Estimation of Time-invariant Effects in Static Panel Data Models*

M. Hashem Pesaran†
University of Southern California, and Trinity College, Cambridge

Qiankun Zhou
University of Southern California

September 5, 2014

Abstract

This paper proposes the Fixed Effects Filtered (FEF) and Fixed Effects Filtered instrumental variable (FEF-IV) estimators for estimation and inference in the case of time-invariant effects in static panel data models when N is large and T is fixed. It is shown that the FEF and FEF-IV estimators are $\sqrt{N}$-consistent, and asymptotically normally distributed. The FEF estimator is compared with the Fixed Effects Vector Decomposition (FEVD) estimator proposed by Plumper and Troeger (2007) and conditions under which the two estimators are equivalent are established. It is also shown that the variance estimator proposed for FEVD estimator is inconsistent and its use could lead to misleading inference. Alternative variance estimators are proposed for both FEF and FEF-IV estimators which are shown to be consistent under fairly general conditions. The small sample properties of the FEF and FEF-IV estimators are investigated by Monte Carlo experiments, and it is shown that FEF has smaller bias and RMSE, unless an intercept is included in the second stage of the FEVD procedure which renders the FEF and FEVD estimators identical. The FEVD procedure, however, results in substantial size distortions since it uses incorrect standard errors. We also compare the FEF-IV estimator with the estimator proposed by Hausman and Taylor (1981), when one of the time-invariant regressors is correlated with the fixed effects. Both FEF and FEF-IV estimators are shown to be robust to error variance heteroskedasticity and residual serial correlation.

Keywords: static panel data models, time-invariant effects, Fixed Effects Filtered estimator, Fixed Effects Filtered instrumental variables estimator

JEL classification: C01, C23, C33

*We would like to thank Alex Chudik and Ron Smith for helpful suggestions.

†Corresponding Author: Department of Economics, University of Southern California, 3620 South Vermont Avenue, Kaprielian Hall 300, Los Angeles, CA 90089, USA, pesaran@usc.edu

1 Introduction

Identification and estimation of the effects of time-invariant regressors, such as the effects of race or gender is often the focus of panel data analysis, yet estimation procedures such as fixed effects (FE), that yield consistent estimates of the coefficients of time-varying regressors under fairly general conditions, cannot be used for estimation of the time-invariant effects, since the FE transformation eliminates all time-invariant regressors. As a result estimation of time-invariant effects has posed a challenge in panel data econometrics - namely how to carry out inference on time-invariant effects without making strong assumptions on the correlation between unobserved individual effects and the time-varying regressors.

For the estimation of time-invariant effects, Plumper and Troeger (2007) (PT) propose the so called Fixed Effects Vector Decomposition (FEVD) through a three-step procedure.[1] As we shall see, whilst the FEVD approach can be modified to yield consistent estimates of the time-invariant effects, the variance estimator proposed by PT for their estimator is not consistent. PT do not provide any formal statistical proofs to support their stated claims about the consistency of their estimator and its variance estimator. See Greene (2011a).

In the case where one or more of the time-invariant regressors are endogenous, an early pioneering contribution by Hausman and Taylor (1981) (HT) propose using instrumental variables in the context of a pooled random coefficient panel data model. The instruments are obtained by assuming that known sub-sets of time-varying and time-invariant regressors are exogenous. HT also assumed that individual-specific effects as well as the idiosyncratic errors of the panel data model under consideration are serially uncorrelated and homoskedastic. Some of these assumptions are relaxed in the subsequent literature, but the main idea that sub-sets of time-varying and time invariant regressors are exogenous is typically maintained. See also Amemiya and MaCurdy (1986), Breusch et al. (1989), Im et al. (1999) and Baltagi and Bresson (2012).

In this paper, we consider a general static panel data model, which allows for an arbitrary degree of correlation between the time-varying covariates and the individual effects, and propose the fixed-effects filtered (FEF) estimation for the coefficients of the time-invariant regressors when the cross-sectional observation, N, is large and the time-series dimension, T, is small and fixed. Our proposed estimator has two simple steps. In the first step FE estimates are computed for the coefficients of the time-varying variables, and these estimates are used to filter out the time-varying effects. The residuals from the first stage panel regression are then averaged over time and used as a dependent variable in a cross-section OLS regression that includes an intercept and the vector of time-invariant regressors. Under the identifying assumption that the time-invariant regressors are uncorrelated with the individual effects and a number of other regularity conditions, it is shown that the FEF estimator is unbiased and consistent for a finite T and as $N \to \infty$. We derive the asymptotic distribution of the FEF estimator and propose a non-parametric estimator of its

[1] Plumper and Troeger (2007)'s FEVD approach is very popular in political science, and there even is a STATA procedure for the implementation of the FEVD estimator.

covariance matrix, not known in the literature, which we show to be consistent in the presence of heteroskedasticity of the individual effects and performs well in the presence of residual serial correlation.

Finally, we consider the case when one or more of the time-invariant variables are endogenous, and develop the FEF-IV estimator assuming there exist valid instruments. It is shown that the FEF-IV estimator is consistent and asymptotically normally distributed. A feasible variance estimator is also proposed for this FEF-IV estimator, which works well under heteroskedasticity and residual serial correlation. The main advantage of the proposed FEF-IV over the HT estimator lies in the fact that it does not require a sub-set of time-varying regressors to be exogenous, whilst at the same can use time averages of the time-varying regressors as instruments if it is known that such time averages are uncorrelated with the fixed effects, and at the same time correlated with the endogenous time-invariant regressors.[2] The second advantage of the FEF-IV estimator of time-invariant effects is its robustness to residual serial correlation and error heteroskedasticity.

We also contribute to the controversy over the FEVD estimator proposed by PT, discussed by Greene (2011a) and Breusch et al. (2011b), and followed up with responses and rejoinders by Plumper and Troeger (2011), Greene (2011b), Breusch et al. (2011a), and Beck (2011). The FEVD estimator of PT is based on a three step procedure, we show that when an intercept is included in the second step of their procedure, then the FEVD estimator is identical to the FEF estimator. But if an intercept is not included in the second stage, the FEVD estimator is in general biased and inconsistent. The extent of the bias of the FEVD estimator depends on the magnitude of intercept and the mean of time-invariant variables. What is more important is that, even if an intercept is included in the second step of the FEVD procedure, inferences based on the FEVD estimators and their variances in the third step of PT's estimation procedure could be misleading since the variance of the FEVD estimator obtained in the third step is biased and most likely will result in over-rejection of the null. This is confirmed by the Monte Carlo simulations.

