Inference on time-invariant variables using panel data: a pretest estimator

Jean-Bernard Chatelain, Kirsten Ralf

PARIS SCHOOL OF ECONOMICS
ECOLE D'ECONOMIE DE PARIS

WORKING PAPER N° 2018 – 07

Inference on time-invariant variables using panel data:A pretest estimator

Jean-Bernard Chatelain
Kirsten Ralf

JEL Codes: C01, C22, C23.
Keywords: Time-invariant variables, panel data, pretest estimator, instrumental vari- ables, Mundlak estimator, Hausman-Taylor estimator.

Funded by a French government subsidy managed by the ANR under the framework of the Investissements d'avenir programme reference ANR-17-EURE-001

Inference on time-invariant variables using panel data: A pretest estimator*

Jean-Bernard Chatelain† and Kirsten Ralf‡

January 26, 2021

Abstract

For static panel data models that include endogenous time-invariant variables correlated with individual effects, exogenous averages over time of time-varying variables can be internal instruments. To pretest their exogeneity, we first estimate a random effects model that includes all averages over time of time-varying variables (Mundlak, 1978; Krishnakumar, 2006). Internal instruments are then selected if their parameter is statistically different from zero (Mundlak, 1978; Hausman and Taylor, 1981). Finally, we estimate a Hausman-Taylor (1981) model using these internal instruments. We then evaluate the biases of currently used alternative estimators in a Monte-Carlo simulation: repeated between, ordinary least squares, two-stage restricted between, Oaxaca-Geisler estimator, fixed effect vector decomposition, and random effects (restricted generalized least squares).

JEL classification numbers: C01, C22, C23.

Keywords: Time-invariant variables, panel data, pretest estimator, instrumental variables, Mundlak estimator, Hausman-Taylor estimator.

1 Introduction

Panel data have a time-series and a cross-sectional dimension, but in some studies one or several relevant explanatory variables are time-invariant. Examples of key time-invariant variables are geographical distance for cross-country data in gravity models of foreign trade and foreign direct investments (Serlenga and Shin, 2007); years of schooling, gender, and race when testing Mincer's wage equation using survey data (Hausman and Taylor, 1981); colonial, legal, or political systems; international conflicts; institutional and governance indicators; and initial gross domestic product per capita when testing the convergence of incomes in growth regressions. These variables are often highly relevant and have a high expected correlation with a cross-sectional and time-varying dependent variable in the cross-sectional dimension.

For static panel data models that include endogenous time-invariant variables correlated with individual effects, exogenous averages over time of time-varying variables can be internal instruments. If the number of exogenous internal instruments is at least equal to the number of endogenous time-invariant variables, Hausman and Taylor's (1981) estimator can be applied (for example, Goh and Tham, 2013; Bouvatier, 2014). This estimator assumes that the practitioner knows which averages over time of time-varying explanatory variables are not correlated with individual effects.

*Acknowledgements: We thank two anonymous referees for very helpful comments. We thank William Greene for helpful comments on a previous draft of this paper, see also Greene (2012), section 11.4.5. This paper is forthcoming in *Economic Modelling*.

†Paris School of Economics, Université Paris I Pantheon Sorbonne, 48 Boulevard Jourdan, 75014 Paris, jean-bernard.chatelain@univ-paris1.fr

‡ESCE International Business School, INSEEC U Research Center, 10 rue Sextius Michel, 75015 Paris, Kirsten.Ralf@esce.fr.

However, practitioners rarely know whether a given time-varying explanatory variable is not correlated with individual random effects. Without a pretest for the exogeneity of internal instruments, the Hausman-Taylor estimator faces potential endogeneity bias by wrongly assuming that all internal instruments are exogenous.

Practitioners may also have too many sets of internal instruments available. For example, Greene (2012, Chapter 11) evaluates the returns to schooling using Mincer's wage equation. The unique endogenous time-invariant variable ($g_2 = 1$) is each individual's years of education. The baseline equation includes $k = 9$ time-varying explanatory variables. Their average over time allows up to $k = 9$ internal instruments for the returns to schooling for the Hausman-Taylor estimate. Without a pretest for selecting which subset of $1 \leq k_1$ of internal instruments includes variables that are not correlated with individual effects, the researcher is left with too many choices of instruments sets. When varying the sets of internal instruments, there are no fewer than $\sum_{i=g_2=1}^{i=k_1=9} \binom{9}{i} = 2^9 - 1 = 511$ Hausman-Taylor estimates of the returns to schooling.

Therefore, this paper proposes an instrument selection procedure using Hausman (1978) pretests of the endogeneity of each time-varying regressor. The practitioner computes the average over time of all time-varying variables. In the first step, he runs a random effects estimator that includes all time-varying variables, all time-invariant variables, and, most importantly, all averages over time of all time-varying variables (Mundlak, 1978; Krishnakumar's (2006) estimator). The practitioner selects as internal instruments the subset of $0 \leq k_1 \leq k$ averages over time of time-varying explanatory variables that do not reject the null hypothesis of exogeneity according to Hausman's (1978) test. If the number of selected instruments is at least equal to the number of endogenous time-invariant variables ($k_1 \geq g_2$), the practitioner runs a Hausman-Taylor (1981) estimation using these k_1 internal instruments.

For the first regression, Mundlak (1978) proved that the estimator of the parameter related to the average over time of time-varying variables corresponds to the difference between the within estimator (which is not biased because of endogeneity) and the between estimator (which may be biased because of endogeneity). Hausman and Taylor (1981) demonstrated that testing the null hypothesis of the equality between the within estimator and the between estimator allows us to test the null hypothesis of the correlation between individual random effects and a given explanatory variable. Greene (2012, Section 11.5.6) presents an example of a Hausman (1978) test using Mundlak's (1978) estimator (p. 381). Hausman and Taylor (1981, p. 1382) and Krishnakumar (2006, Section 5.2) confirms that Hausman specification tests are carried out in the same manner whether time-invariant variables are present or not. Guggenberger (2010) completed the theoretical analysis of the statistical properties of Hausman's (1978) pretest for fixed versus random effects in the case of a single instrument.

A robustness check may use an upward testing procedure for instrument selection (Andrews, 1999; Chatelain, 2007) that tests a sequence of joint null hypotheses of an increasing number of parameters of the average over time of time varying variables. For example, if two parameters are individually not statistically different from zero for each of their individual Hausman tests, the upward testing procedure will also test the joint hypothesis that both parameters are simultaneously equal to zero. It is often the case that this joint null hypothesis will not be rejected if each of the two individual null hypotheses has been rejected. The theoretical background that justifies upward testing procedures for selecting instrumental variables can be found in Andrews (1999).

We compute the theoretical determinants of the biases of the estimated parameters and the estimated standard errors of several alternative estimators used by practitioners: repeated between, ordinary least squares, two-stage restricted between, Oaxaca-Geisler estimator, fixed effect vector decomposition, and restricted generalized least squares (random effects) (Greene, 2010, 2012). First, an omitted variable bias on the estimated parameter may occur due to an omitted average over time of endogenous time-varying variables.

Second, the estimator of the standard error of the parameters of time-invariant variables may be biased because of an excessive weight on within-mean square error and on within-degrees of freedom (depending on the time dimension T). Since a time-invariant variable has no variance in the time direction, it can only explain the variance of a time- and individual-varying variable in its individual direction. The reason it matters for inference is because "*the effect of a random component can only be averaged out if the sample increases in the direction of that random (time or individual) component*" (Kelejian and Stephan, 1983; see also Greene, 2012, pp. 404-411; Hsiao, 2014, pp. 58-64). For example, the pooled ordinary least squares (OLS) estimator with time-invariant explanatory variables uses the dimension NT when drawing inferences on time-invariant variables.

Finally, Guggenberger's (2010) Monte Carlo design for evaluation of the Hausman (1978) pretest for fixed versus random effects with a single instrument is extended to the case of the Hausman (1978) pretest for Hausman and Taylor (1981) in this paper. It presents the results of a Monte Carlo simulation that investigates the small-sample characteristics of the pretest and compares them with other available estimators.

The paper proceeds as follows. Section 2 defines the pretest estimator and compares it with alternative estimators used in panel data estimations. Section 3 describes the design of the Monte Carlo simulation, and Section 4 presents the results of this simulation for nine estimators including the pretest. Section 5 concludes.

2 A pretest estimator including time-invariant variables and correlated individual effects

2.1 The model

The static model of time-series cross-section regression estimates the following equation:

$$y = X\beta + (\mathbf{I}_N \otimes \mathbf{e}_T) Z\gamma + (\mathbf{I}_N \otimes \mathbf{e}_T) \alpha + \varepsilon \ , \tag{1}$$

where the $NT \times 1$ vector $y = [y_{it}]$ denotes the endogenous variable. Observations are ordered first by individual and then (in case of equality) by time. Subscripts indicate variation over individuals ($i = 1, ..., N$) and time ($t = 1, ..., T$). β is a $k \times 1$ vector and γ is a $g \times 1$ vector of coefficients associated with time-varying and time-invariant observable variables, respectively. X is a $NT \times k$ matrix of full rank with rows X_{it} of cross-sections time-series data. $\mathbf{e}_T$ is $T \times 1$ vector of ones and $\otimes$ is the Kronecker product. The vector $(\mathbf{I}_N \otimes \mathbf{e}_T) \alpha$ and each column of $(\mathbf{I}_N \otimes \mathbf{e}_T) Z$ are NT vectors having blocks of T identical entries within each individual $i = 1, ..., N$. Z is a $N \times g$ matrix of full rank with rows Z_i for g time-invariant variables. The constant can be included in the matrix of time-invariant variables, in which case $Z_{1it} = 1$; it is also individual invariant. The disturbances ε_{it} are assumed to be uncorrelated with the columns of (X, Z, α) and have zero mean $E(\varepsilon) = 0$ and constant variance covariance matrix $\sigma_\varepsilon^2 \mathbf{I}_{NT}$. The individual effect α is a $N \times 1$ matrix with rows α_i assumed to be a time-invariant random variable, distributed independently across individuals with zero mean $E(\alpha) = \mathbf{0}$ and variance covariance matrix $\sigma_\alpha^2 \mathbf{I}_N$ and such that $E(\alpha_i \mid X_{it}, Z_i) \neq 0$.

2.2 The pretest estimator based on the Mundlak-Krishnakumar unrestricted GLS estimator

Although $E(\alpha_i \mid X_{it}, Z_i) \neq 0$ need not be linear, Mundlak (1978) and Krishnakumar (2006) introduce an auxiliary equation that assumes a linear relation between the individual random effects and the explanatory variables:

$$\alpha = \overline{X}.\pi + Z\phi + \alpha^M, \tag{2}$$

where $\overline{X}_{\cdot} = \frac{1}{T}\left(\mathbf{I}_N \otimes \mathbf{e}_T^{/}\right) X$ is a $N \times k$ matrix computing the average over time of each time-varying and individual-varying variable for N individuals. π is a k vector of coefficients associated with the average over time of time-varying variables. ϕ is a g vector of coefficients associated with the time-invariant observable variables. The N disturbance terms α^M are normally distributed, $\alpha^M \sim \left(\mathbf{0}, \sigma^2_{\alpha^M} \mathbf{I}_{NN}\right)$.

Let the projection matrix on the column space of a matrix $\mathbf{Q}$ is $\mathbf{Q}\left(\mathbf{Q}'\mathbf{Q}\right)^{-1}\mathbf{Q}'$. The between projection matrix on the subspace spanned by individual dummy variables stacked in the matrix $\mathbf{Q} = \mathbf{I}_N \otimes \mathbf{e}_T$ is $\mathbf{B} = \mathbf{I}_N \otimes \left(\frac{1}{T}\mathbf{e}_T \mathbf{e}_T^{/}\right)$. The within projection matrix is $\mathbf{W} = \mathbf{I}_{NT} - \mathbf{B}$. One has $\mathbf{B}' = \mathbf{B}$, $\mathbf{B}^2 = \mathbf{B}$, $\mathbf{W}^2 = \mathbf{W}$ and $\mathbf{BW} = \mathbf{0}$. The auxiliary equation (which is an N vector) can be repeated T times and written as an NT vector:

$$\left(\mathbf{I}_N \otimes \mathbf{e}_T\right)\alpha = \mathbf{B}X\pi + \left(\mathbf{I}_N \otimes \mathbf{e}_T\right) Z\phi + \left(\mathbf{I}_N \otimes \mathbf{e}_T\right)\alpha^M, \tag{3}$$

where $\mathbf{B}X = \left(\mathbf{I}_N \otimes \mathbf{e}_T\right)\overline{X}_{\cdot}$ is a $NT \times k$ matrix that repeats T times the average over time of each time-varying variable.

