rho$  in case of the first order autocorrelation model (5). Owing to consistency under both hypotheses choosing the ordinary LSDV is natural for this purpose. Conditional on the first step estimate the slope parameters of the panel model under the null and the alternative hypothesis by means of the feasible estimators  $\widehat{\beta}_{GLS}$  and  $\widehat{\beta}_{FE}$ . ( $\widehat{\beta}_{GLS}$  and  $\widehat{\beta}_{FE}$  are the analogues of  $\widehat{\beta}_{GLS}$  and  $\widehat{\beta}_{FE}$  with  $\widehat{\Sigma}$  instead of  $\Sigma$ .)

From the latter step the covariance matrix of  $\widehat{\beta}_{FE} - \widehat{\beta}_{GLS}$  is available and the test statistic  $\widehat{\mathcal{H}}_N$  in (13) can be computed.

2. Using the slope estimates  $\widehat{\beta}_{FE}$ , an implied estimator of the intercept parameter is obtained as  $\widehat{\nu} = \bar{y} - \bar{x}'\widehat{\beta}_{FE}$ , with  $\bar{y} = (NT)^{-1} \sum_{i,t} y_{it}$  and  $\bar{x}$  denoting the  $K \times 1$ -dimensional vector of unconditional means of explanatory variables. From the latter estimator residuals  $\widehat{u} = Y - \widehat{\nu}\mathbb{1}_{NT} - X\widehat{\beta}_{FE}$  are obtained.
3. For  $s = 1, \dots, S$ , with  $S$  sufficiently large,
  - draw bootstrap variables  $u_i^*$  having the second order features of  $\widehat{u}_i$  as

$$u_i^* = \widehat{u}_i \cdot \eta_i, \quad \eta_i \sim (0, 1), \quad (18)$$

where  $\eta_i, i = 1, \dots, N$ , is a sequence of independent random variables, also independent of the variables in the model;

- obtain a bootstrap version of the Hausman statistic  $\widehat{\mathcal{H}}_N^*$  from its quadratic form representation as

$$\mathcal{H}_N^* = \left\| \left( \widehat{B}_N^{-1} - \widehat{A}_N^{-1} \right)^{-1/2} \sum_{i=1}^N \widehat{C}_{N,i} u_i^* \right\|^2, \quad (19)$$

where  $\widehat{A}_N, \widehat{B}_N, \widehat{C}_N$  are the estimated counterparts of  $A_N, B_N, C_N$  defined in connection with (8), (6) and (11), respectively.

4. Decision: Reject  $H_0$  with significance level  $\gamma$  if  $\widehat{\mathcal{H}}_N$  exceeds  $c_\gamma^*$ , the  $(1 - \gamma)$ -quantile of  $\widehat{\mathcal{H}}_N^*$ .

The central ingredient of the bootstrap procedure is the imitation of the first two moments of  $e_i = (e_{i1}, \dots, e_{iT})'$  by means of the quantity

$$u_i^* = (u_{i1}^*, \dots, u_{iT}^*)' = \eta_i (\widehat{u}_{i1}, \dots, \widehat{u}_{iT})' = \eta_i \widehat{u}_i.$$

The basic reason for the wild bootstrap to work for vector-valued random variables can be seen from the relation

$$\frac{1}{N} \sum_{i=1}^N \text{Cov}(u_i^* | X, u_1, \dots, u_N) = \frac{1}{N} \sum_{t=1}^N \widehat{u}_i \widehat{u}_i' = \frac{1}{N} \sum_{i=1}^N u_i u_i' + o_P(1) = \frac{1}{N} \sum_{i=1}^N \text{Cov}(u_i) + o_P(1),$$

such that, on average, the bootstrap reflects the true underlying covariances. Note that the latter are allowed to exhibit some variation over the cross section as formalized by assumption (A1,iii).

Several approaches to draw  $\eta_i$  are available from the literature (Mammen, 1993; Liu, 1988) and differ with respect to the low order moments of  $e_i$  imitated by the bootstrap design. For the Monte Carlo study and empirical exercises in this paper  $\eta_i$  is drawn from the Rademacher distribution (Liu, 1988; Davidson and Flachaire, 2001), i.e.

$$P(\eta_i = 1) = P(\eta_i = -1) = 0.5. \quad (20)$$

Having discussed the implementation of the wild bootstrap scheme we will state the asymptotic features of the bootstrap distribution in the following. For this purpose denote  $\mathcal{X}_N = (X, u_1, \dots, u_N)$ . We will show that the bootstrap counterpart  $\mathcal{H}_N^*$  of the Hausman statistic has the same asymptotic behavior as  $\mathcal{H}_N$ . Since the conditional distribution of  $\mathcal{H}_N^*$  given  $\mathcal{X}_N$  is itself random we obtain convergence of these distributions to their common limit only in probability. To better distinguish between the two random mechanisms involved, we describe weak convergence by a certain metric. Inspired by Bickel and Freedman (1981), we define a distance  $d$  between two distributions  $P$  and  $Q$  on  $\mathbb{R}^p$  as

$$d(P, Q) = \inf_{(X, Y): X \sim P, Y \sim Q} E [\|X - Y\| \wedge 1].$$

It is clear that  $d$  defines a metric and it can be shown that weak convergence of  $(P_N)_{N \in \mathbb{N}}$  to  $P$  is equivalent to  $d(P_N, P) \xrightarrow{N \rightarrow \infty} 0$ . (For the sufficiency part, note that  $P_N \implies P$  implies that one can construct on an appropriate probability space random variables  $(Z_N)_{N \in \mathbb{N}}$  and  $Z$  with respective distributions  $P_N$  and  $P$  such that  $Z_N \xrightarrow{a.s.} Z$ ; see e.g. Theorem IV.3.13 in Pollard (1984, p. 71). This, however, implies  $d(P_N, P) \xrightarrow{N \rightarrow \infty} 0$  by dominated convergence.)

**Proposition 3.** *Suppose that (A1) and (A2) are fulfilled. Then, as  $N \rightarrow \infty$ ,*

$$d(\mathcal{L}(\mathcal{H}_N^* | \mathcal{X}_N), \chi^2(K)) \xrightarrow{P} 0.$$

Propositions 1 and 3 together imply that the bootstrap test has asymptotically the correct size.

**Theorem 1.** *Suppose that (A1) and (A2) are fulfilled. Then*

$$P_{H_0}(\mathcal{H}_N > c_\gamma^*) \xrightarrow{N \rightarrow \infty} \gamma.$$

In the case of incorrectly specified covariances we obtain analogous results: Denote by  $\tilde{\mathcal{H}}_N^*$  the analogue to  $\tilde{\mathcal{H}}_N$  given in (16).

**Proposition 4.** *Suppose that (A1), (A2) and (A3) are fulfilled. Then, as  $N \rightarrow \infty$ ,*

$$d\left(\mathcal{L}(\tilde{\mathcal{H}}_N^* | \mathcal{X}_N), \sum_{i=1}^K \lambda_i Z_i^2\right) \xrightarrow{P} 0,$$

where  $Z_1, \dots, Z_K$  and  $\lambda_1, \dots, \lambda_K$  are as in Proposition 2.

Denote by  $\tilde{c}_\gamma^*$  the  $(1 - \gamma)$ -quantile of  $\mathcal{L}(\tilde{\mathcal{H}}_N^* | \mathcal{X}_N)$ . Propositions 2 and 4 together imply that the bootstrap test has asymptotically the correct size.

**Theorem 2.** *Suppose that (A1), (A2) and (A3) are fulfilled. Then*

$$P_{H_0}(\tilde{\mathcal{H}}_N > \tilde{c}_\gamma^*) \xrightarrow{N \rightarrow \infty} \gamma.$$

## 4 Monte Carlo analysis

As argued above the bootstrap approach to test the null hypothesis that individual effects  $\alpha_i$  and explanatory variables  $x_{it}$  are uncorrelated allows for numerous deviations from standard assumptions typically made in the analysis of stationary panel data models. Error terms  $e_i$  may have a non-diagonal covariance matrix  $\Sigma_{e_i}$ , formalizing e.g. serial correlation. Along its diagonal the covariance matrix  $\Sigma_{e_i}$  might also collect time specific variances. As a particular merit of the more general model setting it might be seen that the covariance matrix  $\Sigma_{e_i}$  is allowed to vary over the cross section. Finally, the unobservable mean zero error components in  $\alpha_i$  are allowed to have second order properties that differ over the cross section.

The Monte Carlo study documented in this section will address the performance of the bootstrap method in finite samples over the possible violations of homogeneity assumptions typically made in stationary panel data models. Moreover, the finite sample properties of bootstrap and standard (E)GLS inference are compared. To compute the Hausman statistic an analyst may start from alternative assumptions on the distributional characteristics of the underlying model disturbances. EGLS estimation of the variance of  $e_{it}$  depends on the analyst's guess concerning the prevalence of serial correlation. Since any presumption on the correlation pattern could be wrong the Monte Carlo analysis also sheds light on the finite sample properties of (E)GLS inference based on a false presumption concerning the serial correlation parameter. In particular we address the effect of neglecting the potential of serial correlation implied by presuming  $\rho = 0$ .

### 4.1 The simulation design

#### 4.1.1 The considered data generating processes

The data generating process (DGP) used for the simulations is the following homogeneous model

$$y_{it} = 1 + x_{it,1} + x_{it,2} + e_{it}, \quad t = 1, \dots, T, i = 1, \dots, N, \quad (21)$$

where the right hand side variables  $x_{it,2}$  are drawn once from a Gaussian distribution,  $x_{it,2} \sim \mathcal{N}(0, 1)$ , and then fixed over all replications of the experiment. Similarly, variables  $x_{it,1}$  are generated according to the model

$$x_{it,1} = \mu_i + \xi_{it}, \quad \xi_{it} \sim \mathcal{N}(0, 9), \quad \mu_i = 6(i - 1)/(N - 1).$$

Owing to the deterministic component  $\mu_i$  the unconditional level of  $x_{it,1}$  is ordered equidistantly over the cross section between values of 0 and 6. Individual effects  $\alpha_i$  are also drawn from the normal distribution. Nesting the null and the alternative hypothesis of the Hausman test the data generating model for the individual effects is

$$\alpha_i = \delta\mu_i + \sigma_{\alpha,i}\zeta_i, \quad \zeta_i \sim \mathcal{N}(0, 1). \quad (22)$$

As formalized in (22) individual effects  $\alpha_i$  and explanatory variables  $x_{it,1}$  will be correlated if  $\delta \neq 0$ . Under the null hypothesis  $\delta = 0$ . Note that the variance of individual effects could be homogeneous,  $\text{Var}(\alpha_i) = \sigma_{\alpha,i}^2 = \sigma_{\alpha}^2$ , or heterogeneous over the cross section,  $\text{Var}(\alpha_i) = \sigma_{\alpha,i}^2$ .