The small sample properties of the FEF and FEF-IV estimators for static panel data model are investigated, using two sets of comprehensive Monte Carlo experiments including error variance heteroskedasticity and residual serial correlation. In one set we generate the time-invariant regressors as exogenous, whilst in the second set we allow one of the time-invariant regressors to be correlated with the fixed effects. In both sets of experiments we allow the time-varying regressors to be correlated with the fixed effects. We compare FEF and FEVD estimators using the first set of experiments only, since these procedures are not appropriate in the case of the second set of experiments where one of the time-invariant regressors is endogenous. We find that our proposed estimator has smaller bias and RMSE, unless an intercept is included in the second stage of the

[2] It is important to note that the assumption of a zero correlation between time averages of time-varying regressors and the fixed effects is less restrictive than the assumption of zero correlation between time-varying regressors and the fixed effects. The latter implies the former but not the reverse. For example, suppose that the time-varying regressor, x_{it}, is related to the fixed effects, α_i, according to $x_{it} = \alpha_i g_t + w_{it}$, for $t = 1, 2, ..., T$, where w_{it} is distributed independently of α_i and $\sum_{t=1}^{T} g_t = 0$. In this setting $Cov(x_{it}, \alpha_i) = g_t Var(\alpha_i) \neq 0$ for each t, but $Cov(\bar{x}_i, \alpha_i) = 0$, where $\bar{x}_i = T^{-1} \sum_{t=1}^{T} x_{it}$.

FEVD procedure which renders the FEF and FEVD estimators identical. However, as predicted by our theoretical derivations, the FEVD procedure results in substantial size distortions since it uses incorrect standard errors. In contrast, the use of the standard errors derived in this paper yields the correct size and satisfactory power in the case of all experiments, illustrating the robustness of our variance formula to heteroskedasticity and residual serial correlation. We also compare the FEF-IV estimator with the HT estimator using the second set of experiments where one of the time-invariant regressors is correlated with the fixed effects. The FEF-IV procedure performs well and has the correct size when an instrument is used for the endogenous time-invariant regressor. It is also robust to error variance heteroskedasticity and residual serial correlation. But a straightforward application of the HT procedure results in biased estimates and size distortions since it incorrectly assumes that one of the time-varying regressors is uncorrelated with the fixed effects. In such cases the HT procedure must be modified so that none of the time-varying regressors are used as instruments.

The rest of the paper is organized as follows: Section 2 sets out the panel data model with time-invariant effects. Section 3 develops the FEF estimator, derives its asymptotic distribution, gives robust variance matrix estimator for the proposed FEF estimator, and provides a comparison of the FEF and FEVD estimators. Section 4 considers the FEF-IV estimator in the case where one or more of the time-invariant regressors are correlated with the errors. Section 5 discusses the HT estimator and derives its covariance matrix under a general specification of the error covariance matrix. The small sample properties of the FEF and FEF-IV estimators are then investigated in Section 6. The paper ends with some concluding remarks in Section 7. Some of the detailed mathematical proofs are provided in the Appendix.

For any real-valued $N \times N$ matrix $\mathbf{A}$, we will use $\|\mathbf{A}\|$ to denote the Frobenius norm of matrix $\mathbf{A}$ defined as $\|\mathbf{A}\| = [tr(\mathbf{A}\mathbf{A}')]^{1/2}$. Throughout, K denotes a generic non-zero positive constant that does not depend on N. The symbols $\rightarrow_p$ and $\rightarrow_d$ are used to denote convergence in probability and in distribution, respectively.

2 Panel data models with time-invariant effects

Consider the following panel data model that contains time-varying as well as time-invariant variables:

$$y_{it} = \alpha_i + \mathbf{z}_i'\boldsymbol{\gamma} + \mathbf{x}_{it}'\boldsymbol{\beta} + \varepsilon_{it}, \quad i = 1, 2, \ldots, N; t = 1, 2, \ldots, T, \tag{1}$$

where

$$\alpha_i = \alpha + \eta_i, \tag{2}$$

$\mathbf{x}_{it}$ is a $k \times 1$ vector of time-varying variables, and $\mathbf{z}_i$ is an $m \times 1$ vector of observed individual-specific variables that only vary over the cross section units, i. In addition to $\mathbf{z}_i$, the outcomes, y_{it}, are also governed by unobserved individual specific effects, α_i. The focus of the analysis is on estimation and inference involving the elements of $\boldsymbol{\gamma}$. It is clear that without further restrictions on

α_i, $\boldsymbol{\gamma}$ cannot be identified even if $\boldsymbol{\beta}$ was known to the researcher. For example consider the simple case where $\boldsymbol{\beta}=\mathbf{0}$, and assume that T is small. Then averaging across t we obtain

$$\bar{y}_i=\alpha+\mathbf{z}_i'\boldsymbol{\gamma}+v_i$$

where $\bar{y}_i=T^{-1}\Sigma_{t=1}^{T}y_{it}$, $v_i=\eta_i+\bar{\varepsilon}_i$, and $\bar{\varepsilon}_i=T^{-1}\Sigma_{t=1}^{T}\varepsilon_{it}$. It is clear that without specifying how v_i and $\mathbf{z}_i$ are related it will not be possible to identify the effects of $\mathbf{z}_i$. To deal with this problem, it is often assumed that there exists instruments that are uncorrelated with v_i but at the same time are sufficiently correlated with $\mathbf{z}_i$. Even if such instruments exist a number of further complications arises if $\boldsymbol{\beta}\neq\mathbf{0}$. In such a case the IV approach must be extended also to deal with the possible dependencies between η_i and $\mathbf{x}_{it}$. In what follows we allow for η_i and $\mathbf{x}_{it}$ to have any degree of dependence, but initially assume that $\mathbf{z}_i$ and v_i are uncorrelated for identification of $\boldsymbol{\gamma}$, and assume that $\mathbf{x}_{it}$ and ε_{is} are uncorrelated for all i,t and s, to identify $\boldsymbol{\beta}$. This approach can be modified in cases where one or more instruments are available for $\mathbf{z}_i$ and/or $\mathbf{x}_{it}$.