Combining the auxiliary equation as an NT vector with the initial equation (1) yields:

$$y = X\beta + \mathbf{B}X\pi + \left(\mathbf{I}_N \otimes \mathbf{e}_T\right) Z\left(\gamma + \phi\right) + \left(\mathbf{I}_N \otimes \mathbf{e}_T\right)\alpha^M + \varepsilon, \tag{4}$$

If $\phi \neq 0$, Krishnakumar (2006) mentions that one can only estimate the sum $\gamma + \phi$ and one cannot identify γ and ϕ separately with her estimator.

Using $\mathbf{W} + \mathbf{B} = \mathbf{I}_{NT}$, one introduces the sum of the between and within transformed dependent variables, explanatory variables and the disturbances $\mathbf{W} + \mathbf{B} = \mathbf{I}_{NT}$ in equation (**??**). Using the property of the within and between as two orthogonal and complementary projectors $\mathbf{WB} = \mathbf{0}$, one finds a system of two equations. In the within subspace of dimension $NT - N$, the first equation is:

$$\mathbf{W}y = \mathbf{W}X\beta + \mathbf{W}\varepsilon \tag{5}$$

with a within estimator, using $\mathbf{W}'\mathbf{W} = \mathbf{W}^2 = \mathbf{W}$.

$$\widehat{\beta}_W = \left(X'\mathbf{W}X\right)^{-1} X'\mathbf{W}y$$

The within estimator can be obtained running ordinary least squares (OLS) on within transformed variables and correcting the degrees of freedom to be equal to $NT - N - k$.

In the between subspace of dimension N, the second equation includes the average over time of time varying variables $\mathbf{B}X$, so that Krisnakumar (2008) label this equation an "extended" between model (one has $\mathbf{B}\left(\mathbf{I}_N \otimes \mathbf{e}_T\right) = \left(\mathbf{I}_N \otimes \mathbf{e}_T\right)$).

$$\mathbf{B}y = \mathbf{B}X\left(\pi + \beta\right) + \left(\mathbf{I}_N \otimes \mathbf{e}_T\right) Z\left(\gamma + \phi\right) + \left(\mathbf{I}_N \otimes \mathbf{e}_T\right)\alpha^M + \mathbf{B}\varepsilon, \tag{6}$$

The extended between estimator $\widehat{\beta}_B = \widehat{\beta + \pi}_B$ (which is potentially biased) and $\widehat{\gamma}_B = \widehat{\gamma + \phi}_B$ (which is potentially biased) can be obtained running ordinary least squares on between transformed variables and correcting the degrees of freedom to be equal to $N - k - g - 1$ (taking into account the intercept.

If $\phi = 0$ for all the g time-invariant variables and if $\pi \neq 0$ for all the k average over time of time-varying variables, Mundlak (1978) and Krisnakumar (2006) demonstrated that the **best linear unbiased estimator** is an unrestricted generalized least-squares estimator (unrestricted random effects) estimator, denoted MK-GLS:

$$\widehat{\beta}_{MK-GLS} = \widehat{\beta}_W,\ \widehat{\pi}_{MK-GLS} = \widehat{\beta}_B - \widehat{\beta}_W,\ \widehat{\gamma}_{MK-GLS} = \widehat{\gamma}_B$$

Mundlak (1978) and Krisnakumar (2006) demonstrated that:

$$var\left(\widehat{\beta}_{MK-GLS}\right) = var\left(\widehat{\beta}_{W}\right), \text{ and } var\left(\widehat{\gamma}_{MK-GLS}\right) = var\left(\widehat{\gamma}_{B}\right), \tag{7}$$

where $\widehat{\beta}_W$ and $\widehat{\beta}_B$ are the within and between estimators for time-varying variables. Because the within and between estimator lie in orthogonal subspaces, the variance of their difference is the sum of their variance:

$$var\left(\widehat{\pi}_{MK-GLS}\right) = var\left(\widehat{\beta}_{B} - \widehat{\beta}_{W}\right) = var\left(\widehat{\beta}_{B}\right) + var\left(\widehat{\beta}_{W}\right),$$

For the time-invariant variables, the estimator $\widehat{\gamma}_{MK-GLS}$ and the estimator of the standard deviations $\widehat{\sigma}_{\widehat{\gamma}_{GLS}}$ are exactly the ones of the extended between estimator.

The usual GLS or random effects estimator is labeled "restricted" GLS estimator by Mundlak (1978), since the vector of parameters of the averages over time of the time-varying variables are restricted to be equal to zero: $\pi = 0$. It will be referred to as R-GLS.

The MK-GLS estimator deals with the endogeneneity of time-varying variables $E(\alpha \mid X) \neq 0$. However, the problem of endogeneity of time-invariant variables $E(\alpha \mid Z) \neq 0$ remains to be solved.

Hausman and Taylor (1981) address the problem of endogeneity of time-invariant variables. They assume that the practitioner has prior information, namely which columns of X and Z are asymptotically uncorrelated with α and which are correlated. Then $[X_{it}] = [X_{1it}, X_{2it}]$ and $[Z_i] = [Z_{1i}, Z_{2i}]$ are split into two sets of variables such that X_1 is $NT \times k_1$, X_2 is $NT \times k_2$, Z_1 is $NT \times g_1$, Z_2 is $NT \times g_2$ with $k_1 + k_2 = k$ and $g_1 + g_2 = g$ where k_1 is the number of exogenous variables and k_2 the number of endogenous variables. X_1 and Z_1 are exogenous and neither correlated with α nor with ε, while X_2 and Z_2 are endogenous due to their correlation with α, but are not correlated with ε. If $k > k_1 \geq g_2$, then the number k_1 of internal instruments is sufficient to identify all g_2 endogenous time-invariant variables.

The pretest estimator proposed in this paper combines Mundlak (1978) and Krishnakumar (2006) approach with Hausman and Taylor (1981) approach when the practitioner has no prior information on the exogenous the time-varying variables. It proceeds in two steps.

The practitioner computes the average over time of all time-varying variables. He runs a random effects estimator including all time varying variables, all averages over time of all time-varying variables and all time-invariant variables.

The practitioner selects as internal instruments the subset of $0 \leq k_1 \leq k$ average over time of time-varying explanatory variables which individually do not reject the null hypothesis of exogeneity. Mundlak (1978) estimation provides the statistical information for Hausman (1978) tests for the endogeneity of time-varying variables X. The MK-GLS regression (equation 4) provides all individual Hausman tests which are equivalent to t-test of each null hypothesis of the parameters of the average over time of time-varying variables:

$$H_{0,m} : \pi_m = \beta_{B,m} - \beta_{W,m} = 0, \text{ for each time varying variable indexed by } m \in \{1, ..., k\}$$

The endogeneity of the single time-varying variable indexed by m corresponds to rejecting the null hypothesis $H_{0,m}$. The outcome of the individual tests is that for k_1 cases, $0 \leq k_1 \leq k$, the null hypothesis $H_{0,m}$ is not rejected. The related k_1 average over time of time varying variables can be selected as internal instruments.

Even though the estimator $\widehat{\beta}_B$ is biased due to endogeneity of all explanatory variables in Mundlak-Krishnakumar estimator, this is not the case from the within estimator $\widehat{\beta}_W$ so that the Hausman's (1978) χ^2 statistic evaluates the magnitude of this bias, which corresponds to $\widehat{\pi}_{MK-GLS}$ in Mundlak (1978) and Krishnakumar (2006) approach.

Greene (2012, section 11.5.6) presents an example of this test (11.9, p.381). According to Krishnakumar, (2006), section 5.2, "*the Hausman specification tests are carried out in the same manner whether time invariant variables are present or not*". Hausman and Taylor (1981),

p.1382 consider specification tests of the null hypothesis $H_0 : E(\alpha \mid X, Z) = \mathbf{0}$ against the alternative $H_1 : E(\alpha \mid X, Z) \neq \mathbf{0}$. Under H_0, $\text{plim}_{N\to+\infty}\widehat{\beta}_B - \widehat{\beta}_W = \mathbf{0}$ and under H_1, $plim_{N\to+\infty}\widehat{\pi}_{MK-GLS} = plim_{N\to+\infty}\widehat{\beta}_B - \beta \neq \mathbf{0}$. Using $\widehat{\pi}_{MK-GLS}$ and $var(\widehat{\pi}_{MK-GLS})$ in Mundlak (1978) and Krishnakumar (2006) allows to compute a χ^2 statistic of the Hausman test denoted $J_N(c)$, similar to a Wald test statistic and, in the case of single variable, similar to the square of a Student t-test statistic:

$$J_N(c) = \left(\widehat{\beta}_{B,c} - \widehat{\beta}_{W,c}\right)' \left(var\left(\widehat{\beta}_{B,c}\right) + var\left(\widehat{\beta}_{W,c}\right)\right)^{-1} \left(\widehat{\beta}_{B,c} - \widehat{\beta}_{W,c}\right), \tag{8}$$

where $c \in \mathbb{R}^k$ is an internal instrument selection vector, which corresponds to a selection of parameters $\widehat{\beta}_{B,c} - \widehat{\beta}_{W,c}$ in the statistic $J_N(c)$. The vector c is a vector of zeros and ones. If the jth element of c is a one, then the jth column vector in the matrix $\overline{X}_{\cdot}$ is included in the instrument set. If the jth element of c is a zero, then the jth column vector in the matrix $\overline{X}_{\cdot}$ is not included in the instrument set. Let $|c| = \sum_{j=1}^{j=k} c_j$ denote the number of instrument selected by c.

Mundlak (1978) and Krishnakumar (2006) regression instantly provides t-tests for each parameter of the average over time of the k time varying variables. This tests each of the k selection vectors c with a unique instrument such that $|c| = 1$. A fast selection procedure selects $|c| = k_1$ averages over time of all the time varying variables which did not reject the null hypothesis for the t-test of their parameter $\widehat{\pi}_{MK-GLS,m}$ or, equivalently, for the Hausman test with critical value $\delta_{N,1} = \chi^2_1(\xi_N)$, where $\chi^2_q(\xi_N)$ denotes the $1 - \xi_N$ quantile of a chi-squared distribution with 1 degree of freedom:

$$J_N(c) \leq \delta_{N,1} = \chi^2_1(\xi_N).$$

For this first step of the selection of internal instruments, if $k_1 > 1$, a robustness check may use an upward testing procedure for instrument selection. For example, if two parameters π_m and π_n respectively related to two average over time of time-varying variables $x_{m,\cdot}$ and $x_{n\cdot}$ are individually not statistically different from zero for each of their individual Hausman tests, the upward testing procedure will also test the joint hypothesis that both parameters are statistically different from zero for a Hausman test:

$$H_{0,m,n} : \pi_m = \beta_{B,m} - \beta_{W,m} = 0 = \pi_n = \beta_{B,n} - \beta_{W,n}$$

This upward testing procedure may be time consuming with little gain for practitioners. In practice, it is often the case that the joint hypothesis $H_{0,m,n}$ will not be rejected if each of the two individual null hypothesis $H_{0,m}$ and $H_{0,n}$ have been rejected. In these cases, the instruments selected with an upward testing procedure will be exactly the same than the k_1 ones which did not reject the null hypothesis $H_{0,m}$, $\forall m \in \{1, ..., k\}$.

More precisely, upward testing procedures are based on the statistic $J_N(c)$ and critical values $\delta_{N,q} = \chi^2_q(\xi_N)$ where q is the degree of freedom equal to the number of joint hypothesis $q \in \{1, ..., k_1\}$. Starting with vectors c which did not rejected the null hypothesis for $|c| = 1$, we carry out tests with progressively larger $|c|$ until we find that all tests with the same value of $|c| = \widehat{k}_{UT} + 1$ reject the null hypothesis $H_{0,c}$ that the set of $|c|$ instruments is exogenous. Given $\widehat{k}_{UT}$, we take the upward testing estimator $\widehat{c}_{UT}$ to be the selection vector that minimizes $J_N(c)$ for all c such that $|c| = \widehat{k}_{UT}$.

The theoretical background for upward testing procedures selecting instrumental variables can be found in Andrews (1999) and Chatelain (2007). Mundlak (1978) and Krishnakumar (2006) linear endogeneity assumption (equation (2)) for a static panel data model grounds Hausman (1978) test and the equivalent Hausman and Taylor (1981) within versus between test.

Andrews (1999) key assumption for the existence of an upward testing procedure estimator

is that there exist a sufficient number of "correct" instrumental variables corresponding to a selection vector c_0 such that $|c_0| = k_{UT} \geq g_2$.

Andrews (1999) assumption "T" is that the critical values $\delta_{N,q} = \chi^2_q(\xi_N) \to \infty$ ($\forall q \in \{1, ..., k_1\}$ where q is the number of joint hypothesis) and $\delta_{N,q} = o(N)$, where $\chi^2_q(\xi_N)$ denotes the $1 - \xi_N$ quantile of a chi-squared distribution with q degrees of freedom. Andrews (1999) mentions that assumption "T" holds with the significance level ξ_N satisfying $\xi_N \to 0$ and $\ln(\xi_N) = o(N)$. Assumption "T" avoids that, when N tends to infinity, all Hausman tests rejects the null hypothesis.

