

Chapter 7

Large-N and Large-T Properties of Panel Data Estimators and the Hausman Test

Seung Chan Ahn and Hyungsik Roger Moon

7.1 Introduction

Error-components models have been widely used to control for unobservable cross-sectional heterogeneity in panel data with a large number of cross-section units (N) and a small number of time-series observations (T). These models assume that stochastic error terms have two components: an unobservable time-invariant individual effect, which captures the unobservable individual heterogeneity, and the usual random noise. The most popular estimation methods for error-components models are the within and the generalized least squares (GLS) estimators. A merit of the within estimator (least squares on data transformed into deviations from individual means) is that it is consistent even if regressors are correlated with the individual effect (fixed effects). A drawback, however, is that it cannot estimate the coefficients of time-invariant regressors.¹ Among various alternative estimators for

¹Estimation of the effect of a certain time-invariant variable on a dependent variable could be an important task in a broad range of empirical research. Examples would be the labor studies about the effects of schooling or gender on individual workers' earnings, and the macroeconomic studies about the effect of a country's geographic location (e.g., whether the country is located in Europe or Asia) on its economic growth. The within estimator is inappropriate for such studies.

S.C. Ahn (✉)

Department of Economics, Arizona State University, Tempe, AZ 85287, USA

School of Economics, Sogang University, 35 Baekbeom-ro, Mapo-gu, Seoul, South Korea

e-mail: miniahn@asu.edu; Seung.Ahn@asu.edu

H.R. Moon

Department of Economics, KAP 300, University of Southern California,

Los Angeles, CA 91208, USA

Department of Economics, Yonsei University, 50 Yonsei-Ro, Seodeamun-Gu,

Seoul 120-749, Korea

e-mail: moonr@usc.edu

the coefficients of time-invariant regressors,² the GLS estimator has been popularly used in the literature due to its efficiency. Its consistency, however, requires a strong assumption that no regressor is correlated with the individual effects (random effects). Because of its restrictiveness, the empirical validity of the random effects assumption should be tested to justify the use of the GLS estimator. The Hausman test statistic (1978) has been popularly used for this purpose (e.g., Hausman and Taylor 1981; Cornwell and Rupert 1988; Baltagi and Khanti-Akom 1990; Ahn and Low 1996; or Guggenberger 2010).

This paper studies the asymptotic and finite-sample properties of the within and GLS estimators and the Hausman statistic for a general panel data error-components model with both large N and T . The GLS estimator has been known to be asymptotically equivalent to the within estimator for the cases with infinite N and T (see, for example, Hsiao 1986, Chap. 3; Mátyás and Sevestre 1992, Chap. 4; Baltagi 1995, Chap. 2). This asymptotic equivalence result has been obtained using a sequential limit method ($T \rightarrow \infty$ followed by $N \rightarrow \infty$) and some strong assumptions such as fixed regressors. This result naturally raises several questions. First, does the equivalence result hold for more general cases? Second, does the equivalence result indicate that the Hausman statistic, which is essentially a distance measure between the within and GLS estimators, should have a degenerating or nonstandard asymptotic distribution under the random effects assumption? Third, does the equivalence result also imply that the Hausman test would have low power to detect any violation of the random effects assumption when T is large? This paper is concerned with answering these questions.

Panel data with a large number of time-series observations have been increasingly more available in recent years in many economic fields such as international finance, finance, industrial organization, and economic growth. Furthermore, popular panel data, such as the Panel Study of Income Dynamics (PSID) and the National Longitudinal Surveys (NLS), contain increasingly more time-series observations as they are updated regularly over the years. Consistent with this trend, some recent studies have examined the large- N and large- T properties of the within and GLS estimators for error-component models.³ For example, Phillips and Moon (1999) and Kao (1999) establish the asymptotic normality of the within estimator for

²For example, if only the time-varying regressors are correlated with the individual effects, all of the coefficients of time-varying and time-invariant regressors can be consistently estimated by a two-step estimation procedure. At the first step, the coefficients of time-varying regressors can be consistently estimated by the within estimator. At the second step, the residuals computed with the within estimator are regressed on the time-invariant regressors. The resulting coefficient estimators are consistent as long as the time-invariant regressors are uncorrelated with the individual effects. We thank an anonymous referee for introducing this estimation procedure to us.

³Some other studies have considered different panel data models with large N and large T . For example, Levin and Lin (1992, 1993), Quah (1994), Im et al. (2003), and Higgins and Zakrajsek (1999) develop unit-root tests for data with large N and large T . Alvarez and Arellano (2003) and Hahn and Kuersteiner (2002) examine the large- N and large- T properties of generalized method of moments (GMM) and within estimators for stationary dynamic panel data models.

the cases in which regressors follow unit root processes. Extending these studies, [Choi \(1998\)](#) considers a general random effects model which contains both unit-root and covariance-stationary regressors. For this model, he derives the asymptotic distributions of both the within and GLS estimators.

This paper is different from the previous studies in three respects. First, the model we consider contains both time-varying and time-invariant regressors. The time-varying regressors are cross-sectionally heterogeneous or homogeneous. Analyzing this model, we study how cross-sectional heterogeneity and the covariance structure between the time-varying and time-invariant regressors, as well as time trends in regressors, would affect the convergence rates of the panel data estimators. Second, we examine how the large- N and large- T asymptotic equivalence of the within and GLS estimators influences the asymptotic and finite-sample performances of the Hausman test. [Ahn and Low \(1996\)](#) have investigated the size and power properties of the Hausman test for the cases with large N and small T . In this paper, we reexamine the asymptotic and finite-sample properties of the test in more detail. In particular, we study how the power of the Hausman test would depend on the size of T and the covariance structure among regressors. Third, and perhaps less importantly, we use the joint limit approach developed by [Phillips and Moon \(1999\)](#).

The main findings of this paper are as follows. First, consistent with the previous studies, we find that the within and GLS estimators of the coefficients of the time-varying regressors are asymptotically equivalent under quite general conditions. However, the convergence rates of the two estimators depend on (i) whether means of time-varying regressors are cross-sectionally heterogeneous or homogeneous and (ii) how the time-varying and time-invariant regressors are correlated. Second, if T is large, the GLS estimators of the coefficients of the time-varying regressors are consistent even if the random effect assumption is violated. This finding implies that the choice between within and GLS is irrelevant for the studies focusing on the effects of time-varying regressors. The choice matters for the studies focusing on the effects of time-invariant regressors. Third, despite the equivalence between the GLS and within estimators, the Hausman statistic has well-defined asymptotic distributions under the random effects assumption and under its local alternatives. We also find that the power of the Hausman test crucially depends on the covariance structure between time-varying and time-invariant regressors, the covariance structure between regressors and the individual effects, and the size of T . The Hausman test has good power to detect non-zero correlation between the individual effects and the permanent (individual-specific and time-invariant) components of time-varying regressors, even if T is small. In contrast, the power of the test is somewhat limited when the effects are correlated with the time-invariant regressors and/or they are only correlated with the transitory (time-varying) components of time-varying regressors. For such cases, the size of T could rather decrease the power of the Hausman test.

This paper is organized as follows. Section [7.2](#) introduces the panel model of interest, and defines the within, between and GLS estimators as well as the Hausman test. For several simple illustrative models, we derive the asymptotic distributions of the panel data estimators and the Hausman test statistic. Section [7.3](#)

reports the results from our Monte Carlo experiments. In Sect. 7.4, we provide our general asymptotic results. Concluding remarks follow in Sect. 7.5. All the technical derivations and proofs are presented in the Appendix and the previous version of this paper.

7.2 Preliminaries

7.2.1 Estimation and Specification Test

The model under discussion here is given:

$$y_{it} = \beta' x_{it} + \gamma' z_i + \zeta + \varepsilon_{it} = \delta' w_{it} + \zeta + \varepsilon_{it}; \varepsilon_{it} = u_i + v_{it}, \quad (7.1)$$

where $i = 1, \dots, N$ denotes cross-sectional (individual) observations, $t = 1, \dots, T$ denotes time, $w_{it} = (x'_{it}, z'_i)'$, and $\delta = (\beta', \gamma)'$. In model (7.1), x_{it} is a $k \times 1$ vector of time-varying regressors, z_i is a $g \times 1$ vector of time-invariant regressors, ζ is an overall intercept term, and the error ε_{it} contains a time-invariant individual effect u_i and random noise v_{it} . We consider the cases with both large numbers of individual and time series observations, so asymptotic properties of the estimators and statistics for model (7.1) apply as $N, T \rightarrow \infty$. The orders of convergence rates of some estimators can depend on whether or not the model contains an overall intercept term. This problem will be addressed later.

We assume that data are distributed independently (but not necessarily identically) across different i , and that the v_{it} are independently and identically distributed (i.i.d.) with $\text{var}(v_{it}) = \sigma_v^2$. We further assume that $u_i, x_{i1}, \dots, x_{iT}$ and z_i are strictly exogenous with respect to v_{it} ; that is, $E(v_{it} \mid u_i, x_{i1}, \dots, x_{iT}) = 0$, for any i and t . This assumption rules out the cases in which the set of regressors includes lagged dependent variables or predetermined regressors. Detailed assumptions about the regressors $x_{i1}, \dots, x_{iT}, z_i$ will be introduced later.

For convenience, we adopt the following notational rule: For any $p \times 1$ vector a_{it} , we denote $\bar{a}_i = \frac{1}{T} \sum_t a_{it}$; $\tilde{a}_{it} = a_{it} - \bar{a}_i$; $\bar{a} = \frac{1}{N} \sum_i \bar{a}_i$; $\tilde{a}_i = \bar{a}_i - \bar{a}$. Thus, for example, for $w_{it} = (x'_{it}, z'_i)'$, we have $\bar{w}_i = (\bar{x}'_i, z'_i)'$; $\tilde{w}_{it} = (\tilde{x}'_{it}, 0_{1 \times g})'$; $\bar{w} = (\bar{x}', \bar{z})$; $\tilde{w}_i = ((\bar{x}_i - \bar{x})', (z_i - \bar{z})')'$.

When the regressors are correlated with the individual effect, the OLS estimator of δ is biased and inconsistent. This problem has been traditionally addressed by the use of the within estimator (OLS on data transformed into deviations from individual means):

$$\hat{\beta}_w = (\sum_{i,t} \tilde{x}_{it} \tilde{x}'_{it})^{-1} \sum_{i,t} \tilde{x}_{it} \tilde{y}'_{it}.$$

Under our assumptions, the variance-covariance matrix of the within estimator is given:

$$\text{Var}(\hat{\beta}_w) = \sigma_v^2 (\sum_{i,t} \tilde{x}_{it} \tilde{x}'_{it})^{-1}. \quad (7.2)$$

Although the within method provides a consistent estimate of β , a serious defect is its inability to identify γ , the impact of time-invariant regressors. A popular treatment of this problem is the random effects (RE) assumption under which the u_i are random and uncorrelated with the regressors:

$$E(u_i \mid x_{i1}, \dots, x_{iT}, z_i) = 0. \quad (7.3)$$

Under this assumption, all of the parameters in model (7.1) can be consistently estimated. For example, a simple but consistent estimator is the between estimator (OLS on data transformed into individual means):

$$\hat{\delta}_b = (\hat{\beta}'_b, \hat{\gamma}'_b)' = (\sum_i \tilde{w}_i \tilde{w}'_i)^{-1} \sum_i \tilde{w}_i \tilde{y}_i.$$

However, as [Balestra and Nerlove \(1966\)](#) suggest, under the RE assumption, an efficient estimator is the GLS estimator of the following form:

$$\hat{\delta}_g = [\sum_{i,t} \tilde{w}_{it} \tilde{w}'_{it} + T\theta_T^2 \sum_i \tilde{w}_i \tilde{w}'_i]^{-1} [\sum_{i,t} \tilde{w}_{it} \tilde{y}_i + T\theta_T^2 \sum_i \tilde{w}_i \tilde{y}_i],$$

where $\theta_T = \sqrt{\sigma_v^2 / (T\sigma_u^2 + \sigma_v^2)}$. The variance-covariance matrix of this estimator is given:

$$Var(\hat{\delta}_g) = \sigma_v^2 [\sum_{i,t} \tilde{w}_{it} \tilde{w}'_{it} + T\theta_T^2 \sum_i \tilde{w}_i \tilde{w}'_i]^{-1}. \quad (7.4)$$

For notational convenience, we assume that σ_u^2 and σ_v^2 are known, while in practice they must be estimated.⁴

An important advantage of the GLS estimator over the within estimator is that it allows researchers to estimate γ . In addition, the GLS estimator of β is more efficient than the within estimator of β because $[Var(\hat{\beta}_w) - Var(\hat{\beta}_g)]$ is positive definite so long as $\theta_T > 0$. Despite these desirable properties, it is important to notice that the consistency of the GLS estimator crucially depends on the RE assumption (7.3). Accordingly, the legitimacy of the RE assumption should be tested to justify the use of GLS. In the literature, a Hausman test ([1978](#)) has been widely used for this purpose. The statistic used for this test is a distance measure between the within and GLS estimators of β :

$$\mathcal{H}_{\mathcal{M}_{NT}} \equiv (\hat{\beta}_w - \hat{\beta}_g)' [Var(\hat{\beta}_w) - Var(\hat{\beta}_g)]^{-1} (\hat{\beta}_w - \hat{\beta}_g). \quad (7.5)$$

⁴There are many different ways to consistently estimate σ_v^2 and σ_u^2 . One way is to use

$$\hat{\sigma}_v^2 = \sum_{i,t} (\tilde{y}_{it} - \tilde{x}_{it} \hat{\beta}_w)^2 / [N(T-1)]; \hat{\sigma}_u^2 = \sum_{i,t} (\tilde{y}_i - \tilde{w}_i \hat{\delta}_{ols})^2 / NT - \hat{\sigma}_v^2,$$

where $\hat{\delta}_{ols}$ is the OLS estimator of δ .

For the cases in which T is fixed and $N \rightarrow \infty$, the RE assumption warrants that the Hausman statistic $\mathcal{H}\mathcal{M}_{NT}$ is asymptotically χ^2 -distributed with degrees of freedom equal to k . This result is a direct outcome of the fact that for fixed T , the GLS estimator $\hat{\beta}_g$ is asymptotically more efficient than the within estimator $\hat{\beta}_w$, and that the difference between the two estimators is asymptotically normal; specifically, as $N \rightarrow \infty$,

$$\sqrt{NT}(\hat{\beta}_w - \hat{\beta}_g) \implies N(0, \text{plim}_{N \rightarrow \infty} NT[\text{Var}(\hat{\beta}_w) - \text{Var}(\hat{\beta}_g)]), \quad (7.6)$$

where “ \implies ” means “converges in distribution.”

An important condition that guarantees (7.6) is that $\theta_T > 0$. If $\theta_T = 0$, then $\hat{\beta}_w$ and $\hat{\beta}_g$ become identical and the Hausman statistic is not defined. Observing $\theta_T \rightarrow 0$ as $T \rightarrow \infty$, we can thus easily conjecture that $\hat{\beta}_w$ and $\hat{\beta}_g$ should be asymptotically equivalent as $T \rightarrow \infty$. This observation naturally raises two issues related to the asymptotic properties of the Hausman test as $T \rightarrow \infty$. First, since the Hausman statistic is asymptotically χ^2 -distributed for any fixed T under the RE assumption, we can conjecture that it should remain asymptotically χ^2 -distributed even if $T \rightarrow \infty$. Thus, we wish to understand the theoretical link between the asymptotic distribution of the Hausman statistic and the equivalence of the within and GLS estimators. Second, it is a well-known fact that the GLS estimator $\hat{\beta}_g$ is a weighted average of the within and between estimators $\hat{\beta}_w$ and $\hat{\beta}_b$ (Maddala 1971). Thus, observing the form of the Hausman statistic, we can conjecture that the Hausman test should have the power to detect any violation of the RE assumption that causes biases in $\hat{\beta}_b$. However, the weight given to $\hat{\beta}_b$ in $\hat{\beta}_g$ decreases with T . Thus, we wish to understand how the power of the Hausman test would be related to the size of T . We will address these two issues in the following sections.

What makes it complex to investigate the asymptotic properties of the Hausman statistic is that its convergence rate crucially depends on data generating processes. The following subsection considers several simple cases to illustrate this point.

7.2.2 Preliminary Results

This section considers several simple examples demonstrating that the convergence rate of the Hausman statistic depends on (i) whether or not the time-varying regressors are cross-sectionally heterogeneous and (ii) how the time-varying and time-invariant regressors are correlated.