3 Fixed effects filtered (FEF) estimator of time-invariant effects

3.1 FEF estimator

Under the assumption that $\mathbf{x}_{it}$ and $\boldsymbol{\varepsilon}_{is}$ are uncorrelated for all i,t and s, as it is well known $\boldsymbol{\beta}$ can be estimated consistently under fairly general assumptions on temporal dependence and cross-sectional heteroskedasticity of ε_{it}, and the distribution of the fixed effects, α_i. Denoting the FE estimator of $\boldsymbol{\beta}$ by $\hat{\boldsymbol{\beta}}$, $\boldsymbol{\gamma}$ can then be estimated by the regression of $\bar{y}_i-\hat{\boldsymbol{\beta}}'\bar{\mathbf{x}}_i$ on an intercept and $\mathbf{z}_i$. We denote this estimator by $\hat{\boldsymbol{\gamma}}_{FEF}$ and refer to it as the fixed effects filtered (FEF) estimator of $\boldsymbol{\gamma}$. Formally, the FEF estimator can be computed using the following two-step procedure:

Step 1: Using model (1), compute the fixed-effects estimator of $\boldsymbol{\beta}$, denoted by $\hat{\boldsymbol{\beta}}$, and the associated residuals $\hat{u}_{it}$ defined by

$$\hat{u}_{it}=y_{it}-\hat{\boldsymbol{\beta}}'\mathbf{x}_{it}. \tag{3}$$

Step 2: Compute the time averages of these residuals, $\bar{\hat{u}}_i=T^{-1}\sum_{t=1}^{T}\hat{u}_{it}$, and regress $\bar{\hat{u}}_i$ on $\mathbf{z}_i$ with an intercept to obtain $\hat{\boldsymbol{\gamma}}_{FEF}$, namely

$$\hat{\boldsymbol{\gamma}}_{FEF}=\left[\sum_{i=1}^{N}\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'\right]^{-1}\sum_{i=1}^{N}\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)\left(\bar{\hat{u}}_i-\bar{\hat{u}}\right), \tag{4}$$

and

$$\hat{\alpha}_{FEF}=\bar{\hat{u}}-\hat{\boldsymbol{\gamma}}_{FEF}'\bar{\mathbf{z}}, \tag{5}$$

where $\bar{\hat{u}}=N^{-1}\sum_{i=1}^{N}\bar{\hat{u}}_i$.

The use of the FE residuals, $\hat{u}_{it}$, for consistent estimation of $\boldsymbol{\gamma}$ is not new and has been used in the literature extensively starting with the pioneering contribution of Hausman and Taylor (1981). The

FEVD procedure proposed by Plumper and Troeger (2007) also makes use of the FE residuals. (see Section 3.4). The main contribution of this paper lies in development of the asymptotic distribution of $\hat{\boldsymbol{\gamma}}_{FEF}$ (and its IV version, $\hat{\boldsymbol{\gamma}}_{FEF-IV}$ introduced in Section 4) under fairly general conditions on the error processes ε_{it}, and η_i, and alternative assumptions concerning the correlation of $\mathbf{z}_i$ and $\eta_i + \bar{\varepsilon}_i$. We also derive conditions under which the covariance matrix of $\hat{\boldsymbol{\gamma}}_{FEF}$ (and $\hat{\boldsymbol{\gamma}}_{FEF-IV}$) can be consistently estimated.

3.2 Asymptotic Properties of the FEF Estimator of $\boldsymbol{\gamma}$

We examine the asymptotic properties of the FEF estimator of $\boldsymbol{\gamma}$, $\hat{\boldsymbol{\gamma}}_{FEF}$, defined by (4), under the following assumptions:

Assumption P1: $E(\varepsilon_{it} \left| \mathbf{x}_{is} \right.) = 0$, for all i, t and s, and $E\left(\varepsilon_{it}^4\right) < K < \infty$, for all i and t.

Assumption P2: $E\left(\varepsilon_{it}\varepsilon_{js} \left| \mathbf{X} \right.\right) = 0$, for all $i \neq j$, and all t and s, where $\mathbf{X} = (\mathbf{x}_{it};\, i = 1, 2, ..., N;\ t = 1, 2, ..., T)$.

Assumption P3: The errors, ε_{it}, are heteroskedastic and temporally dependent, namely

$$E\left(\varepsilon_{it}\varepsilon_{is} \left| \mathbf{X} \right.\right) = \gamma_i(t,s), \text{ for all } t \text{ and } s,$$

where $0 < \gamma_i(t,t) = \sigma_i^2$, and $|\gamma_i(t,s)| < K$, for all i, t and s.

Assumption P4 : The regressors, $\mathbf{x}_{it}$, have either bounded supports, namely $\|\mathbf{x}_{it}\| < K < \infty$, or satisfy the moment conditions $E\left\|\mathbf{x}_{it} - \bar{\mathbf{x}}\right\|^4 < K < \infty$, and $E\left\|\bar{\mathbf{x}}_i - \bar{\mathbf{x}}\right\|^4 < K < \infty$, for all i and t, where $\bar{\mathbf{x}}_i = T^{-1}\sum_{t=1}^{T}\mathbf{x}_{it}$, and $\bar{\mathbf{x}} = N^{-1}\sum_{i=1}^{N}\bar{\mathbf{x}}_i$.

Assumption P5: The $k \times k$ matrices $\mathbf{Q}_{p,NT}$ and $\mathbf{Q}_{FE,NT}$ defined by

$$\mathbf{Q}_{p,NT} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\mathbf{x}_{it} - \bar{\mathbf{x}}\right)\left(\mathbf{x}_{it} - \bar{\mathbf{x}}\right)', \tag{6}$$

$$\mathbf{Q}_{FE,NT} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\mathbf{x}_{it} - \bar{\mathbf{x}}_i\right)\left(\mathbf{x}_{it} - \bar{\mathbf{x}}_i\right)', \tag{7}$$

converge (in probability) to $\mathbf{Q}_{p,T}$ and $\mathbf{Q}_{FE,T}$ for a fixed T and as N tends to infinity, $\lambda_{\min}\left(\mathbf{Q}_{FE,NT}\right) > 1/K$ and $\lambda_{\min}\left(\mathbf{Q}_{p,NT}\right) > 1/K$, for all N and T where K is a finite, non-zero constant.

Assumption P6: The $m \times m$ matrix, $\mathbf{Q}_{zz,N}$, and the $m \times k$ matrix $\mathbf{Q}_{z\bar{x},N}$ defined by

$$\mathbf{Q}_{zz,N} = \frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{z}_i - \bar{\mathbf{z}}\right)\left(\mathbf{z}_i - \bar{\mathbf{z}}\right)', \tag{8}$$

$$\mathbf{Q}_{z\bar{x},N} = \frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{z}_i - \bar{\mathbf{z}}\right)\left(\bar{\mathbf{x}}_i - \bar{\mathbf{x}}\right)', \tag{9}$$

converge (in probability) to the non-stochastic limits $\mathbf{Q}_{zz}$ and $\mathbf{Q}_{z\bar{x}}$, and $\lambda_{\min}\left(\mathbf{Q}_{zz,N}\right) > 1/K$, for all $N > m$.