Cross section specific vectors of error terms  $e_i$  are drawn from a  $T$ -dimensional normal distribution as

$$e_i = G_i'v_i, \quad v_i \sim \mathcal{N}(0, I_T), \quad G_i'G_i = \Sigma_{e_i}, \quad (23)$$

where  $I_T$  is the  $T$ -dimensional identity matrix and  $G_i$  is an upper triangular matrix obtained from a Cholesky decomposition of  $\Sigma_{e_i}$ . The particular choices of  $\Sigma_{e_i}$  having typical elements  $\sigma_{kl}^{(e_i)}$  cover the following cases (DGPs):

- Case 1: Cross sectionally homogeneous patterns of serial correlation with unconditional unit variance. The variance of individual effects is unity over the cross section. Summarizing the latter features we have the following specification:

$$\Sigma_{e_i} = \Sigma_e = ((\sigma_{kl}^{e_i}))_{k,l}, \quad \sigma_{kl}^{(e_i)} = \rho^{|k-l|}, \quad \rho = 0.5, \quad \sigma_{\alpha,i}^2 = \sigma_\alpha^2 = 1. \quad (24)$$

- Case 2: Cross sectionally varying patterns of serial correlation with an unconditional variance of unity, obtaining

$$\sigma_{kl}^{(e_i)} = \rho_i^{|k-l|}, \quad \rho_i = 0.3 + 0.4(i-1)/(N-1). \quad (25)$$

As in case 1 the variance of individual effects is  $\sigma_{\alpha,i}^2 = \sigma_\alpha^2 = 1$ .

- Case 3: Homoskedastic individual effects  $\sigma_{\alpha,i}^2 = 1$  and homogeneous patterns of serial correlation,  $\rho_i = \rho$ , coupled with time dependent second order properties of  $e_{it}$ . Formally the latter scenario reads as

$$\sigma_{kl}^{(e_i)} = \begin{cases} \sigma_{e,k}^2 & \text{if } k = l, \\ \rho^{|k-l|} \sqrt{\sigma_{e,k}^2 \sigma_{e,l}^2}, \rho = 0.5, & \text{if } k \neq l \end{cases}, \quad (26)$$

with  $\sigma_{e,k}^2 = 0.6 + 0.8(k-1)/(T-1)$ . Note that by generation of the time varying second order moments it holds that  $\text{var}(e_{it}) = 1$ .

- Case 4: Cross sectionally homogeneous patterns of serial correlation,  $\rho_i = \rho$ , combined with an error variance of  $e_{it}$  that depends on the cross section,  $e_{it} \sim (0, \sigma_{e,i}^2)$ . The variance of the individual effects is  $\sigma_{\alpha,i}^2 = 1$  throughout. For a typical element of the covariance matrix of  $e_i$  we have in this case

$$\sigma_{kl}^{(e_i)} = \begin{cases} \sigma_{e,k}^2 & \text{if } k = l, \\ \rho^{|k-l|} \sigma_{e,i}^2, \rho = 0.5, & \text{if } k \neq l \end{cases}, \quad (27)$$

where  $\sigma_{e,k}^2 = 0.6 + 0.8(k-1)/(N-1)$ .

- Case 5: Cross sectionally homogeneous covariance of  $e_{it}$  as given in case 1 combined with a cross section specific variance of individual effects. For the variance of individual effects we choose:

$$\sigma_{\alpha,i}^2 = 0.6 + 0.8(i-1)/(N-1). \quad (28)$$

All particular choices of model parameters imply that, unconditionally, the autocorrelation parameter is  $\bar{\rho} = 0.5$ . Moreover,  $E[\alpha_i^2] = 1$  and  $\text{var}(e_{it}) = 1$ . For the Monte Carlo experiments we will take the perspective of an analyst regarding all processes as stemming from a cross sectionally homogeneous first order autocorrelation model as given in (5). Thus, the presumed covariance matrix  $\Sigma$  depends on the parameters  $\rho$ ,  $\sigma_\alpha^2$  and  $\sigma_\epsilon^2$ . Unconditional moments  $\text{var}(e_{it}) = 1$  and  $\bar{\rho} = 0.5$  imply for the latter variance parameter an unconditional choice of  $\sigma_\epsilon^2 = \text{var}(e_{it})(1 - \bar{\rho}^2) = 0.75$  required when implementing  $\mathcal{H}$  or  $\tilde{\mathcal{H}}$ . Note that a parameter selection  $\sigma_\alpha^2 = 1$ ,  $\sigma_\epsilon^2 = 0.75$ ,  $\rho = 0.5$  only corresponds to the true underlying covariance pattern ( $\mathcal{H}$ ) if the actual data stem from a DGP given under case 1 above. In

all remaining cases the presumed covariance structure is in some direction misspecified such that the DGPs 2 to 5 are suitable to investigate the empirical size features of  $\widehat{\mathcal{H}}$  type statistics. For the DGPs 2 to 5 it will be of interest to which extent the empirical properties of GLS inference are adversely affected if parametric assumptions are postulated that hold on average but fail when conditioning on either the time or cross section dimension. In analogy to these scenarios, implementing the Hausman statistic with estimated covariance parameters will obtain statistics of type  $\widehat{\mathcal{H}}$  (DGP 1) and  $\widetilde{\mathcal{H}}$  (DGPs 2 to 5).

All considered scenarios except for DGP 5 formalize invariant second order features of individual effects  $\alpha_i$ . For this particular model it is worthwhile to point out that along common lines of panel data modelling the true underlying parameters  $\sigma_{\alpha,i}^2$  cannot be estimated consistently. In this case the bootstrap approach is particularly promising since it allows robust inference even under a false presumption concerning the individual effects' variances.

#### 4.1.2 Alternative test statistics

Instead of a-priori stating specific parameter settings feasible versions of the Hausman statistic will require some first step parameter estimates. For the Monte Carlo analysis EGLS based test statistics  $\widehat{\mathcal{H}}$  and  $\widetilde{\mathcal{H}}$  are computed along two alternative strategies that are both in an analyst's opportunity set.

- Estimation of all unknown parameters

In the first place an analyst may proceed from the assumption that model disturbances show serial correlation which is invariant over the cross section. Then, the following strategy could be adopted to obtain parameter estimates entering EGLS inference. Firstly a LSDV regression obtains estimated model disturbances  $\widehat{e}_{it}$  and estimates of the individual effects  $\widehat{\alpha}_i$ . The unconditional variance of the serially correlated error terms is estimated as

$$\widehat{\text{var}}(e_{it}) = \widehat{\sigma}_\epsilon^2 / (1 - \widehat{\rho}^2) = \frac{1}{(N(T-1) - K)} \sum_{i=1}^N \sum_{t=1}^T \widehat{e}_{it}^2. \quad (29)$$

The autocorrelation parameter  $\rho$  is also estimated from LSDV residuals by means of a pooled OLS regression, i.e.

$$\widehat{\rho} = \frac{\sum_{i=1}^N \sum_{t=2}^T \widehat{e}_{i,t} \widehat{e}_{i,t-1}}{\sum_{i=1}^N \sum_{t=2}^T \widehat{e}_{i,t-1}^2}. \quad (30)$$

From (29) an estimator for the variance parameter of the disturbances  $e_{it}$  is obtained as

$$\widehat{\sigma}_\epsilon^2 = \widehat{\text{var}}(\widehat{e}_{it})(1 - \widehat{\rho}^2). \quad (31)$$

Finally a variance estimator (Nerlove, 1971) for the individual effects might be determined as

$$\widehat{\sigma}_\alpha^2 = \frac{1}{N-1} \sum_{i=1}^N (\widehat{\alpha}_i - \bar{\alpha})^2, \quad \bar{\alpha} = \frac{1}{N} \sum_{i=1}^N \widehat{\alpha}_i. \quad (32)$$

Since  $\sigma_\alpha^2$  is estimated as the empirical variance of estimated fixed effects  $\widehat{\sigma}_\alpha^2$  its estimate does not depend on the a-priori choice or estimate of  $\rho$ . The estimator in (32) is



consistent as  $N \rightarrow \infty$  and nonnegative by construction. Owing to the latter property it might be particularly useful in Monte Carlo experiments.

Apart from the estimators given in (29), (32) and (31) one may also evaluate the variance parameters by means of other approaches going back to Swamy and Arora (1972), Wallace and Hussain (1969) or Amemiya (1971). The focus of this paper lies, however, on the characterization of finite sample features offered by alternative venues to obtain critical values for the Hausman statistic. Therefore, we expect that the relative merits of resampling as an alternative to (falsely) stating an asymptotic  $\chi^2$ -distribution will not depend on the particular implementation of EGLS inference. The a-priori assumption of time or cross sectionally invariant parameters might also be relaxed when adopting EGLS methods. A-priori information necessary to justify conditional parameter estimation is, however, often not available. From this perspective the generality of the bootstrap methodology is immediately clear as it obtains asymptotically correct critical values even under some misspecification of the actual covariance pattern.

- Conditional estimation of variance parameters

In the second place an analyst might proceed by imposing some a-priori restrictions on the model parameters and will then, accordingly, estimate the remaining unrestricted parameters conditional on the former choice. Along these lines it is here of particular interest how alternative a-priori assumptions concerning the autoregressive parameter affect the outcome of inference. We will consider three alternative a-priori restrictions made for the autoregressive parameter, namely  $\rho = 0, 0.25, 0.5$ . The first choice resembles the widespread situation where the potential of serial correlation is neglected. The second parameter selection implies that the analyst underestimates the true (unconditional) level of the autoregressive parameter. Choosing  $\rho = 0.5$  mirrors the (unrealistic) scenario where an analyst has access to the true underlying unconditional correlation parameter.