For model (7.1), we can easily show that

$$\hat{\beta}_w - \beta = A_{NT}^{-1} a_{NT}; \quad (7.7)$$

$$\hat{\beta}_b - \beta = (B_{NT} - C_{NT} H_N^{-1} C'_{NT})^{-1} [b_{NT} - C_{NT} H_N^{-1} c_{NT}]; \quad (7.8)$$

$$\hat{\beta}_g - \beta = [A_{NT} + T\theta_T^2(B_{NT} - C_{NT}H_N^{-1}C'_{NT})]^{-1} \times [A_{NT}(\hat{\beta}_w - \beta) + T\theta_T^2(B_{NT} - C_{NT}H_N^{-1}C'_{NT})(\hat{\beta}_b - \beta)]; \quad (7.9)$$

$$\hat{\beta}_w - \hat{\beta}_g = [A_{NT} + T\theta_T^2(B_{NT} - C_{NT}H_N^{-1}C_{NT})]^{-1} \times T\theta_T^2(B_{NT} - C_{NT}H_N^{-1}C'_{NT})[(\hat{\beta}_w - \beta) - (\hat{\beta}_b - \beta)]; \quad (7.10)$$

$$\text{Var}(\hat{\beta}_w) - \text{Var}(\hat{\beta}_g) = A_{NT}^{-1} - [A_{NT} + T\theta_T^2(B_{NT} - C_{NT}H_N^{-1}C'_{NT})]^{-1}, \quad (7.11)$$

where,

$$A_{NT} = \sum_{i,t} \tilde{x}_{it}\tilde{x}'_{it}; B_{NT} = \sum_i \tilde{x}_i\tilde{x}'_i; C_{NT} = \sum_i \tilde{x}_i\tilde{z}'_i; H_N = \sum_i \tilde{z}_i\tilde{z}'_i; \\ a_{NT} = \sum_{i,t} \tilde{x}_{it}v_{it}; b_{NT} = \sum_i \tilde{x}_i(u_i + \bar{v}_i); c_{NT} = \sum_i \tilde{z}_i(u_i + \bar{v}_i).$$

Equation (7.10) provides some insight into the convergence rate of the Hausman test statistic. Note that $(\hat{\beta}_w - \hat{\beta}_g)$ depends on both $(\hat{\beta}_w - \beta)$ and $(\hat{\beta}_b - \beta)$. Apparently, the between estimator $\hat{\beta}_b$ exploits only N between-individual variations, while the within estimator $\hat{\beta}_w$ is computed based on $N(T - 1)$ within-individual variations. Accordingly, $(\hat{\beta}_b - \beta)$ converges to a zero vector in probability much slower than $(\hat{\beta}_w - \beta)$ does. Thus, we can conjecture that the convergence rate of $(\hat{\beta}_w - \hat{\beta}_g)$ will depend on that of $(\hat{\beta}_b - \beta)$, not $(\hat{\beta}_w - \beta)$. Indeed, we below justify this conjecture.

In this subsection, we only consider a simple model that has a single time-varying regressor (x_{it}) and a single time-invariant regressor (z_i). Accordingly, all of the terms defined in (7.7)–(7.11) are scalars. The within and GLS estimators and the Hausman test are well defined even if there is no time-invariant regressor. However, we consider the cases with both time-varying and time-invariant regressors because the correlation between the two regressors plays an important role in determining the convergence rate of the Hausman statistic. Asymptotic results for the cases with a single time-varying regressor only can be easily obtained by setting $C_{NT} = 0$.

We consider asymptotics under the RE assumption (7.3). The power property of the Hausman test will be discussed in the following subsection. To save space, we only consider the estimators of β and the Hausman test. The asymptotic distributions of the estimators of γ will be discussed in Sect. 7.4. Throughout the examples below, we assume that the z_i are *i.i.d.* over different i with $N(0, \sigma_z^2)$. In addition, we introduce a notation e_{it} to denote a white noise component in the time-varying regressor x_{it} . We assume that the e_{it} are *i.i.d.* over different i and t with $N(0, \sigma_e^2)$, and are uncorrelated with the z_i .

We consider two different cases separately: the cases in which x_{it} and z_i are uncorrelated (CASE A), and the cases in which the regressors are correlated (CASE B).

CASE A: We here consider a case in which the time-varying regressor x_{it} is stationary without trend. Specifically, we assume:

$$x_{it} = \Theta_{a,i} + \Psi_{a,t} + e_{it}, \quad (7.12)$$

where $\Theta_{a,i}$ and $\Psi_{a,t}$ are fixed individual-specific and time-specific effects, respectively. Define $\bar{\Theta}_a = \frac{1}{N} \sum_i \Theta_{a,i}$ and $\bar{\Psi}_a = \frac{1}{N} \sum_i \Psi_i$; and let

$$p_{a,1} = \lim_{N \rightarrow \infty} \bar{\Theta}_a; p_{a,2} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_i (\Theta_{a,i} - \bar{\Theta}_a)^2;$$

$$q_{a,1} = \lim_{N \rightarrow \infty} \bar{\Psi}_a; q_{a,2} = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_t (\Psi_{a,t} - \bar{\Psi}_a)^2.$$

We can allow them to be random without changing our results, but at the cost of analytical complexity. We consider two possible cases: one in which the parameters $\Theta_{a,i}$ are heterogeneous, and the other in which they are constant over different individuals. Allowing the $\Theta_{a,i}$ to be different across different individuals, we allow the means of x_{it} to be cross-sectionally heterogeneous. In contrast, if the $\Theta_{a,i}$ are constant over different i , the means of x_{it} become cross-sectionally homogeneous. As we show below, the convergence rates of the between estimator and Hausman test statistic are different in the two cases.

To be more specific, consider the three terms B_{NT} , C_{NT} , and b_{NT} defined below (7.11). Straightforward algebra reveals that

$$B_{NT} = \sum_i (\Theta_{a,i} - \bar{\Theta}_a)^2 + 2 \sum_i (\Theta_{a,i} - \bar{\Theta}_a)(\bar{e}_i - \bar{e}) + \sum_i (\bar{e}_i - \bar{e})^2;$$

$$C_{NT} = \sum_i (\Theta_{a,i} - \bar{\Theta}_a)(\bar{z}_i - \bar{z}) + \sum_i (\bar{e}_i - \bar{e})(\bar{z}_i - \bar{z});$$

$$b_{NT} = \sum_i (\Theta_{a,i} - \bar{\Theta}_a)u_i + \sum_i (\bar{e}_i - \bar{e})u_i$$

$$+ \sum_i (\Theta_{a,i} - \bar{\Theta}_a)\bar{v}_i + \sum_i (\bar{e}_i - \bar{e})\bar{v}_i.$$

It can be shown that the terms including $(\Theta_{a,i} - \bar{\Theta}_a)$ will be the dominant factors determining the asymptotic properties of B_{NT} , C_{NT} , and b_{NT} . However, if the parameters $\Theta_{a,i}$ are constant over different individuals so that $\Theta_{a,i} - \bar{\Theta}_a = 0$, none of B_{NT} , C_{NT} , and b_{NT} depend on $(\Theta_{a,i} - \bar{\Theta}_a)$. For this case, the asymptotic properties of the three terms depend on $(\bar{e}_i - \bar{e})$. This result indicates that the asymptotic distribution of the between estimator $\hat{\beta}_b$, which is a function of B_{NT} , C_{NT} , and b_{NT} , will depend on whether the parameters $\Theta_{a,i}$ are cross-sectionally heterogeneous or homogeneous.⁵

We now consider the asymptotic distributions of the within, between, GLS estimators and the Hausman statistic under the two alternative assumptions about the parameters $\Theta_{a,i}$.

⁵Somewhat interestingly, however, the distinction between these two cases becomes unimportant when the model has no intercept term ($\zeta = 0$) and is estimated with this restriction. For such a case, B_{NT} , C_{NT} and b_{NT} depend on \bar{x}_i instead of \tilde{x}_i . With \bar{x}_i , the terms $(\Theta_{a,i} - \bar{\Theta}_a)$ and $(\bar{e}_i - \bar{e})$ in B_{NT} , C_{NT} , and b_{NT} are replaced by $\Theta_{a,i}$ and \bar{e}_i , respectively. Thus, the terms containing the $\Theta_{a,i}$ remain as a dominating factor whether or not the $\Theta_{a,i}$ are heterogeneous.

CASE A.1: Assume that the $\Theta_{a,i}$ vary across different i ; that is, $p_{a,2} \neq 0$. With this assumption, we can easily show that as $(N, T \rightarrow \infty)$ ⁶:

$$\sqrt{NT}(\hat{\beta}_w - \beta) \implies N \left(0, \frac{\sigma_v^2}{(q_{a,2} + \sigma_e^2)} \right); \quad (7.13)$$

$$\sqrt{N}(\hat{\beta}_b - \beta) \implies N \left(0, \frac{\sigma_u^2}{p_{a,2}} \right); \quad (7.14)$$

$$\begin{aligned} \sqrt{NT}(\hat{\beta}_g - \beta) &= \sqrt{NT}(\hat{\beta}_w - \beta) \\ &+ \frac{1}{\sqrt{T}} \frac{\sigma_v^2}{\sigma_u^2} \frac{p_{a,2}}{(q_{a,2} + \sigma_e^2)} \sqrt{N}(\hat{\beta}_b - \beta) + o_p \left(\frac{1}{\sqrt{T}} \right); \end{aligned} \quad (7.15)$$

$$plim_{N,T \rightarrow \infty} NT^2 [Var(\hat{\beta}_w) - Var(\hat{\beta}_g)] = \frac{\sigma_v^4}{\sigma_u^2} \frac{p_{a,2}}{(q_{a,2} + \sigma_e^2)^2}. \quad (7.16)$$

Three remarks follow. First, consistent with previous studies, we find from (7.15) that the within and GLS estimators, $\hat{\beta}_w$ and $\hat{\beta}_g$, are \sqrt{NT} -equivalent in the sense that $(\hat{\beta}_w - \hat{\beta}_g)$ is $o_p(\sqrt{NT})$. This is so because the second term in the right-hand side of (7.15) is $O_p(1/\sqrt{T})$. At the same time, (7.15) also implies that $(\hat{\beta}_w - \hat{\beta}_g)$ is $O_p(\sqrt{NT}^2)$ and asymptotically normal. These results indicate that the within and GLS estimators are equivalent to each other by the order of \sqrt{NT} , but not by the order of \sqrt{NT}^2 . Second, from (7.15) and (7.16), we can see that the Hausman statistic is asymptotically χ^2 -distributed with the convergence rate equal to \sqrt{NT}^2 .⁷ In particular, (7.15) indicates that the asymptotic distribution of the Hausman statistic is closely related to the asymptotic distribution of the between estimator $\hat{\beta}_b$.

Finally, the above asymptotic results imply some simplified GLS and Hausman test procedures. From (7.13) and (7.14), it is clear that

$$\sqrt{N}(\hat{\beta}_b - \hat{\beta}_w) = \sqrt{N}(\hat{\beta}_b - \beta) - \frac{1}{\sqrt{T}} \sqrt{NT}(\hat{\beta}_w - \beta) = \sqrt{N}(\hat{\beta}_b - \beta) + o_p(1).$$

With (7.15), this result indicates that the Hausman test based on the difference between $\hat{\beta}_w$ and $\hat{\beta}_g$ is asymptotically equivalent to the Wald test based on the between estimator $\hat{\beta}_b$ for the hypothesis that the true β equals the within

⁶These result are obtained utilizing the fact that limits of $\frac{1}{NT}A_{NT}$, $\frac{1}{N}B_N$, $\frac{1}{N}C_N$, and $\frac{1}{N}H_N$ are finite, while $\frac{1}{\sqrt{NT}}a_{NT}$, $\frac{1}{\sqrt{N}}b_N$, and $\frac{1}{\sqrt{N}}c_{NT}$ are asymptotically normal. More detailed calculations can be found from an earlier version of this paper.

⁷As we can see clearly from (7.15) and (7.16), the Hausman statistic does not depend on \sqrt{NT}^2 because it cancels out. Nonetheless, we say that the convergence rate of the Hausman test statistic equals \sqrt{NT}^2 because the asymptotic χ^2 result for the statistic is obtained based on the fact that $(\hat{\beta}_w - \hat{\beta}_g)$ is \sqrt{NT}^2 -consistent.

estimator $\hat{\beta}_w$. We obtain this result because the convergence rate of $\hat{\beta}_w$ is faster than that of $\hat{\beta}_b$. In addition, the GLS estimator of $\hat{\gamma}_g$ can be obtained by the between regression treating $\hat{\beta}_w$ as the true β , that is, the regression on the model $\tilde{y}_i - \tilde{x}_i \hat{\beta}_w = \tilde{z}_i \gamma + error$. An advantage of these alternative procedures is that GLS and Hausman tests can be conducted without estimating θ_T . The alternative procedures would be particularly useful for the analysis of unbalance panel data. For such data, θ_T is different over different cross-sectional units. When T is sufficiently large for individual i 's, we do not need to estimate these different θ_T 's for GLS. The alternative procedures work out for all of the other cases analyzed below.

CASE A.2: Now, we consider the case in which the $\Theta_{a,i}$ are constant over different i (Θ_a); that is, $p_{a,2} = 0$. It can be shown that the asymptotic distributions of the within and GLS estimators are the same under both CASEs A.1 and A.2. However, the asymptotic distributions of the between estimator $\hat{\beta}_b$ and the Hausman statistic are different under CASEs A.1 and A.2.⁸ Specifically, for CASE A.2, we can show that as $(N, T \rightarrow \infty)$,

$$\sqrt{\frac{N}{T}}(\hat{\beta}_b - \beta) \implies N\left(0, \frac{\sigma_u^2}{\sigma_e^2}\right); \tag{7.17}$$

$$\begin{aligned} \sqrt{NT^3}(\hat{\beta}_w - \hat{\beta}_g) &= -\frac{\sigma_v^2}{\sigma_u^2} \frac{\sigma_e^2}{(q_{a,2} + \sigma_e^2)} \sqrt{\frac{N}{T}}(\hat{\beta}_b - \beta) + o_p(1) \\ &\implies N\left(0, \frac{\sigma_v^4}{\sigma_u^2} \frac{\sigma_e^2}{(q_{a,2} + \sigma_e^2)^2}\right); \end{aligned} \tag{7.18}$$

$$plim_{N,T \rightarrow \infty} NT^3 [Var(\hat{\beta}_w) - Var(\hat{\beta}_g)] = \frac{\sigma_v^4 \sigma_e^2}{\sigma_u^2 (q_{a,2} + \sigma_e^2)^2}, \tag{7.19}$$

Several comments follow. First, observe that differently from CASE A.1, the between estimator $\hat{\beta}_b$ is now $\sqrt{N/T}$ -consistent. An interesting result is obtained when $N/T \rightarrow c < \infty$. For this case, the between estimator is inconsistent although it is still asymptotically unbiased. This implies that the between estimator is an inconsistent estimator for the analysis of cross-sectionally homogeneous panel data unless N is substantially larger than T . Second, the convergence rate of $(\hat{\beta}_w - \hat{\beta}_g)$, as well as that of the Hausman statistic, is different between CASEs A.1 and A.2. Notice that the convergence rate of $(\hat{\beta}_w - \hat{\beta}_g)$ is $\sqrt{NT^3}$ for CASE A.2, while it is $\sqrt{NT^2}$ for CASE A.1. Thus, $(\hat{\beta}_w - \hat{\beta}_g)$ converges in probability to zero

⁸If the model contains no intercept term ($\zeta = 0$) and it is estimated with this restriction, all of the results (7.13)–(7.16) are still valid with Θ_a^2 replacing $p_{a,2}$. Thus, the convergence rates of the within, between, GLS estimators and the Hausman statistic are the same under both CASEs A.1 and A.2.

much faster in CASE A.2 than in CASE A.1. Nonetheless, the Hausman statistic is asymptotically χ^2 -distributed in both cases, though with different convergence rates.

Even if the time-varying regressor x_{it} contains a time trend, we can obtain the similar results as in CASES A.1 and A.2. For example, consider a case in which the time-varying regressor x_{it} contains a time trend of order m :

$$x_{it} = \Theta_{b,i}t^m + e_{it}, \quad (7.20)$$

where the parameters $\Theta_{b,i}$ are fixed.⁹ Not surprisingly, for this case, we can show that the within and GLS estimators are superconsistent and $T^m\sqrt{NT}$ -equivalent. However, the convergence rates of the between estimator $\hat{\beta}_b$ and the Hausman test statistic crucially depend on whether the parameters $\Theta_{b,i}$ are heterogeneous or homogeneous. When, the parameters $\Theta_{b,i}$ are heterogeneous over different i , the between estimator $\hat{\beta}_b$ is $T^m\sqrt{N}$ -consistent, while the convergence rate of the Hausman statistic equals $T^m\sqrt{NT^2}$. In contrast, somewhat surprisingly, when the parameters $\Theta_{b,i}$ are homogeneous over different i , the estimator $\hat{\beta}_b$ is no longer superconsistent. Instead, it is $\sqrt{N/T}$ -consistent as in CASE A.2. The convergence rate of the Hausman statistic changes to $T^{2m}\sqrt{NT^3}$.¹⁰ This example demonstrates that the convergence rates of the between estimator and the Hausman statistic crucially depend on whether means of time-varying regressors are cross-sectionally heterogeneous or not.