Assumption P7: The time-invariant regressors, $\mathbf{z}_i$, are independently distributed of $v_j = \eta_j + \bar{\varepsilon}_j$, for all i and j, and η_i and $\bar{\varepsilon}_i$ are independently distributed. Also, $\mathbf{z}_i$ either have bounded support, namely $\|\mathbf{z}_i\| < K$, or satisfy the moment conditions $E\left\|(\mathbf{z}_i - \bar{\mathbf{z}})\right\|^4 < K$, for all i.

Remark 1 *Note that since*

$$\|\mathbf{x}_{it} - \bar{\mathbf{x}}_i\| = \|\mathbf{x}_{it} - \bar{\mathbf{x}} - (\bar{\mathbf{x}}_i - \bar{\mathbf{x}})\| \leq \|\mathbf{x}_{it} - \bar{\mathbf{x}}\| + \|\bar{\mathbf{x}}_i - \bar{\mathbf{x}}\|,$$

then any order moment conditions on $\|\bar{\mathbf{x}}_i - \bar{\mathbf{x}}\|$ and $\|\bar{\mathbf{x}}_i - \bar{\mathbf{x}}\|$ imply the same order moment conditions on $\|\mathbf{x}_{it} - \bar{\mathbf{x}}_i\|$. The boundedness of $\|\mathbf{x}_{it} - \bar{\mathbf{x}}\|$ and $\|\bar{\mathbf{x}}_i - \bar{\mathbf{x}}\|$ are also sufficient for the boundedness of $\|\mathbf{x}_{it} - \bar{\mathbf{x}}_i\|$.

Remark 2 *Assumptions P5 and P6 ensure that there exists a finite N_0 such that for all $N > N_0$, $\mathbf{Q}_{zz,N}$ and $\mathbf{Q}_{FE,NT}$ are positive definite and converge in probability to the fixed matrices $\mathbf{Q}_{zz}$ and $\mathbf{Q}_{FE}$, respectively. But using the results in lemma A.1 in the Appendix, one can then relax the conditions $\lambda_{\min}(\mathbf{Q}_{zz,N}) > 1/K$ and $\lambda_{\min}(\mathbf{Q}_{FE,NT}) > 1/K$ by requiring $\lambda_{\min}(\mathbf{Q}_{zz}) > 2/K$ and $\lambda_{\min}(\mathbf{Q}_{FE}) > 2/K$. Under our assumptions the latter conditions ensure that the former conditions hold with probability approaching one.*

Remark 3 *Although, our focus is on fixed T and N large panels, we shall also discuss conditions under which our analysis will be valid when both T and N are large.*

To derive the asymptotic distribution of $\hat{\boldsymbol{\gamma}}_{FEF}$, we first note that the FE estimator of $\boldsymbol{\beta}$ is given by

$$\hat{\boldsymbol{\beta}} = \left[\sum_{t=1}^{T}\sum_{i=1}^{N}(\mathbf{x}_{it} - \bar{\mathbf{x}}_i)(\mathbf{x}_{it} - \bar{\mathbf{x}}_i)'\right]^{-1}\sum_{t=1}^{T}\sum_{i=1}^{N}(\mathbf{x}_{it} - \bar{\mathbf{x}}_i)(y_{it} - \bar{y}_i). \tag{10}$$

Under the above assumptions, $\hat{\boldsymbol{\beta}}$ is unbiased and consistent for any fixed T and as $N \to \infty$, and

$$Var\left(\hat{\boldsymbol{\beta}}\,|\mathbf{X}\right) = \frac{1}{NT}\mathbf{Q}_{FE,NT}^{-1}\mathbf{V}_{FE,NT}\mathbf{Q}_{FE,NT}^{-1}, \tag{11}$$

where

$$\mathbf{V}_{FE,NT} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\sigma_i^2(\mathbf{x}_{it} - \bar{\mathbf{x}}_i)(\mathbf{x}_{it} - \bar{\mathbf{x}}_i)' + \frac{1}{NT}\sum_{i=1}^{N}\sum_{t\neq s}^{T}\gamma_i(t,s)(\mathbf{x}_{it} - \bar{\mathbf{x}}_i)(\mathbf{x}_{is} - \bar{\mathbf{x}}_i)'. \tag{12}$$

In the standard case where $\varepsilon_{it} \sim IID(0, \sigma^2)$, we obtain the more familiar expression $Var\left(\hat{\boldsymbol{\beta}}\,|\mathbf{X}\right) = (NT)^{-1}\sigma^2\mathbf{Q}_{FE,NT}^{-1}$. Also, for a fixed T and as $N \to \infty$, we have the following limiting distribution

$$\sqrt{N}\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) \to_d N(\mathbf{0}, T^{-1}\boldsymbol{\Omega}_{\hat{\beta}}), \tag{13}$$

where

$$\boldsymbol{\Omega}_{\hat{\beta}} = \mathbf{Q}_{FE,T}^{-1}\mathbf{V}_{FE,T}\mathbf{Q}_{FE,T}^{-1}, \tag{14}$$

and $\mathbf{Q}_{FE,T}$ is defined in Assumption P5, and $\mathbf{V}_{FE,T} = p\lim_{N\to\infty}(\mathbf{V}_{FE,NT})$.

Consider now the FEF estimator of $\boldsymbol{\gamma}$ defined by (4) and note that

$$\bar{\hat{u}}_i - \bar{\hat{u}} = (\eta_i - \bar{\eta}) + (\bar{\varepsilon}_i - \bar{\varepsilon}) + (\mathbf{z}_i - \bar{\mathbf{z}})' \boldsymbol{\gamma} - (\bar{\mathbf{x}}_i - \bar{\mathbf{x}})' \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right).$$

Using this result in (4) we now have (noting that $N^{-1}\sum_{i=1}^{N}(\mathbf{z}_i - \bar{\mathbf{z}})(\bar{\varepsilon} + \bar{\eta}) = \mathbf{0}$)

$$\hat{\boldsymbol{\gamma}}_{FEF} - \boldsymbol{\gamma} = \mathbf{Q}_{zz,N}^{-1}\left[N^{-1}\sum_{i=1}^{N}(\mathbf{z}_i - \bar{\mathbf{z}})\zeta_i\right], \tag{15}$$

where $\mathbf{Q}_{zz,N}$ is defined by (8) and

$$\zeta_i = \eta_i + \bar{\varepsilon}_i - (\bar{\mathbf{x}}_i - \bar{\mathbf{x}})' \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right). \tag{16}$$