### 4.1.3 Further remarks

To implement the bootstrap procedure we use the so-called Rademacher distribution given in (20). The number of bootstrap replications is set to  $S = 199$ . Investigating size and power properties the parameter  $\delta$  in (22) is chosen as  $\delta = 0$  and  $\delta = 0.1$ , respectively. With regard to the power features of the alternative testing strategies we consider the case of a relatively weak violation of the null hypothesis. When choosing higher values for  $\delta$  it might become difficult to distinguish (trivial) power properties across alternative methods of inference. The considered time series dimensions are  $T = 5, 10, 20, 50$ . Since the asymptotic theory outlined in Sections 2 and 3 has been set out under the assumption of a fixed time dimension  $T$  and  $N \rightarrow \infty$  the cross section dimensions are chosen as  $N = 5, 10, 20, 50, 100$ . Each panel data model is generated 2000 times.

## 4.2 Monte Carlo results

### 4.2.1 Provision of size estimates

Empirical rejection frequencies obtained for alternative implementations of the Hausman test are provided in Table 1 and Table 2. For space considerations detailed simulation results are

given for selected panel dimensions  $N, T$ . The nominal level corresponding to most results listed in these tables is  $\gamma = 5\%$ . On some aggregate level empirical test features obtained at the 10% significance level are also given. To facilitate the interpretation of the empirical size estimates bold entries indicate that the nominal and empirical size differ with 5% significance. Significant size distortions are diagnosed in case the empirical rejection frequencies obtained under the null hypothesis are not covered by a confidence interval constructed around the nominal level as  $\gamma \pm 1.96\sqrt{\gamma(1-\gamma)/2000}$ . Setting, for instance  $\gamma = 0.05$ , 4.04% and 5.96% are obtained as lower and upper bound of the latter confidence interval, respectively. When varying the nominal level of the tests, it turns out that in relative terms the empirical performance of alternative test procedures is very similar. For the latter reason we do not provide detailed results obtained from inference at the 10% nominal level.

Each row of Table 1 or Table 2 collects rejection frequencies for a given implementation of the Hausman statistic over the five alternative DGPs. Since critical values for each test are taken from the  $\chi^2(2)$ -distribution and, alternatively, determined by means of a bootstrap scheme the overall number of rejection frequencies in each row of these tables is 10. Each panel in Table 1 or Table 2 consists of five rows which in turn correspond to alternative implementations of the Hausman statistic: A GLS version ( $\mathcal{H}$  or  $\tilde{\mathcal{H}}$ ) using the unconditional parameter choices ( $\sigma_\epsilon^2 = 0.75, \sigma_\alpha^2 = 1, \rho = 0.5$ ) and four alternative test statistics ( $\hat{\mathcal{H}}$  and  $\tilde{\hat{\mathcal{H}}}$ ) building on first step estimates of all or some parameters entering the covariance matrix  $\hat{\Sigma}$ .

To facilitate the overall comparison of alternative tools of inference for a given DGP or over alternative specifications of the underlying DGPs the last two columns of Table 1 and Table 2 and the last four rows of Table 2 provide row and columnwise summary measures of the absolute frequencies of significant over- or undersizing. For the latter aggregated results we distinguish between inference using critical values taken from the  $\chi^2(2)$ -distribution ( $\mathcal{H}$ ) and generated by means of the bootstrap scheme ( $\mathcal{H}^*$ ). Aggregating over all Monte Carlo experiments we also give the latter statistics as obtained from inference at the 10% significance level. Since the asymptotic results given in Sections 2 and 3 apply for the case of a fixed time dimension and  $N \rightarrow \infty$  empirical violations of the nominal test levels should not be overvalued whenever the cross section dimension does not exceed the time dimension. For the latter reason Table 2 also provides columnwise summary statistics conditional on Monte Carlo experiments with  $N > T$ .

#### 4.2.2 An overall assessment of empirical size features

Before discussing empirical size properties of alternative tests in more detail it is tempting to take a view at some aggregated results. In sum, we perform Monte Carlo experiments for 20 alternative combinations of the panel dimensions  $N$  and  $T$ . For one half of these settings we have  $T < N$ , such that one may expect asymptotic results to better apply for this latter subset in comparison with experiments where  $N \leq T$ . Recall that each panel in Table 1 or Table 2 contains empirical results for 5 DGPs and 5 alternative implementations of the Hausman test. In total  $20 \cdot 5 \cdot 5 = 500$  ( $10 \cdot 5 \cdot 5 = 250$ ) experiments are conducted for both venues to obtain critical values (for experiments with  $T < N$ ). To evaluate the overall performance of a particular test implementation ( $\mathcal{H}$  or  $\tilde{\mathcal{H}}$  implemented with  $\chi^2(2)$  critical values, say) one may consider the absolute number of empirical size estimates violating the nominal level with 5% significance. Since 100 experiments are performed for each test implementation one would expect (as  $N \rightarrow \infty$ ) about five significant size distortions under independence of the sampling experiments. For the subset of experiments with  $N > T$  about

3 significant size violations are to be expected, accordingly.

Absolute frequencies of violations of the nominal test levels are listed in the bottom rows of Table 2 and are also collected in Table 3. Aggregating over all Monte Carlo experiments we obtain that testing at a nominal level of 5% (10%)  $\chi^2(2)$  and bootstrap critical values involve 378 and 229 (370 and 237) significant violations of the nominal size, respectively. Concentrating on experiments with  $T < N$  the absolute frequencies of size violations obtained at the 5% (10%) nominal level are 101 and 34 (168 and 37) for inference with  $\chi^2(2)$  and bootstrap critical values, respectively. Although the latter views at the simulation results might suffer from aggregation over heterogeneous items it is apparent from the summary statistics given in Table 3, that the bootstrap method delivers more accurate critical values for the respective test statistics as relying on the  $\chi^2(2)$ -distribution. For the interpretation of the latter result it is worthwhile to recall that the  $\chi^2(2)$ -distribution is hardly the true asymptotic distribution for most simulated test statistics. Simulation results for cases with  $N > T$  are likely to reflect the adverse effects of falsely presuming an asymptotic  $\chi^2$ -distribution more precisely. Conditional on experiments with  $N > T$  GLS and EGLS inference obtain similar results regarding the prevalence of empirical size distortions. In particular, two bootstrap EGLS versions turn out to show empirical features coming closer to the expected frequency of size violations. The latter either require to estimate all three model parameters or condition the estimation of the variance parameters on the presumption  $\rho = 0.25$ . For both implementations of the test only 6 and 7 significant violations of the nominal 5% level are documented, respectively. Aggregating in the same way over alternative DGPs it turns out that most favorable results are obtained for bootstrap inference performed at the 10% level for DGPs belonging to Case 2 or Case 4, which obtain 3 and 5 significant size violations, respectively. Both classes of DGPs are characterized by cross sectionally varying serial correlation features. In comparison to the latter DGPs the bootstrap approach performs somewhat worse for DGP 5 (cross section specific variance of individual effects) which obtains 13 and 11 significant size violations when the nominal test level is 5% and 10%, respectively.

Interestingly, over all performed experiments size violations of the bootstrap venue are overrejections throughout. Taking critical values from the  $\chi^2$ -distribution significant under- as well as oversizing may result depending on the underlying DGP and the implemented test statistic.

#### 4.2.3 Empirical test features for given panel dimensions

Having discussed the performance of competing test procedures at an aggregated level we turn next to an evaluation of single Monte Carlo experiments. GLS based inference shows large significant oversizing irrespective of the underlying DGP when both panel dimensions are rather small ( $N = T = 5$ ). The bootstrap approach attains empirical rejection frequencies which are about half of the corresponding quantities obtained when critical values are taken from the  $\chi^2$ -distribution. EGLS based inference with critical values from the  $\chi^2(2)$ -distribution yields empirical size estimates that are closer to the nominal level and, in addition, cannot be distinguished with 5% significance from the former for some experiments. Owing to more accurate size features it appears at the first sight that bootstrap inference is preferable to common inference using  $\chi^2$ -quantiles as critical values. Note, however, that none of the competing procedures can be justified on the basis of asymptotic arguments since the small data dimensions are unlikely sufficient to invoke convergence to the postulated limit distributions. The different performance of GLS and EGLS based inference shows that in case of a very small cross sectional dimension conclusions obtained from Hausman testing

should be treated with care.

For experiments performed with  $N = 10$  and  $T = 5$  GLS based inference using quantiles of the  $\chi^2$ -distribution still suffers from significant oversizing with empirical rejection frequencies exceeding twice the nominal test level, throughout. Implementing the Hausman test with variance parameters estimated unrestrictedly or conditional on the false presumption  $\rho = 0$  obtains empirical rejection frequencies about 1% over all considered DGPs. The bootstrap procedure shows significant oversizing for some DGPs but its size estimates are closer to the nominal level than corresponding estimates obtained when using  $\chi^2(2)$ -quantiles as critical values. As it might be expected the imposition of ‘realistic’ a-priori restrictions ( $\rho = 0.5$ ) improves the empirical features of standard inference drastically such that the empirical significance levels cannot be distinguished from their nominal counterparts with 5% significance.

When further increasing the cross section dimension for a fixed  $T$  the effects of nuisance parameters invalidating asymptotic pivotalness of the Hausman statistic become more apparent. Aggregating the outcomes of  $\chi^2(2)$  based inference over the 5 considered DGPs it turns out that for  $T = 5$  and  $N = 50, 100$  at least 22 out of 25 experiments show significant size distortions. In contrast, the corresponding resampling schemes do not give any significant size distortion.

Monte Carlo experiments with time series dimension  $T = 10$  offer analogous results as discussed for the cases where  $T = 5$ . Size distortions diagnosed for standard inference based on critical values from the  $\chi^2(2)$ -distribution persist over alternative dimensions of the cross section  $N = 5, 10, 50$ . Bootstrap based empirical size estimates differ insignificantly from the nominal level for all experiments with cross section dimension  $N = 50$ . Interestingly, conditioning EGLS inference on the false presumption  $\rho = 0$  obtains accurate empirical size features in case  $T = 10$  and  $N = 100$  even when using quantiles of the  $\chi^2(2)$ -distribution as critical values.