CASE B: So far, we have considered the cases in which the time-varying regressor x_{it} and the time-invariant regressor z_i are uncorrelated. We now examine the cases in which this assumption is relaxed. The degree of the correlation between the x_{it} and z_i may vary over time. As we demonstrate below, the asymptotic properties of the panel data estimators and the Hausman test statistic depend on how the correlation varies over time. The basic model we consider here is given by

$$x_{it} = \Pi_i z_i / t^m + e_{it}, \quad (7.21)$$

where the Π_i are individual-specific fixed parameters, and m is a non-negative real number.¹¹ Observe that because of the presence of the Π_i , the x_{it} are not i.i.d. over

⁹We can consider a more general case: for example, $x_{it} = a_i t^m + \Theta_i + \Theta_t + b_i z_i + e_{it}$. However, the same asymptotic results apply to this general model. This is so because the trend term (t^m) dominates asymptotically.

¹⁰Detailed asymptotic results can be found from an earlier version of this paper.

¹¹We can consider a more general model:

$$x_{it} = \Theta_{b,i} + \Psi_{b,t} + \Pi_i z_i / t^m + e_{it},$$

where the $\Theta_{b,i}$ and $\Psi_{b,t}$ are individual- and time-specific fixed parameters, respectively. When the parameters $\Theta_{b,i}$ are cross-sectional heterogeneous, the asymptotic results are essentially the same

different i .¹² The correlation between x_{it} and z_i decreases over time if $m > 0$. In contrast, $m = 0$ implies that the correlation remains constant over time. For CASE B, the within and GLS estimators are always \sqrt{NT} -consistent regardless of the size of m . Thus, we only report the asymptotic results for the between estimator $\hat{\beta}_b$ and the Hausman statistic.

We examine three possible cases: $m \in (0.5, \infty]$, $m = 0.5$, and $m \in [0, 0.5)$. We do so because, depending on the size of m , one (or both) of the two terms e_{it} and $\Pi_i z_i / t^m$ in x_{it} becomes a dominating factor in determining the convergence rates of the between estimator $\hat{\beta}_b$ and the Hausman statistic \mathcal{HM}_{TN} .

CASE B.1: Assume that $m \in (0.5, \infty]$. This is the case where the correlation between x_{it} and z_i fades away quickly over time. Thus, one could expect that the correlation between x_{it} and z_i (through the term $\Pi_i z_i / t^m$) would not play any important role in asymptotics. Indeed, straightforward algebra, which is not reported here, justifies this conjecture: The term e_{it} in x_{it} dominates $\Pi_i z_i / t^m$ in asymptotics, and, thus, this is essentially the same case as CASE A.2.¹³

CASE B.2: We now assume $m = 0.5$. For this case, define $\bar{\Pi} = \frac{1}{N} \sum_i \Pi_i$;

$$p_{b,1} = \lim_{N \rightarrow \infty} \bar{\Pi}; p_{b,2} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_i (\Pi_i - \bar{\Pi})^2,$$

and $q_b = \lim_{T \rightarrow \infty} \frac{1}{T^{1-m}} \int_0^1 r^{-m} dr = \frac{1}{1-m}$ for $m \leq 0.5$. With this notation, a little algebra shows that as $(N, T \rightarrow \infty)$,

$$\sqrt{\frac{N}{T}} (\hat{\beta}_b - \beta) \implies N \left(0, \frac{\sigma_u^2}{p_{b,2} q_b^2 \sigma_z^2 + \sigma_e^2} \right).$$

Observe that the asymptotic variance of the between estimator $\hat{\beta}_b$ depends on both the terms σ_e^2 and $p_{b,2} q_b^2 \sigma_z^2$. That is, both the terms e_{it} and $\Pi_i z_i / t^m$ in x_{it} are important in the asymptotics of the between estimator $\hat{\beta}_b$. This implies that the correlation between the x_{it} and z_i , when it decreases reasonably slowly over time, matters for the asymptotic distribution of the between estimator $\hat{\beta}_b$.

as those we obtain for CASE A.1. This is so because the terms $\Theta_{b,i}$ dominate and the terms $\Pi_i z_i / t^m$ become irrelevant in asymptotics. Thus, we set $\Theta_{b,i} = 0$ for all i . In addition, we set $\Psi_{b,t} = 0$ for all t , because presence of the time effects is irrelevant for convergence rates of panel data estimators and the Hausman statistic.

¹²We here assume that the Π_i are cross-sectionally heterogeneous. For the cases in which the Π_i are the same for all i , $\hat{\beta}_b$ does not depend on $\Pi_i z_i / t^m$, and we obtain exactly the same asymptotic results as those for CASE A.2. This is due to the fact that the individual mean of the time-varying regressor \bar{x}_i becomes a linear function of the time invariant regressor z_i if the Π_i are the same for all i .

¹³We can obtain this result using the fact that $\lim_{T \rightarrow \infty} \frac{1}{\sqrt{T}} \sum_t t^{-m} = 0$, if $m > 0.5$.

Nonetheless, the convergence rate of $\hat{\beta}_b$ is the same as that of $\hat{\beta}_b$ for CASES A.2 and B.1. We can also show

$$\begin{aligned}\sqrt{NT^3}(\hat{\beta}_w - \hat{\beta}_g) &= -\frac{\sigma_v^2 p_{b,2} q_b^2 \sigma_z^2 + \sigma_e^2}{\sigma_u^2} \sqrt{\frac{N}{T}} (\hat{\beta}_b - \beta) + o_p(1) \\ &\implies N \left(0, \frac{\sigma_v^4 p_{b,2} q_b^2 \sigma_z^2 + \sigma_e^2}{\sigma_u^4} \right); \\ \text{plim}_{N,T \rightarrow \infty} NT^3 [\text{Var}(\hat{\beta}_w) - \text{Var}(\hat{\beta}_g)] &= \frac{\sigma_v^4 p_{b,2} q_b^2 \sigma_z^2 + \sigma_e^2}{\sigma_u^4},\end{aligned}$$

both of which imply that the Hausman statistic is asymptotically χ^2 -distributed.

CASE B.3: Finally, we consider the case in which $m \in [0, 0.5)$, where the correlation between x_{it} and z_i decays over time slowly. Note that the correlation remains constant over time if $m = 0$. We can show

$$\sqrt{\frac{N}{T^{2m}}} (\hat{\beta}_b - \beta) \implies N \left(0, \frac{\sigma_u^2}{p_{b,2} q_b^2 \sigma_z^2} \right).$$

Observe that the asymptotic distribution of $\hat{\beta}_b$ no longer depends on σ_e^2 . This implies that the term $\Pi_i z_i / t^m$ in x_{it} dominates e_{it} in the asymptotics for $\hat{\beta}_b$. Furthermore, the convergence rate of $\hat{\beta}_b$ now depends on m . Specifically, so long as $m < 0.5$, the convergence rate increases as m decreases. In particular, when the correlation between x_{it} and z_i remains constant over time ($m = 0$), the between estimator $\hat{\beta}_b$ is \sqrt{N} -consistent as in CASE A.1. Finally, the following results indicate that the convergence rate of the Hausman statistic $\mathcal{H}\mathcal{M}_{NT}$ also depends on m :

$$\begin{aligned}\sqrt{NT^{2m+2}}(\hat{\beta}_w - \hat{\beta}_g) &= -\frac{\sigma_v^2 p_{b,2} q_b^2 \sigma_z^2}{\sigma_u^2} \sqrt{\frac{N}{T^{2m}}} (\hat{\beta}_b - \beta) + o_p(1) \\ &\implies N \left(0, \frac{\sigma_v^4 p_{b,2} q_b^2 \sigma_z^2}{\sigma_u^4} \right); \\ \text{plim}_{N,T \rightarrow \infty} NT^{2m+2} [\text{Var}(\hat{\beta}_w) - \text{Var}(\hat{\beta}_g)] &= \frac{\sigma_v^4 p_{b,2} q_b^2 \sigma_z^2}{\sigma_u^4}.\end{aligned}$$

7.2.3 Asymptotic Power Properties of the Hausman Test

In this section, we consider the asymptotic power properties of the Hausman test for the special cases discussed in the previous subsection. To do so, we need to

Table 7.1 Local alternatives and asymptotic results

Case	Local alternatives	Convergence rate	Noncentral parameter
A.1	$E(u_i \tilde{x}_i, \tilde{z}_i) = \tilde{x}_i \frac{\lambda_x}{\sqrt{N}} + \tilde{z}_i \frac{\lambda_z}{\sqrt{N}}$ ^a	$\sqrt{NT^2}$	$\frac{p_{a,2}\lambda_x^2}{\sigma_u^2}$
A.2	$E(u_i \tilde{x}_i, \tilde{z}_i) = \sqrt{T}\tilde{x}_i \frac{\lambda_x}{\sqrt{N}} + \tilde{z}_i \frac{\lambda_z}{\sqrt{N}}$	$\sqrt{NT^3}$	$\frac{\sigma_e^2\lambda_x^2}{\sigma_u^2}$
B.1	$E(u_i \tilde{x}_i, \tilde{z}_i) = \sqrt{T}\tilde{x}_i \frac{\lambda_x}{\sqrt{N}} + \tilde{z}_i \frac{\lambda_z}{\sqrt{N}}$	$\sqrt{NT^3}$	$\frac{\sigma_e^2\lambda_x^2}{\sigma_u^2}$
B.2	$E(u_i \tilde{x}_i, \tilde{z}_i) = \sqrt{T}\tilde{x}_i \frac{\lambda_x}{\sqrt{N}} + \tilde{z}_i \frac{\lambda_z}{\sqrt{N}}$	$\sqrt{NT^3}$	$(4pb_2\sigma_z^2 + \sigma_e^2)\lambda_x^2$
B.3	$E(u_i \tilde{x}_i, \tilde{z}_i) = T^m\tilde{x}_i \frac{\lambda_x}{\sqrt{N}} + \tilde{z}_i \frac{\lambda_z}{\sqrt{N}}$	$\sqrt{NT^{2m+2}}$	$\frac{\sigma_u^2}{\left(\frac{1-m}{2}\right)^2 pb_2\sigma_z^2}\lambda_x^2$

^aThis sequence of local alternatives can be replaced by

$$E(u_i|\bar{x}_i, \bar{z}_i) = \lambda_o + \bar{x}_i \frac{\lambda_x}{\sqrt{N}} + \bar{z}_i \frac{\lambda_z}{\sqrt{N}},$$

where λ_o is any constant scalar. The asymptotic results remain the same with this replacement.

specify a sequence of local alternative hypotheses for each case. Among many, we consider the alternative hypotheses under which the conditional mean of u_i is a linear function of the regressors \tilde{x}_i and \tilde{z}_i .

We list in Table 7.1 our local alternatives and asymptotic results for CASES A.1–B.3. For all of the cases, we assume that $var(u_i|\tilde{x}_i, \tilde{z}_i) = \sigma_u^2$ for all i . The parameters λ_x and λ_z are nonzero real numbers. Notice that for CASES A.2–B.3, we use $\sqrt{T}\tilde{x}_i$ or $T^m\tilde{x}_i$ instead of \tilde{x}_i . We do so because, for those cases, $plim_{T \rightarrow \infty} \tilde{x}_i = 0$. The third column indicates the convergence rates of the Hausman test, which are the same as those obtained under the RE assumption. Under the local alternatives, the Hausman statistic asymptotically follows a noncentral χ^2 distribution. The noncentral parameters for individual cases are listed in the fourth column of Table 7.1.

A couple of comments follow. First, although the noncentral parameter does not depend on λ_z for any case reported in Table 7.1, it does not mean that the Hausman test has no power to detect nonzero correlation between the effect u_i and the time-invariant regressor z_i . The Hausman test comparing the GLS and within estimators is not designed to directly detect the correlations between the time-invariant regressors and the individual effects. Nonetheless, the test has power as long as the individual effect u_i is correlated with the time-varying regressors conditionally on the time-invariant regressors. To see this, consider a model in which x_{it} and z_i have a common factor f_i ; that is, $x_{it} = f_i + e_{it}$ and $z_i = f_i + \eta_i$. (This is the case discussed below in Assumption 5.) Assume $E(u_i | f_i, \eta_i, \bar{e}_i) = c\eta_i/\sqrt{N}$. Also assume that f_i , η_i and e_{it} are normal, mutually independent, and i.i.d. over different i and t with zero means, and variances σ_f^2 , σ_η^2 , and σ_e^2 , respectively. Note that x_{it} is not correlated with u_i , while z_i is. For this case, however, we can show that

$$E(u_i|\bar{x}_i, \bar{z}_i) = \bar{x}_i \frac{\lambda_x}{\sqrt{N}} + \bar{z}_i \frac{\lambda_z}{\sqrt{N}},$$

where $d = (\sigma_f^2 + \sigma_e^2/T)(\sigma_f^2 + \sigma_\eta^2) - \sigma_f^4$, $\lambda_x = -c\sigma_f^2\sigma_\eta^2/d$, $\lambda_z = c(\sigma_f^2 + \sigma_e^2/T)\sigma_\eta^2/d$. Observe that λ_x is functionally related to λ_z : $\lambda_x = 0$ if and only if $\lambda_z = 0$. For this case, it can be shown that the Hausman test has the power to detect non-zero correlation between u_i and \tilde{z}_i .

Second and more importantly, the results for CASE A show that the large- T and large- N power properties of the Hausman test may depend on (i) what components of x_{it} are correlated with the effect u_i and (ii) whether the mean of x_{it} is cross-sectionally heterogeneous. For CASE A, the time-invariant and time-varying parts of x_{it} , $\Theta_{a,i}$ and e_{it} , can be viewed as permanent and transitory components, respectively. For fixed T , it does not matter to the Hausman test which of these two components of x_{it} is correlated with the individual effect u_i . The Hausman test has power to detect any kind of correlations between x_{it} and u_i . In contrast, for the cases with large T , the same test can have power for specific correlations only. To see why, observe that for CASE A.1, the noncentral parameter of the Hausman statistic depends only on the variations of the permanent components Θ_i , not on those of the transitory components e_{it} . This implies that for CASE A.1 (where the permanent components $\Theta_{a,i}$ are cross-sectionally heterogeneous), the Hausman test has power for nonzero correlation between the effect u_i and the permanent component $\Theta_{a,i}$, but no power for nonzero-correlation between the effect and the temporal component e_{it} . In contrast, for CASE A.2 (where the permanent components $\Theta_{a,i}$ are the same for all i), the noncentral parameter depends on the variations in e_{it} . That is, for CASE A.2, the Hausman test does have power to detect nonzero-correlation between the effect and the temporal component of x_{it} .¹⁴

Similar results are obtained from the analysis of CASE B. The results reported in the fourth column of Table 7.1 show that when the correlation between x_{it} and z_i decays slowly over time ($m \leq 0.5$), the Hausman test has power to detect nonzero-correlation between the individual effect and the transitory component of time-varying regressors, even if T is large. In contrast, when $m > 0.5$, the same test has no power to detect such nonzero correlations if T is large. These results indicate that the asymptotic power of the Hausman test can depend on the size of T . That is, the Hausman test results based on the entire data set with large T could be different from those based on subsamples with small T . These findings will be more elaborated in Sect. 7.3.

¹⁴This point can be better presented if we choose the following sequence of local alternative hypotheses for CASE A:

$$E(u_i | \tilde{x}_i, \tilde{z}_i) = E(\tilde{x}_i) \frac{\lambda_{x,1}}{\sqrt{N}} + (\tilde{x}_i - E(\tilde{x}_i)) \frac{\lambda_{x,2}}{\sqrt{N}} + \tilde{z}_i \frac{\lambda_z}{\sqrt{N}}.$$

Observe that for CASE A, $E(\tilde{x}_i) = \Theta_i$, and $(\tilde{x}_i - E(\tilde{x}_i)) = \tilde{e}_i$. Under the local alternative hypotheses, it can be shown that the noncentral parameter of the Hausman test depends on either $\lambda_{x,1}$ or $\lambda_{x,2}$, but not both. When $E(\tilde{x}_i) \neq 0$ (CASE A.1), the noncentral parameter of the test depends only on $\lambda_{x,1}$. In contrast, when $E(\tilde{x}_i) = 0$, the noncentral parameter depends only on $\lambda_{x,2}$.

7.3 Monte Carlo Experiments

In this section, we investigate the finite-sample properties of the within and GLS estimators, as well as those of the Hausman test. Consistent with the previous section, we consider a model with one time-varying regressor x_{it} and one time-invariant regressor z_i . The foundations of our experiments are CASES A and B discussed above. For all of our simulations, both the individual effects u_i and random errors v_{it} are drawn from $N(0, 1)$.

For the cases in which the two regressors are uncorrelated (CASE A), we generate x_{it} and z_i as follows:

$$x_{it} = \Theta_i + \xi_i \phi_t + e_{it}, \quad (7.22)$$

$$z_i = \rho_{zu} u_i + \sqrt{1 - \rho_{zu}^2} f_i, \quad (7.23)$$

where $\Theta_i = \rho_{xu,1} u_i + \sqrt{1 - \rho_{xu,1}^2} \Theta_i^c$, $\xi_i = \rho_{xu,2} u_i + \sqrt{1 - \rho_{xu,2}^2} \xi_i^c$, the Θ_i^c and ξ_i^c are random variables from $N(0, 1)$, the e_{it} and f_i are drawn from a uniform distribution in the range $(-2, 2)$. The term $\xi_i \phi_t + e_{it}$ is the transitory component of x_{it} . The term $\xi_i \phi_t$ is introduced to investigate the cases with non-zero correlations between the individual effect and the transitory component of x_{it} . The degrees of correlations between the individual effects and regressors are controlled by $\rho_{xu,1}$, $\rho_{xu,2}$ and ρ_{zu} . For each of the simulation results reported below, 5,000 random samples are used.