Let $\mathbf{Z} = (\mathbf{z}_1, ..., \mathbf{z}_N)'$, $\mathbf{X} = (\mathbf{x}_{it}; i = 1, 2, ..., N; t = 1, 2, ..., T)$, and $\boldsymbol{\eta} = (\eta_1, \eta_2, ..., \eta_N)'$, and note that

$$\begin{aligned} E\left[(\mathbf{z}_i - \bar{\mathbf{z}})\zeta_i \,|\mathbf{Z}, \mathbf{X}, \boldsymbol{\eta}\right] &= (\mathbf{z}_i - \bar{\mathbf{z}})\eta_i - (\mathbf{z}_i - \bar{\mathbf{z}})(\bar{\mathbf{x}}_i - \bar{\mathbf{x}})' \left[E\left(\hat{\boldsymbol{\beta}}\,|\mathbf{X}\right) - \boldsymbol{\beta}\right] \\ &= (\mathbf{z}_i - \bar{\mathbf{z}})\eta_i. \end{aligned}$$

Also under Assumption P7 we have $E\left[(\mathbf{z}_i - \bar{\mathbf{z}})\eta_i\right] = \mathbf{0}$ for all i, and using (15) it follows that $E(\hat{\boldsymbol{\gamma}}_{FEF}) = \boldsymbol{\gamma}$, which establishes that $\hat{\boldsymbol{\gamma}}_{FEF}$ is an unbiased estimator of $\boldsymbol{\gamma}$.

Consider now the consistency and the asymptotic distribution of $\hat{\boldsymbol{\gamma}}_{FEF}$. To this end we first note that

$$\begin{aligned} N^{-1}\sum_{i=1}^{N}(\mathbf{z}_i - \bar{\mathbf{z}})\zeta_i &= N^{-1}\sum_{i=1}^{N}(\mathbf{z}_i - \bar{\mathbf{z}})(\eta_i + \bar{\varepsilon}_i) \\ &\quad - \left[N^{-1}\sum_{i=1}^{N}(\mathbf{z}_i - \bar{\mathbf{z}})(\bar{\mathbf{x}}_i - \bar{\mathbf{x}})'\right]\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right). \end{aligned}$$

Also under Assumptions P6 and P7, $N^{-1}\sum_{i=1}^{N}(\mathbf{z}_i - \bar{\mathbf{z}})(\bar{\mathbf{x}}_i - \bar{\mathbf{x}})' \to_p \mathbf{Q}_{z\bar{x}}$,

$$N^{-1/2}\sum_{i=1}^{N}(\mathbf{z}_i - \bar{\mathbf{z}})(\eta_i + \bar{\varepsilon}_i) \to_d N(\mathbf{0}, \omega_{iT}^2 \mathbf{Q}_{zz}), \tag{17}$$

where

$$\omega_{iT}^2 = \sigma_\eta^2 + \frac{\sigma_i^2}{T} + \frac{1}{T^2}\sum_{s\neq t}\gamma_i(s,t) \tag{18}$$

with $\gamma_i(s,t) = E(\varepsilon_{is}\varepsilon_{it})$, and

$$N^{-1}\sum_{i=1}^{N}(\mathbf{z}_i - \bar{\mathbf{z}})\zeta_i = O_p\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) + O_p\left(N^{-1/2}\right).$$

Now using (15), and since $\mathbf{Q}_{zz,N} \rightarrow_p \mathbf{Q}_{zz}$, which is a non-singular matrix, then we also have

$$\hat{\boldsymbol{\gamma}}_{FEF} - \boldsymbol{\gamma} = O_p\left(N^{-1/2}\right) + O_p\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right). \tag{19}$$

Therefore, in view of (19) we obtain

$$\hat{\boldsymbol{\gamma}}_{FEF} - \boldsymbol{\gamma} = O_p\left(N^{-1/2}\right),$$

which establishes that $\hat{\boldsymbol{\gamma}}_{FEF}$, is a $\sqrt{N}$ consistent estimator of $\boldsymbol{\gamma}$.

To derive the asymptotic distribution of $\hat{\boldsymbol{\gamma}}_{FEF}$, we first note that

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\mathbf{z}_i - \bar{\mathbf{z}})\zeta_i = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\mathbf{z}_i - \bar{\mathbf{z}})\eta_i + \frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\mathbf{z}_i - \bar{\mathbf{z}})\left[\bar{\varepsilon}_i - (\bar{\mathbf{x}}_i - \bar{\mathbf{x}})'\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)\right], \tag{20}$$

and consider the limiting distribution of the two terms of (20) and their covariance. We first note that the second term of the above can be written as

$$\begin{aligned}
\bar{\varepsilon}_i - (\bar{\mathbf{x}}_i - \bar{\mathbf{x}})'\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) &= \frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it} - \frac{1}{T}\sum_{t=1}^{T}(\bar{\mathbf{x}}_i - \bar{\mathbf{x}})'\mathbf{Q}_{FE,NT}^{-1}\frac{1}{NT}\sum_{t=1}^{T}\sum_{j=1}^{N}(\mathbf{x}_{jt} - \bar{\mathbf{x}}_j)\varepsilon_{jt} \\
&= \frac{1}{T}\sum_{t=1}^{T}\left(\varepsilon_{it} - \frac{1}{N}\sum_{j=1}^{N}w_{ij,t}\varepsilon_{jt}\right),
\end{aligned}$$

where[3]

$$w_{ij,t} = (\bar{\mathbf{x}}_i - \bar{\mathbf{x}})'\mathbf{Q}_{FE,NT}^{-1}(\mathbf{x}_{jt} - \bar{\mathbf{x}}_j). \tag{21}$$

Hence,

$$\begin{aligned}
&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\mathbf{z}_i - \bar{\mathbf{z}})\left[\bar{\varepsilon}_i - (\bar{\mathbf{x}}_i - \bar{\mathbf{x}})'\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)\right] \\
&= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\mathbf{z}_i - \bar{\mathbf{z}})\frac{1}{T}\sum_{t=1}^{T}\left(\varepsilon_{it} - \frac{1}{N}\sum_{j=1}^{N}w_{ij,t}\varepsilon_{jt}\right) \\
&= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\bar{\boldsymbol{\xi}}_{i,N},
\end{aligned}$$

[3]Note that $w_{ij,t} \neq w_{ji,t}$.