In the second and final step, the choice of the estimation procedure depends on the number of exogenous time-varying variables found in the first step with respect to the number of endogenous time-invariant variables:

- If $k_1 < g_2$, the number k_1 of internal instruments is not sufficient to identify all g_2 endogenous time-invariant variables. If there are not enough internal instruments and no $g_2 - k_1$ external instruments available, it is not possible to identify γ and ϕ. In this case, we cannot go further than the first step MK-GLS estimate. Unfortunately, this first step MK-GLS leads to potentially biased parameters for endogenous time-invariant variables.

- If $g_2 \leq k_1 < k$, then the number k_1 of internal instruments is sufficient to identify all g_2 endogenous time-invariant variables, and then the pretest estimates for time-invariant variables are the ones obtained running an unrestricted Hausman-Taylor (denoted U-HT k_1) estimation using the k_1 averages over time of exogenous time-varying variables as internal instruments and **including** the averages over time of the $k_2 = k - k_1$ endogenous time-varying variables in the regression. The usual restricted Hausman-Taylor (denoted R-HT) may face a bias because of **omitting** k_2 endogenous variables of the matrix $\overline{X}$. as compared to the unrestricted Hausman-Taylor estimator of pretest proposed here.

- If $k_1 = k$, all time-varying variables are exogenous, the Hausman-Taylor approach is not applicable, and a R-GLS estimation should be done.

The properties of this pretest estimator are investigated in a Monte-Carlo simulation, including the R-GLS estimator, the U-HT estimator, and the MK-GLS estimator since the pretest chooses among these three possibilities. But first, in the next section, the theoretical properties of alternative estimators and their ability to draw inference are discussed.

2.3 Alternative estimators of time-invariant variables

2.3.1 Unrestricted and restricted OLS estimators (U-OLS and R-OLS)

The first approach to estimate equations (1) and (3) is to use a standard ordinary least-squares estimator. As defined before an estimator is restricted if the parameter values of the averages-over-time of the time-varying variables have to be equal to zero and unrestricted otherwise.

The unrestricted pooled OLS estimator, denoted U-OLS, however, ignores the random effects α_i: it is the best linear unbiased estimator when $\alpha_i = 0$. It leads to the *same parameter estimates* as the MK-GLS estimator, but *not* to the same standard error estimates since it uses the irrelevant $NT - k - g - 1$ degrees of freedom for the time-invariant variables. The $RMSE$ is larger because it includes the sum of squares of the error of the within model:

$$\begin{aligned}\widehat{\sigma}^{U-OLS}_{\widehat{\gamma}} &= \frac{\sqrt{\frac{T \cdot SSE_B + SSE_W}{NT-k-g-1}}}{\sqrt{T \cdot CSS(z_i)}} \frac{1}{\sqrt{1 - R^2_A(Z_j)}} \\ &= \sqrt{1 + \frac{SSE_W}{T \cdot SSE_B}} \sqrt{\frac{N-k-g-1}{NT-k-g-1}} \widehat{\sigma}^B_{\widehat{\gamma}},\end{aligned}$$

where subscripts W and B refer again to the within and between estimations. $R^2_A(Z_j)$ is the coefficient of determination of the U-OLS auxiliary regression of the variable Z_j on all the other

regressors: it is the same for the between and U-OLS estimators.

The restricted pooled OLS (R-OLS) estimator faces an omitted variables bias, $X.\left(\widehat{\beta}_B - \widehat{\beta}_W\right)$, on the estimated parameters when estimating the Mundlak Krishnakumar model (equation 4). For details see e.g. Oaxaca and Geisler [2003] who evaluate the consistency of the R-OLS estimator of a parameter of a time-invariant explanatory variable.

Both OLS and R-OLS estimators are included in the Monte Carlo simulation.

2.3.2 Restricted random effects (R-GLS) and restricted Hausman Taylor (R-HT) estimators

The restricted random effect estimator (R-GLS) is one of the most used estimators by applied econometricians when their variables of interest are time-invariant variables in panel data. It uses the same weight $\widehat{\theta}$ for computing quasi demeaned variables as the MK-GLS, when it is based on the Swamy and Arora method (1972):

$$y_{it} - \overline{y}_{i.} + \widehat{\theta}\overline{y}_{i.} \text{ with } \widehat{\theta} = \frac{RMSE_W}{\sqrt{T} \cdot RMSE_B}.$$

The between regression includes the averages-over-time of all explanatory variables. Hence, the root mean squared error of the between estimator $RMSE_B$ is the same in both cases. R-GLS, however, faces an omitted variable bias when the null hypothesis $\widehat{\beta}_B = \widehat{\beta}_W$ is rejected for at least one time-varying variable. In this case, the bias of the random effect estimator $\widehat{\gamma}_{R-GLS}$ as compared to the Mundlak-Krishnakumar $\widehat{\gamma}_B$ estimator (which is equal to the between estimator) is:

$$\widehat{\gamma}_{R-GLS} = \widehat{\gamma}_B + \sum_{j=1}^{j=k} \left(\widehat{\beta}_B - \widehat{\beta}_W\right) \widehat{\beta}_{\widehat{\theta}\overline{x}_{j.}/\widehat{\theta}z_j}$$

with $\widehat{\beta}_{\widehat{\theta}\overline{x}_{j.}/\widehat{\theta}z_j}$ given by the auxiliary regression using OLS on quasi demeaned variables:

$$\widehat{\theta}\overline{x}_{i.} = \beta_{\widehat{\theta}\overline{x}_{j.}/\widehat{\theta}z_j}\widehat{\theta}z_i + \beta_{\overline{x}_{i.}/x_{it}}\left(\widehat{\theta}\overline{x}_{i.} + x_{it} - \overline{x}_{i.}\right) + \widehat{\theta}\beta_0 + \varepsilon_{it}.$$

The restricted Hausman and Taylor estimator (R-HT) uses different weights $\widehat{\theta}_{R-HT}$ in the term $y_{it} - \overline{y}_{i.} + \widehat{\theta}_{R-HT}\overline{y}_{i.}$ because of the choice of internal instrumental variables. A bias similar to the bias of the R-GLS due to the omission of the average-over-time of endogenous time-varying variables may offset the correction of the endogeneity bias of the time-invariant variables using internal instrumental variables.

Both estimators are included in the Monte Carlo simulation.

2.3.3 T times repeated-between estimator (T-BE)

Another widely used estimator is the T-times-between estimator which is not a good choice. We present it because repeating T times the time-invariant observations turns out to be the major component of the increase of the t-statistic of time-invariant variables using panel data for several other estimators. Kelejian and Stephan (1983), however, argue that *"the effect of a random component can only be averaged out if the sample increases in the direction of that random (time or individual) component"*. By contrast, Oaxaca and Geisler (2003) assume that the consistency of the estimator of a parameter of a time-invariant explanatory variable *"depends on the time series observations approaching infinity"*. The estimated parameters of time-invariant variables are the same for the repeated-between and the between estimator ($\widehat{\gamma}^B = \widehat{\gamma}^{T-BE}$), (the subscript

for this estimator is T-BE). The estimator of the standard error is equal to:

$$\widehat{\sigma}_{\widehat{\gamma}}^{T-BE} = \frac{\sqrt{\frac{T \cdot SSE_B}{NT-k-g-1}}}{\sqrt{T \cdot CSS(z_j)}} \frac{1}{\sqrt{1-R_A^2\left(Z_j\right)}} = \sqrt{\frac{N-k-g-1}{NT-k-g-1}} \widehat{\sigma}_{\widehat{\gamma}}^{B}.$$

Inference uses $NT-k-g-1$ degrees of freedom instead of $N-k-g-1$. The coefficient of determination of the auxiliary between regression $R_A^2\left(Zz_j\right)$ of the variable Z_j on all the other regressors does not change in the between or repeated-between samples. As the parameter estimator is the same as $\widehat{\gamma}^B$, the repeated-between t^{T-BE}-statistic amounts to multiply the between t^B statistics by the following factor:

$$t^{T-BE} = \frac{\widehat{\gamma}^B}{\widehat{\sigma}_{\widehat{\gamma}}^{T-BE}} = \sqrt{\frac{NT-k-g-1}{N-k-g-1}} \frac{\widehat{\gamma^B}}{\widehat{\sigma}_{\widehat{\gamma}}^{B}} = \sqrt{\frac{NT-k-g-1}{N-k-g-1}} t^B.$$

When N is large, the t-statistic of the repeated-between model is multiplied by around $\sqrt{T}$ as compared to the between model. Due to these shortcomings the estimator is not included in the Monte-Carlo simulations.

2.3.4 Two-stage restricted between (R-BE) and Oaxaca Geisler (2003) estimator

A common practice consists of using a **two-stage restricted between** (denoted R-BE) estimator of time-invariant variables regress the average over time of the first stage residuals of the within estimation. For a N vector of group means estimated from the within group residuals, one expands this expression:

$$\overline{y}_{\cdot} - \overline{X}_{\cdot}.\widehat{\beta}_W = Z\gamma + \alpha + \varepsilon_{i\cdot} + \overline{X}_{\cdot}.\left(\beta - \widehat{\beta}_W\right)$$

with $\text{plim}_{N\to+\infty}\widehat{\beta}_W - \beta = \mathbf{0}$ when T is finite. This equation can be used for estimating γ because the last two terms can be treated as unobservable mean zero disturbances. For example, the last two terms can be written as the following NT vector:

$$\left(\mathbf{I}_N \otimes \mathbf{e}_T\right)\left[\varepsilon_{i\cdot} + \overline{X}_{\cdot}.\left(\beta - \widehat{\beta}_W\right)\right] = \left[B + BX_{\cdot}.\left(XWX\right)^{-1}X'W\right]\varepsilon$$

This common practice sets the **restriction** $\widehat{\beta}_B = \widehat{\beta}_W$ for the average-over-time of time-varying variables in the extended between regression. Because of this restriction, the estimated standard error of the parameter of the time-invariant variables is necessarily smaller than the estimated standard error of the between estimator:

$$\widehat{\sigma}_{\widehat{\gamma}_{R-BE}} = \sqrt{\frac{SSE_{R-BE}}{N-g-1}} \frac{1}{\sqrt{CSS(Z_i)}\sqrt{1-R_{A,R-BE}^2\left(Z_j\right)}} \neq$$

$$\widehat{\sigma}_{\widehat{\gamma}_B} = \sqrt{\frac{SSE_B}{N-k-g-1}} \frac{1}{\sqrt{CSS(Z_i)}\sqrt{1-R_A^2\left(Z_j\right)}}.$$

This stems from the fact that the sum of squares of errors SSE_{R-BE} is larger than the SSE_B because the model constrains the parameters of the time-varying variables to their within estimate which may not minimize the between sum of squares of errors. On the other hand, the increase of the estimated standard error may be offset for two reasons due to the fact the averages of k time-varying variables are on the left hand side of the equation. First, the degrees of freedom increase by k. This decreases the root mean squared error. Second, the variance

inflation factor decreases because $R^2_{A,R-BE}(Z_j) < R^2_A(Z_j)$. $R^2_A(Z_j)$ is the coefficient of determination of an auxiliary regression where the time-invariant variable is correlated with the other $g-1$ time-invariant explanatory variables on the right hand side of the equation. In the between estimator, $R^2_A(Z_j)$ is the coefficient of determination of an auxiliary regression where the time-invariant variable is correlated with the other $g-1$ time-invariant explanatory variables and k averages of time-varying explanatory variables. As the number of explanatory variables increases by k, one has $R^2_{A,R-BE}(Z_j) < R^2_A(Z_j)$.

Oaxaca and Geisler's (2003) two-stage estimator deals with two-stage restricted between. The second stage estimates the equation (**??**) with a generalized least-squares estimator using a covariance matrix that takes into account that the disturbances of the restricted between include the omitted term: $\overline{X}_{\cdot}\left(\beta - \widehat{\beta}_W\right)$ for a finite sample. They claim that the consistency of their estimator ($\widehat{\gamma}$) "*depends on time series approaching infinity so that* $T_i \to +\infty, \forall i$" for fixed N, which does not hold for an estimator of the parameters of time-invariant variables. This suggest their estimator depends on observations repeated T times. For these reasons only the two-stage restricted between estimator, but not the Oaxaca-Geisler estimator is included in the Monte Carlo simulation.