Table 7.2 reports the simulation results from CASE A.1. When regressors are uncorrelated with the effect u_i , both the GLS and within estimators have only small biases. The Hausman test is reasonably well sized although it is somewhat oversized when both N and T are small. When the permanent component of x_{it} is correlated with the effect (Panel I), the GLS estimator of β is biased. However, the size of bias decreases with T , as we expected. The bias in the within estimator of β is always small regardless of the sizes of N and T . The Hausman test has great power to detect non-zero correlation between the permanent component of x_{it} and u_i regardless of sample size. The power increases with T while the bias in the GLS estimator of β decreases.

Panel II of Table 7.2 shows the results from the cases in which the effect u_{it} and the transitory component of x_{it} are correlated. Our asymptotic results predict that this type of correlation does not bias the GLS estimates when T is large and is not detected by the Hausman test. The results reported in Panel II are consistent with this prediction. Even if T is small, we do not see substantial biases in the GLS estimates. When T is small, the Hausman test has some limited power to detect the non-zero correlation between the effect and the transitory component of the time-varying regressor. However, the power decreases with T .

Table 7.3 reports the results from the cases in which the time-varying regressor does not have a permanent component (CASE A.2). Similar to those that are

Table 7.2 Monte Carlo simulation results for CASE A.1

N	T	Panel I: $\rho_{xu,2} = 0$					Hausman rejection rate	Panel II: $\rho_{xu,1} = 0$				
		$\rho_{xu,1}$	Bias		β	γ		Bias		β	γ	Hausman rejection rate
			GLS	Within				GLS	Within			
50	5	0.0	0.001	0.000	0.001	0.063	0.0	0.001	0.000	0.001	0.063	
		0.0	0.000	0.001	0.000	0.054	0.0	0.000	0.001	0.000	0.054	
	100	0.0	0.000	-0.001	0.000	0.055	0.0	0.000	-0.001	0.000	0.055	
100	5	0.0	0.000	0.000	-0.001	0.049	0.0	0.000	0.000	-0.001	0.049	
		0.0	0.000	-0.002	0.000	0.053	0.0	0.000	-0.002	0.000	0.053	
	100	0.0	0.000	-0.002	0.000	0.053	0.0	0.000	-0.002	0.000	0.053	
50	5	0.5	0.046	0.000	0.001	0.753	0.5	0.001	0.000	0.000	0.181	
		0.5	0.005	0.001	0.000	0.960	0.5	0.000	0.001	0.000	0.081	
	100	0.5	0.003	-0.001	0.000	0.967	0.5	0.000	-0.001	0.000	0.069	
100	5	0.5	0.045	0.000	-0.001	0.956	0.5	0.000	0.000	-0.001	0.286	
		0.5	0.005	-0.002	0.000	1.000	0.5	0.000	-0.002	0.000	0.111	
	100	0.5	0.003	-0.002	0.000	1.000	0.5	0.000	-0.002	0.000	0.080	
50	5	1.0	0.110	0.000	0.001	0.999	1.0	0.000	0.000	0.000	0.433	
		1.0	0.013	0.001	0.000	1.000	1.0	0.000	0.001	0.000	0.163	
	100	1.0	0.006	-0.001	0.000	1.000	1.0	0.000	-0.001	0.000	0.112	
100	5	1.0	0.109	0.000	-0.001	1.000	1.0	0.000	0.000	-0.001	0.577	
		1.0	0.012	-0.002	0.000	1.000	1.0	0.000	-0.002	0.000	0.249	
	100	1.0	0.006	-0.002	0.000	1.000	1.0	0.000	-0.002	0.000	0.164	

reported in Panel II of Table 7.2, there is no sign that non-zero correlation between the effect and the transitory component of x_{it} causes a substantial bias in the GLS estimator. However, differently from the results reported in Panel II of Table 7.2, the Hausman test now has better power to detect non-zero correlation between the effect and the transitory component of x_{it} . The power increases as either N or T increases.

We now consider the cases in which x_{it} and z_i are correlated (CASE B). For these cases, the x_{it} are generated by

$$x_{it} = \pi_i f_i / t^m + \xi_i \phi_t + e_{it}, \tag{7.24}$$

where the π_i are drawn from a uniform distribution in the range (0, 1). As in (7.22), the term $\xi_i \phi_t$ is introduced to investigate the cases with non-zero correlations between the individual effect and the transitory component of x_{it} . Observe that in (7.24), we use f_i , not z_i . As we have discussed in the previous section, the Hausman test would not have any power to detect non-zero correlation between z_i and u_i for the cases with large T if z_i instead of f_i were used for (7.24). We use f_i to investigate the power properties of the Hausman test under more general cases than the cases we have considered in the previous sections. Under (7.24), the Hausman test can have power to detect non-zero correlation between z_i and u_i .

Table 7.3 Monte Carlo simulation results for CASE A.2

N	T	$\rho_{xu,2}$	Bias		Within	Hausman rejection rate
			GLS			
			β	γ	β	
50	5	0.0	0.000	0.001	-0.001	0.052
	50	0.0	0.000	0.001	0.000	0.051
	100	0.0	0.000	0.001	0.000	0.052
100	5	0.0	0.000	0.002	0.000	0.050
	50	0.0	0.000	-0.001	0.000	0.049
	100	0.0	0.000	-0.002	0.000	0.055
50	5	0.5	-0.001	0.001	-0.001	0.396
	50	0.5	0.000	0.001	0.000	0.448
	100	0.5	0.000	0.001	0.000	0.461
100	5	0.5	0.000	0.002	0.000	0.569
	50	0.5	0.000	-0.001	0.000	0.626
	100	0.5	0.000	-0.002	0.000	0.620
50	5	1.0	-0.002	0.001	-0.002	0.999
	50	1.0	0.000	0.001	0.000	1.000
	100	1.0	0.000	0.001	0.000	1.000
100	5	1.0	0.000	0.002	0.000	1.000
	50	1.0	0.000	-0.001	0.000	1.000
	100	1.0	0.000	-0.002	0.000	1.000

Table 7.4 shows our simulation results from the cases in which the time-varying regressor is correlated with the time-invariant regressor. Panel I reports the results when the time-invariant regressor z_i is correlated with the effect u_i . Regardless of how fast the correlation between x_{it} and z_i decays over time, the GLS estimator of γ shows some signs of biases. While the biases reported in Panel I appear only mild, the biases become substantial if we increase the size of ρ_{zu} further. When the correlation between x_{it} and z_i remains constant over time ($m = 0$), the GLS estimator of β is mildly biased. However, the bias becomes smaller as the size of T increases. When the correlation between x_{it} and z_i decays over time ($m \geq 0.5$), no substantial bias is detected in the GLS estimator of β even if T is small. The Hausman test has some power to detect non-zero correlation between the time-invariant regressor z_i and the effect u_i . However, the power appears to be limited in our simulation exercises: Its power never exceeds 61%. The power increases with T when $m \leq 0.5$, but the power decreases with T if $m > 0.5$.

Panel II of Table 7.4 reports the results for the cases in which the transitory component of x_{it} is correlated with the effect u_i . Similarly to those reported in Table 7.3, both the GLS estimators of β and γ show no signs of significant biases. For the cases in which the correlation between x_{it} and z_i decays only mildly over time ($m < 0.5$), the power of the Hausman test to detect nonzero-correlation

Table 7.4 Monte Carlo simulation results for CASE B

<i>m</i>	<i>N</i>	<i>T</i>	Panel I: $\rho_{xu,2} = 0$					Panel II: $\rho_{zu} = 0$					
			ρ_{zu}	Bias		β	Hausman rejection rate	$\rho_{xu,2}$	Bias		β	Hausman rejection rate	
				GLS	γ				GLS	γ			
0.0	50	5	0.0	0.001	-0.001	0.001	0.059	0.0	0.001	-0.001	0.001	0.059	
		50	0.0	0.000	-0.001	0.000	0.056	0.0	0.000	-0.001	0.000	0.056	
		100	0.0	0.000	-0.001	0.000	0.057	0.0	0.000	-0.001	0.000	0.057	
	100	5	0.0	-0.001	0.000	-0.001	0.048	0.0	-0.001	0.000	-0.001	0.048	
		50	0.0	0.000	0.000	0.000	0.051	0.0	0.000	0.000	0.000	0.051	
		100	0.0	0.000	0.000	0.000	0.052	0.0	0.000	0.000	0.000	0.052	
	50	5	5	0.5	-0.023	0.046	0.001	0.226	0.5	0.001	-0.001	0.000	0.136
			50	0.5	-0.002	0.034	0.000	0.344	0.5	0.000	-0.001	0.000	0.067
		50	100	0.5	-0.001	0.033	0.000	0.363	0.5	0.000	-0.001	0.000	0.067
			5	0.5	-0.024	0.046	-0.001	0.385	0.5	-0.001	0.000	-0.001	0.226
		100	50	0.5	-0.003	0.034	0.000	0.578	0.5	0.000	0.000	0.000	0.085
			100	0.5	-0.001	0.033	0.000	0.596	0.5	0.000	0.000	0.000	0.070
0.5	50	5	0.0	0.000	-0.001	0.000	0.054	0.0	0.000	-0.001	0.000	0.054	
		50	0.0	0.000	-0.001	0.000	0.047	0.0	0.000	-0.001	0.000	0.047	
		100	0.0	0.000	-0.001	0.000	0.053	0.0	0.000	-0.001	0.000	0.053	
	100	5	0.0	-0.001	0.000	0.000	0.050	0.0	-0.001	0.000	0.000	0.050	
		50	0.0	0.000	0.000	0.000	0.049	0.0	0.000	0.000	0.000	0.049	
		100	0.0	0.000	0.000	0.000	0.049	0.0	0.000	0.000	0.000	0.049	
	50	5	5	0.5	-0.014	0.038	0.000	0.199	0.5	0.000	0.000	0.000	0.211
			50	0.5	0.000	0.032	0.000	0.258	0.5	0.000	-0.001	0.000	0.214
		50	100	0.5	0.000	0.032	0.000	0.267	0.5	0.000	-0.001	0.000	0.214
			5	0.5	-0.015	0.038	0.000	0.348	0.5	-0.001	0.000	0.000	0.350
		100	50	0.5	-0.001	0.033	0.000	0.451	0.5	0.000	0.000	0.000	0.335
			100	0.5	0.000	0.033	0.000	0.466	0.5	0.000	0.000	0.000	0.331
2.0	50	5	0.0	0.000	0.000	0.000	0.056	0.0	0.000	0.000	0.000	0.056	
		50	0.0	0.000	-0.001	0.000	0.054	0.0	0.000	-0.001	0.000	0.054	
		100	0.0	0.000	-0.001	0.000	0.050	0.0	0.000	-0.001	0.000	0.050	
	100	5	0.0	0.000	0.000	0.000	0.052	0.0	0.000	0.000	0.000	0.052	
		50	0.0	0.000	0.000	0.000	0.048	0.0	0.000	0.000	0.000	0.048	
		100	0.0	0.000	0.000	0.000	0.051	0.0	0.000	0.000	0.000	0.051	
	50	5	5	0.5	-0.005	0.033	0.000	0.113	0.5	0.000	0.000	0.000	0.323
			50	0.5	0.000	0.032	0.000	0.065	0.5	0.000	-0.001	0.000	0.437
		50	100	0.5	0.000	0.032	0.000	0.058	0.5	0.000	-0.001	0.000	0.442
			5	0.5	-0.005	0.033	0.000	0.183	0.5	0.000	0.000	0.000	0.512
		100	50	0.5	0.000	0.033	0.000	0.076	0.5	0.000	0.000	0.000	0.603
			100	0.5	0.000	0.033	0.000	0.064	0.5	0.000	0.000	0.000	0.604

between the individual effect and the transitory component of the time-varying regressor is extremely low, especially when T is large. In contrast, when $m \geq 0.5$, the power of the Hausman test increases with T . These results are consistent with what the asymptotic results derived in the previous section have predicted.

Our simulation results can be summarized as follows. First, the finite-sample properties of the GLS estimators and the Hausman test are generally consistent with their asymptotic properties. Second, even if time-varying or time-invariant regressors are correlated with the unobservable individual effects, the GLS estimators of the coefficients of the time-varying regressors do not suffer from substantial biases if T is large, although the GLS estimators of the coefficients on time-invariant regressors could be biased regardless of the size of T . Third, the Hausman test has great power to detect non-zero correlation between the unobservable individual effects and the permanent components of time-varying regressors. In contrast, it has only limited power to detect non-zero correlations between the effects and transitory components of the time-varying regressors and between the effects and time-invariant regressors. The power of the Hausman test crucially depends on both the size of T and the covariance structure among regressors and the effects.

Both our asymptotic and Monte Carlo results provide empirical researchers with practical guidance. For the studies that focus on the effects of time-varying regressors on the dependent variable, the choice between the GLS and within estimators is an irrelevant issue when T is as large as N . Both the GLS and within estimators are consistent. For the studies that focus on the effects of time-invariant regressors, some cautions are required for correct interpretations of the Hausman test results. One important reason to prefer GLS over within is that it allows estimation of the effects of time-invariant regressors on dependent variables. For the consistent estimation of the effects of the time-invariant regressors, however, it is important to test endogeneity of the regressors. Our results indicate that large- T data do not necessarily improve the power property of the Hausman test. When the degrees of correlations between time-varying and time-invariant variables decrease quickly over time, the Hausman test generally lacks the power to detect endogeneity of time-invariant regressors. The tests based on subsamples with small T could provide more reliable test results. The different test results from a large- T sample and its small- T subsamples may provide some information about how the individual effect might be correlated with time-varying regressors. The rejection by large- T data but acceptance by small- T data would indicate that the effect is correlated with the permanent components of the time-varying regressor, but the degrees of the correlations are low. In contrast, the acceptance by large- T data but rejection by small- T data may indicate that the effect is correlated with the temporal components of the time-varying regressors.

So far, we have considered several simple cases to demonstrate how the convergence rates of the popular panel data estimators and the Hausman test are sensitive to data generating processes. For these simple cases, all of the relevant asymptotics can be obtained in a straightforward manner. In the following section, we will show that the main results obtained from this section apply to more general cases in which regressors are serially dependent with arbitrary covariance structures.

7.4 General Case

This section derives for the general model (7.1) the asymptotic distributions of the within, between, GLS estimators and the Hausman statistic. In Sect. 7.2, we have considered independently several simple models in which regressors are of particular characteristics. The general model we consider in this section contains all of the different types of regressors analyzed in Sect. 7.2. More detailed assumptions are introduced below.

From now on, the following notation is repeatedly used. The expression “ \rightarrow_p ” means “converges in probability,” while “ \Rightarrow ” means “converges in distribution” as in Sect. 7.2.2. For any matrix A , the norm $\|A\|$ signifies $\sqrt{\text{tr}(AA')}$. When B is a random matrix with $E\|B\|_p < \infty$, then $\|B\|_p$ denotes $(E\|B\|^p)^{1/p}$. We use $E_{\mathcal{F}}(\cdot)$ to denote the conditional expectation operator with respect to a sigma field \mathcal{F} . We also define $\|B\|_{\mathcal{F},p} = (E_{\mathcal{F}}\|B\|^p)^{1/p}$. The notation $x_N \sim a_N$ indicates that there exists n and finite constants d_1 and d_2 such that $\inf_{N \geq n} \frac{x_N}{a_N} \geq d_1$ and $\sup_{N \geq n} \frac{x_N}{a_N} \leq d_2$. We also use the following notation for relevant sigma-fields: $\mathcal{F}_{x_i} = \sigma(x_{i1}, \dots, x_{iT})$; $\mathcal{F}_{z_i} = \sigma(z_i)$; $\mathcal{F}_z = \sigma(\mathcal{F}_{z_1}, \dots, \mathcal{F}_{z_N})$; $\mathcal{F}_{w_i} = \sigma(\mathcal{F}_{x_i}, \mathcal{F}_{z_i})$; and $\mathcal{F}_w = \sigma(\mathcal{F}_{w_1}, \dots, \mathcal{F}_{w_N})$. The x_{it} and z_i are now $k \times 1$ and $g \times 1$ vectors, respectively.

As in Sect. 7.2, we assume that the regressors $(x'_{i1}, \dots, x'_{iT}, z'_i)'$ are independently distributed across different i . In addition, we make the following the assumption about the composite error terms u_i and v_{it} :

Assumption 1 (about u_i and v_{it}): For some $q > 1$,

- (i) The u_i are independent over different i with $\sup_i E|u_i|^{4q} < \infty$.
- (ii) The v_{it} are i.i.d. with mean zero and variance σ_v^2 across different i and t , and are independent of x_{is} , z_i and u_i , for all i , t , and s . Also, $\|v_{it}\|_{4q} \equiv \kappa_v$ is finite.