where

$$\bar{\boldsymbol{\xi}}_{i,N} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{d}_{z,it}\varepsilon_{it}, \tag{22}$$

and

$$\mathbf{d}_{z,it} = (\mathbf{z}_i - \bar{\mathbf{z}}) - \frac{1}{N}\sum_{j=1}^{N}(\mathbf{z}_j - \bar{\mathbf{z}})\, w_{ji,t}. \tag{23}$$

Using these results in (20) now yield

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\mathbf{z}_i - \bar{\mathbf{z}})\,\zeta_i = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\mathbf{z}_i - \bar{\mathbf{z}})\,\eta_i + \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\bar{\boldsymbol{\xi}}_{i,N}. \tag{24}$$

However,

$$\begin{aligned} & Cov\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\mathbf{z}_i - \bar{\mathbf{z}})\,\eta_i, \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\bar{\boldsymbol{\xi}}_{i,N}\right) \\ = & \frac{1}{N}\sum_{i,j}E\left[(\mathbf{z}_i - \bar{\mathbf{z}})\,\eta_i\bar{\boldsymbol{\xi}}_{j,N}\right] = \frac{1}{N}\sum_{i,j}E\left\{E\left[\eta_i\,(\mathbf{z}_i - \bar{\mathbf{z}})\,\bar{\boldsymbol{\xi}}'_{j,N}|\mathbf{Z},\mathbf{X},\boldsymbol{\eta}\right]\right\} = \mathbf{0}, \end{aligned}$$

and it is sufficient to derive the asymptotic distributions of the two terms in (24), separately. To this end we note that under Assumptions P6 and P7, and using standard central limit theorems it follows that

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\mathbf{z}_i - \bar{\mathbf{z}})\,\eta_i \rightarrow_d N\left(\mathbf{0}, \sigma_\eta^2\mathbf{Q}_{zz}\right). \tag{25}$$

Consider now the second term in (24) and note that under Assumption P1-P3 and P7, $w_{ji,t}$ and $\mathbf{z}_i$ are distributed independently of ε_{is}, for all i, j, t, and s, and hence conditional on $\mathbf{Z}$ and $\mathbf{X}$, $\bar{\boldsymbol{\xi}}_{i,N}$ have zero means, and are cross sectionally independently distributed (noting that by Assumption P2, ε_{it} are assumed to be cross-sectionally independent). But since the terms, $\bar{\boldsymbol{\xi}}_{i,N}$, in (24) vary with N it suffices to show that the following Liapunov condition, (see Davidson (1994), p. 373) is satisfied.

$$\lim_{N\to\infty}\sum_{i=1}^{N}E\left\|N^{-1/2}\bar{\boldsymbol{\xi}}_{i,N}\right\|^{2+\delta} = 0 \text{ for some } \delta > 0. \tag{26}$$

The validity of this condition is established under Assumptions P1-P7 in Section A.1 of the Appendix. Hence

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\bar{\boldsymbol{\xi}}_{i,N} = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(T^{-1}\sum_{t=1}^{T}\mathbf{d}_{z,it}\varepsilon_{it}\right) \rightarrow_d N\left(\mathbf{0}, \boldsymbol{\Omega}_{\bar{\xi}}\right) \tag{27}$$

where $\boldsymbol{\Omega}_{\bar{\xi}} = \lim_{N\to\infty}\boldsymbol{\Omega}_{\bar{\xi},N}$, and (since ε_{it} are assumed to be cross-sectionally independent)

$$\boldsymbol{\Omega}_{\bar{\xi},N} = N^{-1}\sum_{i=1}^{N}\left[T^{-2}\sum_{t,s=1}^{T}\mathbf{d}_{z,it}\mathbf{d}'_{z,is}E\left(\varepsilon_{it}\varepsilon_{is}\right)\right]. \tag{28}$$

The above results are summarized in the following theorem:

Theorem 1 *Consider the FEF estimator* $\hat{\boldsymbol{\gamma}}_{FEF}$ *of* $\boldsymbol{\gamma}$ *in the panel data model (1) defined by (4), and suppose that Assumptions P1-P7 hold. Then* $\hat{\boldsymbol{\gamma}}_{FEF}$ *is an unbiased and a consistent estimator of* $\boldsymbol{\gamma}$*, and*

$$\sqrt{N}\left(\hat{\boldsymbol{\gamma}}_{FEF}-\boldsymbol{\gamma}\right)\rightarrow_{d} N\left(\mathbf{0},\boldsymbol{\Omega}_{\hat{\gamma}_{FEF}}\right), \tag{29}$$

where

$$\boldsymbol{\Omega}_{\hat{\gamma}_{FEF}}=\mathbf{Q}_{zz}^{-1}\left(\sigma_{\eta}^{2}\mathbf{Q}_{zz}+\boldsymbol{\Omega}_{\bar{\xi}}\right)\mathbf{Q}_{zz}^{-1}. \tag{30}$$

$\mathbf{Q}_{zz}$ *is defined in Assumption P6,* $\boldsymbol{\Omega}_{\bar{\xi}}=\lim_{N\rightarrow\infty}\boldsymbol{\Omega}_{\bar{\xi},N}$*, with* $\boldsymbol{\Omega}_{\bar{\xi},N}$ *is defined by (28), and* σ_{η}^{2} *is the variance of the fixed effects defined by (2).*

3.3 Consistent estimation of $Var\left(\hat{\boldsymbol{\gamma}}_{FEF}\right)$

In order to estimate $\boldsymbol{\Omega}_{\hat{\gamma}_{FEF}}$, it is helpful to begin with the following proposition regarding $\Omega_{\bar{\xi}}$, defined by (28), which enters the expression for $\boldsymbol{\Omega}_{\hat{\gamma}_{FEF}}$.