#### 4.2.4 Asymptotic properties

To briefly sketch the asymptotic features of alternative implementations of the Hausman statistic under different DGPs Table 4 gives empirical size estimates obtained for a fixed time dimension  $T = 5$  combined with rather large cross section dimensions  $N = 500, 1000$ . Obviously size distortions are the rule when employing quantiles of the  $\chi^2(2)$ -distribution as critical values for the Hausman statistic. Whereas the GLS version, assuming the underlying model parameters to equal their true unconditional values, amounts in most cases to significant oversizing EGLS statistics yield empirical size estimates that are below their nominal counterparts.

#### 4.2.5 Power estimates

Selected power results ( $T = 5, 10, N = 5, 10, 50$ ) are displayed in Table 5. Since some implementations of the Hausman test turn out to show large size distortions we report unadjusted power estimates. Apparently all implemented test procedures have some power against the null hypothesis of no correlation between individual effects and the explanatory variables. Regarding, for instance, the case with  $T = 10$  and  $N = 50$  almost all empirical rejection frequencies estimated under the alternative hypothesis are about twice the corresponding quantities obtained under the null hypothesis. From the latter relation one may draw two

conclusions: Firstly, reliance on robust resampling methods does not go along with power losses. Secondly, similar to empirical size, power features are more or less invariant with respect to the sources of misspecification underlying DGP2 to DGP5 in comparison with DGP1.

## 5 Empirical examples

To illustrate some cases for robust Hausman testing in empirical practice we consider three data sets that have been discussed in the literature before and which are available from the net <http://www.wiley.com/legacy/wileychi/baltagi/datasets.html> (Baltagi, 2001). The empirical examples cover (see Appendix 7.2 for detailed representations):

- an investment equation for US firms (Grunfeld, 1958),
- a Cobb-Douglas production function modelling the impact of public capital on the GDP of US states (Munnell, 1990; Baltagi and Pinnoi, 1995),
- a model of gasoline demand in the OECD (Baltagi and Griffin, 1983).

For each empirical model three EGLS based Hausman statistics are provided in Table 6. In analogy to the discussion in Section 4.1.2 these are obtained from estimating the variance parameters jointly with the serial correlation coefficient  $\rho$  and conditional on the presumptions of no ( $\rho = 0.0$ ) and moderate serial correlation ( $\rho = 0.5$ ). Critical values for standard inference are taken from  $\chi^2(K)$ -distributions, where  $K = 2, 4, 3$  when modelling US firm investment, productivity of public capital and gasoline demand, respectively. Alternatively, critical values are determined by means of the wild bootstrap.

Estimates of the autoregressive parameter are at least  $\hat{\rho} = .664$  and likely significant. For all Hausman test statistics 'uniform' test decisions are obtained when using  $\chi^2(K)$  critical values. The null hypothesis of no correlation between individual effects and the explanatory variables is rejected at conventional significance levels when modelling gasoline demand or productivity of public capital. Regarding the investment function of US firms the null hypothesis cannot be rejected such that for this model GLS estimation seems appropriate.

Adopting the bootstrap to generate critical values for the Hausman statistic the null hypothesis is rejected with 5% significance when modelling investment behavior of US firms. In addition, with respect to gasoline demand in the OECD bootstrap based inference at the 5% level obtains conclusions which are at odds with the highly significant statistics evaluated by means of the  $\chi^2(3)$ -quantile. Modelling the productivity of public capital both conditional versions of the Hausman statistic are not significant when employing critical values from the bootstrap. Both choices  $\rho = 0, 0.5$ , however, are clearly distinct from the estimated serial correlation coefficient  $\hat{\rho} = .801$  such that conclusions based on the conditional EGLS statistics may suffer from power loss. Estimating the correlation parameter from the data the Hausman statistic is significant at the 5% level according to the bootstrap distribution.

As argued throughout the paper specification testing in panel data models relies heavily on the presumption of time or cross sectional homogeneity of the distributional features of error terms. To underscore that actual error distributions are heterogeneous over the time dimension Table 6 also gives Hausman statistics for 4 separate subsamples of the example data. Owing to space considerations we focus on results obtained from EGLS inference where

all covariance parameters are estimated. For all these subsamples the actual time dimension is either  $T = 4$  or  $T = 5$  such that the cross sectional dimension exceeds the former by a factor of 2 (investment behavior) to 9 (production function). Therefore one may expect that the asymptotic theory as outlined in this work better applies for the Hausman statistics obtained from the subsample information. Time dependent estimates of the parameters entering the presumed covariance structure are also shown in Table 6. From the simulation results documented in Table 1 we have the impression that bootstrap methods are to be preferred for inference in case of such panel dimensions since they show correct empirical size features and offer in the same time some power against correlation between individual effects and explanatory variables. Estimation results for the correlation as well as for the variance parameters indicate that the covariance structure underlying the panel models varies over time. All subsample specific estimates  $\hat{\rho}$  are substantially smaller than the overall estimates discussed above such that the latter might be spurious owing to the neglected time variation.

Regarding inference using quantiles of  $\chi^2$ -distributions as critical values the evidence offered from subsample modelling is mostly in line with the conclusions discussed above for the ‘large’ sample. Since the estimated covariance patterns lack time homogeneity, however, the latter results might not hold at the presumed 5% significance level. Bootstrap inference in subsamples reveals that correlation between unobservable heterogeneity and explanatory variables is more typical for particular subperiods in case of investment (1950-54) and gasoline demand (1975-78) modelling. Regarding the productivity of public capital three out of four subsample test statistics are significant at conventional levels according to bootstrap based critical values.

## 6 Conclusions

In this paper we address the issue of testing for correlation between unobserved panel heterogeneity and explanatory variables under general covariance structures of underlying error distributions. We consider the case of a finite time dimension while  $N \rightarrow \infty$ . The considered second order features cover (cross sectionally varying patterns of) serial correlation or time heteroskedasticity. For the determination of critical values we propose a wild bootstrap scheme that retains its validity even in case the presumed covariance structure differs from the true second order features of error terms. In the latter case nuisance parameters are likely to invalidate asymptotic pivotalness of the generalized Hausman statistic.

Finite sample features involved when critical values for the Hausman statistic are taken from the  $\chi^2$ -distribution or estimated alternatively by means of the bootstrap are examined. We find that the bootstrap approach is characterized by more accurate empirical size features at least when the cross section dimension exceeds the time dimension. Standard inference by means of critical values from the  $\chi^2$ -distribution suffers from both, weaker empirical size features in small samples under correct covariance specification, and nonpivotalness in case second order features are misspecified.

Summarizing the empirical test results it is apparent that panel covariance homogeneity is likely exceptional at least when modelling longitudinal data. In the light of potential heterogeneity robust critical values promise actual significance levels which are close to the nominal test levels. The considered subsamples underscore, in addition, that correlation between individual effects and explanatory variables might also undergo some form of time variation. In the latter case it is important to have tools of inference at hand that show

accurate empirical features in case of a small time dimension.

Throughout our analysis proceeds under the (common) assumption of cross sectional independence which might be at odds with macroeconomic or spatial panel data. Recent contributions to spatial econometrics or panel unit root testing allow for cross sectional error correlation. Immunizing the Hausman statistic against cross sectional error correlation is an important issue of further research.

## 7 Appendix

### 7.1 Proofs

*Proof of Proposition 1.* We obtain from (11) that

$$\text{Cov} \left( \sum_{i=1}^N C_{N,i} u_i \middle| X \right) = \text{Cov} \left( \sqrt{N}(\hat{\beta}_{FE} - \hat{\beta}_{GLS}) \middle| X \right) = B_N^{-1} - A_N^{-1} \xrightarrow{P} B^{-1} - A^{-1}.$$

Furthermore, we obtain from the uniform integrability of  $(\|X'_i X_i\|)_{i \in \mathbb{N}}$  that, for arbitrary  $c > 0$ ,

$$\begin{aligned} P \left( \max_{1 \leq i \leq N} \|X'_i X_i\| > cN \right) &\leq \sum_{i=1}^N P(\|X'_i X_i\| > cN) \\ &\leq \frac{1}{cN} \sum_{i=1}^N E[\|X'_i X_i\| I(\|X'_i X_i\| > cN)] \xrightarrow{N \rightarrow \infty} 0. \end{aligned}$$

In other words, we have that  $\max_{1 \leq i \leq N} \|X_i\| = o_P(\sqrt{N})$ , which implies that  $c_N = \max_{1 \leq i \leq N} \|C_{N,i}\| = o_P(1)$ . Hence, we obtain by (A1,iv) that, for arbitrary  $\epsilon > 0$ ,

$$\begin{aligned} &\sum_{i=1}^N E \left( \|C_{N,i} u_i\|^2 I(\|C_{N,i} u_i\| > \epsilon) \middle| X \right) \\ &\leq \sum_{i=1}^N \|C_{N,i}\|^2 E \left( u_i^2 I(|u_i| > \epsilon/c_N) \middle| X \right) \\ &= o_P(1) \cdot \sum_{i=1}^N \|C_{N,i}\|^2 = o_P(1), \end{aligned} \tag{33}$$

that is, a *conditional* Lindeberg condition is fulfilled. Now we obtain by the Lindeberg-Feller central limit theorem that

$$(B_N^{-1} - A_N^{-1})^{-1/2} \sqrt{N}(\hat{\beta}_{FE} - \hat{\beta}_{GLS}) \xrightarrow{d} \mathcal{N}(0_K, I_K),$$

which implies by the continuous mapping theorem

$$\mathcal{H}_N = N(\hat{\beta}_{FE} - \hat{\beta}_{GLS})' (B_N^{-1} - A_N^{-1})^{-1} (\hat{\beta}_{FE} - \hat{\beta}_{GLS}) \xrightarrow{d} \chi^2(K).$$

The second assertion (15) follows immediately from  $\|\hat{A}_N - A_N\| \xrightarrow{P} 0$  and  $\|\hat{B}_N - B_N\| \xrightarrow{P} 0$ .  $\square$