Assumption 1(i) is a standard regularity condition for error-components models. Assumption 1(ii) indicates that all of the regressors and individual effects are strictly exogenous with respect to the error terms v_{it} .¹⁵

We now make the assumptions about regressors. We here consider three different types of time-varying regressors: We partition the $k \times 1$ vector x_{it} into three subvectors, $x_{1,it}$, $x_{2,it}$, and $x_{3,it}$, which are $k_1 \times 1$, $k_2 \times 1$, and $k_3 \times 1$, respectively. The vector $x_{1,it}$ consists of the regressors with deterministic trends. We may think of three different types of trends: (i) cross-sectionally heterogeneous nonstochastic trends in mean (but not in variance or covariances); (ii) cross-sectionally homogeneous nonstochastic trends; and (iii) stochastic trends (trends in variance) such as unit-root time series. In Sect. 7.2, we have considered the first two cases while discussing the cases related with (7.20). The latter case is materially similar to CASE A.2, except that the convergence rates of estimators and test statistics are different under these

¹⁵As discussed in Sect. 7.2.1, this assumption rules out weakly exogenous or predetermined regressors.

two cases. Thus, we here only consider the case (i). We do not cover the cases of stochastic trends (iii), leaving the analysis of such cases to future study.

The two subvectors $x_{2,it}$ and $x_{3,it}$ are random regressors with no trend in mean. The partition of $x_{2,it}$ and $x_{3,it}$ is made based on their correlation with z_i . Specifically, we assume that the $x_{2,it}$ are not correlated with z_i , while the $x_{3,it}$ are. In addition, in order to accommodate CASEs A.1 and A.2, we also partition the subvector $x_{2,it}$ into $x_{21,it}$ and $x_{22,it}$, which are $k_{21} \times 1$ and $k_{22} \times 1$, respectively. Similarly to CASE A.1, the regressor vector $x_{21,it}$ is heterogeneous over different i , as well as different t , with different means $\Theta_{21,it}$. In contrast, $x_{22,it}$ is homogeneous cross-sectionally with means $\Theta_{22,t}$ for all i for given t (Case A.2). We also incorporate CASEs B.1–B.3 into the model by partitioning $x_{3,it}$ into $x_{31,it}$, $x_{32,it}$, and $x_{33,it}$, which are $k_{31} \times 1$, $k_{32} \times 1$, and $k_{33} \times 1$, respectively, depending on how fast their correlations with z_i decay over time. The more detailed assumptions on the regressors x_{it} and z_i follow:

Assumption 2 (about $x_{1,it}$):

- (i) For some $q > 1$, $\kappa_{x_1} \equiv \sup_{i,t} \|x_{1,it} - Ex_{1,it}\|_{4q} < \infty$.
- (ii) Let $x_{h,1,it}$ be the h^{th} element of $x_{1,it}$. Then, $Ex_{h,1,it} \sim t^{m_{h,1}}$ for all i and $h = 1, \dots, k_1$, where $m_{h,1} > 0$.

Assumption 3 (about $x_{2,it}$): For some $q > 1$,

- (i) $E(x_{21,it}) = \Theta_{21,it}$ and $E(x_{22,it}) = \Theta_{22,t}$, where $\sup_{i,t} \|\Theta_{21,it}\|$, $\sup_t \|\Theta_{22,t}\| < \infty$, and $\Theta_{21,it} \neq \Theta_{21,jt}$ if $i \neq j$.
- (ii) $\kappa_{x_2} \equiv \sup_{i,t} \|x_{2,it} - Ex_{2,it}\|_{4q} < \infty$.

Assumption 4 (about $x_{3,it}$): For some $q > 1$,

- (i) $E(x_{3,it}) = \Theta_{3,it}$, where $\sup_{i,t} \|\Theta_{3,it}\| < \infty$.
- (ii) $E\left(\sup_{i,t} \|x_{3,it} - E_{\mathcal{F}_{z_i}} x_{3,it}\|_{\mathcal{F}_{z_i}, 4q}^{8q}\right) < \infty$.
- (iii) Let $x_{h,3k,it}$ be the h^{th} element of $x_{3k,it}$, where $k = 1, 2, 3$. Then, conditional on z_i ,

$$\text{(iii.1)} \quad (E_{\mathcal{F}_{z_i}} x_{h,31,it} - Ex_{h,31,it}) \sim t^{-m_{h,31}} \text{ a.s., where } \frac{1}{2} < m_{h,31} \leq \infty \text{ for } h = 1, \dots, k_{31}$$

(here, $m_{h,31} = \infty$ implies that $E_{\mathcal{F}_{z_i}} x_{h,31,it} - Ex_{h,31,it} = 0$ a.s.);

$$\text{(iii.2)} \quad (E_{\mathcal{F}_{z_i}} x_{h,32,it} - Ex_{h,32,it}) \sim t^{-\frac{1}{2}} \text{ a.s. for } h = 1, \dots, k_{32};$$

$$\text{(iii.3)} \quad (E_{\mathcal{F}_{z_i}} x_{h,33,it} - Ex_{h,33,it}) \sim t^{-m_{h,33}} \text{ a.s., where } 0 \leq m_{h,33} < \frac{1}{2} \text{ for } h = 1, \dots, k_{33}.$$

Assumption 5 (about z_i): $\{z_i\}_i$ is i.i.d. over i with $E(z_i) = \Theta_z$, and $\|z_i\|_{4q} < \infty$ for some $q > 1$.

Panel data estimators of individual coefficients have different convergence rates depending on the types of the corresponding regressors. To address these differences, we define:

$$D_{x,T} = \text{diag}(D_{1T}, D_{2T}, D_{3T});$$

$$D_T = \text{diag}(D_{x,T}, I_g),$$

where

$$\begin{aligned} D_{1T} &= \text{diag}(T^{-m_1}, \dots, T^{-m_{k_1}}); \\ D_{2T} &= \text{diag}(D_{21T}, D_{22T}) = \text{diag}\left(I_{k_{21}}, \sqrt{T}I_{k_{22}}\right); \\ D_{3T} &= \text{diag}(D_{31T}, D_{32T}, D_{33T}) \\ &= \text{diag}\left(\sqrt{T}I_{k_{31}}, \sqrt{T}I_{k_{32}}, T^{m_{1,33}}, \dots, T^{m_{k_{33},33}}\right). \end{aligned}$$

Observe that D_{1T} , D_{2T} , and D_{3T} are conformable to regressor vectors $x_{1,it}$, $x_{2,it}$, and $x_{3,it}$, respectively, while D_T and I_g are to x_{it} and z_i , respectively. The diagonal matrix D_T is chosen so that $\text{plim}_{N \rightarrow \infty} \frac{1}{N} \sum_i D_T \tilde{w}_i \tilde{w}_i' D_T$ is well defined and finite. For future use, we also define

$$\begin{aligned} G_{x,T} &= \text{diag}(D_{1T}, I_{k_{21}}, I_{k_{22}}, I_{k_3}); \\ J_{x,T} &= \text{diag}(I_{k_1}, I_{k_{21}}, D_{22T}, D_{3T}), \end{aligned}$$

so that

$$D_{x,T} = G_{x,T} J_{x,T}.$$

Using this notation, we make the following regularity assumptions on the unconditional and conditional means of regressors:

Assumption 6 (convergence as $T \rightarrow \infty$): Defining $t = [Tr]$, we assume that the following restrictions hold as $T \rightarrow \infty$.

- (i) Let $\tau_1(r) = \text{diag}(r^{m_{1,1}}, \dots, r^{m_{k_1,1}})$, where $m_{h,1}$ is defined in Assumption 2. Then,

$$D_{1T} E(x_{1,it}) \rightarrow \tau_1(r) \Theta_{1,i}$$

uniformly in i and $r \in [0, 1]$, for some $\Theta_{1,i} = (\Theta_{1,1,i}, \dots, \Theta_{k_1,1,i})'$ with $\sup_i \|\Theta_{1,i}\| < \infty$.

- (ii) $\Theta_{21,it} \rightarrow \Theta_{21,i}$ and $\Theta_{3,it} \rightarrow \Theta_{3,i}$ uniformly in i with $\sup_i \|\Theta_{21,i}\| < \infty$ and $\sup_i \|\Theta_{3,i}\| < \infty$.

- (iii) Uniformly in i and $r \in [0, 1]$,

$$\begin{aligned} D_{31T} (E_{\mathcal{F}_{z_i}} x_{31,it} - E x_{31,it}) &\rightarrow 0_{k_{31} \times 1} \text{ a.s.}; \\ D_{32T} (E_{\mathcal{F}_{z_i}} x_{32,it} - E x_{32,it}) &\rightarrow \frac{1}{\sqrt{r}} I_{k_{32}} g_{32,i}(z_i) \text{ a.s.}; \\ D_{33T} (E_{\mathcal{F}_{z_i}} x_{33,it} - E x_{33,it}) &\rightarrow \tau_{33}(r) g_{33,i}(z_i) \text{ a.s.}, \end{aligned}$$

where

$$g_{32,i} = (g_{1,32,i}, \dots, g_{k_{32},32,i})'; g_{33,i} = (g_{1,33,i}, \dots, g_{k_{33},33,i})',$$

and $g_{32,i}(z_i)$ and $g_{33,i}(z_i)$ are zero-mean functions of z_i with

$$0 < E \sup_i \|g_{3k,i}(z_i)\|^{4q} < \infty, \text{ for some } q > 1,$$

and $g_{3k,i} \neq g_{3k,j}$ for $i \neq j$, and $\tau_{33}(r) = \text{diag}(r^{-m_{1,33}}, \dots, r^{-m_{k_{33},33}})$.

(iv) There exist $\tilde{\tau}(r)$ and $\tilde{G}_i(z_i)$ such that

$$\|D_{3T}(E_{\mathcal{F}_{z_i}} x_{3,it} - E x_{3,it})\| \leq \tilde{\tau}(r) \tilde{G}_i(z_i),$$

where $\int \tilde{\tau}(r)^{4q} dr < \infty$ and $E \sup_i \tilde{G}_i(z_i)^{4q} < \infty$ for some $q > 1$.

(v) Uniformly in (i, j) and $r \in [0, 1]$;

$$\begin{aligned} D_{31T}(E x_{31,it} - E x_{31,jt}) &\rightarrow 0_{k_{31} \times 1}, \\ D_{32T}(E x_{32,it} - E x_{32,jt}) &\rightarrow \frac{1}{\sqrt{r}} I_{k_{32}} (\mu_{g_{32i}} - \mu_{g_{32j}}), \\ D_{33T}(E x_{33,it} - E x_{33,jt}) &\rightarrow \tau_{33}(r) (\mu_{g_{33i}} - \mu_{g_{33j}}), \end{aligned}$$

with $\sup_i \|\mu_{g_{32i}}\|, \sup_i \|\mu_{g_{33i}}\| < \infty$.

Some remarks would be useful to understand Assumption 6. First, to have an intuition about what the assumption implies, we consider, as an illustrative example, the simple model in CASE 3 in Sect. 7.2.2, in which $x_{3,it} = \Pi_i z_i / t^m + e_{it}$, where e_{it} is independent of z_i and *i.i.d.* across i . For this case,

$$\begin{aligned} D_{3T}(E_{\mathcal{F}_{z_i}} x_{3,it} - E x_{3,it}) &= D_{3T} \Pi_i (z_i - E z_i) / t^m; \\ D_{3T}(E x_{3,it} - E x_{3,jt}) &= D_{3T} (\Pi_i E z_i - \Pi_j E z_j) / t^m. \end{aligned}$$

Thus,

$$\begin{aligned} g_{3k,i}(z_i) &= \Pi_i (z_i - E z_i); \\ \mu_{g_{3k,i}} &= \Pi_i E z_i. \end{aligned}$$

Second, Assumption 6(iii) makes the restriction that $E \sup_i \|g_{3k,i}(z_i)\|^{4q}$ is strictly positive, for $k = 2, 3$. This restriction is made to warrant that $g_{3k,i}(z_i) \neq 0$ a.s. If $g_{3k,i}(z_i) = 0$ a.s.,¹⁶ then

$$D_{3kT} E_{\mathcal{F}_{z_i}} (x_{3k,it} - E x_{3k,it}) \sim \tau_{3k}(r) g_{32,i}(z_i) = 0 \text{ a.s.},$$

and the correlations between $x_{3,it}$ and z_i no longer play any important role in asymptotics. Assumption 6(iii) rules out such cases.

¹⁶An example is the case in which $x_{3,it} = e_{it} \Pi_i z_i / t^m$, where e_{it} is independent of z_i with mean zero.

Assumption 6 is about the asymptotic properties of means of regressors as $T \rightarrow \infty$. We also need additional regularity assumptions on the means of regressors that apply as $N \rightarrow \infty$. Define

$$H_1 = \int_0^1 \tau_1(r)dr; H_{32} = \left(\int_0^1 \frac{1}{\sqrt{r}}dr \right) I_{k_{32}}; H_{33} = \int_0^1 \tau_{33}(r)dr;$$

and

$$E \left[\begin{pmatrix} H_{32}g_{32,i}(z_i) \\ H_{33}g_{33,i}(z_i) \\ z_i - E z_i \end{pmatrix} \begin{pmatrix} H_{32}g_{32,i}(z_i) \\ H_{33}g_{33,i}(z_i) \\ z_i - E z_i \end{pmatrix}' \right] = \begin{pmatrix} \Gamma_{g_{32},g_{32},i} & \Gamma_{g_{32},g_{33},i} & \Gamma_{g_{32},z,i} \\ \Gamma'_{g_{32},g_{33},i} & \Gamma_{g_{33},g_{33},i} & \Gamma_{g_{33},z,i} \\ \Gamma'_{g_{32},z,i} & \Gamma'_{g_{33},z,i} & \Gamma_{zz,i} \end{pmatrix}.$$

With this notation, we assume the followings:

Assumption 7 (convergence as $N \rightarrow \infty$): Define $\tilde{\Theta}_{1,i} = \Theta_{1,i} - \frac{1}{N} \sum_i \Theta_{1,i}$; $\tilde{\Theta}_{21,i} = \Theta_{21,i} - \frac{1}{N} \sum_i \Theta_{21,i}$; $\tilde{\mu}_{g_{32},i} = \mu_{g_{32},i} - \frac{1}{N} \sum_i \mu_{g_{32},i}$; and $\tilde{\mu}_{g_{33},i} = \mu_{g_{33},i} - \frac{1}{N} \sum_i \mu_{g_{33},i}$. As $N \rightarrow \infty$,

- (i) $\frac{1}{N} \sum_i \begin{pmatrix} H_1 \tilde{\Theta}_{1,i} \\ \tilde{\Theta}_{21,i} \\ H_{32} \tilde{\mu}_{g_{32},i} \\ H_{33} \tilde{\mu}_{g_{33},i} \end{pmatrix} \begin{pmatrix} H_1 \tilde{\Theta}_{1,i} \\ \tilde{\Theta}_{21,i} \\ H_{32} \tilde{\mu}_{g_{32},i} \\ H_{33} \tilde{\mu}_{g_{33},i} \end{pmatrix}' \rightarrow \begin{pmatrix} \Gamma_{\Theta_{1,\Theta_1}} & \Gamma_{\Theta_{1,\Theta_{21}}} & \Gamma_{\Theta_{1,\mu_{32}}} & \Gamma_{\Theta_{1,\mu_{33}}} \\ \Gamma'_{\Theta_{1,\Theta_{21}}} & \Gamma_{\Theta_{21,\Theta_{21}}} & \Gamma_{\Theta_{21,\mu_{32}}} & \Gamma_{\Theta_{21,\mu_{33}}} \\ \Gamma'_{\Theta_{1,\mu_{32}}} & \Gamma'_{\Theta_{21,\mu_{32}}} & \Gamma_{\mu_{32},\mu_{32}} & \Gamma_{\mu_{32},\mu_{33}} \\ \Gamma'_{\Theta_{1,\mu_{33}}} & \Gamma'_{\Theta_{21,\mu_{33}}} & \Gamma'_{\mu_{32},g_{33}} & \Gamma_{\mu_{33},\mu_{33}} \end{pmatrix}.$
- (ii) $\frac{1}{N} \sum_i \begin{pmatrix} \Gamma_{g_{32},g_{32},i} & \Gamma_{g_{32},g_{33},i} & \Gamma_{g_{32},z,i} \\ \Gamma'_{g_{32},g_{33},i} & \Gamma_{g_{33},g_{33},i} & \Gamma_{g_{33},z,i} \\ \Gamma'_{g_{32},z,i} & \Gamma'_{g_{33},z,i} & \Gamma_{zz,i} \end{pmatrix} \rightarrow \begin{pmatrix} \Gamma_{g_{32},g_{32}} & \Gamma_{g_{32},g_{33}} & \Gamma_{g_{32},z} \\ \Gamma'_{g_{32},g_{33}} & \Gamma_{g_{33},g_{33}} & \Gamma_{g_{33},z} \\ \Gamma'_{g_{32},z} & \Gamma'_{g_{33},z} & \Gamma_{z,z} \end{pmatrix}.$
- (iii) The limit of $\frac{1}{N} \sum_i \Theta_{1,i} \Theta'_{1,i}$ exists.