2.3.5 Three-stage FEVD restricted estimator

As demonstrated by Greene (2011), the FEVD estimator (Plümper and Troeger (2007)) of time-invariant variables is not valid. Nonetheless, we present this estimator because it turns out to be the extreme case, where the root mean square error of the within regression is used for doing inference on time invariant variables in panel data instead of the root mean square error of the between regression as in the Mundlak-Krishnakumar estimator. Assuming exogeneity ($E(\alpha \mid Z) = \phi_2 = \mathbf{0}$ and $E(\alpha \mid X) = \pi = \mathbf{0}$), a restricted FEVD estimator calculated the time-invariant residuals of stage-two restricted-between, denoted $\hat{\eta}_{2R-BE}$:

$$\begin{aligned}\hat{\eta}_{2R-BE} &= \hat{d} - Z\hat{\gamma}_{R-BE} = \overline{y}_{\cdot} - \overline{X}_{\cdot}\hat{\beta}_W - Z\hat{\gamma}_{R-BE} \\ &= \alpha^M + \left(B - \overline{X}_{\cdot}\left(X'WX\right)^{-1}X'W\right)\varepsilon_{\cdot}\end{aligned}$$

The third stage of a restricted FEVD estimator is an OLS regression which includes the time-invariant residuals of the second stage, $\hat{\eta}_{2R-BE}$, as an additional regressor with a parameter δ:

$$y = X\beta + (\mathbf{I}_N \otimes \mathbf{e}_T) Z\gamma + (\mathbf{I}_N \otimes \mathbf{e}_T)\hat{\eta}_{2R-BE}.\delta + \eta_{FEDV}$$

where the third stage disturbances are denoted η_{FEDV}. This is equivalent to running OLS for the following equation including only demeaned and time-invariant variables:

$$\begin{aligned}y - \overline{y}_{\cdot} = &\left(X - \overline{X}_{\cdot}\right)\beta - (\mathbf{I}_N \otimes \mathbf{e}_T)\overline{X}_{\cdot}\left(\beta - \hat{\beta}_W\right) + (\mathbf{I}_N \otimes \mathbf{e}_T) Z\left(\gamma - \hat{\gamma}_{R-BE}\right) \\ &+ (\mathbf{I}_N \otimes \mathbf{e}_T)\hat{\eta}_{2R-BE}(\delta - 1) + \eta_{FEVD}.\end{aligned}$$

Because of the orthogonality of demeaned (within transformed) variables with time-invariant variables ($cov(y - \overline{y}_{\cdot}, z_i) = cov(y - \overline{y}_{\cdot}, \overline{x}_{i\cdot}) = cov(y - \overline{y}_{\cdot}, \hat{\eta}_{2,RBE\cdot}) = 0$) and by definition of the within estimator (OLS on demeaned variables), the OLS estimates of third stage FEVD equation are $\hat{\delta}_{FEVD} = 1$, $\hat{\beta}_{FEVD} = \hat{\beta}_W$, $\hat{\gamma}_{FEVD} = \hat{\gamma}_{RBE}$. Hence, the restricted third stage FEVD estimator has the same parameter estimates as the two-stage restricted-between estimator. The only difference is that the third stage residuals $\hat{\eta}_{3,FEVD}$ are the within-regression residuals:

$$\hat{\eta}_{FEVD} = \hat{\varepsilon_{it}} - \hat{\varepsilon_{i\cdot}} = y - \overline{y}_{\cdot} - \left(X - \overline{X}_{\cdot}\right)\hat{\beta}_W.$$

Hence, with respect to the two-stage restricted-between estimator, the only change of the third stage restricted FEVD estimator is the estimated standard error of the parameters γ of time-invariant variables Z which no longer depends on the **mean squared error** of the second-stage time-invariant residuals of the restricted-between estimator $\hat{\eta}_{2R-BE}$ (observed N times, with $MSE_{RBE} = SSE_{RBE}/(N-k-g-1)$), but depends now on the mean square error of the within residuals $\hat{\eta}_{FEVD}$, observed NT times, with $MSE_W = SSE_W/(NT-N-k)$.

The estimated parameters $\hat{\gamma}_{R-BE}$ are related to the projection in the between subspace of observations, but their estimated standard errors are related to residuals in the within subspace of observations, which is orthogonal to the between subspace. This contradicts the Frisch and Waugh (1933) and Lovell (1963) theorem: the same orthogonal projection matrix has to be used in the parameter estimator and in the estimator of the standard error of the estimated parameter.

As observed by Breusch et al. (2011) in Monte Carlo simulations, the FEVD estimated standard error of a time-invariant variable appears "abnormally" small. The restricted FEVD estimator of the standard error of a parameter of a time invariant variable is related to one the best (lowest variance) linear unbiased estimator (BLUE) of the unrestricted Mundlak Krishnakumar model (U-OLS) as follows:

$$\hat{\sigma}_{\hat{\gamma}_{FEVD}} = \frac{\sqrt{\frac{SSE_W}{NT-N-k}}}{\sqrt{T \cdot CSS(z_i)}} \frac{1}{\sqrt{1-R_A^2(z_j)}} = \sqrt{\frac{N-k-g-1}{NT-N-k}} \sqrt{\frac{SSE_W}{T \cdot SSE_B}} \hat{\sigma}_{\hat{\gamma}}^{BE}.$$

Therefore, $\hat{\sigma}_{\hat{\gamma}_{FEVD}}$ is biased downward for two reasons:

- It uses $NT-N-k$ degrees of freedom (with N the number of individuals, k the number of time-varying explanatory variables, with T the number of periods) instead of $N-k-g-1$ degrees of freedom.

- It multiplies the repeated-between estimator of the standard error by a positive factor $\sqrt{\frac{SSE_W}{T \cdot SSE_B}}$ which can be much smaller than one when T increases.

For these reasons the FEVD estimator has not been included in the Monte Carlo simulation. Its large number of citations in google scholar database (around 900 citations during 11 years) is correlated with the researchers' demand for statistically significant parameters for time-invariant variable in panel data in order to be published in academic journals (Wasserstein and Lazar, 2016). In order to reach a large number of citations, theoretical econometricians may be tempted to propose new computations of *downward-biased* standard errors of estimated parameters to applied econometricians. By contrast, stating that the correct degrees of freedom that should be used is $N-k-g-1 << NT-N-k$ for inference of the effect of time-invariant variables is on the opposite side of this popular demand.

Using $NT-N-k$ degrees of freedom instead $N-k-g-1$ is a "magical solution" to the following issue. If the full population is small (e.g. the cross-section of $N=35$ OECD countries), so that the sample size N is small with a p-value larger than 0.05 and nonetheless the cross-section correlation of the dependent variable is large with a time-invariant variable, then the p-value below 5% may not be a useful criterion with respect to a loss function of decision makers (Wasserstein and Lazar (2016), Chatelain (2010) section 2). For a detailed discussion of the FEVD estimator see also Pesaran and Zhou (2018).

3 A Monte Carlo simulation: The model

3.1 Design

To analyze the finite sample properties of the pretest estimator and compare it to the other methods mentioned in the previous section, our Monte Carlo simulations extend a Hausman-Taylor

world using Guggenberger's (2010) design. The reason why we selected Guggenberger's (2010) complete design of the correlation matrix between regressors is that Plümper and Troeger's (2007) FEVD Monte-Carlo simulations presented some surprising results that could only be theoretically explained if some other correlations between variables where changing at the same time they were changing one parameter in their simulations.

Guggenberger's design controls the correlation matrix of all explanatory variables, including the individual effect, while assuming multi-normality. Additional simulations assuming specific cases of non-normality are available upon request from the authors. The simulated model includes three time-varying exogenous variables (X_{1it},X_{2it},X_{3it}), two time-invariant variables ($Z_{1,i}$ and $Z_{2,i}$) and an individual effect α_i explaining an endogenous time-varying variable y_{it}:

$$y_{it} = \beta_1 X_{1it} + \beta_2 X_{2it} + \beta_3 X_{3it} + \gamma_1 Z_{1,i} + \gamma_2 Z_{2,i} + \alpha_i + \varepsilon_{it},\ t = 1 \dots T,\ i = 1 \dots N. \quad (9)$$

The disturbance ε_{it} is independently normally distributed with mean 0 and standard deviation $\sigma_\varepsilon > 0$ uncorrelated with the other endogenous variables. The exogenous time invariant variable $Z_{1,i}$ is a constant, $Z_{1,i} = 1$; it is also individual invariant.

Following Guggenberger (2010), firstly, the time-varying explanatory variables, indexed by $m = 1, 2, 3$ and denoted $X_{mit}, t = 1, \dots T$, and $T = 5$, follow an auto-regressive process of order one, AR(1). Then, in the multi-normal setting, only the correlation matrix has to be defined:

$$\mathbf{R}\left(X_{m,it}\right) = \begin{pmatrix} 1 & r & r^2 & r^3 & r^4 \\ r & 1 & r & r^2 & r^3 \\ r^2 & r & 1 & r & r^2 \\ r^3 & r^2 & r & 1 & r \\ r^4 & r^3 & r^2 & r & 1 \end{pmatrix}$$

with r being the coefficient of autocorrelation between adjacent periods for a time varying variable X_{mit}. It is the same for the three time varying variables.

Secondly, the correlation matrix of all variables corresponds to a Hausman-Taylor world. By construction, the correlations of the regressors X_{1it}, X_{2it}, X_{3it}, Z_{2i} with the constant Z_{1i} are equal to zero. The full correlation matrix $\mathbf{R}$ without the constant variable Z_{1i} is a 17×17 matrix (block matrices are in bold and the matrix is symmetric so that the lower triangular matrix is not reported). The constraint $\det(\mathbf{R}) \geq 0$ restricts the range of values of simple correlation coefficients.

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}\left(X_1, r\right) & \mathbf{0} & \mathbf{0} & \rho_{X_1 Z_2} & \rho_{X_1,\alpha} \\ & \mathbf{R}\left(X_2, r\right) & \mathbf{0} & \rho_{X_2 Z_2} & \rho_{X_2,\alpha} \\ & & \mathbf{R}\left(X_3, r\right) & \rho_{X_3 Z_2} & \rho_{X_3 \alpha} \\ & & & 1 & \rho_{Z_2 \alpha} \\ & & & & 1 \end{pmatrix}$$

Thirdly, the time- and individual-varying variables X_{mit}, the time-invariant variable Z_{2i} and the individual random effects α are drawn from a standardized *multi-normal* distribution with mean zero and standard deviations for all variables $\sigma_{X_{m,it}} = \sigma_{Z_{2ii}} = 1$ with the exception of the individual random effects α_i, for which we assume a variance σ_α^2. In the simulations, we vary the share of the individual effect disturbances $\sigma_\alpha^2 / \left(\sigma_\alpha^2 + \sigma_\varepsilon^2\right)$. The overall variance of the disturbances is fixed to 3 ($\sigma_\alpha^2 + \sigma_\varepsilon^2 = 3$), as in Im, Ahn, Schmidt and Wooldridge (1999) and Baltagi, Bresson and Pirotte (2003).

Finally, the parameter values of the regression equation are chosen as $\beta_1 = \beta_2 = \beta_3 = \gamma_1 = \gamma_2 = 1$. Setting these five parameters is equivalent to setting the simple correlations between the dependent variable and the explanatory variables, once the correlation matrix and the variances of explanatory variables are given.

Guggenberger's (2010) multi-normal design has four advantages:

Firstly, it controls all the correlations between the four explanatory variables and the random individual term for all periods.

Secondly, it is checked that the correlation matrix has a strictly positive determinant

$(\det(\mathbf{R}) \geq 0)$, so that the matrix is positive definite for all simulations.

Thirdly, in the benchmark simulation, the time-invariant variables are *not* drawn from a uniform distribution instead of a normal distribution (the distributions of regressors do not have excess kurtosis in the between space). We nevertheless controlled for non-normality of the individual effect and the random disturbance term in an additional simulation.

Fourthly, and most importantly, this design allows to *separate* the distinct effects of two deviations from the classical model of uncorrelated variables on the size of the Hausman pretest as detailed in Guggenberger (2010). A first deviation is the correlation $\rho_{Z_2\alpha}$ between the random individual effect and the endogenous time-invariant variable. A second deviation is the relative variance of the individual random term with respect to the variance of the time-varying disturbances $\sigma_\alpha^2/\left(\sigma_\alpha^2+\sigma_\varepsilon^2\right)$. In the design of Im *at al.* (1999) and of Plümper and Troeger (2007), it is not possible to change $\sigma_\alpha^2/\left(\sigma_\alpha^2+\sigma_\varepsilon^2\right)$ without changing $\rho_{Z_2\alpha}$ and $\rho_{X_3\alpha}$ at the same time. Furthermore, in the multi-normal design, the simulations allows to check the theoretical result that the OLS endogeneity bias is a linear function of σ_α when $\rho_{Z_2\alpha}$ is fixed.

3.2 Parameter values and estimation

In the benchmark simulation, the variance of random effects is $\sigma_\alpha^2 = 1.5$ so that it represents 50% of the overall variance: $\sigma_\alpha^2/\left(\sigma_\alpha^2+\sigma_\varepsilon^2\right) = 1.5/3 = 1/2$. The correlation coefficients are $\rho_{X_1Z_2} = \rho_{X_2Z_2} = \rho_{X_3Z_2} = 0.4$, $\rho_{X_1,\alpha} = 0$, $\rho_{X_2,\alpha} = 0$, $\rho_{X_3\alpha} = 0.75$, $\rho_{Z_2\alpha} = 0.52$.