*Proof of Proposition 2.* Analogous to the proof of Proposition 1. □

*Proof of Proposition 3.* We will first show that

$$d \left( \mathcal{L} \left( (B_N^{-1} - A_N^{-1})^{-1/2} \sum_{i=1}^N C_{N,i} u_i^* \middle| \mathcal{X}_N \right), \mathcal{N}(0_K, I_K) \right) \xrightarrow{P} 0, \quad (34)$$

which implies by the continuous mapping theorem that

$$d(\mathcal{L}(\mathcal{H}_N^* | \mathcal{X}_N), \chi^2(K)) \xrightarrow{P} 0.$$

(34) will actually follow from

$$d \left( \mathcal{L} \left( (B_N^{-1} - A_N^{-1})^{-1/2} \sum_{i=1}^N C_{N,i} u_i \eta_i \middle| \mathcal{X}_N \right), \mathcal{N}(0_K, I_K) \right) \xrightarrow{P} 0 \quad (35)$$

and

$$T_N := \sum_{i=1}^N C_{N,i} (\hat{u}_i - u_i) \eta_i \xrightarrow{P} 0. \quad (36)$$

It follows from (33) that there exists a null sequence  $(\epsilon_N)_{N \in \mathbb{N}}$  such that

$$E \left( \sum_{i=1}^N \|C_{N,i} u_i\|^2 I(\|C_{N,i} u_i\|^2 > \epsilon_N) \middle| X \right) \xrightarrow{P} 0.$$

Let  $\gamma_{N,i} = C_{N,i} u_i I(\|C_{N,i} u_i\| \leq \epsilon_N)$ . It follows from the latter display that

$$\sum_{i=1}^N C_{N,i} u_i u_i' C_{N,i}' = \sum_{i=1}^N \gamma_{N,i} \gamma_{N,i}' + o_P(1). \quad (37)$$

Using  $E(\sum_{i=1}^N C_{N,i} u_i u_i' C_{N,i}' | X) = B_N^{-1} - A_N^{-1}$  we obtain that

$$\begin{aligned} & \left\| E \left( \sum_{i=1}^N \gamma_{N,i} \gamma_{N,i}' \middle| X \right) - (B_N^{-1} - A_N^{-1}) \right\| \\ & \leq \sum_{i=1}^N E(\|C_{N,i} u_i\|^2 I(\|C_{N,i} u_i\| > \epsilon_N) | X) \xrightarrow{P} 0. \end{aligned} \quad (38)$$

For the  $(k, l)$ -th entry of the matrix  $\sum_{i=1}^N \gamma_{N,i} \gamma_{N,i}'$ , we have

$$\begin{aligned} & E \left( \left[ \sum_{i=1}^N (\gamma_{N,i})_k (\gamma_{N,i})_l - E((\gamma_{N,i})_k (\gamma_{N,i})_l | X) \right]^2 \middle| X \right) \\ & = \sum_{i=1}^N E \left( [(\gamma_{N,i})_k (\gamma_{N,i})_l - E((\gamma_{N,i})_k (\gamma_{N,i})_l | X)]^2 \middle| X \right) \\ & \leq \sum_{i=1}^N E \left( [(\gamma_{N,i})_k (\gamma_{N,i})_l]^2 \middle| X \right) \\ & \leq \epsilon_N^2 \sum_{i=1}^N E \left( (\gamma_{N,i})_k^2 \middle| X \right) \\ & \leq \epsilon_N^2 \sum_{i=1}^N E \left( (C_{N,i} u_i)_k^2 \middle| X \right) = o_P(1), \end{aligned}$$



which implies that

$$\sum_{i=1}^N \gamma_{N,i} \gamma'_{N,i} = E \left( \sum_{i=1}^N \gamma_{N,i} \gamma'_{N,i} \middle| X \right) + o_P(1). \quad (39)$$

From (37), (38) and (39) we conclude that

$$\sum_{i=1}^N C_{N,i} u_i u_i' C_{N,i}' = B_N^{-1} - A_N^{-1} + o_P(1),$$

which implies that

$$\begin{aligned} & \text{Cov} \left( (B_N^{-1} - A_N^{-1})^{-1/2} \sum_{i=1}^N C_{N,i} u_i \eta_i \middle| \mathcal{X}_N \right) \\ &= (B_N^{-1} - A_N^{-1})^{-1/2} \sum_{i=1}^N C_{N,i} u_i u_i' C_{N,i}' (B_N^{-1} - A_N^{-1})^{-1/2} \xrightarrow{P} I_K. \end{aligned} \quad (40)$$

Moreover, since  $\sum_{i=1}^N \|C_{N,i} u_i\|^2 \xrightarrow{P} \text{tr}(B^{-1} - A^{-1})$  and, according to (33),  $P(\max_{1 \leq i \leq N} \|C_{N,i} u_i\| > c \mid X) \leq (1/c^2) \sum_{i=1}^N E[\|C_{N,i} u_i\|^2 I(\|C_{N,i} u_i\| > c) \mid X] \xrightarrow{P} 0$  we obtain, for arbitrary  $\epsilon > 0$ , that

$$\begin{aligned} & \sum_{i=1}^N E^* (\|C_{N,i} u_i \eta_i\|^2 I(\|C_{N,i} u_i \eta_i\| > \epsilon) \mid \mathcal{X}_N) \\ & \leq \sum_{i=1}^N \|C_{N,i} u_i\|^2 E^* (\eta_i^2 I(\|C_{N,i} u_i\| \|\eta_i\| > \epsilon) \mid \mathcal{X}_N) \xrightarrow{P} 0, \end{aligned}$$

that is, we have again a conditional Lindeberg condition being fulfilled. Therefore, (35) follows from (40) by the Lindeberg-Feller central limit theorem.

Now it remains to prove (36). We have that

$$T_N = - \sum_{i=1}^N C_{N,i} X_i (\hat{\beta}_{FE} - \beta) \eta_i.$$

Since  $E^*(\eta_i^2 \mid \mathcal{X}_N) = 1$  we obtain that

$$\begin{aligned} E^* (\|T_N\|^2 \mid \mathcal{X}_N) &= \sum_{i=1}^N (\hat{\beta}_{FE} - \beta)' X_i' C_{N,i}' C_{N,i} X_i (\hat{\beta}_{FE} - \beta) \\ &\leq \|\hat{\beta}_{FE} - \beta\|^2 \cdot \max_{1 \leq i \leq N} \|X_i' X_i\| \cdot \sum_{i=1}^N C_{N,i}' C_{N,i} \\ &= O_P(N^{-1}) \cdot o_P(N) \cdot O_P(1) = o_P(1), \end{aligned}$$

that is, (36) holds also true. □

*Proof of Proposition 4.* Analogous to the proof of Proposition 3. □

*Proof of Theorem 1.* Since  $\chi^2(K)$  is a continuous distribution we conclude from Proposition 1 that

$$\sup_{-\infty < z < \infty} |P(\mathcal{H}_N \leq z) - P(\chi^2(K) \leq z)| \xrightarrow{N \rightarrow \infty} 0, \quad (41)$$

and from Proposition 3

$$\sup_{-\infty < z < \infty} |P(\mathcal{H}_N^* \leq z | \mathcal{X}_N) - P(\chi^2(K) \leq z)| \xrightarrow{P} 0. \quad (42)$$

The assertion of the theorem follows now directly from (41) and (42).  $\square$

*Proof of Theorem 2.* Analogous to the proof of Theorem 1.  $\square$

## 7.2 The empirical models

1. Investment equation (Grunfeld, 1958)

$$I_{it} = \nu + \beta_1 F_{it} + \beta_2 C_{it} + u_{it},$$

where

$$\begin{aligned} I_{it} &= \text{annual real gross investment,} \\ F_{it} &= \text{real value of the firm (shares outstanding),} \\ C_{it} &= \text{real value of the capital stock.} \end{aligned}$$

The panel consists of  $N = 10$  large US manufacturing firms over  $T = 20$  years (1935-1954).

2. Cobb-Douglas production function (Munnell, 1990; Baltagi and Pinnoi, 1995).

$$y_{it} = \nu + \beta_1 k_{it}^{(1)} + \beta_2 k_{it}^{(2)} + \beta_3 l_{it} + \beta_4 U_{it} + u_{it},$$

where

$$\begin{aligned} y_{it} &= \text{log gross state product,} \\ k_{it}^{(1)} &= \text{log public capital including highways and streets,} \\ &\quad \text{water and sewer facilities and other public buildings and structures,} \\ k_{it}^{(2)} &= \text{log private capital stock,} \\ l_{it} &= \text{log labor input measured as employment in nonagricultural payrolls,} \\ U_{it} &= \text{state unemployment rate.} \end{aligned}$$

The panel consists of annual observations for  $N = 48$  contiguous states over the period 1970-1986.

3. Gasoline demand (Baltagi and Griffin, 1983)

$$\log \left( \frac{\text{Gas}}{\text{Car}} \right)_{it} = \nu + \beta_1 \log \left( \frac{Y}{\text{Pop}} \right)_{it} + \beta_2 \log \left( \frac{P_{mg}}{P_{gdp}} \right)_{it} + \beta_3 \log \left( \frac{\text{car}}{\text{pop}} \right)_{it} + u_{it},$$

where

$$\begin{aligned}\frac{\text{Gas}}{\text{Car}} &= \text{motor gasoline consumption per car,} \\ \frac{Y}{\text{Pop}} &= \text{real per capita income,} \\ \frac{P_{mg}}{P_{gdp}} &= \text{real motor gasoline price,} \\ \frac{\text{Car}}{\text{Pop}} &= \text{stock of cars per capita.}\end{aligned}$$

The panel consists of annual observations across  $N = 18$  OECD countries over  $T = 19$  years (1960 - 1978).