Apparently, by Assumptions 6 and 7, we assume the sequential convergence of the means of regressors as $T \rightarrow \infty$ followed by $N \rightarrow \infty$. However, this by no means implies that our asymptotic analysis is a sequential one. Instead, the uniformity conditions in Assumption 6 allow us to obtain our asymptotic results using the joint limit approach that applies as $(N, T \rightarrow \infty)$ simultaneously.¹⁷ Joint limit results can be obtained under an alternative set of conditions that assume uniform limits of the means of regressors sequentially as $N \rightarrow \infty$ followed by $T \rightarrow \infty$. Nonetheless, we adopt Assumptions 6 and 7 because they are much more convenient to handle the trends in regressors $x_{1,it}$ and $x_{3,it}$ for asymptotics.

The following notation is for conditional or unconditional covariances among time-varying regressors. Define

$$\Gamma_i(t, s) = [\Gamma_{jl,i}(t, s)]_{jl},$$

¹⁷For the details on the relationship between the sequential and joint approaches, see [Apostol \(1974, Theorems 8.39 and 9.16\)](#) for the cases of double indexed real number sequences, and [Phillips and Moon \(1999\)](#) for the cases of random sequences.

where $\Gamma_{j,l,i}(t,s) = E(x_{j,it} - E_{\mathcal{F}_{z_i}} x_{j,it})(x_{l,is} - E_{\mathcal{F}_{z_i}} x_{l,is})$, for $j, l = 2, 3$. Essentially, the Γ_i is the unconditional mean of the conditional variance-covariance matrix of $(x'_{2,it}, x'_{3,it})'$. We also define the unconditional variance-covariance matrix of $(x'_{1,it}, x'_{2,it}, x'_{3,it})'$ by

$$\tilde{\Gamma}_i(t,s) = [\tilde{\Gamma}_{j,l,i}(t,s)]_{j,l},$$

where $\tilde{\Gamma}_{j,l,i}(t,s) = E(x_{j,it} - E x_{j,it})(x_{l,is} - E x_{l,is})$, for $j, l = 1, 2, 3$. Observe that $\Gamma_{22,i}(t,s) = \tilde{\Gamma}_{22,i}(t,s)$, since $x_{2,it}$ and z_i are independent. With this notation, we make the following assumption on the convergence of variances and covariances:

Assumption 8 (convergence of covariances): As $(N, T \rightarrow \infty)$,

- (i) $\frac{1}{N} \sum_i \frac{1}{T} \sum_t \sum_s \begin{pmatrix} \Gamma_{22,i}(t,s) & \Gamma_{23,i}(t,s) \\ \Gamma'_{23,i}(t,s) & \Gamma_{33,i}(t,s) \end{pmatrix} \rightarrow \begin{pmatrix} \Gamma_{22} & \Gamma_{23} \\ \Gamma'_{23} & \Gamma_{33} \end{pmatrix}$.
- (ii) $\frac{1}{N} \sum_i \frac{1}{T} \sum_t \tilde{\Gamma}_i(t,t) \rightarrow \Phi$.

Note that the variance matrix $[\Gamma_{jl}]_{j,l=2,3}$ is the cross section average of the long-run variance-covariance matrix of $(x'_{2,it}, x'_{3,it})'$. For future use, we partition the two limits in the assumption conformably to $(x'_{21,it}, x'_{22,it}, x'_{31,it}, x'_{32,it}, x'_{33,it})'$ as follows:

$$\begin{pmatrix} \Gamma_{22} & \Gamma_{23} \\ \Gamma'_{23} & \Gamma_{33} \end{pmatrix} = \begin{pmatrix} \Gamma_{21,21} & \Gamma_{21,22} & \Gamma_{21,31} & \Gamma_{21,32} & \Gamma_{21,33} \\ \Gamma'_{21,22} & \Gamma_{22,22} & \Gamma_{22,31} & \Gamma_{22,32} & \Gamma_{22,33} \\ \Gamma'_{21,31} & \Gamma'_{22,31} & \Gamma_{31,31} & \Gamma_{31,32} & \Gamma_{31,33} \\ \Gamma'_{21,32} & \Gamma'_{22,32} & \Gamma'_{31,32} & \Gamma_{32,32} & \Gamma_{32,33} \\ \Gamma'_{21,33} & \Gamma'_{22,33} & \Gamma'_{31,33} & \Gamma'_{32,33} & \Gamma_{33,33} \end{pmatrix};$$

$$\Phi = \begin{pmatrix} \Phi_{11} & \Phi_{12} & \Phi_{13} \\ \Phi'_{12} & \Phi_{22} & \Phi_{23} \\ \Phi'_{13} & \Phi'_{23} & \Phi_{33} \end{pmatrix}.$$

Assumption 9 Let $\mathcal{F}_z^\infty = \sigma(z_1, \dots, z_N, \dots)$. For a generic constant M that is independent of N and T , the followings hold:

- (i) $\sup_{i,T} \left\| \frac{1}{T} \sum_t (x_{1,it} - E x_{1,it}) \right\|_4 < M$,
- (ii) $\sup_{i,T} \left\| \frac{1}{\sqrt{T}} \sum_t (x_{2,it} - E x_{2,it}) \right\|_4 < M$,
- (iii) $\sup_{i,T} \left\| \frac{1}{\sqrt{T}} \sum_t (x_{3,it} - E_{\mathcal{F}_{z_i}} x_{3,it}) \right\|_4 < M$.

Assumption 9 assumes that the fourth moments of the sums of the regressors $x_{1,it}$, $x_{2,it}$, and $x_{3,it}$ are uniformly bounded. This assumption is satisfied under mild restrictions on the moments of x_{it} and on the temporal dependence of $x_{2,it}$ and $x_{3,it}$. For sufficient conditions for Assumption 9, refer to [Ahn and Moon \(2001\)](#).

Finally, we make a formal definition of the random effects assumption, which is a more rigorous version of (7.3).

Assumption 10 (*random effects*): Conditional on \mathcal{F}_w , $\{u_i\}_{i=1,\dots,N}$ is i.i.d. with mean zero, variance σ_u^2 and finite $\kappa_u \equiv \|u_i\|_{\mathcal{F}_w,4}$.

To investigate the power property of the Hausman test, we also need to define an alternative hypothesis that states a particular direction of model misspecification. Among many alternatives, we here consider a simpler one. Specifically, we consider an alternative hypothesis under which the conditional mean of u_i is a linear function of $D_T \tilde{w}_i$. Abusing the conventional definition of fixed effects (that indicates nonzero-correlations between $w_i = (x'_{it}, z'_i)'$ and u_i), we refer to this alternative as the fixed effects assumption:

Assumption 11 (*fixed effects*): Conditional on \mathcal{F}_w , the $\{u_i\}_{i=1,\dots,N}$ is i.i.d. with mean $\tilde{w}'_i D_T \lambda$ and variance σ_u^2 , where λ is a $(k + g) \times 1$ nonrandom nonzero vector.

Here, $D_T \tilde{w}_i = [(D_{x,T} \tilde{x}_i)', \tilde{z}_i]$ can be viewed as a vector of detrended regressors. Thus, Assumption 11 indicates non-zero correlations between the effect u_i and detrended regressors. The term $\tilde{w}'_i D_T \lambda$ can be replaced by $\lambda_o + \tilde{w}'_i D_T \lambda$, where λ_o is any constant scalar. We use the term $\tilde{w}'_i D_T \lambda$ instead of $\lambda_o + \tilde{w}'_i D_T \lambda$ simply for convenience.

A sequence of local versions of the fixed effects hypothesis is given:

Assumption 12 (*local alternatives to random effects*): Conditional on \mathcal{F}_w , the sequence $\{u_i\}_{i=1,\dots,N}$ is i.i.d. with mean $\tilde{w}'_i D_T \lambda / \sqrt{N}$, variance σ_u^2 , and $\kappa_u^4 = E_{\mathcal{F}_w} (u_i - E_{\mathcal{F}_w} u_i)^4 < \infty$, where $\lambda \neq 0_{(k+g) \times 1}$ is a nonrandom vector in \mathbb{R}^{k+g} .

Under this Assumption, $E(D_T \tilde{w}_i u_i) = \frac{1}{\sqrt{N}} E(D_T \tilde{w}_i \tilde{w}'_i D_T) \lambda \rightarrow 0_{(k+g) \times 1}$, as $(N, T \rightarrow \infty)$. Observe that these local alternatives are of the forms introduced in Table 7.1

The following assumption is required for identification of the within and between estimators of β and γ .

Assumption 13 *The matrices Ψ_x and Ξ are positive definite.*

Two remarks on this assumption follow. First, this assumption is also sufficient for identification of the GLS estimation. Second, while the positive definiteness of the matrix Ξ is required for identification of the between estimators, it is not a necessary condition for the asymptotic distribution of the Hausman statistic obtained below. We can obtain the same asymptotic results for the Hausman test even if we alternatively assume that within estimation can identify β (positive definite Ψ_x) and between estimation can identify γ given β (the part of Ξ corresponding to \tilde{z}_i is positive definite).¹⁸ Nonetheless, we assume that Ξ is invertible for convenience.

¹⁸This claim can be checked with the following simple example. Consider a simple model with one time-varying regressor x_{it} and one time invariant regressor z_i . Assume that $x_{it} = az_i + e_{it}$, where the e_{it} are i.i.d. over different i and t . For this model, it is straightforward to show that the

We now consider the asymptotic distributions of the within, between and GLS estimators of β and γ :

Theorem 1 (asymptotic distribution of the within estimator): Under Assumptions 1–8 and 13, as $(N, T \rightarrow \infty)$,

$$\sqrt{NT}G_{x,T}^{-1}(\hat{\beta}_w - \beta) \Rightarrow N(0, \sigma_v^2\Psi_x^{-1}).$$

Theorem 2 (asymptotic distribution of the between estimator): Suppose that Assumptions 1–8 and 13 hold. As $(N, T \rightarrow \infty)$,

(a) Under Assumption 10 (random effects),

$$D_T^{-1}\sqrt{N}\begin{pmatrix} \hat{\beta}_b - \beta \\ \hat{\gamma}_b - \gamma \end{pmatrix} = \begin{pmatrix} D_{x,T}^{-1}\sqrt{N}(\hat{\beta}_b - \beta) \\ \sqrt{N}(\hat{\gamma}_b - \gamma) \end{pmatrix} \Rightarrow N(0, \sigma_u^2\Xi^{-1});$$

(b) Under Assumption 12 (local alternatives to random effects),

$$D_T^{-1}\sqrt{N}\begin{pmatrix} \hat{\beta}_b - \beta \\ \hat{\gamma}_b - \gamma \end{pmatrix} = \begin{pmatrix} D_{x,T}^{-1}\sqrt{N}(\hat{\beta}_b - \beta) \\ \sqrt{N}(\hat{\gamma}_b - \gamma) \end{pmatrix} \Rightarrow N(\Xi\lambda, \sigma_u^2\Xi^{-1}).$$

Theorem 3 (asymptotic distribution of the GLS estimator of β): Suppose that Assumptions 1–8 and 13 hold.

(a) Under Assumption 12 (local alternatives to random effects),

$$\sqrt{NT}G_{x,T}^{-1}(\hat{\beta}_g - \beta) = \sqrt{NT}G_{x,T}^{-1}(\hat{\beta}_w - \beta) + o_p(1),$$

as $(N, T \rightarrow \infty)$.

(b) Suppose that Assumption 11 (fixed effects) holds. Partition $\lambda = (\lambda'_x, \lambda'_z)'$ conformably to the sizes of x_{it} and z_i . Assume that $\lambda_x \neq 0_{k \times 1}$. If $N/T \rightarrow c < \infty$ and the included regressors are only of the $x_{22,it}$ - and $x_{3,it}$ -types (no trends and no cross-sectional heterogeneity in x_{it}), then

$$\sqrt{NT}G_{x,T}^{-1}(\hat{\beta}_g - \beta) = \sqrt{NT}G_{x,T}^{-1}(\hat{\beta}_w - \beta) + o_p(1).$$

Theorem 4 (asymptotic distribution of the GLS estimator of γ): Suppose that Assumptions 1–8 and 13 hold. Define $l'_z = (0_{g \times k}; I_g)$. Then, the following statements hold as $(N, T \rightarrow \infty)$.

matrix Ξ fails to be invertible. Nonetheless, under the random effects assumption, the Hausman statistic can be shown to follow a χ^2 distribution with the degree of freedom equal to one.

(a) Under Assumption 12 (local alternatives to random effects),

$$\begin{aligned}\sqrt{N}(\hat{\gamma}_g - \gamma) &= \left(\frac{1}{N} \sum_i \tilde{z}_i \tilde{z}_i' \right)^{-1} \left(\frac{1}{\sqrt{N}} \sum_i \tilde{z}_i \tilde{u}_i \right) + o_p(1) \\ &\Rightarrow N \left((l_z' \Xi l_z)^{-1} l_z' \Xi \lambda, \sigma_u^2 (l_z' \Xi l_z)^{-1} \right).\end{aligned}$$

(b) Under Assumption 11 (fixed effects),

$$(\hat{\gamma}_g - \gamma) \rightarrow_p (l_z' \Xi l_z)^{-1} l_z' \Xi \lambda.$$

Several remarks follow. First, all of the asymptotic results given in Theorems 1–4 except for Theorem 3(b) hold as $(N, T \rightarrow \infty)$, without any particular restriction on the convergence rates of N and T . The relative size of N and T does not matter for the results, so long as both N and T are large. Second, one can easily check that the convergence rates of the panel data estimates of individual β coefficients (on the $x_{2,it}$ - and $x_{3,it}$ -type regressors) reported in Theorems 1–4 are consistent with those from Sect. 7.2.2. Third, Theorem 2 shows that under Assumption 10 (random effects), the between estimator of γ , $\hat{\gamma}_b$, is \sqrt{N} -consistent regardless of the characteristics of time-varying regressors. Fourth, both the between estimators of β and γ are asymptotically biased under the sequence of local alternatives (Assumption 12). Fifth, as Theorem 3(a) indicates, the within and GLS estimators of β are asymptotically equivalent not only under the random effects assumption, but also under the local alternatives. Furthermore, the GLS estimator of β is asymptotically unbiased under the local alternatives, while the between estimator of β is not. The asymptotic equivalence between the within and GLS estimation under the random effects assumption is nothing new. Previous studies have shown this equivalence based on a naive sequential limit method ($T \rightarrow \infty$ followed by $N \rightarrow \infty$) and some strong assumptions such as fixed regressors. Theorem 3(a) and (b) confirm the same equivalence result but with more a rigorous joint limit approach as $(N, T \rightarrow \infty)$ simultaneously. It is also intriguing to see that the GLS and within estimators are equivalent even under the local alternative hypotheses.

Sixth, somewhat surprisingly, as Theorem 3(b) indicates, even under the fixed effects assumption (Assumption 11), the GLS estimator of β could be asymptotically unbiased (and consistent) and equivalent to the within counterpart, (i) if the size (N) of the cross section units does not dominate excessively the size (T) of time series in the limit ($N/T \rightarrow c < \infty$), and (ii) if the model does not contain trended or cross-sectionally heterogeneous time-varying regressors. This result indicates that when the two conditions are satisfied, the biases in GLS caused by fixed effects are generally much smaller than those in the between estimator. If at least one of these two conditions is violated, that is, if $N/T \rightarrow \infty$, or if the other types of regressors

are included, the limit of $(\hat{\beta}_g - \hat{\beta}_w)$ is determined by how fast $N/T \rightarrow \infty$ and how fast the trends in the regressors increase or decrease.¹⁹

Finally, Theorem 4(a) indicates that under the local alternative hypotheses, the GLS estimator $\hat{\gamma}_g$ is \sqrt{N} -consistent and asymptotically normal, but asymptotically biased. The limiting distribution of $\hat{\gamma}_g$, in this case, is equivalent to the limiting distribution of the OLS estimator of γ in the panel model with the known coefficients of the time-varying regressors x_{it} (OLS on $\tilde{y}_{it} - \beta' \tilde{x}_{it} = \gamma' \tilde{z}_i + (u_i + \tilde{v}_{it})$). Clearly, the GLS estimator $\hat{\gamma}_g$ is asymptotically more efficient than the between estimator $\hat{\gamma}_b$. On the other hand, under the fixed effect assumption, unlike the GLS estimator of β , $\hat{\beta}_g$, the GLS estimator $\hat{\gamma}_g$ is not consistent as $(N, T \rightarrow \infty)$. The asymptotic bias of $\hat{\gamma}_g$ is given in Theorem 4(b).