Proposition 1 *Let*

$$\mathbf{V}_{zz,N}=\frac{1}{N}\sum_{i=1}^{N}\omega_{iT}^{2}\left(\mathbf{z}_{i}-\bar{\mathbf{z}}\right)\left(\mathbf{z}_{i}-\bar{\mathbf{z}}\right)', \tag{31}$$

where $\omega_{iT}^{2}=\sigma_{\eta}^{2}+\kappa_{iT}^{2}$*,*

$$\kappa_{iT}^{2}=\frac{\sigma_{i}^{2}}{T}+\frac{1}{T^{2}}\sum_{s\neq t}\gamma_{i}\left(s,t\right), \tag{32}$$

$\gamma_{i}\left(t,s\right)=E\left(\varepsilon_{it}\varepsilon_{is}\right)$*, and* $\sigma_{i}^{2}=\gamma_{i}\left(t,t\right)$*. Then* $\sigma_{\eta}^{2}\mathbf{Q}_{zz}+\Omega_{\bar{\xi},N}$*, with* $\Omega_{\bar{\xi},N}$ *defined by (28), can be written as*

$$\sigma_{\eta}^{2}\mathbf{Q}_{zz}+\Omega_{\bar{\xi},N}=\mathbf{V}_{zz,N}+\mathbf{Q}_{z\bar{x},N}Var\left(\sqrt{N}\hat{\boldsymbol{\beta}}\right)\mathbf{Q}_{z\bar{x},N}'-\left(\boldsymbol{\Delta}_{\bar{\xi},N}+\boldsymbol{\Delta}_{\bar{\xi},N}'\right), \tag{33}$$

where $\mathbf{Q}_{z\bar{x},N}$ *is defined in (9), and*

$$\boldsymbol{\Delta}_{\bar{\xi},N}=\mathbf{Q}_{z\bar{x},N}\mathbf{Q}_{FE,NT}^{-1}\left\{\frac{1}{N}\sum_{i=1}^{N}\left[T^{-2}\sum_{t,s=1}^{T}\gamma_{i}(t,s)\left(\mathbf{x}_{it}-\bar{\mathbf{x}}_{i}\right)\left(\mathbf{z}_{i}-\bar{\mathbf{z}}\right)'\right]\right\}. \tag{34}$$

Proof. A proof is provided in Section A.2 in the Appendix. ■

Proposition 2 *Under Assumptions P1-P7, and if*

$$\frac{1}{NT^{2}}\sum_{i=1}^{N}\sum_{t,s=1}^{T}\gamma_{i}(t,s)\left(\mathbf{x}_{it}-\bar{\mathbf{x}}_{i}\right)\left(\mathbf{z}_{i}-\bar{\mathbf{z}}\right)'=o_{p}\left(1\right), \tag{35}$$

then the variance of the FEF estimator (4), can be consistently estimated for a fixed T and as $N \to \infty$, by

$$\hat{\mathbf{\Omega}}_{\hat{\gamma}_{FEF}} = N\ \widehat{Var}\left(\hat{\boldsymbol{\gamma}}_{FEF}\right) = \mathbf{Q}_{zz,N}^{-1}\left[\hat{\mathbf{V}}_{zz,N} + \mathbf{Q}_{z\bar{x},N}\left(N\widehat{Var}(\hat{\boldsymbol{\beta}})\right)\mathbf{Q}_{z\bar{x},N}'\right]\mathbf{Q}_{zz,N}^{-1}, \tag{36}$$

where $\mathbf{Q}_{zz,N}$ and $\mathbf{Q}_{z\bar{x},N}$ are defined by (8) and (9), respectively,

$$\widehat{Var}(\hat{\boldsymbol{\beta}}) = \left(\sum_{i=1}^{N}\mathbf{x}_{i\cdot}'\mathbf{x}_{i\cdot}\right)^{-1}\left(\sum_{i=1}^{N}\mathbf{x}_{i\cdot}'\mathbf{e}_i\mathbf{e}_i'\mathbf{x}_{i\cdot}\right)\left(\sum_{i=1}^{N}\mathbf{x}_{i\cdot}'\mathbf{x}_{i\cdot}\right)^{-1}, \tag{37}$$

where $\mathbf{x}_{i\cdot}' = (\mathbf{x}_{i1}-\bar{\mathbf{x}}_i, \mathbf{x}_{i2}-\bar{\mathbf{x}}_i, \ldots, \mathbf{x}_{iT}-\bar{\mathbf{x}}_i)$ denotes the demeaned vector of $\mathbf{x}_{it}$, and the t-th element of $\mathbf{e}_i$ is given by $e_{it} = y_{it} - \bar{y}_i - (\mathbf{x}_{it}-\bar{\mathbf{x}}_i)'\hat{\boldsymbol{\beta}}$, and

$$\hat{\mathbf{V}}_{zz,N} = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{\varsigma}_i - \bar{\hat{\varsigma}}\right)^2(\mathbf{z}_i-\bar{\mathbf{z}})(\mathbf{z}_i-\bar{\mathbf{z}})', \tag{38}$$

where

$$\hat{\varsigma}_i - \bar{\hat{\varsigma}} = \bar{y}_i - \bar{y} - (\bar{\mathbf{x}}_i - \bar{\mathbf{x}})'\hat{\boldsymbol{\beta}} - (\mathbf{z}_i-\bar{\mathbf{z}})'\hat{\boldsymbol{\gamma}}_{FEF}. \tag{39}$$

Proof. A proof is provided in Section A.3 of the Appendix. ■

Condition (35) is not as restrictive as it may appear at first, and holds under a number of still fairly general assumptions regarding the error processes, ε_{it}. To see this, first not that

$$\frac{1}{NT^2}\sum_{t,s=1}^{T}\sum_{i=1}^{N}\gamma_i(t,s)\left(\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right)\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)' = T^{-1}\left[N^{-1}\sum_{i=1}^{N}\left(\frac{\mathbf{X}_i'\mathbf{M}_T\mathbf{\Gamma}_i\boldsymbol{\tau}_T\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'}{T}\right)\right],$$

where $\mathbf{X}_i = (\mathbf{x}_{i1},\ldots,\mathbf{x}_{iT})'$, $\mathbf{\Gamma}_i = (\gamma_i(t,s))$, and $\boldsymbol{\tau}_T$ is a $T\times 1$ vector of ones. Also, $T^{-1}\mathbf{\Gamma}_i\boldsymbol{\tau}_T = (\bar{\gamma}_{i1},\bar{\gamma}_{i2},...,\bar{\gamma}_{iT})'$, where $\bar{\gamma}_{it} = T^{-1}\sum_{s=1}^{T}\gamma_i(t,s)$. Then condition (35) is met exactly if $\bar{\gamma}_{it} = c_i$ for all t. Since in such a case $T^{-1}\mathbf{\Gamma}_i\boldsymbol{\tau}_T = c_i\boldsymbol{\tau}_T$, and $T^{-1}\mathbf{X}_i'\mathbf{M}_T\mathbf{\Gamma}_i\boldsymbol{\tau}_T\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)' = c_i\mathbf{X}_i'\mathbf{M}_T\boldsymbol{\tau}_T\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)' = 0$. Condition $\bar{\gamma}_{it} = c_i$ is clearly met if $\gamma_i(t,s) = 0$ for all $t \neq s$, and $\gamma_i(t,t) = E(\varepsilon_{it}^2) = \sigma_i^2$. After extensive simulations including cases where there are significant variations over time in $\bar{\gamma}_{it}$, we find that the effect of $\mathbf{\Delta}_{\bar{\xi},N}$ is negligible and the use of (36) for inference seems to be justified more generally.[4] Furthermore, the quality of approximating the variance of $\hat{\boldsymbol{\gamma}}_{FEF}$ by (36) tends to improve with T so long as $T^{-1}\sum_{t,s=1}^{T}|\gamma_i(t,s)| < K$.