We have $k_1 = 2$ exogenous time-varying variables, $k_2 = 1$ endogenous time-varying variable, $g_1 = 1$ exogenous time-invariant variable, and $g_2 = 1$ endogenous time-invariant variable. In the first step of the pretest procedure, the following Mundlak-Krishakumar model is estimated using GLS:

$$y_{it} = \beta_1 X_{1it} + \beta_2 X_{2it} + \beta_3 X_{3it} + \pi_1 \overline{X}_{1i\cdot} + \pi_2 \overline{X}_{2i\cdot} + \pi_3 \overline{X}_{3i\cdot} + \gamma_1 Z_{1i} + \gamma_2 Z_{2i} + \alpha_i + \varepsilon_{it}. \quad (10)$$

According to the estimated values of π_1, π_2, and π_3, the second step chooses one of the following 8 estimators:

(1) MK-GLS: the outcome of individual tests is that π_1, π_2, and π_3 are statistically different from zero. Then the pretest stops at the first step with the MK-GLS estimator.

(2) U-HT1: π_1 is not different from zero and π_2, and π_3 are different from zero. Then we use an unrestricted Hausman-Taylor (U-HT1) estimator with X_1 as the exogenous variable and X_2 and X_3 as endogenous variables.

(3) U-HT2: π_2 is not different from zero and π_1, and π_3 are different from zero. Then we use the U-HT2 estimator, with X_2 as the exogenous variable and X_1 and X_3 as endogenous variables.

(4) U-HT3: π_3 is not different from zero and π_1, and π_2 are different from zero. Then we use the U-HT3 estimator, with X_3 as the exogenous variable and X_1 and X_2 as endogenous variables.

(5) U-HT12: π_1 and π_2 are not different from zero and and π_3 is different from zero. Then we use the U-HT12 estimator with both X_1 and X_2 as exogenous variables and X_3 as the endogenous variable.[1]

(6) U-HT13: π_1 and π_3 are not different from zero and π_2 is different from zero. Then we use the U-HT13 estimator with both X_1 and X_3 as exogenous variables and X_2 as the endogenous variable.

(7) U-HT23: π_2 and π_3 are not different from zero and π_1 is different from zero. Then we use the U-HT23 estimator with both X_2 and X_3 as exogenous variables and X_1 as the endogenous

[1] Additionally, a pooled Hausman test for the contrast of within versus between parameters of (X_{1it}, X_{2it}) could be carried out. If this pooled Hausman test rejects the null joint hypothesis, we revert to the U-HT1 or U-HT2 taking as internal instrument the variable among (X_{1it}, X_{2it}) which has the *lowest p-value* of the test $\pi_1 = 0$ and $\pi_2 = 0$ (both p-values being below the 5% threshold).

variable.

(8) R-GLS: π_1, π_2, and π_3 are all not different from zero. Then we simply run a R-GLS estimation.

Given the above parameterization of our model, the pretest procedure should choose the U-HT12 estimator and use both X_1 and X_2 as exogenous internal instruments. Alternative U-HT1 and U-HT2 estimators are not that bad since at least one of the variables X_1 and X_2 is chosen as an exogenous variables. The Baltagi et al. (2003) pretest considers three possible options instead of the 8 cases dealt with the above pretest (1, 5 and 8): (1) fixed effects (FE) for β but no estimator for γ (this paper pretest reports the Between estimator using MK-GLS), (5) with restricted Hausman Taylor (R-HT12) (this paper pretest reports the unrestricted Hausman Taylor (U-HT12)) and (8) R-GLS (same choice for Baltagi et al. (2003 pretest and for this paper pretest). The simulations report the results of four *restricted* estimators (R-OLS, R-GLS, R-BE, R-HT12), of four *unrestricted* estimators (U-OLS, MK-GLS, BE, U-HT12) and finally of the pretest estimator. Following the advice of Greene (2011) and Breusch et al. (2011), the FEVD estimator is not reported.

The number of individuals N was chosen to be equal to 100, the number of periods T equal to 5, and the experiment war repeated 1000 times. For each estimator, we report the bias of the parameter β_3 and γ_2, the corresponding simulated root mean square error, and the 5% size, which is the frequency of rejections in 1000 replications of the null hypotheses $\beta_3 = 1$ and $\gamma_2 = 1$, respectively at the 5% significance level. The Monte-Carlo simulation then aims at answering two questions: How does the pretest perform as compared to other methods, i.e. how biased are the results? And does the pretest find the correct alternative, i.e. does it choose both, X_1 and X_2 as endogenous internal instruments?

3.3 Theoretical results for the bias

Before the simulation results are presented, some theoretical issues are discussed. For the OLS estimator the bias can be calculated as follows:

$$\begin{aligned}\lim \hat{\beta}_{OLS} &= \lim \left(\frac{X'X}{NT}\right)^{-1} \left(\frac{X'Y}{NT}\right) = \lim \left(\frac{X'X}{NT}\right)^{-1} \left(\frac{X'\left[X\beta+\alpha\right]}{NT}\right) \\ &= \beta + \lim \left(\frac{X'X}{NT}\right)^{-1} \lim \left(\frac{X'\alpha}{NT}\right)\end{aligned}$$

that is, in our case with $\beta_i = 1$ and $\gamma_2 = 1$:

$$\begin{pmatrix} \widehat{\beta}_1 - 1 \\ \widehat{\beta}_2 - 1 \\ \widehat{\beta}_3 - 1 \\ \widehat{\gamma}_2 - 1 \end{pmatrix} = \frac{\sigma_\alpha}{1 - \rho^2_{X_1Z} - \rho^2_{X_2Z} - \rho^2_{X_3Z}} \begin{pmatrix} \rho_{X_1Z}\left(\rho_{ZX_3}\rho_{X_3\alpha} - \rho_{Z\alpha}\right) \\ \rho_{X_2Z}\left(\rho_{ZX_3}\rho_{X_3\alpha} - \rho_{Z\alpha}\right) \\ \left(1 - \rho^2_{X_1Z} - \rho^2_{X_2Z}\right)\rho_{X_3\alpha} - \rho_{X_3Z}\rho_{Z\alpha} \\ \rho_{Z\alpha} - \rho_{ZX_3}\rho_{X_3\alpha} \end{pmatrix}.$$

The biases of the OLS estimator due to endogeneity increase linearly with σ_α for all parameters. The bias for γ increases with $\rho_{Z\alpha}$ and decreases with $\rho_{X_3\alpha}$ when $\rho_{ZX_3} > 0$. The bias for β_m increases with $\rho_{X_3\alpha}$ and decreases with $\rho_{Z\alpha}$ when $\rho_{ZX_m} > 0$. In what follows, we vary only one parameter with respect to the benchmark case, keeping the other parameters unchanged. We focus on the coefficients of the endogenous regressors X_3 and Z, i.e. β_3 and γ_2, respectively. Results on the other coefficients are available upon request from the authors.

4 Results of the simulation

4.1 Changing the variance σ_α^2 of the individual effect

In a first set of simulations we changed the variance of the individual effect keeping the overall variance constant. The values of σ_α were chosen as in Baltagi et al. (2003, p.368, table 4), namely σ_α^2 varying from 0 to 2.75 which corresponds to a variation of $\sigma_\alpha^2/(\sigma_\alpha^2+\sigma_\epsilon^2)$ from 0 to 11/12. Table 1a (all results are reported in appendix B) reports the choice of the pretest estimator for these different values of σ_α. With $\sigma_\alpha^2 = 0$ the pretest estimator chooses the R-GLS (RE) estimator in 87.4% of all replications and the U-HT12 estimator (unrestricted Hausman Taylor estimator with X_1 and X_2 as internal instruments, the correct choice given the parametrization of the model) in 3.4% of all replications. With $0.5 \leq \sigma_\alpha^2 \leq 2.75$, the pretest procedure chooses the U-HT12 estimator in at least 54.4% and at most 86.8% replications.

Table 1b reports the bias, the RMSE, and the 5% size for various estimators of β_3 and γ_2. For σ_α^2 between 0 and 1, the restricted OLS estimator has the lowest RMSE for β_3. Increasing σ_α reduces the RMSE for the restricted and unrestricted Hausman Taylor estimator, the within estimator (used at the first step of the two-step restricted-between estimator), and the pretest for β_3. For γ_2, R-OLS and R-GLS have the lowest RMSE (0.112), but their value is questionable as they use NT observations in their degrees for freedom. By contrast, between (U-OLS) and restricted and unrestricted Hausman and Taylor estimators have the largest RMSE (0.178), but they use N instead of NT in their degrees of freedom. When σ_α increases, the restricted Hausman Taylor RMSE for γ_2 *increases*, whereas the between RMSE *decreases*, as well as the unrestricted Hausman Taylor (U-HT12), because more variance of the dependent variable is explained when including the average-over-time of endogenous time-varying variables.

The R-OLS, between and R-between biases for $\widehat{\beta_3}$ and $\widehat{\gamma_2}$ are linear increasing functions of σ_α, see figure 1a and figure 1b. The R-GLS (the usual "random effect" model) bias for $\widehat{\beta_3}$ increases for value of $\sigma_\alpha \leq 2$ and then decreases as in Baltagi et al. (2003, p.368). The between estimator has the largest bias for $\widehat{\beta_3}$ and the smallest bias for $\widehat{\gamma_2}$ of those 3 estimators. The endogeneity bias for $\widehat{\beta_3}$ is fully corrected when including $X_{3i.}$ in the three unrestricted estimators (U-OLS, MK-GLS, U-HT12). The unrestricted OLS and GLS provide the between estimate for the parameter $\widehat{\gamma_2}$. The restricted between is related to a large omitted variable bias for $\widehat{\gamma_2}$. The biases for $\widehat{\beta_3}$ and for $\widehat{\gamma_2}$ using the restricted Hausman Taylor are negligible. However, they are between 4 to 10 times larger than the one obtained when using the three unrestricted estimators: OLS, GLS and Hausman Taylor. The bias for γ_2 of the pretest estimator increases first because it selects the R-GLS (RE) estimator for $\sigma_\alpha = 0.5$ in 14.5% of replications, then it falls and increases again when $\sigma_\alpha > 1.41$, because it increasingly selects the MK-GLS estimator (between estimator for γ_2), up to 16.8% of replications with $\sigma_\alpha = 1.66$.

The 5% size columns in tables 1b report the frequency of rejections in 1000 replications of $\beta_3 = 1$ and $\gamma_2 = 1$, respectively at the 5% significance level. Since the null hypothesis is always true, this represents the empirical size of the test. As expected, R-OLS and R-GLS (RE) estimators perform badly; they reject the (true) null hypothesis frequently, especially when σ_α is large. On the other hand, R-HT12 and U-HT12 perform well for β_3 and γ_2, giving the required 5% size, while MK-GLS (FE for β_3 and BE for γ_2) do well for β_3 and not so well for γ_2, with 5% size increasing steadily from around 6% for $\sigma_\alpha = 0.5$ to 59% for $\sigma_\alpha = 1.66$. The MK-GLS (BE) for γ_2 5% sizes are nonetheless smaller than R-OLS, R-GLS and R-BE with 5% size at 100% when $\sigma_\alpha > 1$. The pretest exceeds the 5% size (reaching 11.2%) for small values of $\sigma_\alpha = 0.5$ (when it selects the R-GLS (RE) in 14.5% of replications) and then for large values of $\sigma_\alpha > 1.41$ (reaching 8.5% to 19.8%) because it selects the MK-GLS (BE for γ_2) in 6.2% to 16.8% of replications.

4.2 Changing the degree of endogeneity of the time invariant variable $\rho_{Z_2\alpha}$

The correlation between the time-invariant variable and the individual error term indicating the presence of endogeneity has to be contained in the interval $0.426 < \rho_{Z_2\alpha} < 0.597$ so that $det(\mathbf{R}) \geq 0$. Table 2a shows that in our simulations the pretest never selected the wrong internal instrument X_3. The pretest chooses the correct alternative U-HT12 in 88.6% of all cases for $\rho_{Z_2\alpha} = 0.5$. For large correlations or small correlations also the Mundlak/Krishnakumar estimator is chosen.

The biases, RMSE and 5% size are presented in table 2b and the biases are shown in figures 2a and 2b. The restricted OLS estimator, the between estimator and the restricted GLS estimator for β_3 are biased and the bias is decreasing when the correlation increases. The other estimators for β_3 are performing well with a RMSE and 5% size approximately the same. The bias for γ_2 increases linearly for R-OLS, R-between, between and non-linearly for R-GLS and the pretest. Restricted and unrestricted Hausman Taylor estimators perform best, even though the unrestricted estimator is slightly better. The less good performance of the pretest comes from the fact that MK-GLS is sometimes chosen. The 5% size of the pretest remains below 19%.