## References

- Ahn, S.C. and Moon, H.R. (2001). On Large- $N$  and Large- $T$  properties of panel data estimators and the Hausman test. *mimeo, University of Southern California*.
- Amemiya, T. (1971). The estimation of variances in a variance-components model. *International Economic Review*, **12**, 1–13.
- Anselin, L., Florax, R.J.G.M. and Rey, S.J. (Eds.) (2004). *Advances in Spatial Econometrics*. Springer, Berlin.
- Baltagi, B.H. (2001). *Econometric Analysis of Panel Data*. John Wiley, Chichester.
- Baltagi, B.H. and Griffin, J.M. (1983). Gasoline demand in the OECD: An application of pooling and testing procedures. *European Economic Review*, **29**, 745–753.
- Baltagi, B.H. and Kao, C. (2000). Nonstationary panels, Cointegration in panels and dynamic panels. A survey. in: Baltagi (Ed.): Nonstationary panels, Cointegration in panels and dynamic panels. *Advances in Econometrics*, **15**, JAI Press, Amsterdam, 7–52.
- Baltagi, B.H. and Pinnoi, N. (1995). Public capital stock and state productivity growth: Further evidence from an error components model. *Empirical Economics*, **20**, 351–359.
- Bickel, P. and Freedman, D.A. (1981). Some asymptotic theory for the bootstrap. *Annals of Statistics*, **9**, 1196–1217.
- Bollerslev, T., Chou, R.Y. and Kroner, K. F. (1992). ARCH modelling in finance: A review of the theory and empirical evidence. *Journal of Econometrics*, **52**, 5–59.
- Davidson, R. and Flachaire, E. (2001). The wild bootstrap, tamed at last. *GREQAM Document de Travail 99A32*
- Grunfeld, Y. (1958). The determinants of corporate investment, unpublished Ph.D. dissertation (University of Chicago, Chicago).
- Hausman, J.A. (1978) Specification Tests in Econometrics, *Econometrica*, **46**, 1251–1271.
- Herwartz, H. and Neumann, M.H. (2005). Bootstrap inference in single equation error correction models. *Journal of Econometrics*, **128**, 165–193.

- Lillard, L.A. and Willis, R.J. (1979). Components of variation in panel earnings data: American scientists 1960-1970, *Econometrica*, **47**, 437–454.
- Liu, R.Y. (1988). Bootstrap procedures under some non-i.i.d. models. *Annals of Statistics*, **16**, 1696–1708.
- Mammen, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models. *Annals of Statistics*, **21**, 255–285.
- Munnell, A. (1990). Why has productivity growth declined? Productivity and public investment. *New England Economic Review*, January/February, 3–22.
- Nerlove, M. (1971). Further evidence on the estimation of dynamic economic relations from a time series of cross sections. *Econometrica*, **39**, 359–382.
- Pollard, D. (1984). *Convergence of Stochastic Processes*. Springer, New York.
- Swamy, P.A.V.B. and Arora, S.S. (1972). The exact finite sample properties of the estimators of coefficients in the error components regression model. *Econometrica*, **40**, 261–275.
- Wallace, T.D. and Hussain, A. (1969). The use of error components models in combining cross section and time series data, *Econometrica*, **37**, 55–72.
- Wu, C.F.J. (1986). Jackknife, bootstrap, and other resampling methods in regression analysis (with discussion). *Annals of Statistics*, **14**, 1261–1343.

	[N; T]	DGP1		DGP2		DGP3		DGP4		DGP5		$\Sigma$	
		$\mathcal{H}$	$\mathcal{H}^*$	$\mathcal{H}$	$\mathcal{H}^*$	$\mathcal{H}$	$\mathcal{H}^*$	$\mathcal{H}$	$\mathcal{H}^*$	$\mathcal{H}$	$\mathcal{H}^*$	$\mathcal{H}$	$\mathcal{H}^*$
GLS	[5;	<b>13.5</b>	<b>6.20</b>	<b>13.4</b>	5.95	<b>13.8</b>	<b>6.80</b>	<b>13.1</b>	<b>6.10</b>	<b>16.6</b>	<b>8.05</b>	5	4
EGLS( $\hat{\rho}$ )	5]	<b>3.30</b>	<b>7.70</b>	<b>4.85</b>	<b>6.95</b>	5.20	<b>7.75</b>	4.60	<b>7.05</b>	<b>3.80</b>	<b>8.80</b>	2	5
( $\rho = 0$ )		<b>0.00</b>	<b>6.85</b>	<b>0.00</b>	<b>6.25</b>	<b>0.00</b>	<b>7.15</b>	<b>0.00</b>	<b>7.25</b>	<b>0.00</b>	<b>8.80</b>	5	5
( $\rho = 0.25$ )		<b>0.00</b>	<b>6.90</b>	<b>11.8</b>	<b>6.20</b>	<b>10.9</b>	<b>7.80</b>	<b>11.6</b>	<b>6.75</b>	<b>13.4</b>	<b>9.35</b>	5	5
( $\rho = 0.5$ )		<b>0.80</b>	<b>6.40</b>	<b>1.20</b>	5.95	<b>1.00</b>	<b>6.50</b>	<b>1.25</b>	<b>6.10</b>	<b>1.75</b>	<b>8.25</b>	5	4
GLS	[10;	<b>11.1</b>	5.85	<b>11.4</b>	5.55	<b>12.4</b>	<b>6.60</b>	<b>11.3</b>	5.60	<b>13.4</b>	<b>6.55</b>	5	2
EGLS( $\hat{\rho}$ )	5]	<b>1.20</b>	<b>6.10</b>	<b>1.15</b>	5.60	<b>0.85</b>	<b>6.00</b>	<b>1.20</b>	5.50	<b>1.00</b>	5.95	5	2
( $\rho = 0$ )		<b>1.25</b>	5.95	<b>0.70</b>	5.40	<b>0.60</b>	5.95	<b>0.80</b>	5.40	<b>0.70</b>	5.75	5	0
( $\rho = 0.25$ )		<b>1.30</b>	<b>6.80</b>	<b>8.00</b>	5.40	<b>8.00</b>	<b>6.40</b>	<b>7.85</b>	5.30	<b>8.60</b>	<b>6.20</b>	5	3
( $\rho = 0.5$ )		4.70	5.85	4.30	5.45	4.75	<b>6.65</b>	4.20	5.60	4.85	<b>6.60</b>	0	2
GLS	[50;	<b>6.45</b>	4.70	<b>6.25</b>	4.45	<b>6.10</b>	5.05	5.95	4.45	<b>7.10</b>	5.40	4	0
EGLS( $\hat{\rho}$ )	5]	<b>3.25</b>	5.30	<b>3.15</b>	5.00	<b>3.00</b>	4.90	<b>3.30</b>	4.75	<b>3.20</b>	5.45	5	0
( $\rho = 0$ )		<b>3.25</b>	5.30	<b>3.30</b>	5.35	<b>3.15</b>	5.10	<b>3.55</b>	4.95	<b>3.30</b>	5.35	5	0
( $\rho = 0.25$ )		<b>2.95</b>	5.05	<b>3.85</b>	4.65	4.20	4.95	<b>3.80</b>	4.55	4.30	5.20	3	0
( $\rho = 0.5$ )		<b>3.00</b>	4.70	<b>3.15</b>	4.40	<b>2.75</b>	5.00	<b>3.10</b>	4.40	<b>3.05</b>	5.55	5	0
GLS	[100;	<b>6.85</b>	5.50	<b>6.35</b>	4.25	<b>6.40</b>	4.50	<b>6.15</b>	4.25	<b>6.60</b>	4.75	5	0
EGLS( $\hat{\rho}$ )	5]	<b>3.80</b>	5.35	<b>3.70</b>	4.85	<b>3.25</b>	4.30	<b>3.70</b>	4.75	<b>3.25</b>	4.95	5	0
( $\rho = 0$ )		<b>4.00</b>	5.30	<b>3.70</b>	4.80	<b>3.40</b>	4.20	<b>3.70</b>	4.75	<b>3.40</b>	5.05	5	0
( $\rho = 0.25$ )		<b>3.60</b>	5.70	<b>3.35</b>	4.85	<b>3.45</b>	4.45	<b>3.35</b>	4.55	<b>3.65</b>	4.85	5	0
( $\rho = 0.5$ )		<b>3.95</b>	5.40	<b>3.65</b>	4.25	<b>3.35</b>	4.60	<b>3.50</b>	4.20	<b>3.50</b>	4.70	5	0
GLS	[5;	<b>37.1</b>	<b>9.50</b>	<b>38.1</b>	<b>9.00</b>	<b>37.5</b>	<b>9.35</b>	<b>38.2</b>	<b>9.15</b>	<b>40.0</b>	<b>8.30</b>	5	5
EGLS( $\hat{\rho}$ )	10]	<b>18.0</b>	<b>9.25</b>	<b>17.1</b>	<b>9.40</b>	<b>20.6</b>	<b>9.70</b>	<b>17.0</b>	<b>8.75</b>	<b>18.9</b>	<b>9.70</b>	5	5
( $\rho = 0$ )		<b>0.00</b>	<b>9.05</b>	<b>0.00</b>	<b>9.00</b>	<b>0.00</b>	<b>8.80</b>	<b>0.00</b>	<b>9.35</b>	<b>0.00</b>	<b>8.35</b>	5	5
( $\rho = 0.25$ )		<b>0.00</b>	<b>9.25</b>	<b>11.8</b>	<b>9.25</b>	<b>11.6</b>	<b>10.3</b>	<b>11.4</b>	<b>9.75</b>	<b>12.8</b>	<b>10.6</b>	5	5
( $\rho = 0.5$ )		<b>35.7</b>	<b>9.20</b>	<b>36.5</b>	<b>8.85</b>	<b>36.0</b>	<b>9.30</b>	<b>36.8</b>	<b>8.65</b>	<b>38.0</b>	<b>8.15</b>	5	5
GLS	[10;	<b>11.8</b>	<b>7.00</b>	<b>12.6</b>	<b>6.60</b>	<b>11.5</b>	<b>7.85</b>	<b>11.9</b>	<b>6.60</b>	<b>15.6</b>	<b>8.10</b>	5	5
EGLS( $\hat{\rho}$ )	10]	<b>1.95</b>	<b>7.45</b>	<b>2.20</b>	<b>6.70</b>	<b>2.00</b>	<b>7.90</b>	<b>2.25</b>	<b>6.60</b>	<b>2.90</b>	<b>8.35</b>	5	5
( $\rho = 0$ )		<b>1.40</b>	<b>7.45</b>	<b>1.55</b>	<b>6.10</b>	<b>1.30</b>	<b>8.55</b>	<b>1.65</b>	<b>6.60</b>	<b>1.95</b>	<b>8.45</b>	5	5
( $\rho = 0.25$ )		<b>1.45</b>	<b>7.40</b>	<b>8.30</b>	<b>6.70</b>	<b>7.50</b>	<b>8.35</b>	<b>8.10</b>	<b>6.45</b>	<b>9.80</b>	<b>8.55</b>	5	5
( $\rho = 0.5$ )		5.85	<b>7.00</b>	5.50	<b>6.55</b>	5.40	<b>7.85</b>	5.60	<b>6.60</b>	<b>8.10</b>	<b>8.10</b>	1	5
GLS	[50;	<b>6.40</b>	4.90	<b>6.10</b>	5.35	<b>6.75</b>	5.60	5.90	5.45	<b>7.00</b>	5.55	4	0
EGLS( $\hat{\rho}$ )	10]	<b>3.05</b>	5.35	<b>3.35</b>	5.75	<b>3.45</b>	5.50	<b>3.40</b>	5.80	<b>3.85</b>	5.60	5	0
( $\rho = 0$ )		<b>3.70</b>	5.35	<b>4.00</b>	5.40	<b>3.75</b>	5.55	<b>3.75</b>	5.55	<b>3.90</b>	5.45	5	0
( $\rho = 0.25$ )		<b>3.00</b>	5.35	4.85	5.65	5.15	5.50	4.80	5.55	5.65	5.30	1	0
( $\rho = 0.5$ )		<b>2.90</b>	4.90	<b>3.65</b>	5.35	<b>3.90</b>	5.55	<b>3.65</b>	5.50	<b>4.00</b>	5.55	5	0
GLS	[100;	<b>6.55</b>	5.20	<b>7.05</b>	5.60	<b>6.80</b>	5.25	<b>6.85</b>	5.60	<b>7.10</b>	5.05	5	0
EGLS( $\hat{\rho}$ )	10]	<b>3.70</b>	5.50	4.15	5.60	4.20	5.45	4.05	5.50	4.05	4.90	1	0
( $\rho = 0$ )		4.50	5.65	4.80	5.80	4.85	5.60	4.70	5.85	5.20	5.35	0	0
( $\rho = 0.25$ )		<b>3.70</b>	5.60	5.05	5.65	4.70	5.55	4.85	5.55	5.05	5.05	1	0
( $\rho = 0.5$ )		4.05	5.30	4.05	5.60	<b>3.90</b>	5.20	4.10	5.55	<b>3.95</b>	5.05	2	0