Lastly, the following theorem finds the asymptotic distribution of the Hausman test statistic under the random effect assumption and the local alternatives:

Theorem 5 *Suppose that Assumptions 1–8 and 13 hold. Corresponding to the size of (\bar{x}'_i, \bar{z}'_i) , partition Ξ and λ , respectively, as follows:*

$$\Xi = \begin{pmatrix} \Xi_{xx} & \Xi_{xz} \\ \Xi'_{xz} & \Xi_{zz} \end{pmatrix}; \lambda = \begin{pmatrix} \lambda_x \\ \lambda_z \end{pmatrix}.$$

Then, as $(N, T \rightarrow \infty)$,

(a) Under Assumption 10 (random effects),

$$\mathcal{HM}_{NT} \Rightarrow \chi^2_k;$$

(b) Under Assumption 12 (local alternatives to random effects),

$$\mathcal{HM}_{NT} \Rightarrow \chi^2_k(\eta),$$

where $\eta = \lambda'_x (\Xi_{xx} - \Xi_{xz} \Xi_{zz}^{-1} \Xi'_{xz}) \lambda_x / \sigma_u^2$ is the noncentral parameter.

The implications of the theorem are discussed in Sect. 7.2.3.

7.5 Conclusion

This paper has considered the large- N and large- T asymptotic properties of the within, between and random effects GLS estimators, as well as those of the Hausman test statistic. The convergence rates of the between estimator and the

¹⁹In this case, without specific assumptions on the convergence rates of N/T and the trends, it is hard to generalize the limits of the difference of the within and the GLS estimators.

Hausman test statistic are closely related, and the rates crucially depend on whether regressors are cross-sectionally heterogeneous or homogeneous. Nonetheless, the Hausman test is always asymptotically χ^2 -distributed under the random effects assumption. Our simulation results indicate that our asymptotic results are generally consistent with the finite-sample properties of the estimators and the Hausman test even if N and T are small.

Under certain local alternatives (where the conditional means of unobservable individual effects are linear in the regressors), we also have investigated the asymptotic power properties of the Hausman test. Regardless of the size of T , the Hausman test has power to detect non-zero correlations between unobservable individual effects and the permanent components of time-varying regressors. In contrast, the test has no power to detect non-zero correlations between the effects and the transitory components of time-varying-regressors if T is large and if the time-varying regressors do have permanent components. The Hausman test has some (although limited) power to detect non-zero correlations between the effects and time-invariant regressors when the correlations between time-varying and time-invariant regressors remain high over time. However, when the correlations decay quickly over time, the test loses its power.

In this paper, we have restricted our attention to the asymptotic and finite-sample properties of the existing estimators and tests when panel data contain both large numbers of cross section and time series observations. No new estimator or test is introduced. However, this paper makes several contributions to the literature. First, we have shown that the GLS and within estimators, as well as the Hausman test, can be used without any adjustment for the data with large T . Second, for the cases with both large N and T , we provide a theoretical link between the asymptotic equivalence of the within and GLS estimator and the asymptotic distribution of the Hausman test. Third, we have shown that cross-sectional heterogeneity in regressors can play an important role in asymptotics. Previous studies have often assumed that data are cross-sectionally i.i.d. Our findings suggest that future studies should pay attention to cross-sectional heterogeneity. Fourth, we find that the power of the Hausman test depends on T .

Fifth and finally, our results also provide empirical researchers with some useful guidance. Different Hausman test results from large- T and small- T data can provide some information about how the individual effect is correlated with time-varying regressors. The rejection by large- T data but acceptance by small- T data would indicate that the effect is correlated with the permanent components of the time-varying regressor, but the degrees of the correlations are low. In contrast, the acceptance by large- T data but rejection by small- T data may indicate that the effect is correlated with the temporal components of the time-varying regressors. Whether the individual effect is correlated with temporal or permanent components of time-varying regressors is important to determine what instruments should be used to estimate the coefficients of time-invariant regressors when the random effects assumption is rejected. For example, as an anonymous referee pointed out, a key identification requirement of the instrumental variables proposed by [Breusch et al. \(1989\)](#) is that only the permanent components of the time-varying regressors

are correlated with the individual effects. If the Hausman test indicates that the individual effects are only temporally correlated with the time-varying regressors, the BMS instruments need not be used.

Needless to say, the model we have considered is a restrictive one. Extensions of our approach to more general models would be useful future research agendas. First, we have not considered the cases with more general errors; e.g., hetroskedastic and/or serially correlated errors. It would be useful to extend our approach to such general cases. Second, we have focused on the large- N and large- T properties of the panel data estimators and tests that are designed for the models with large N and small T . For the models with large N and large T , it may be possible to construct the estimators and test methods based on large- N and large- T asymptotics that may have better properties than the estimators and the tests analyzed here. Developing alternative estimators based on large- T and large- N asymptotics and addressing the issue of unit roots would be important research agendas. Third, another possible extension would be the instrumental variables estimation of Hausman and Taylor (1981), Amemiya and MaCurdy (1986), and Breusch et al. (1989). For an intermediate model between fixed effects and random effects, these studies propose several instrumental variables estimators by which both the coefficients on time-varying and time-invariant regressors can be consistently estimated. It would be interesting to investigate the large- N and large- T properties of these instrumental variables estimators as well as those of the Hausman test and other GMM tests based on these estimators.

Appendix

First, we provide some preliminary lemmas that are useful in proving the main results in Sect. 7.4. Due to space limitation, we omit the proofs of the lemmas.²⁰

In this section, notation M denotes a generic constant that is finite. Recall that $w_{it} = (x'_{1,it}, x'_{2,it}, x'_{3,it}, z'_i)'$. We also repeatedly use the diagonal matrix D_T defined in Sect. 7.4.

Lemma 6 *Under Assumptions 1–8, we obtain the following results as $(N, T \rightarrow \infty)$.*

- (a) $\frac{1}{N} \sum_i \frac{1}{T} \sum_t G_{x,T} \tilde{x}_{it} \tilde{x}'_{it} G_{x,T} \rightarrow_p \Psi_x$;
 - (b) $\frac{1}{\sqrt{N}} \sum_i \frac{1}{\sqrt{T}} \sum_t G_{x,T} \tilde{x}_{it} \tilde{v}_{it} \Rightarrow N(0, \sigma_v^2 \Psi_x)$,
- where

$$\Psi_x = \begin{pmatrix} \int_0^1 (\tau_1 - \int \tau_1) \left(\lim_N \frac{1}{N} \sum_i \Theta_{1,i} \Theta'_{1,i} \right) (\tau_1 - \int \tau_1)' dr & 0 & 0 \\ 0 & \Phi_{22} & \Phi_{23} \\ 0 & \Phi_{32} & \Phi_{33} \end{pmatrix}.$$

²⁰Detailed proofs are available from the authors upon request.

Lemma 7 Suppose that Assumptions 1–8 hold. Define $\Xi = \Xi_1 + \Xi_2$, where

$$\Xi_1 = \text{diag} \left(0_{k_1}, 0_{k_{21}}, \begin{pmatrix} \Gamma_{22,22} & \Gamma_{22,31} & \Gamma_{22,32} \\ \Gamma'_{22,31} & \Gamma_{31,31} & \Gamma_{31,32} \\ \Gamma'_{22,32} & \Gamma'_{31,32} & \Gamma_{32,32} \end{pmatrix}, 0_{k_{33}}, 0_{k_z} \right);$$

$$\Xi_2 = \begin{pmatrix} \Gamma_{\Theta_1, \Theta_1} & \Gamma_{\Theta_1, \Theta_{21}} & 0 & 0 & \Gamma_{\Theta_1, \mu_{32}} & \Gamma_{\Theta_1, \mu_{33}} & 0 \\ \Gamma'_{\Theta_1, \Theta_{21}} & \Gamma_{\Theta_{21}, \Theta_{21}} & 0 & 0 & \Gamma_{\Theta_{21}, \mu_{32}} & \Gamma_{\Theta_{21}, \mu_{33}} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \Gamma'_{\Theta_1, \mu_{32}} & \Gamma'_{\Theta_{21}, \mu_{32}} & 0 & 0 & \Gamma_{g_{32}, g_{32}} & \Gamma_{g_{32}, g_{33}} & \Gamma_{g_{32}, z} \\ & & & & +\Gamma_{\mu_{32}, \mu_{32}} & +\Gamma_{\mu_{32}, \mu_{33}} & \\ \Gamma'_{\Theta_1, \mu_{33}} & \Gamma'_{\Theta_{21}, \mu_{33}} & 0 & 0 & \Gamma'_{g_{32}, g_{33}} & \Gamma_{g_{33}, g_{33}} & \Gamma_{g_{33}, z} \\ & & & & +\Gamma'_{\mu_{32}, \mu_{32}} & +\Gamma_{\mu_{33}, \mu_{33}} & \\ 0 & 0 & 0 & 0 & \Gamma'_{g_{32}, z} & \Gamma'_{g_{33}, z} & \Gamma_{z, z} \end{pmatrix}.$$

Then, under Assumption 12, as $(N, T \rightarrow \infty)$, the followings hold.

- (a) $\frac{1}{N} \sum_i D_T \tilde{w}_i \tilde{w}_i' D_T \rightarrow_p \Xi$.
 (b) $\sup_{N, T} \sup_{1 \leq i \leq N} E \|D_T \tilde{w}_i\|^4 < M$, for some constant $M < \infty$.
 (c) $\frac{1}{\sqrt{N}} \sum_i D_T \tilde{w}_i \tilde{v}_i \rightarrow_p 0$.

Lemma 8 Under Assumptions 1–8 and 12 (local alternatives to random effects), as $(N, T \rightarrow \infty)$,

$$\frac{1}{\sqrt{N}} \sum_i D_T \tilde{w}_i \tilde{u}_i \Rightarrow N (\Xi \lambda, \sigma_u^2 \Xi).$$

Lemma 9 Under Assumptions 1–8 and 11 (fixed effects),

$$\frac{1}{N} \sum_i D_T \tilde{w}_i \tilde{u}_i \rightarrow_p \Xi \lambda,$$

as $(N, T \rightarrow \infty)$.

Proof of Theorem 1

Theorem 1 follows by Lemma 6(a) and (b). ■

Proof of Theorem 2

Theorem 2 holds by Lemmas 7(a), (c) and 8. ■

Before we prove the rest of the theorems given in Sect. 7.4, we introduce the following notation:

$$\begin{aligned}
 A_1 &= \frac{1}{N} \sum_i \frac{1}{T} \sum_t (x_{it} - \bar{x}_i) (x_{it} - \bar{x}_i)' ; \\
 A_2 &= \frac{1}{N} \sum_i \frac{1}{T} \sum_t (x_{it} - \bar{x}_i) (v_{it} - \bar{v}_i) ; \\
 A_3 &= \frac{1}{N} \sum_i \tilde{x}_i \tilde{x}_i' ; A_4 = \frac{1}{N} \sum_i \tilde{x}_i \tilde{u}_i ; A_5 = \frac{1}{N} \sum_i \tilde{x}_i \tilde{v}_i ; \\
 B_3 &= \frac{1}{N} \sum_i \tilde{z}_i \tilde{z}_i' ; B_4 = \frac{1}{N} \sum_i \tilde{z}_i \tilde{u}_i ; B_5 = \frac{1}{N} \sum_i \tilde{z}_i \tilde{v}_i ; \\
 C &= \frac{1}{N} \sum_i \tilde{x}_i \tilde{z}_i' ; \\
 F_1 &= A_3 - CB_3^{-1}C' ; F_2 = A_4 + A_5 - CB_3^{-1} (B_4 + B_5) .
 \end{aligned} \tag{7.25}$$

Proof of Theorem 3

Using the notation given in (7.25), we can express the GLS estimator $\hat{\beta}_g$ by

$$\begin{aligned}
 &\sqrt{NT}G_{x,T}^{-1} (\hat{\beta}_g - \beta) \\
 &= [G_{x,T}A_1G_{x,T} + \theta_T^2G_{x,T} \{A_3 - CB_3^{-1}C'\} G_{x,T}]^{-1} \\
 &\quad \times \sqrt{NT}G_{x,T} \{A_2 + \theta_T^2 [(A_4 + A_5) - CB_3^{-1} (B_4 + B_5)]\} , \tag{7.26}
 \end{aligned}$$

where $\theta_T = \sqrt{\sigma_v^2 / (T\sigma_u^2 + \sigma_v^2)}$.

Part (a): Using Lemma 7(a), we can show that

$$\theta_T^2G_{x,T} \{A_3 - CB_3^{-1}C'\} G_{x,T} = O_p(\theta_T^2) = o_p(1) . \tag{7.27}$$

Next, from Lemmas 7(c) and 8, under the local alternatives to random effects (Assumption 12), it is possible to show that

$$\sqrt{NT}G_{x,T} \{\theta_T^2 [(A_4 + A_5) - CB_3^{-1} (B_4 + B_5)]\} = O_p\left(\frac{1}{\sqrt{T}}\right) = o_p(1) . \tag{7.28}$$

Substituting (7.27) and (7.28) into (7.26), we have

$$\begin{aligned}
 \sqrt{NT}(\hat{\beta}_g - \beta) &= [G_{x,T}A_1G_{x,T} + o_p(1)]^{-1} [\sqrt{NT}G_{x,T}A_2 + o_p(1)] \\
 &= \sqrt{NT}(\hat{\beta}_w - \beta) + o_p(1) .
 \end{aligned}$$

The last equality results from Lemma 6(a), (b) and Theorem 1. ■

Part (b): Similarly to Part (a), we can easily show that under the assumptions given in Part (b), the denominator in (7.26) is

$$\frac{1}{NT} \sum_i \sum_t G_{x,T} (x_{it} - \bar{x}_i) (x_{it} - \bar{x}_i)' G_{x,T} + o_p(1) . \tag{7.29}$$

Consider the second term of the numerator of (7.26):

$$\theta^2 \sqrt{T} G_{x,T} \left\{ \begin{array}{c} \frac{1}{\sqrt{N}} \sum_i \tilde{x}_i (\tilde{u}_i + \tilde{v}_i) \\ - \left(\frac{1}{N} \sum_i \tilde{x}_i \tilde{z}'_i \right) \left(\frac{1}{N} \sum_i \tilde{z}_i \tilde{z}'_i \right)^{-1} \left(\frac{1}{\sqrt{N}} \sum_i \tilde{z}_i (\tilde{u}_i + \tilde{v}_i) \right) \end{array} \right\}. \quad (7.30)$$

Notice that by Lemmas 6, 7, and 9, under the fixed effect assumption (Assumption 11), the first term of (7.30) is

$$\begin{aligned} \theta^2 \sqrt{NT} G_{x,T} \frac{1}{N} \sum_i \tilde{x}_i (\tilde{u}_i + \tilde{v}_i) &= \theta^2 \sqrt{NT} G_{x,T} D_{x,T}^{-1} \frac{1}{N} \sum_i D_{x,T} \tilde{x}_i (\tilde{u}_i + \tilde{v}_i) \\ &= \theta^2 \sqrt{NT} G_{x,T} D_{x,T}^{-1} \{ \Xi_x \lambda + o_p(1) \}, \end{aligned}$$

where we partition $\Xi = \begin{pmatrix} \Xi_{xx} & \Xi_{xz} \\ \Xi_{zx} & \Xi_{zz} \end{pmatrix}$ conformably to the sizes of x_{it} and z_i , and set $\Xi_x = (\Xi_{xx}, \Xi_{xz})$. Similarly, by Lemmas 6, 7, and 9, under the fixed effect assumption, the second term of (7.30) is

$$\begin{aligned} &\theta^2 \sqrt{NT} G_{x,T} D_{x,T}^{-1} \left\{ \left(\frac{1}{N} \sum_i D_{x,T} \tilde{x}_i \tilde{z}'_i \right) \left(\frac{1}{N} \sum_i \tilde{z}_i \tilde{z}'_i \right)^{-1} \left(\frac{1}{N} \sum_i \tilde{z}_i (\tilde{u}_i + \tilde{v}_i) \right) \right\} \\ &= \theta^2 \sqrt{NT} G_{x,T} D_{x,T}^{-1} \{ \Xi_{xz} \Xi_{zz}^{-1} \Xi_z \lambda + o_p(1) \}. \end{aligned}$$

Therefore, the limit of (7.30) is

$$\left(\theta^2 \sqrt{NT} G_{x,T} D_{x,T}^{-1} \right) \left[\left(\Xi_{xx} - \Xi_{xz} \Xi_{zz}^{-1} \Xi_{zx} : 0 \right) \lambda + o_p(1) \right].$$

Recall that it is assumed that $\frac{N}{T} \rightarrow c < \infty$. Also, recall that under the restrictions given in the theorem, $G_{x,T} = \text{diag}(I_{k_{22}}, I_{k_3})$ and $D_{x,T} = \text{diag}(\sqrt{T} I_{k_{22}}, D_{3T}^{-1})$. Then, letting $\lambda_{\max}(A)$ denote the maximum eigenvalue of matrix A , we can have

$$\lambda_{\max} \left(\theta^2 \sqrt{NT} G_{x,T} D_{x,T}^{-1} \right) = O(1) \sqrt{\frac{N}{T}} \lambda_{\max} \left(\frac{1}{\sqrt{T}} I_{k_{22}}, D_{3T}^{-1} \right) \rightarrow 0.$$