4.3 Shifting from weak to strong internal instruments

Results when increasing **simultaneously** $\rho_{X_1Z_2}$ **and** $\rho_{X_2Z_2}$ ($-0.461 < \rho_{X_1Z_2} = \rho_{X_2Z_2} < 0.461$, so that $det(R) \geq 0$)) are presented in tables 3a and 3b. Since the problem is symmetric only the results for positive values, $0 < \rho_{X_1Z_2} = \rho_{X_2Z_2} < 0.461$, are stated.

The percentage of times the pretest selects the different alternatives when moving from weak to strong internal instruments is shown in table 3a. With the given parametrization the pretest never chooses the wrong instrument. With increasing correlation, however, the pretest chooses also the MK-GLS estimator on takes only one of the averages of the time- and individual varying variables as instrument. This is due to the fact that the overall correlation matrix is nearly non-invertible because of near multi-collinearity. This will be also reflected in the biases.

Biases of the estimators, their RMSE and 5% size are shown in table 3b and figures 3a and 3b. When the internal instruments are weak ($\rho_{X_1Z_2} = \rho_{X2Z_2} = 0$), the restricted Hausman Taylor bias is identical to the restricted between bias for $\widehat{\gamma}$ (0.62) which is an intermediate step in the restricted Hausman Taylor estimator. By contrast, the unrestricted Hausman bias for $\widehat{\gamma}$ (0.12) is closer to the bias (0.05) of the between estimator. As a consequence, the pretest estimator has a smaller bias than the restricted Hausman Taylor for weak instruments up to the level: $\rho_{X_1Z_2} = \rho_{X_2Z_2} = 0.25$. For large levels of these correlation coefficients ($\rho_{X_1Z_2} = \rho_{X_2Z_2} = 0.45$), whose square terms also decrease the denominator of the R-OLS bias, the overall correlation matrix $\mathbf{R}$ is close to be non invertible because of near-multicollinearity. Then, the bias for γ_2 tends to increase faster (non linearly) up to the level of the restricted between (0.62) for R-OLS, R-GLS and between, except for the unbiased R-HT12 and U-HT12 estimators which benefit from strong internal instruments. It is unfortunately also the case for the bias for β_1 and β_2 in R-OLS, R-GLS and BE, so that the pretest rejects both null hypothesis $\pi_1 = \beta_{1,BE} - 1$ and $\pi_2 = \beta_{2,BE} - 1$ and selects the MK-GLS (BE for γ) estimator in 24.9% of replications. Its bias is then 0.238 with a 5% size equal to 27% in this limit case.

4.4 Other simulations

To test the robustness of our results, different levels of significance and different population size were simulated. Furthermore, the assumptions of normality were relaxed, first a uniform distribution for the time-varying error was assumed and secondly a uniform distribution for the individual random disturbance, keeping the overall variance constant.

Results when increasing the size of the sample and the significance threshold are presented

in table 4. The number of individuals went from $N = 100$ to 500 to 1000. When the sample increases, the absolute value of the t-statistics increases, so that the null hypothesis is more easily rejected, which is also the case when the significance threshold increases from 2% to 5% and to 10%. Then, the pretest selects more often the MK-GLS estimator. In these simulations, the endogenous variable X_3 was never selected as exogenous. We can therefore conclude that the pretest performs well for relatively small samples.

When changing the distribution of the disturbances ε_{it} from normal to uniform (with identical variance but more kurtosis) the results are almost the same. Details are available from the authors upon request. Chatelain and Ralf (2010) used these various estimators on a wage equation evaluating the returns to schooling on the data set used by Greene (2012), section 11.4.5.

5 Conclusion

When a researcher does not know which are the exogenous internal instruments in order to correct the endogeneity bias of some time-invariant regressors, a pretest estimator based upon the Mundlak-Krishnakumar and an unrestricted Hausman-Taylor estimator is a viable alternative to other estimators in terms of bias, RMSE and inference. The procedures are easy to program since it is only required to compute the average over time of time-varying variables and merge them with the initial data set, and then to use random effects and Hausman-Taylor procedures available e.g. in STATA.

An alternative to Hausman and Taylor (1981) estimator is the FEF-IV estimator for the two-way error component model (Chen, Yue and Wong (2020)) and for the one-way error component (Pesaran and Zhou (2018)). Further research may consider pre-tests for instrument selection for these estimators.

Further research may consider the case where the econometricain does not know which time-invariant variables are endogenous. A final cross-section instrumental variables (IV) Hausman test upward testing procedure can be tried. In a first step, the pre-test Hausman and Taylor (1981) estimator including the best selection of a number of internal instruments at least equal to the number of *all time-invariant variables* if one first assume all of time-invariant variables are endogenous. Hausman and Taylor (1981) provides IV estimates for the set of parameters γ_{HT} of all time-invariant variables. We have the set of parameters $(\gamma + \phi)_{MK-GLS}$ of all time-invariant variables with MK-GLS estimator which may be biased because it assumes that all time-invariant variables are exogenous: $\phi = 0$. If the null hypothesis of an IV Hausman test $H_0 : \phi = 0$ (if and only if $H_0 : \gamma_{HT} = (\gamma + \phi)_{MK-GLS}$) is not rejected, then all time-invariant variables are exogenous and MK-GLS is the best linear unbiased estimator. Else, a more precise upward testing procedure (Andrews (1999)) of IV Hausman tests may find out if a subset of time-invariant variables are nonetheless exogenous. The remaining subset may have driven the rejection of the null hypothesis when pooling the estimates for all the time-invariant variables in the Hausman (1978) test statistics.

References

[1] Andrews D. (1999). "Consistent Moment Selection Procedures for Generalized Method of Moments Estimation". *Econometrica* 67(3). pp.543-564.

[2] Baltagi B. H., Bresson G. and Pirotte A. (2003). Fixed Effects, Random Effects or Hausman-Taylor? A Pre-test Estimator. *Economics Letters.* 79, pp. 361-369.

[3] Bouvatier, V. (2014). Heterogeneous bank regulatory standards and the cross-border supply of financial services. Economic modelling, 40, 342-354.

[4] Breusch T., Ward M.B., Nguyen H. and Kompas T. (2011). On the fixed effects vector decomposition. *Political Analysis,* 19, 123-134.

[5] Chatelain, J. B. (2007). Improving consistent moment selection procedures for generalized method of moments estimation. *Economics Letters*, 95(3), 380-385.

[6] Chatelain, J. B. (2010). Can statistics do without artefacts? *Centre Cournot Prisme 19.*

[7] Chatelain, J. B., and Ralf, K. (2010). Inference on time-invariant variables using panel data: A pre-test estimator with an application to the returns to schooling. CES working paper.

[8] Chen, J., Yue, R., & Wu, J. (2020). Testing for individual and time effects in the two-way error component model with time-invariant regressors. Economic Modelling, 92, 216-229.

[9] Frisch R. and Waugh F.V. (1933). Partial time regression as compared with individual trends. *Econometrica* 1, pp. 387-401.

[10] Goh, S. K., & Tham, S. Y. (2013). Trade linkages of inward and outward FDI: Evidence from Malaysia. *Economic Modelling,* 35, 224-230.

[11] Greene W. (2011). Fixed Effects Vector Decomposition: A Magical Solution to the Problem of Time Invariant Variables in Fixed Effects Models? *Political Analysis,* 19, pp. 135-146.

[12] Greene W. (2012). *Econometric Analysis.* 7th edition. Cambridge University Press. Cambridge.

[13] Guggenberger P. (2010). The Impact of a Hausman pretest on the size of a hypothesis test: The Panel data case. *Journal of Econometrics* 156, pp. 337-343.

[14] Hausman J.A. (1978). Specification Tests in Econometrics. *Econometrica* 46, pp. 1251-1271.

[15] Hausman J.A. and Taylor W.E. (1981). Panel data and unobservable individual effects. *Econometrica* 49, pp. 1377-1398.

[16] Hsiao C. (2014). *Analysis of Panel Data.* 3rd edition. Cambridge University Press. Cambridge.

[17] Im K.S., Ahn S.C., Schmidt P. and Wooldridge J.M. (1999). Efficient estimation of panel data models with strictly exogenous explanatory variables. *Journal of Econometrics* 93, pp. 177-203.

[18] Kelejian H.K. and Stephan S.W. (1983). Inference in Random Coefficient Panel Data Models: a Correction and Clarification of the Literature. *International Economic Review* 24, pp. 249-254.

[19] Krishnakumar J. (2006.) Time Invariant Variables and Panel Data Models: A Generalised Frisch-Waugh Theorem and its Implications. in B. Baltagi (editor), *Panel Data Econometrics: Theoretical Contributions and Empirical Applications,* published in the series "Contributions to Economic Analysis", North Holland (Elsevier Science), Amsterdam, Chapter 5, 119-132.

[20] Mundlak Y. (1978). On the pooling of time series and cross section data. *Econometrica* 46, pp. 69-85.

[21] Oaxaca R.L. and Geisler I. (2003). Fixed effects models with time invariant variables: a theoretical note. *Economics Letters* 80, pp. 373-377.

[22] Pesaran M. H. and Zhou Q. (2018). Estimation of time-invariant effects in static panel data models. *Econometric Reviews* 37, pp. 1137-1171.

[23] Plümper D. and Troeger V. (2007). Efficient Estimation of Time Invariant and Rarely Changing Variables in Finite Sample Panel Analyses with Unit Fixed Effects. *Political Analysis* 15, pp. 124-139.

[24] Serlenga L. and Shin Y. (2007). Gravity models of intra-EU trade: application of the CCEP-HT estimation in heterogeeous panels with unobserved common time-specific factors. *Journal of Applied Econometrics* 22, pp. 361-381.

[25] 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 models. *Econometrica* 40, pp. 261-275

[26] Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's statement on p-values: context, process, and purpose. *The American Statistician*, 70:2, 129-133.

Appendix A. Unrestricted Hausman and Taylor (U-HT) estimator

If $k_2 \geq 1$, to obtain consistent estimators for both β and γ in the second stage, let

$$\widehat{d} = \overline{y}_{\cdot} - \overline{X}_{\cdot}\widehat{\beta}_W = \left(B - \overline{X}_{\cdot}\left(X'WX\right)^{-1}X'W\right)y$$

be the NT vector of group means estimated from the within-groups residuals. The only difference of the unrestricted Hausman and Taylor with respect to the usual restricted Hausman and Taylor is to keep the average over time of the endogenous time varying variables $\overline{X}_{2\cdot}$ in each of the steps derived by Hausman and Taylor (1981). Expanding this expression using equation 4 including only $\overline{X}_{2\cdot}$leads to:

$$\widehat{d} = Z_1\gamma_1 + Z_2\left(\gamma_2 + \phi_2\right) + \overline{X}_{2\cdot}\pi_3 + \alpha^M + \left(B - \overline{X}_{\cdot}\left(X'WX\right)^{-1}X'W\right)\varepsilon. \tag{11}$$

Treating the last two terms as an unobservable mean zero disturbance, we estimate γ from the above equation using N observations. If α is correlated with the columns of Z_2, $E\left(\alpha \mid Z_2\right) = \phi_2 \neq 0$, according to prior information, both OLS and GLS will be inconsistent estimators for γ. Consistent estimation is possible, however, if the columns of X_1, uncorrelated with α according to the non rejection of the null hypothesis of preliminary tests, provide sufficient instruments for the columns of Z in equation (11). The two stage least squares (2SLS) estimator for γ in equation (11) is:

$$\widehat{\gamma}_{II} = \left(\left[Z', \overline{X}_{2\cdot}\right]P_A\left[Z', \overline{X}_{2\cdot}\right]\right)^{-1}\left[Z', \overline{X}_{2\cdot}\right]'P_A\widehat{d} \tag{12}$$

where $A = \left[\overline{X}_{1\cdot}, Z_1\right]$ and P_A is the orthogonal projection operator onto its column space. The sampling error is given by

$$\widehat{\gamma}_{II} - \gamma = \left(\left[Z', \overline{X}_{2\cdot}\right]P_A\left[Z', \overline{X}_{2\cdot}\right]\right)^{-1}\left[Z', \overline{X}_{2\cdot}\right]'P_A\left(\alpha^M + \left(B - \overline{X}_{\cdot}\left(X'WX\right)^{-1}X'W\right)\varepsilon\right)$$

and under the usual assumptions governing X and Z, the 2SLS estimator is consistent for γ, since for fixed T, $\operatorname{plim}_{N\to\infty}\frac{1}{N}A'\alpha = 0$ and $\operatorname{plim}_{N\to\infty}\frac{1}{N}X'\varepsilon = 0$.