Table 1: Selected size estimates for GLS ( $\mathcal{H}, \tilde{\mathcal{H}}$ ) and EGLS based versions ( $\hat{\mathcal{H}}, \hat{\tilde{\mathcal{H}}}$ ) of the Hausman test over various DGPs introduced in Section 4.1.1.  $\mathcal{H}$  and  $\mathcal{H}^*$  indicate if  $\chi^2(2)$  or bootstrap critical values are used, respectively. Bold faced entries correspond to empirical rejection frequencies that are not covered by a 95% confidence band around the nominal level  $\gamma = 5\%$ . The last two columns report how often the empirical size over 5 DGPs differs from the nominal level with 5% significance.

	$[N; T]$	DGP1		DGP2		DGP3		DGP4		DGP5		$\Sigma$	
		$\mathcal{H}$	$\mathcal{H}^*$	$\mathcal{H}$	$\mathcal{H}^*$	$\mathcal{H}$	$\mathcal{H}^*$	$\mathcal{H}$	$\mathcal{H}^*$	$\mathcal{H}$	$\mathcal{H}^*$	$\mathcal{H}$	$\mathcal{H}^*$
GLS	[10; 20]	<b>8.05</b>	<b>7.75</b>	<b>8.55</b>	<b>7.15</b>	<b>8.75</b>	<b>7.00</b>	<b>8.25</b>	<b>7.20</b>	<b>10.7</b>	<b>7.45</b>	5	5
EGLS( $\hat{\rho}$ )		<b>1.80</b>	<b>7.50</b>	<b>2.00</b>	<b>6.45</b>	<b>2.30</b>	<b>6.95</b>	<b>2.05</b>	<b>6.75</b>	<b>2.50</b>	<b>7.60</b>	5	5
( $\rho = 0$ )		<b>1.15</b>	<b>6.75</b>	<b>1.25</b>	5.85	<b>1.60</b>	<b>7.50</b>	<b>1.30</b>	5.75	<b>1.80</b>	<b>8.15</b>	5	3
( $\rho = 0.25$ )		<b>1.00</b>	<b>7.30</b>	<b>7.55</b>	<b>6.10</b>	<b>7.70</b>	<b>7.40</b>	<b>7.25</b>	<b>6.00</b>	<b>9.25</b>	<b>7.85</b>	5	5
( $\rho = 0.5$ )		<b>3.45</b>	<b>7.75</b>	<b>3.65</b>	<b>7.10</b>	<b>3.45</b>	<b>7.05</b>	<b>3.65</b>	<b>7.20</b>	<b>3.70</b>	<b>7.45</b>	5	5
GLS	[50; 20]	5.45	5.25	5.35	5.00	<b>6.10</b>	5.55	5.40	4.90	<b>6.20</b>	5.65	2	0
EGLS( $\hat{\rho}$ )		<b>3.35</b>	5.30	<b>3.75</b>	4.90	<b>3.45</b>	5.50	<b>3.85</b>	4.85	<b>3.80</b>	5.50	5	0
( $\rho = 0$ )		<b>3.95</b>	5.15	<b>4.00</b>	5.05	4.20	5.45	<b>3.90</b>	5.00	4.15	5.05	3	0
( $\rho = 0.25$ )		<b>3.60</b>	5.25	5.20	4.90	5.15	5.50	5.15	4.80	5.40	5.65	1	0
( $\rho = 0.5$ )		<b>3.50</b>	5.25	<b>3.80</b>	5.00	<b>3.55</b>	5.60	<b>3.75</b>	4.90	<b>3.85</b>	5.65	5	0
GLS	[100; 20]	5.50	5.15	<b>6.00</b>	5.05	5.30	5.00	5.90	5.00	5.00	4.75	1	0
EGLS( $\hat{\rho}$ )		<b>4.00</b>	5.10	4.10	5.05	<b>3.10</b>	5.05	4.05	4.90	<b>3.65</b>	4.80	3	0
( $\rho = 0$ )		4.25	4.95	4.65	5.05	4.10	5.20	4.55	4.70	4.10	4.90	0	0
( $\rho = 0.25$ )		<b>3.90</b>	5.50	5.25	5.10	4.05	5.05	5.15	5.00	4.40	5.10	1	0
( $\rho = 0.5$ )		<b>3.95</b>	5.15	<b>4.00</b>	5.05	<b>2.95</b>	4.95	4.15	5.00	<b>3.45</b>	4.65	4	0
GLS	[50; 50]	4.80	4.85	4.95	4.90	5.40	5.70	4.80	5.10	5.45	5.55	0	0
EGLS( $\hat{\rho}$ )		<b>3.55</b>	4.80	<b>3.75</b>	4.95	<b>3.70</b>	5.85	<b>3.80</b>	5.00	<b>3.80</b>	5.50	5	0
( $\rho = 0$ )		<b>3.80</b>	4.75	4.10	5.20	4.20	5.50	4.10	4.95	4.15	5.15	1	0
( $\rho = 0.25$ )		<b>3.70</b>	4.70	4.90	5.15	5.35	5.80	4.90	5.05	5.60	5.30	1	0
( $\rho = 0.5$ )		<b>3.50</b>	4.85	<b>3.75</b>	4.90	<b>3.70</b>	5.70	<b>3.75</b>	5.10	<b>3.70</b>	5.55	5	0
GLS	[100; 50]	5.00	4.75	5.80	5.00	5.35	4.90	5.90	5.10	5.20	4.70	0	0
EGLS( $\hat{\rho}$ )		<b>3.85</b>	4.70	4.65	5.00	4.40	4.90	4.85	5.00	4.45	4.75	1	0
( $\rho = 0$ )		4.30	4.80	4.95	5.55	4.95	4.85	4.95	5.50	4.50	5.00	0	0
( $\rho = 0.25$ )		4.10	4.80	5.90	5.20	5.90	4.70	5.95	5.20	5.60	4.90	0	0
( $\rho = 0.5$ )		<b>3.85</b>	4.75	4.60	5.00	4.35	4.90	4.75	5.15	4.50	4.70	1	0
$\Sigma$	$\gamma=5\%$	86	45	74	37	74	51	69	42	75	54		
	$\gamma=10\%$	83	49	76	43	72	49	76	45	63	51		
$\Sigma_{T < N}$	$\gamma=5\%$	40	5	31	3	31	8	27	5	32	13		
	$\gamma=10\%$	39	9	33	3	33	9	34	5	29	11		

Table 2: Selected size estimates for GLS ( $\mathcal{H}, \tilde{\mathcal{H}}$ ) and EGLS based versions ( $\hat{\mathcal{H}}, \hat{\tilde{\mathcal{H}}}$ ) of the Hausman test over various DGPs introduced in Section 4.1.1. See also Table 1. The last rows show how often the empirical and nominal size differ with 5% significance when aggregating over all 20 alternative sample sizes used for Monte Carlo analysis, i.e.  $T = 5, 10, 20, 50$ ,  $N = 5, 10, 20, 50, 100$  for a given DGP and testing device. The last two rows provide the latter summary measures obtained over simulation experiments with ( $T < N$ ).