Thus, under the assumptions of Part (b), the probability limit of the numerator of (7.26) is

$$\frac{1}{\sqrt{NT}} \sum_i \sum_t G_{x,T} \tilde{x}_{it} \tilde{v}_{it} + o_p(1). \quad (7.31)$$

Combining (7.29) and (7.31), we can obtain Part (b). ■

Proof of Theorem 4

Using the notation in (7.25), we can express the GLS estimator $\hat{\gamma}_g$ by

$$\begin{aligned} \hat{\gamma}_g - \gamma &= \left[B_3 - C' \left(\frac{1}{\theta_T^2} A_1 + A_3 \right)^{-1} C \right]^{-1} \\ &\times \left[(B_4 + B_5) - C' \left(\frac{1}{\theta_T^2} A_1 + A_3 \right)^{-1} \left(\frac{1}{\theta_T^2} A_2 + (A_4 + A_5) \right) \right]. \end{aligned} \quad (7.32)$$

Part (a): Using Lemmas 6(a) and 7(a), we can show that $C' \left(\frac{1}{\theta_T^2} A_1 + A_3 \right)^{-1} C = o_p(1)$, which implies that, as $(N, T \rightarrow \infty)$, the denominator of (7.32) is

$$B_3 + o_p(1). \quad (7.33)$$

Next, under both the random effects assumption (Assumption 10) and the local alternatives (Assumption 12), it follows from Lemmas 6–8 that the second term in the numerator of (7.32) is

$$C' \left(\frac{1}{\theta^2} A_1 + A_3 \right)^{-1} \left(\frac{1}{\theta^2} \sqrt{N} A_2 + \sqrt{N} (A_4 + A_5) \right) = o_p(1).$$

Also, by Lemma 7(c), $\sqrt{N} B_5 = o_p(1)$. Therefore, the numerator of (7.32) is

$$\sqrt{N} B_4 + o_p(1), \quad (7.34)$$

as $(N, T \rightarrow \infty)$. In view of (7.32)–(7.34), we have

$$\sqrt{N} (\hat{\gamma}_g - \gamma) = \left(\frac{1}{N} \sum_i \tilde{z}_i \tilde{z}'_i \right)^{-1} \left(\frac{1}{\sqrt{N}} \sum_i \tilde{z}_i \tilde{u}_i \right) + o_p(1),$$

as $(N, T \rightarrow \infty)$. Finally, by Lemmas 7(a) and 8, as $(N, T \rightarrow \infty)$,

$$\begin{aligned} \sqrt{N} (\hat{\gamma}_g - \gamma) &= \left(\frac{1}{N} \sum_i \tilde{z}_i \tilde{z}'_i \right)^{-1} \left(\frac{1}{\sqrt{N}} \sum_i \tilde{z}_i \tilde{u}_i \right) \\ &\Rightarrow N \left((l'_z \Xi l_z)^{-1} (l'_z \Xi \lambda), (l'_z \Xi l_z)^{-1} \right), \end{aligned}$$

as required. ■

Part (b): Under the assumptions in Part (b), as shown for the denominator of (7.32),

$$B_3 - C' \left(\frac{1}{\theta^2} A_1 + A_3 \right)^{-1} C = \frac{1}{N} \sum_i \tilde{z}_i \tilde{z}_i' + o_p(1) \rightarrow_p l_z' \Xi l_z, \quad (7.35)$$

as $(N, T \rightarrow \infty)$. Next, consider the numerator of (7.32),

$$\begin{aligned} & \left[(B_4 + B_5) - C' \left(\frac{1}{\theta^2} A_1 + A_3 \right)^{-1} \left(\frac{1}{\theta^2} A_2 + (A_4 + A_5) \right) \right] \\ &= (B_4 + B_5) \\ & \quad - T \theta^2 (C' D_{x,T}) \frac{J_{x,T}^{-1}}{\sqrt{T}} \left(G_{x,T} A_1 G_{x,T} + \frac{J_{x,T}^{-1}}{\sqrt{T}} D_{x,T} A_3 D_{x,T} \frac{J_{x,T}^{-1}}{\sqrt{T}} \right)^{-1} \\ & \quad \times \left(\frac{1}{\sqrt{NT} \theta^2} \sqrt{NT} G_{x,T} A_2 + \frac{J_{x,T}^{-1}}{\sqrt{T}} D_{x,T} (A_4 + A_5) \right). \end{aligned}$$

By Lemmas 6 and 7, $B_5 = o_p(1)$, $\sqrt{NT} G_{x,T} A_2 = O_p(1)$, and $D_{x,T} A_5 = O_p(1)$. Under the fixed effect assumption (Assumption 11), Lemma 9 implies that $D_{x,T} A_4 = O_p(1)$, as $(N, T \rightarrow \infty)$. Since $\frac{1}{\sqrt{NT} \theta^2} = o(1)$ and $\frac{J_{x,T}^{-1}}{\sqrt{T}} = o(1)$, and

$$T \theta^2 (C' D_{x,T}) \frac{J_{x,T}^{-1}}{\sqrt{T}} \left(G_{x,T} A_1 G_{x,T} + \frac{J_{x,T}^{-1}}{\sqrt{T}} D_{x,T} A_3 D_{x,T} \frac{J_{x,T}^{-1}}{\sqrt{T}} \right)^{-1} = o_p(1)$$

(as shown in Part (a)), we have

$$(B_4 + B_5) - C' \left(\frac{1}{\theta^2} A_1 + A_3 \right)^{-1} \left(\frac{1}{\theta^2} A_2 + (A_4 + A_5) \right) = B_4 + o_p(1). \quad (7.36)$$

But, according to Lemma 9,

$$B_4 = \frac{1}{N} \sum_i \tilde{z}_i \tilde{u}_i \rightarrow_p l_z' \Xi \lambda. \quad (7.37)$$

Therefore, (7.32) and (7.35)–(7.37) imply

$$\hat{\gamma}_g \rightarrow_p \gamma + (l_z' \Xi l_z)^{-1} l_z' \Xi \lambda,$$

as $(N, T \rightarrow \infty)$. ■

Proof of Theorem 5

Using the notation in (7.25), we can express the Hausman test statistic by

$$\begin{aligned} \mathcal{H}\mathcal{M}_{NT} &= \left[(A_1 + \theta_T^2 F_1)^{-1} \sqrt{NT} (A_2 + \theta_T^2 F_2) - A_1^{-1} \sqrt{NT} A_2 \right]' \\ &\quad \times \left[\sigma_v^2 A_1^{-1} - \sigma_v^2 (A_1 + \theta_T^2 F_1)^{-1} \right]^{-1} \\ &\quad \times \left[(A_1 + \theta_T^2 F_1)^{-1} \sqrt{NT} (A_2 + \theta_T^2 F_2) - A_1^{-1} \sqrt{NT} A_2 \right]. \end{aligned}$$

Write

$$(A_1 + \theta_T^2 F_1)^{-1} - A_1^{-1} = -\theta_T^2 A_1^{-1} F_1 A_1^{-1} + \theta_T^4 R_1, \quad (7.38)$$

where $R_1 = (A_1 + \theta_T^2 F_1)^{-1} F_1 A_1^{-1} F_1 A_1^{-1}$. Define $\mathcal{Q} = (A_1 + \theta_T^2 F_1)^{-1} \sqrt{NT} (A_2 + \theta_T^2 F_2) - A_1^{-1} \sqrt{NT} A_2$. Then, we can deduce that

$$\mathcal{Q} = -\theta_T^2 \sqrt{NT} [A_1^{-1} F_1 A_1^{-1} A_2 - A_1^{-1} F_2] - \theta_T^4 \sqrt{NT} R_2, \quad (7.39)$$

where $R_2 = A_1^{-1} F_1 A_1^{-1} F_2 - R_1 \{A_2 + \theta_T^2 F_2\}$. Using (7.38) and (7.39), we now can rewrite the Hausman statistic $\mathcal{H}\mathcal{M}_{NT} = \mathcal{Q}' \left[\sigma_v^2 A_1^{-1} - \sigma_v^2 (A_1 + \theta_T^2 F_1)^{-1} \right]^{-1} \mathcal{Q}$ as

$$\begin{aligned} \mathcal{H}\mathcal{M}_{NT} &= \theta_T \sqrt{NT} \begin{bmatrix} G_1 (J_{x,T}^{-1} D_{x,T} F_1 D_{x,T} J_{x,T}^{-1}) G_1 G_{x,T} A_2 \\ -G_1 J_{x,T}^{-1} D_{x,T} F_2 + \theta_T^2 G_{x,T}^{-1} R_2 \end{bmatrix}' \\ &\quad \times \left[\sigma_v^2 G_1 (J_{x,T}^{-1} D_{x,T} F_1 D_{x,T} J_{x,T}^{-1}) G_1 + \sigma_v^2 \theta_T^2 (G_{x,T}^{-1} R_1 G_{x,T}^{-1}) \right]^{-1} \\ &\quad \times \theta_T \sqrt{NT} \begin{bmatrix} G_1 (J_{x,T}^{-1} D_{x,T} F_1 D_{x,T} J_{x,T}^{-1}) G_1 G_{x,T} A_2 \\ -G_1 J_{x,T}^{-1} D_{x,T} F_2 + \theta_T^2 G_{x,T}^{-1} R_2 \end{bmatrix}, \end{aligned}$$

where $G_1 = G_{x,T}^{-1} A_1^{-1} G_{x,T}^{-1}$. Using Lemmas 6(a), 7(a), and Assumption 13, we can show that

$$\sigma_v^2 \theta_T^2 (G_{x,T}^{-1} R_1 G_{x,T}^{-1}) = O_p(\theta_T^2) = o_p(1). \quad (7.40)$$

Also, under the local alternatives (Assumption 12), from Lemmas 6–8 we may deduce that

$$\theta_T \sqrt{NT} \theta_T^2 G_{x,T}^{-1} R_2 = O(1) [\theta_T^2 O_p(1) + o_p(1)] = o_p(1). \quad (7.41)$$

From (7.40) and (7.41), we now can approximate the Hausman statistic as follows:

$$\begin{aligned}
& \mathcal{H}\mathcal{M}_{NT} \\
&= \frac{\theta_T}{\sigma_v} \sqrt{NT} \left[(J_{x,T}^{-1} D_{x,T} F_1 D_{x,T} J_{x,T}^{-1}) (G_1 G_{x,T} A_2) - J_{x,T}^{-1} D_{x,T} F_2 + o_p(1) \right]' \\
&\quad \times (J_{x,T}^{-1} D_{x,T} F_1 D_{x,T} J_{x,T}^{-1} + o_p(1))^{-1} \\
&\quad \times \frac{\theta_T}{\sigma_v} \sqrt{NT} \left[(J_{x,T}^{-1} D_{x,T} F_1 D_{x,T} J_{x,T}^{-1}) (G_1 G_{x,T} A_2) - J_{x,T}^{-1} D_{x,T} F_2 + o_p(1) \right] \\
&= \frac{\theta_T}{\sigma_v} \sqrt{NT} (D_{x,T} F_2)' (D_{x,T} F_1 D_{x,T})^{-1} \frac{\theta_T}{\sigma_v} \sqrt{NT} (D_{x,T} F_2) \\
&\quad + o_p(1),
\end{aligned}$$

where the last line holds because under the local alternative hypotheses, $J_{x,T}^{-1} D_{x,T} F_1 D_{x,T} J_{x,T}^{-1} = O_p(1)$, and $\frac{\theta_T}{\sigma_v} G_1 (\sqrt{NT} G_{x,T} A_2) = O_p(\theta_T) = o_p(1)$ by Lemma 6(a), (b) and Assumption 13. Finally, by Lemma 7(a), as $(N, T \rightarrow \infty)$, $D_{x,T} F_1 D_{x,T} = D_{x,T} A_3 D_{x,T} - D_{x,T} C B_3^{-1} C' D_{x,T} \rightarrow_p \Xi_{xx} - \Xi_{xz} \Xi_{zz}^{-1} \Xi_{zx} > 0$. Also, Lemmas 7(c) and 8 imply that under the local alternative hypotheses, as $(N, T \rightarrow \infty)$,

$$\begin{aligned}
& \sqrt{N} D_{x,T} F_2 \\
&= \sqrt{N} D_{x,T} A_4 + \sqrt{N} D_{x,T} A_5 - D_{x,T} C B_3^{-1} \sqrt{N} (B_4 + B_5) \\
&= \frac{1}{\sqrt{N}} \sum_{i=1}^N D_{x,T} \tilde{x}_i \tilde{u}_i \\
&\quad - \left(\frac{1}{N} \sum_{i=1}^N D_{x,T} \tilde{x}_i \tilde{z}_i' \right) \left(\frac{1}{N} \sum_{i=1}^N \tilde{z}_i \tilde{z}_i' \right)^{-1} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N \tilde{z}_i \tilde{u}_i \right) + o_p(1) \\
&\Rightarrow (I_k - \Xi_{xz} \Xi_{zz}^{-1}) N (\Xi \lambda, \sigma_u^2 \Xi).
\end{aligned}$$

As $(N, T \rightarrow \infty)$, $\frac{\theta_T}{\sigma_v} \sqrt{T} \rightarrow \frac{1}{\sigma_u}$. Therefore, under the hypothesis of random effects, $\mathcal{H}\mathcal{M}_{NT} \Rightarrow \chi_k^2$, a χ^2 distribution with the degrees of freedom equal to k . In contrast, under the local alternative hypotheses, $\mathcal{H}\mathcal{M}_{NT} \Rightarrow \chi_k^2(\eta)$, where η is the noncentral parameter. ■

Acknowledgements We would like to thank two referees for valuable comments and suggestions and Geert Ridder for helpful discussions. We also appreciate the comments of seminar participants at Arizona State University, the University of British Columbia, and the University of California, Davis.

The first author gratefully acknowledges the financial support of the College of Business and Dean's Council of 100 at Arizona State University, the Economic Club of Phoenix, and the alumni of the College of Business.

The second author gratefully acknowledges the financial supports of USC via FDA and of Yonsei University.

References

- Ahn SC, Low S (1996) A reformulation of the Hausman test for regression models with cross-section-time-series data. *J Econom* 71:291–307
- Ahn SC, Moon HR (2001) Large-N and large-T properties of panel data estimators and the Hausman test. CLEO discussion papers, C01-20, University of Southern California
- Alvarez J, Arellano M (2003) The time series and cross-section asymptotics of dynamic panel data estimators. *Econometrica* 71:1121–1159
- Amemiya T, MaCurdy TE (1986) Instrumental-variables estimation of an error components model. *Econometrica* 54:869–880
- Apostol T (1974) *Mathematical analysis*, 2nd edn. Addison-Welsey, Reading
- Balestra P, Nerlove M (1966) Pooling cross-section and time-series data in the estimation of a dynamic model: the demand for natural gas. *Econometrica* 34:585–612
- Baltagi B (1995) *Econometric analysis of panel data*. Wiley, New York
- Baltagi B, Khanti-Akom S (1990) On efficient estimation with panel data: an empirical comparison of instrumental variables estimators. *J Appl Econom* 5:401–406
- Breusch TS, Mizon GE, Schmidt P (1989) Efficient estimation using panel data. *Econometrica* 57:695–700
- Choi I (1998) Asymptotic analysis of a nonstationary error component model. Mimeo, Kookmin University, Korea
- Cornwell C, Rupert P (1988) Efficient estimation with panel data: an empirical comparison of instrumental variables estimators. *J Appl Econom* 3:149–155
- Guggenberger P (2010) The impact of a Hausman pretest on the asymptotic size of a hypothesis test. *Econom Theory* 26:369–382
- Hahn J, Kuersteiner G (2002) Asymptotically unbiased inference for a dynamic panel model with fixed effects when both n and T are large. *Econometrica* 70:1639–1657
- Hausman JA (1978) Specification tests in econometrics. *Econometrica* 46:1251–1272
- Hausman JA, Taylor WE (1981) Panel data and unobservable individual effects. *Econometrica* 49:1377–1398
- Higgins M, Zakrajsk E (1999) Purchasing power parity: three stakes through the heart of the unit root null. NBER working paper, Federal Reserve Bank of New York, New York
- Hsiao C (1986) *Analysis of panel data*. Cambridge University Press, New York
- Im KS, Pesaran MH, Shin Y (2003) Testing for unit roots in heterogeneous panels. *J Econom* 115:53–74
- Kao C (1999) Spurious regression and residual-based tests for cointegration in panel data. *J Econom* 90:1–44
- Levin A, Lin C-F (1992) Unit root tests in panel data: asymptotic and finite-sample properties. Mimeo, University of California, San Diego
- Levin A, Lin C-F (1993) Unit root tests in panel data: new results. Mimeo, University of California, San Diego
- Maddala GS (1971) The use of variance components models in pooling cross section and time series data. *Econometrica* 39:341–358
- Mátyás L, Sevestre P (1992) *The econometrics of panel data*. Kluwer Academic, Dordrecht
- Phillips PCB, Moon H (1999) Linear regression limit theory for nonstationary panel data. *Econometrica* 67:1057–1111
- Quah D (1994) Exploiting cross-section variations for unit root inference in dynamic data. *Econ Lett* 44:9–19