Having consistent estimators of β and, under the condition $k_1 \geq g_2$, γ, we can construct consistent estimators for the variance components. A consistent estimator of σ_ε^2 can be derived from the within-group residuals in the first step $\widehat{\sigma_\varepsilon^2} = MSE_W$. Whenever we have consistent estimators for both β and γ, a consistent estimator for σ_α^2 can be obtained. Let

$$s^2 = (1/N)\left(\overline{y}_{\cdot} - \overline{X}_{\cdot}\widehat{\beta}_W - Z\widehat{\gamma}_{II} - \overline{X}_{2\cdot}\widehat{\pi}_{2,II}\right)'\left(\overline{y}_{\cdot} - \overline{X}_{\cdot}\widehat{\beta}_W - Z\widehat{\gamma}_{II} - \overline{X}_{2\cdot}\widehat{\pi}_{2,II}\right)$$

then

$$\operatorname*{plim}_{N\to\infty} s^2 = \operatorname*{plim}_{N\to\infty}\frac{1}{N}\left(\alpha + \varepsilon\right)'\left(\alpha + \varepsilon\right) = \sigma_\alpha^2 + \frac{1}{T}\sigma_\varepsilon^2$$

so that $s_a^2 = s^2 - (1/T)\, s_\varepsilon^2$ is consistent for s_a^2.

Appendix B

Table 1a: Percentage of times the pretest selects the internal instruments when changing σ_α

$\sigma_\alpha^2/(\sigma_\alpha^2+\sigma_\varepsilon^2)$	σ_α	MK-GLS	U-HT1	U-HT2	U-HT12	R-GLS	U-HT13	U-HT23
0	0	0	0.3	0.2	3.4	**87.4**	3.2	4.7
1/12	0.5	0.5	3.7	3.3	**75.1**	**14.5**	1.2	1.6
1/6	0.71	0	5	5.2	**88.1**	0.3	0.3	-
1/4	0.87	1	6.1	5.6	**86.8**	-	-	-
1/3	1	1.5	7.7	6.5	**84.1**	-	-	-
1/2	1.22	1.7	8.2	8.6	**79.8**	-	-	-
2/3	1.41	3.4	10.8	10.3	**72.7**	-	-	-
3/4	1.5	6.2	12.4	11.4	**68.2**	-	-	-
10/12	1.58	**11.2**	13.6	12.1	**63.1**	-	-	-
11/12	1.66	**16.8**	16.2	12.6	**54.4**	-	-	-

Figure 1a: Bias of $\widehat{\beta}_3$ when changing σ_α

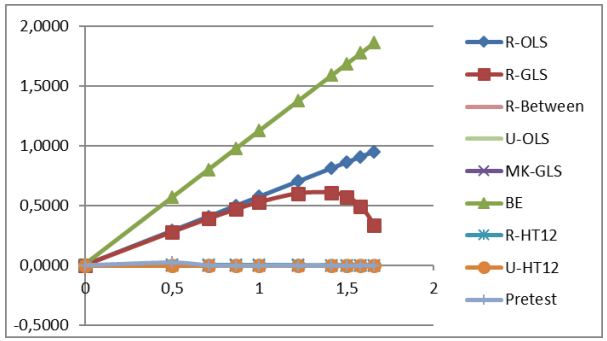

Figure 1b: Bias of $\widehat{\gamma}_2$ when changing σ_α

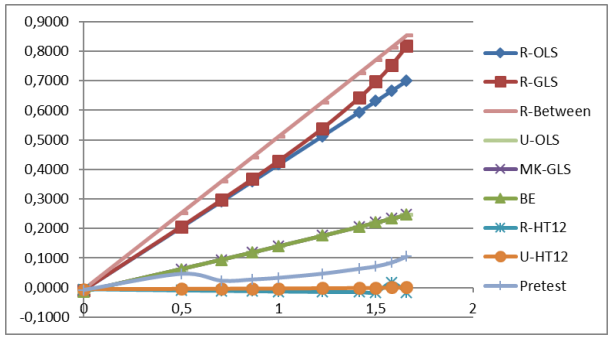

Table 1b: Bias. RMSE and 5% size test for $\widehat{\beta}_3$ and $\widehat{\gamma}_2$ in a Hausman-Taylor world, 1000 replications, N=100 and T=5, changing the variance of the individual effect

		R-OLS			R-OLS			R-GLS			R-GLS			R-Between			R-Between		
σ^2_α	σ_α	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size
0	0	0.0020	0.094	6.2%	-0.0089	0.112	4.8%	0.0020	0.093	5.7%	-0.0090	0.112	4.2%	-0.0005	0.124	4.30%	-0.0064	0.120	18.1%
0.25	0.5	0.2889	0.093	91.2%	0.2035	0.111	49%	0.2792	0.094	88%	0.2060	0.111	47%	-0.0005	0.119	4.30%	0.2522	0.122	75%
0.5	0.71	0.4077	0.093	99.1%	0.2917	0.109	79%	0.3894	0.094	99%	0.2965	0.110	79%	-0.0005	0.113	4.30%	0.3594	0.124	91%
0.75	0.87	0.4988	0.092	99.9%	0.3595	0.108	93%	0.4687	0.095	100%	0.3674	0.108	93%	-0.0005	0.108	4.30%	0.4418	0.127	97%
1	1	0.5756	0.092	100.0%	0.4168	0.107	99%	0.5287	0.097	100%	0.4290	0.108	99%	-0.0004	0.101	4.30%	0.5114	0.129	99%
1.5	1.22	0.7044	0.091	100.0%	0.5131	0.104	100%	0.6025	0.100	100%	0.5387	0.107	100%	-0.0004	0.088	4.30%	0.6282	0.134	100%
2	1.41	0.8129	0.090	100.0%	0.5946	0.101	100%	0.6078	0.100	100%	0.6432	0.107	100%	-0.0003	0.072	4.30%	0.7269	0.139	100%
2.25	1.5	0.8620	0.089	100.0%	0.6317	0.099	100%	0.5729	0.098	100%	0.6970	0.109	100%	-0.0003	0.062	4.30%	0.7717	0.141	100%
2.5	1.58	0.9084	0.089	100.0%	0.6669	0.098	100%	0.4948	0.091	100%	0.7543	0.113	100%	-0.0002	0.051	4.30%	0.8142	0.144	100%
2.75	1.66	0.9524	0.089	100.0%	0.7007	0.096	100%	0.3378	0.071	100%	0.8183	0.123	100%	-0.0002	0.036	4.30%	0.8549	0.146	100%
		U-OLS			U-OLS			MK-GLS			MK-GLS			Between			Between		
σ^2_α	σ_α	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size
0	0	-0.0005	0.124	4.4%	-0.014	0.177	5.9%	-0.0005	0.124	4.60%	-0.0140	0.177	5.5%	0.0070	0.159	6.3%	-0.0140	0.177	6.1%
0.25	0.5	-0.0005	0.119	4.2%	0.0623	0.172	8%	-0.0005	0.119	4.60%	0.0623	0.172	7%	0.5677	0.154	97%	0.0623	0.172	8%
0.5	0.71	-0.0005	0.113	4.1%	0.0943	0.167	11%	-0.0005	0.113	4.60%	0.0943	0.167	10%	0.7998	0.150	100%	0.0943	0.167	10%
0.75	0.87	-0.0005	0.108	3.9%	0.1190	0.162	14%	-0.0005	0.108	4.60%	0.1190	0.162	12%	0.9778	0.145	100%	0.1190	0.162	12%
1	1	-0.0004	0.101	3.8%	0.1400	0.157	18%	-0.0004	0.101	4.50%	0.1400	0.157	15%	1.1277	0.140	100%	0.1400	0.157	15%
1.5	1.22	-0.0004	0.088	3.7%	0.1756	0.146	32%	-0.0005	0.124	4.60%	0.1756	0.15	23%	1.3791	0.130	100.0%	0.1756	0.146	23%
2	1.41	-0.0003	0.072	3.3%	0.2061	0.134	51%	-0.0005	0.119	4.60%	0.2061	0.13	35%	1.5908	0.120	100%	0.2061	0.134	35%
2.25	1.5	-0.0003	0.062	2.9%	0.2202	0.128	62%	-0.0005	0.113	4.60%	0.2202	0.13	42%	1.6865	0.114	100%	0.2202	0.128	42%
2.5	1.58	-0.0002	0.051	2.1%	0.2338	0.122	73%	-0.0005	0.108	4.60%	0.2338	0.12	51%	1.7769	0.108	100%	0.2338	0.122	51%
2.75	1.66	-0.0002	0.036	1.0%	0.2471	0.116	85%	-0.0004	0.101	4.50%	0.2471	0.12	59%	1.8627	0.102	100%	0.2471	0.116	59%
		R-HT12			R-HT12			U-HT12			U-HT12			Pretest			Pretest		
σ^2_α	σ_α	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size
0	0	0.0001	0.124	4.4%	-0.0053	0.178	3.5%	-0.0003	0.124	4.4%	-0.0056	0.176	4.0%	0.0014	0.100	6.20%	-0.0090	0.118	3.9%
0.25	0.5	0.0030	0.119	4.7%	-0.0095	0.186	4.5%	-0.0003	0.119	4.4%	-0.0047	0.169	3.9%	0.0265	0.152	16.80%	0.0461	0.185	11.2%
0.5	0.71	0.0036	0.114	4.7%	-0.0111	0.193	4.4%	-0.0003	0.114	4.4%	-0.0041	0.164	3.6%	0.0006	0.115	4.70%	0.0224	0.164	3.2%
0.75	0.87	0.0037	0.108	4.7%	-0.0123	0.199	4.6%	-0.0003	0.108	4.4%	-0.0036	0.155	3.8%	-0.0002	0.108	4.50%	0.0268	0.155	3.3%
1	1	0.0035	0.102	4.6%	-0.0132	0.205	4.4%	-0.0003	0.102	4.2%	-0.0032	0.148	3.7%	-0.0003	0.102	4.30%	0.0326	0.149	3.8%
1.5	1.22	0.0029	0.088	4.7%	-0.0146	0.217	4.1%	-0.0003	0.088	4.3%	-0.0023	0.133	3.8%	-0.0003	0.088	4.40%	0.0467	0.148	5.6%
2	1.41	0.0019	0.072	4.7%	-0.0156	0.228	4.2%	-0.0002	0.072	4.4%	-0.0013	0.115	3.8%	-0.0003	0.072	4.50%	0.0633	0.141	8.5%
2.25	1.5	0.0014	0.062	4.5%	-0.0160	0.234	4.3%	-0.0002	0.062	4.4%	-0.0008	0.106	3.9%	-0.0002	0.062	4.50%	0.0716	0.139	10.6%
2.5	1.58	0.0009	0.051	4.6%	0.0162	0.239	4.8%	0.0001	0.051	4.4%	0.0002	0.095	3.8%	0.0002	0.051	4.50%	0.0846	0.143	13.8%
2.75	1.66	0.0004	0.036	4.5%	-0.0162	0.244	5.1%	-0.0001	0.036	4.3%	0.0006	0.083	3.7%	-0.0001	0.036	4.30%	0.1052	0.150	19.8%

Table 2a: Percentage of times the pretest selects the internal instruments when changing $\rho_{Z_2\alpha}$

$\rho_{Z_2\alpha}$	MK-GLS	U-HT1	U-HT2	U-HT12	R-GLS	U-HT13	U-HT23
0.45	4.2	10.7	11.4	**73.74**	-	-	-
0.5	0.7	5.7	5	**88.6**	-	-	-
0.52	3.4	8.2	8.6	**79.8**			
0.55	14.2	14.2	16.2	**51.1**	-	-	-