$\gamma$					$T < N$			
	$\mathcal{H}$	$\mathcal{H}^*$	$\mathcal{H}$	$\mathcal{H}^*$	$\mathcal{H}$	$\mathcal{H}^*$	$\mathcal{H}$	$\mathcal{H}^*$
	5%		10%		5%		10%	
GLS	67	47	73	48	32	8	38	8
EGLS( $\hat{\rho}$ )	87	46	85	46	40	6	42	6
( $\rho = 0$ )	77	43	70	49	32	5	28	9
( $\rho = 0.25$ )	65	46	64	45	20	7	21	5
( $\rho = 0.5$ )	82	47	78	49	37	8	39	9
$\Sigma$	378	229	370	237	161	34	168	37

Table 3: Summary measures for empirical size estimates for GLS ( $\mathcal{H}, \tilde{\mathcal{H}}$ ) and EGLS based versions ( $\hat{\mathcal{H}}, \hat{\tilde{\mathcal{H}}}$ ) of the Hausman test over various DGPs. See also Table 1. The table shows the number of violations of the nominal test levels  $\gamma = 5\%, 10\%$  that are significant at the 5% level. The summation is performed for alternative test devices over all 20 ( $10, T < N$ ) considered sample sizes and 5 DGPs. Results for  $\chi^2(2)$  and bootstrap based critical values are distinguished. Summary statistics for experiments with ( $T < N$ ) are also provided.

	$[N; T]$	DGP1		DGP2		DGP3		DGP4		DGP5	
		$\mathcal{H}$	$\mathcal{H}^*$	$\mathcal{H}$	$\mathcal{H}^*$	$\mathcal{H}$	$\mathcal{H}^*$	$\mathcal{H}$	$\mathcal{H}^*$	$\mathcal{H}$	$\mathcal{H}^*$
GLS	[500; 5]	<b>6.05</b>	4.80	<b>7.30</b>	5.45	<b>6.75</b>	5.10	<b>6.60</b>	5.10	<b>7.15</b>	5.20
EGLS( $\hat{\rho}$ )		<b>3.50</b>	4.95	4.45	5.95	4.10	5.60	4.35	5.90	4.15	5.30
( $\rho = 0$ )		<b>3.60</b>	4.85	4.65	5.90	4.25	5.80	4.60	5.80	4.35	5.35
( $\rho = 0.25$ )		<b>3.05</b>	4.75	4.20	5.80	4.10	5.25	4.15	5.55	4.00	5.45
( $\rho = 0.5$ )		<b>3.15</b>	4.85	4.10	5.40	<b>4.00</b>	5.15	<b>4.00</b>	5.10	<b>3.75</b>	5.25
GLS	[1000; 5]	<b>6.85</b>	5.50	<b>6.70</b>	5.25	<b>6.25</b>	5.10	<b>6.65</b>	5.10	<b>6.15</b>	5.10
EGLS( $\hat{\rho}$ )		<b>3.95</b>	5.30	4.05	4.90	<b>3.55</b>	4.95	4.05	5.45	<b>3.60</b>	4.75
( $\rho = 0$ )		<b>4.00</b>	5.35	4.15	5.00	<b>3.95</b>	4.80	4.10	5.45	<b>3.75</b>	4.80
( $\rho = 0.25$ )		<b>3.60</b>	5.85	4.05	5.40	<b>3.55</b>	4.85	<b>3.95</b>	5.35	<b>3.85</b>	4.60
( $\rho = 0.5$ )		<b>4.00</b>	5.40	4.15	5.25	<b>3.30</b>	5.10	<b>3.80</b>	5.15	<b>3.65</b>	5.05

Table 4: Size estimates for GLS ( $\mathcal{H}, \tilde{\mathcal{H}}$ ) and EGLS based versions ( $\hat{\mathcal{H}}, \hat{\tilde{\mathcal{H}}}$ ) of the Hausman test over various DGPs,  $T = 5$ , and large cross section dimensions  $N = 500, 1000$ . See also Table 1.

	[ $N; T$ ]	DGP1		DGP2		DGP3		DGP4		DGP5	
		$\mathcal{H}$	$\mathcal{H}^*$	$\mathcal{H}$	$\mathcal{H}^*$	$\mathcal{H}$	$\mathcal{H}^*$	$\mathcal{H}$	$\mathcal{H}^*$	$\mathcal{H}$	$\mathcal{H}^*$
GLS	[5;	15.6	7.40	14.5	6.15	15.4	6.65	14.2	6.60	17.9	8.30
EGLS( $\hat{\rho}$ )	5]	3.45	8.60	4.65	7.25	5.10	7.90	4.20	7.45	3.25	9.90
( $\rho = 0$ )		0.00	7.25	0.00	6.90	0.00	7.85	0.00	7.05	0.00	9.50
( $\rho = 0.25$ )		0.00	7.75	12.4	6.25	10.5	8.05	12.4	7.40	13.0	9.20
( $\rho = 0.5$ )		1.00	7.80	1.25	6.10	0.70	6.55	1.10	6.60	0.90	8.45
GLS	[10;	13.5	6.70	12.7	6.00	14.1	7.35	12.7	6.10	15.0	7.70
EGLS( $\hat{\rho}$ )	5]	1.35	7.70	1.50	7.75	1.25	6.90	1.45	7.70	1.40	7.15
( $\rho = 0$ )		1.50	8.30	1.05	7.20	1.20	7.55	1.05	7.60	1.25	7.30
( $\rho = 0.25$ )		1.55	7.90	8.65	6.30	8.20	7.85	8.85	6.45	9.25	7.70
( $\rho = 0.5$ )		5.35	6.95	4.55	6.10	5.50	7.15	4.50	6.00	5.65	7.55
GLS	[50;	11.0	8.65	11.2	7.05	11.3	9.25	10.7	7.25	11.9	8.65
EGLS( $\hat{\rho}$ )	5]	5.90	9.70	4.95	8.10	5.85	9.30	4.95	8.05	6.60	9.65
( $\rho = 0$ )		6.25	9.65	5.15	8.25	5.90	9.40	5.20	7.95	6.70	9.55
( $\rho = 0.25$ )		5.40	9.30	6.75	8.00	6.70	9.15	6.80	8.00	7.55	9.50
( $\rho = 0.5$ )		5.35	8.70	5.25	7.20	5.65	9.35	5.10	7.25	6.20	8.75
GLS	[5;	36.4	10.6	37.1	10.1	36.1	10.2	36.5	9.65	37.4	8.65
EGLS( $\hat{\rho}$ )	10]	17.9	9.50	17.1	10.3	19.0	10.2	17.2	9.65	18.8	9.75
( $\rho = 0$ )		0.00	9.35	0.00	10.8	0.00	10.1	0.00	10.3	0.00	9.60
( $\rho = 0.25$ )		0.00	10.3	12.0	10.4	11.7	11.1	11.8	10.9	12.7	10.7
( $\rho = 0.5$ )		34.5	10.9	34.8	9.65	33.9	9.80	34.2	10.0	35.7	9.05
GLS	[10;	13.5	7.65	13.9	6.55	13.0	7.60	13.5	6.60	17.8	8.40
EGLS( $\hat{\rho}$ )	10]	1.80	8.10	2.55	6.30	1.90	8.15	2.30	6.65	2.80	8.55
( $\rho = 0$ )		1.35	7.95	1.65	6.60	1.45	8.00	1.60	6.80	2.00	8.05
( $\rho = 0.25$ )		1.65	7.90	9.00	6.40	7.55	8.10	8.65	6.20	10.2	8.25
( $\rho = 0.5$ )		6.15	7.65	5.95	6.55	5.60	7.55	5.85	6.60	7.85	8.35
GLS	[50;	13.4	11.9	13.8	11.5	13.6	12.6	13.4	11.7	14.5	11.8
EGLS( $\hat{\rho}$ )	10]	7.30	12.0	8.35	12.1	8.25	13.3	8.35	12.3	8.10	13.0
( $\rho = 0$ )		8.65	11.9	9.65	13.2	9.45	13.4	9.70	13.4	9.55	12.7
( $\rho = 0.25$ )		7.60	11.8	11.0	12.4	10.7	13.3	11.0	12.6	11.0	13.0
( $\rho = 0.5$ )		7.35	11.9	7.85	11.5	8.05	12.7	7.60	11.8	8.25	11.9

Table 5: Selected power estimates for GLS ( $\mathcal{H}, \tilde{\mathcal{H}}$ ) and EGLS based versions ( $\hat{\mathcal{H}}, \hat{\tilde{\mathcal{H}}}$ ) of the Hausman test over various DGPs. See also Table 1.



	$\hat{\rho}$	$\hat{\sigma}_\alpha$	$\hat{\sigma}_e$	$\mathcal{H}$	$\chi^2$	$\mathcal{H}^*$	
Investment behavior							
EGLS	$\hat{\rho}$	.664	85.73	39.46	2.604	.272	.000
	$\rho = 0$		85.73	52.77	2.058	.357	.006
	$\rho = 0.5$		85.73	45.70	2.383	.304	.000
Production function							
EGLS	$\hat{\rho}$	.801	0.091	0.023	43.60	.000	.000
	$\rho = 0$		0.091	0.038	8.409	.015	.404
	$\rho = 0.5$		0.091	0.033	15.06	.001	.046
Gasoline demand							
EGLS	$\hat{\rho}$	.778	0.348	0.058	10.18	.006	.028
	$\rho = 0$		0.348	0.092	10.54	.005	.102
	$\rho = 0.5$		0.348	0.080	10.05	.007	.050
Specification testing for subperiods (EGLS using $\hat{\rho}$ )							
Investment behavior							
1935-39		.003	26.06	76.26	2.663	.264	.152
40-44		.171	26.01	89.70	2.005	.367	.521
45-49		.011	34.05	93.52	1.618	.445	.130
50-54		.479	45.28	160.4	3.724	.155	.018
Production function							
1970-73		.122	0.018	0.256	25.72	.000	.000
74-77		.273	0.018	0.256	25.72	.000	.000
78-81		.004	0.017	0.121	14.32	.000	.154
82-86		.196	0.019	0.234	21.41	.000	.004
Gasoline demand							
1960-64		.131	0.049	0.308	7.39	.028	.228
65-69		.257	0.045	0.277	5.22	.074	.353
70-74		.034	0.042	0.304	9.67	.008	.124
75-78		.118	0.330	0.381	10.2	.006	.046

Table 6: Empirical examples for Hausman testing; Upper panel: Estimated variance, (unconditional) autocorrelation parameters and Hausman test results for the entire sample period.  $\mathcal{H}$  indicates the values of the test statistics with corresponding p-values provided in the last two columns (critical values from  $\chi^2(K)$ -distribution, wild bootstrap ( $\mathcal{H}^*$ )). Lower panel: Hausman test results over subperiods covering  $T = 4$  to  $T = 5$  time points.