Figure 2a: Bias of $\widehat{\beta}_3$ when changing $\rho_{Z_2\alpha}$

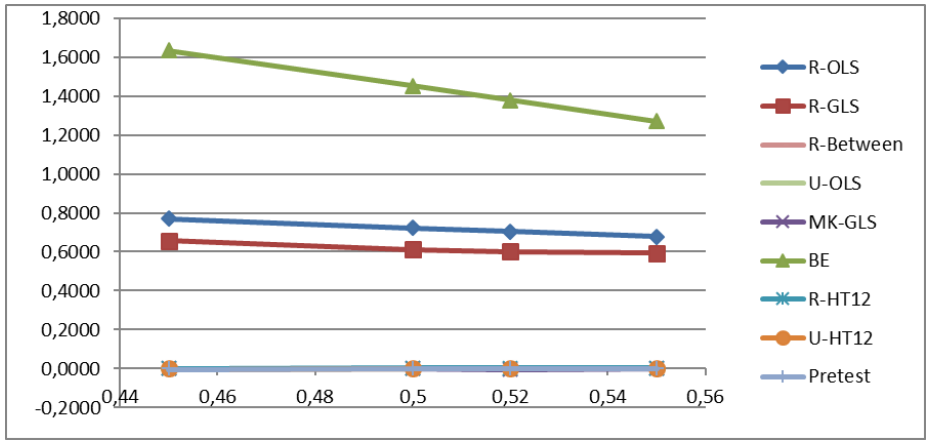

Figure 2b: Bias of $\widehat{\gamma}_2$ when changing $\rho_{Z_2\alpha}$

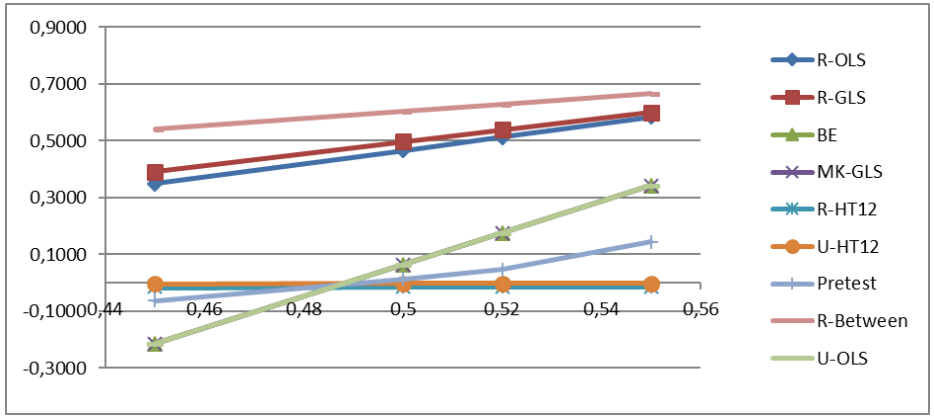

Table 2b: Bias. RMSE and 5% size test for $\widehat{\beta}_3$ and $\widehat{\gamma}_2$ in a Hausman-Taylor world. 1000 replications, N=100 and T=5, changing the variance of the individual effect

	R-OLS			R-OLS			R-GLS			R-GLS			R-Between			R-Between		
$\rho_{z2\alpha}$	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size
0.45	0.7694	0.095	100%	0.3483	0.108	95%	0.6555	0.107	100%	0.3903	0.111	97%	-0.0004	0.088	4.3%	0.5411	0.137	98%
0.5	0.7230	0.092	100%	0.4661	0.105	100%	0.6126	0.102	100%	0.4978	0.108	100%	-0.0004	0.088	4.3%	0.6034	0.351	100%
0.52	0.7044	0.091	100%	0.5131	0.104	100%	0.6025	0.100	100%	0.5387	0.107	100%	-0.0004	0.088	4.3%	0.6282	0.134	100%
0.55	0.6766	0.089	100%	0.5834	0.102	100%	0.5939	0.097	100%	0.5993	0.104	100%	-0.0004	0.088	4.3%	0.6652	0.133	100%

	U-OLS			U-OLS			MK-GLS			MK-GLS			Between			Between		
$\rho_{z2\alpha}$	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size
0.45	-0.0004	0.088	3.6%	-0.2143	0.147	40%	-0.0004	0.088	4.3%	-0.2143	0.147	34%	1.6329	0.129	100%	-0.2143	0.147	34%
0.5	-0.0004	0.088	3.7%	0.0643	0.147	12%	-0.0004	0.088	4.4%	0.0643	0.147	8%	1.4516	0.131	100%	0.0643	0.147	8%
0.52	-0.0004	0.088	3.7%	0.1756	0.146	32%	-0.0005	0.124	4.6%	0.1756	0.146	23%	1.3791	0.130	100%	0.1756	0.146	23%
0.55	-0.0004	0.088	3.7%	0.3425	0.142	76%	-0.0004	0.088	4.4%	0.3425	0.142	69%	1.2703	0.128	100%	0.3425	0.142	69%

	R-HT12			R-HT12			U-HT12			U-HT12			Pretest			Pretest		
$\rho_{z2\alpha}$	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size
0.45	0.0028	0.088	4.3%	-0.0150	0.022	4.4%	-0.0004	0.088	4.2%	-0.0026	0.131	3.5%	-0.0004	0.088	4.4%	-0.0635	0.156	6.7%
0.5	0.0029	0.088	4.7%	-0.0146	0.217	4.0%	-0.0002	0.088	4.2%	-0.0022	0.132	3.7%	-0.0002	0.088	4.4%	0.0138	0.131	3.0%
0.52	0.0029	0.088	4.7%	-0.0146	0.217	4.1%	-0.0003	0.088	4.3%	-0.0023	0.133	3.8%	-0.0003	0.088	4.4%	0.0467	0.148	5.6%
0.55	0.0029	0.088	4.7%	-0.0149	0.217	4.3%	-0.0002	0.088	4.3%	-0.0025	0.133	3.9%	-0.0004	0.088	4.3%	0.1444	0.192	17.8%

Table 3a: Percentage of times the pretest selects the internal instruments when changing $\rho_{X_1Z_2} = \rho_{X_2Z_2}$

$\rho_{X_1Z_2} = \rho_{X_2Z_2}$	MK-GLS	U-HT1	U-HT2	U-HT12	R-GLS	U-HT13	U-HT23
0	0.3	4	5.1	**90.6**	-	-	-
0.1	0.3	4.1	5.3	**90.3**	-	-	-
0.25	0.4	4.7	5.7	**89.2**	-	-	-
0.35	0.8	7	7	**85.2**	-	-	-
0.4	3.4	8.2	8.6	**79.8**	-	-	-
0.45	24.9	13.5	13.2	**48.4**	-	-	-

Figure 3a: Bias of $\widehat{\beta}_3$ when changing $\rho_{X_1Z_2} = \rho_{X_2Z_2}$

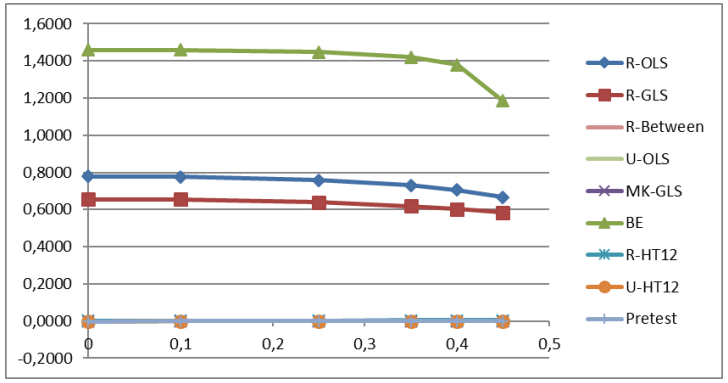

Figure 3b: Bias of $\widehat{\gamma}_2$ when changing $\rho_{X_1Z_2} = \rho_{X_2Z_2}$

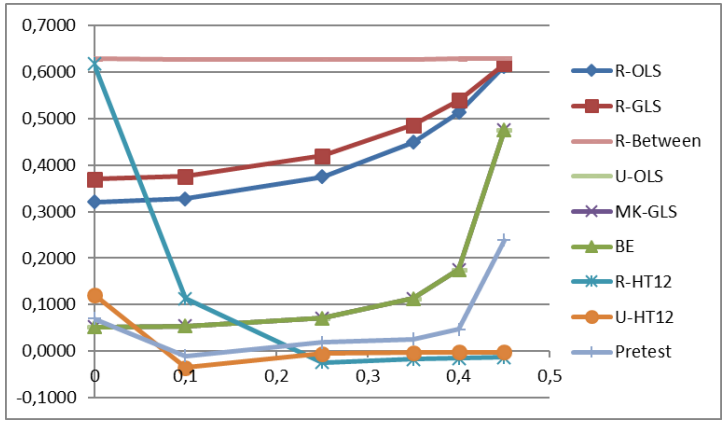

Table 3: Bias. RMSE and 5% size test for $\widehat{\beta}_3$ and $\widehat{\gamma}_2$ in a Hausman Taylor world, 1000 replications, N=100 and T=5: changing from weak to strong internal instruments

	R-OLS			R-OLS			R-GLS			R-GLS			R-Between			R-Between		
$\rho_{x_1z_2}$	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size
0	0.7791	0.091	100.0%	0.3207	0.090	98.3%	0.6561	0.105	100.0%	0.3694	0.097	99.3%	-0.0004	0.088	4.3%	0.6283	0.126	99.8%
0.1	0.7763	0.092	100.0%	0.3280	0.091	98.4%	0.6539	0.105	100.0%	0.3763	0.098	99.1%	-0.0004	0.088	4.3%	0.6275	0.126	99.8%
0.25	0.7582	0.092	100.0%	0.3749	0.096	99.3%	0.6398	0.104	100.0%	0.4200	0.101	99.4%	-0.0004	0.088	4.3%	0.6272	0.129	99.8%
0.35	0.7293	0.091	100.0%	0.4491	0.101	99.8%	0.6189	0.102	100.0%	0.4856	0.105	99.8%	-0.0004	0.088	4.3%	0.6277	0.132	99.7%
0.4	0.7044	0.091	100.0%	0.5131	0.104	99.8%	0.6025	0.100	100.0%	0.5387	0.107	99.9%	-0.0004	0.088	4.3%	0.6282	0.134	99.6%
0.45	0.6655	0.089	100.0%	0.6129	0.107	100.0%	0.5833	0.096	0.0%	0.6169	0.109	100.0%	-0.0004	0.088	4.3%	0.6290	0.081	99.6%
	U-OLS			U-OLS			MK-GLS			MK-GLS			Between			Between		
$\rho_{x_1z_2}$	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size
0	-0.0004	0.088	3.7%	0.0519	0.080	15.2%	-0.0004	0.088	4.4%	0.0519	0.080	9.7%	1.4598	0.103	100.0%	0.0519	0.080	9.7%
0.1	-0.0004	0.088	3.7%	0.0542	0.081	15.2%	-0.0004	0.088	4.4%	0.0542	0.081	9.7%	1.4583	0.104	100.0%	0.0542	0.081	9.8%
0.25	-0.0004	0.088	3.7%	0.0716	0.094	17.6%	-0.0004	0.088	4.4%	0.0716	0.094	11.5%	1.4470	0.108	100.0%	0.0716	0.094	11.6%
0.35	-0.0004	0.088	3.7%	0.1127	0.117	23.3%	-0.0004	0.088	4.4%	0.1127	0.117	16.2%	1.4201	0.117	100.0%	0.1127	0.117	16.3%
0.4	-0.0004	0.088	3.7%	0.1756	0.146	32.3%	-0.0004	0.088	4.4%	0.1756	0.146	23.2%	1.3791	0.130	100.0%	0.1756	0.146	23.4%
0.45	-0.0004	0.088	3.7%	0.4746	0.233	62.7%	-0.0004	0.088	4.4%	0.4746	0.233	53.3%	1.1842	0.176	100.0%	0.4746	0.233	53.3%
	R-HT12			R-HT12			U-HT12			U-HT12			Pretest			Pretest		
$\rho_{x_1z_2}$	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size	$\widehat{\beta}_3$ bias	RMSE	5% size	$\widehat{\gamma}_2$ bias	RMSE	5% size
0	0.0015	0.088	44.0%	0.6170	2.498	1.7%	-0.0004	0.088	4.2%	0.1207	1.949	0.1%	-0.0004	0.088	4.3%	0.0702	2.165	0.1%
0.1	0.0025	0.088	4.8%	0.1136	1.515	3.2%	-0.0003	0.088	4.2%	-0.0348	0.801	0.4%	-0.0003	0.088	4.4%	-0.0099	1.024	0.2%
0.25	0.0028	0.088	4.6%	-0.0239	0.364	3.8%	-0.0002	0.088	4.2%	-0.0047	0.218	2.8%	-0.0003	0.088	4.4%	0.0191	0.207	1.6%
0.35	0.0029	0.088	4.6%	-0.0168	0.251	4.2%	-0.0002	0.088	4.2%	-0.0028	0.152	3.5%	-0.0003	0.088	4.4%	0.0262	0.146	2.7%
0.4	0.0029	0.088	4.7%	-0.0146	0.217	4.1%	-0.0002	0.088	4.3%	-0.0023	0.133	3.8%	-0.0003	0.088	4.4%	0.0467	0.142	5.6%
0.45	0.0029	0.088	4.7%	-0.0130	0.192	4.6%	-0.0002	0.088	4.3%	-0.0018	0.118	3.8%	-0.0004	0.088	4.4%	0.2380	0.314	27.4%

Table 4: Percentage of times the pretest selects the internal instruments when changing the level of significance (in %) and the size of the sample (N)

	2%	2%	2%	5%	5%	5%	10%	10%	10%
N	100	500	1000	100	500	1000	100	500	1000
MK-GLS	1.1	11.4	**34.7**	3.4	23.1	**52.9**	6.9	**33.3**	**68.8**
U-HT1	5.2	16	19.5	8.2	17.9	15.2	12.8	19.2	11.1
U-HT2	4.9	17.8	20.4	8.6	20.6	17.6	13.8	21.7	12.5
U-HT12	**88.8**	**54.8**	25.4	**79.8**	**38.4**	14.3	**66.5**	25.8	7.6