3.4 Comparison of FEF and FEVD estimators

In this section, we will compare the FEF estimator with the FEVD proposed by Plumper and Troeger (2007). The FEVD procedure is based on the following three steps:

Step 1: The fixed effects approach is applied to (1), to compute the FE residuals, $\hat{u}_{it}$, defined by (3).

[4] The results of these simulations are available upon request.

Step 2: In the second step, PT regress $\overline{\hat{u}}_i$ on $\mathbf{z}_i$ where $\overline{\hat{u}}_i = T^{-1}\sum_{t=1}^{T}\hat{u}_{it} = \bar{y}_i - \bar{\mathbf{x}}_i'\hat{\boldsymbol{\beta}}$, $\bar{y}_i = T^{-1}\sum_{t=1}^{T} y_{it}$ and $\bar{\mathbf{x}}_i = T^{-1}\sum_{t=1}^{T}\mathbf{x}_{it}$. To obtain equivalence between the FEVD and FEF estimators of $\boldsymbol{\gamma}$, we modify this regression by also including an intercept in the regression and hence define the residuals from the second stage by

$$\hat{h}_i = \overline{\hat{u}}_i - \hat{a} - \mathbf{z}_i'\hat{\boldsymbol{\gamma}}, \tag{40}$$

where $\hat{a} = \overline{\hat{u}} - \bar{\mathbf{z}}'\hat{\gamma}$, $\overline{\hat{u}} = N^{-1}\sum_{i=1}^{N}\overline{\hat{u}}_i$, and

$$\hat{\boldsymbol{\gamma}} = \left[\sum_{i=1}^{N}(\mathbf{z}_i - \bar{\mathbf{z}})(\mathbf{z}_i - \bar{\mathbf{z}})'\right]^{-1}\sum_{i=1}^{N}(\mathbf{z}_i - \bar{\mathbf{z}})\left(\overline{\hat{u}}_i - \overline{\hat{u}}\right). \tag{41}$$

which is exactly the same as our FEF estimator. Using the above results we now have

$$\begin{aligned}\hat{h}_i - \overline{\hat{h}} &= \left(\overline{\hat{u}}_i - \overline{\hat{u}}\right) - (\mathbf{z}_i - \bar{\mathbf{z}})'\hat{\boldsymbol{\gamma}} \\ &= (\bar{y}_i - \bar{y}) - (\bar{\mathbf{x}}_i - \bar{\mathbf{x}})'\hat{\boldsymbol{\beta}} - (\mathbf{z}_i - \bar{\mathbf{z}})'\hat{\boldsymbol{\gamma}},\end{aligned} \tag{42}$$

where $\overline{\hat{h}} = N^{-1}\sum_{i=1}^{N}\hat{h}_i$. Also, from the normal equations of this step, note that

$$\overline{\hat{h}} = 0,\ N^{-1}\sum_{i=1}^{N}\left(\hat{h}_i - \overline{\hat{h}}\right)(\mathbf{z}_i - \bar{\mathbf{z}})' = \mathbf{0}. \tag{43}$$

Step 3: The third step uses $\hat{h}_i$ computed in the earlier stage, as defined by (40), and estimates the following panel regression by pooled OLS

$$y_{it} = a + \mathbf{x}_{it}'\boldsymbol{\beta} + \mathbf{z}_i'\boldsymbol{\gamma} + \delta\hat{h}_i + \tilde{\varepsilon}_{it}. \tag{44}$$

These estimators are the modified FEVD estimators which we shall denote by $\tilde{\boldsymbol{\gamma}}$, $\tilde{\boldsymbol{\delta}}$ and $\tilde{\boldsymbol{\beta}}$, and as before we denote the FE estimator of $\boldsymbol{\beta}$ by $\hat{\boldsymbol{\beta}}$, and the estimator of $\boldsymbol{\gamma}$ obtained in the second step of FEVD approach by $\hat{\boldsymbol{\gamma}}$ (which is identical to the FEF estimator if an intercept is included in the second step). The original FEVD estimators proposed by PT are based on the same pooled OLS regression, but do not include an intercept in the second stage regression that computes the $\hat{h}_i$.[5] As we shall see this makes a great deal of difference to the resultant estimators.

To investigate the relationship between FEF and FEVD estimators we first introduce the fol-

[5] For example, equations (4) and (5) on p128 of Plumper and Troeger (2007).

lowing notations

$$\mathbf{q}_{p,NT} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\mathbf{x}_{it}-\bar{\mathbf{x}}\right)\left(y_{it}-\bar{y}\right),\ \mathbf{Q}_{z\bar{y},N} = \frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{z}_{i}-\bar{\mathbf{z}}\right)\left(\bar{y}_{i}-\bar{y}\right), \tag{45}$$

$$\mathbf{Q}_{h\bar{x},N} = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{h}_i-\bar{\hat{h}}\right)\left(\bar{\mathbf{x}}_{i}-\bar{\mathbf{x}}\right)',\ \mathbf{Q}_{hz,N} = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{h}_i-\bar{\hat{h}}\right)\left(\mathbf{z}_{i}-\bar{\mathbf{z}}\right)', \tag{46}$$

$$\mathbf{Q}_{h\bar{y},N} = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{h}_i-\bar{\hat{h}}\right)\left(\bar{y}_{i}-\bar{y}\right),\ \mathbf{Q}_{hh,N} = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{h}_i-\bar{\hat{h}}\right)^2, \tag{47}$$

where $\bar{\hat{h}}$, $\bar{\mathbf{x}}$, $\bar{\mathbf{x}}_i$, $\bar{\mathbf{z}}$, $\bar{y}$, and $\bar{y}_i$ are defined as before. Using these additional notations, the normal equations of the pooled OLS regressions for the panel data model defined by (44) are given by

$$\begin{aligned}
\mathbf{q}_{p,NT} &= \mathbf{Q}_{p,NT}\tilde{\boldsymbol{\beta}} + \mathbf{Q}'_{z\bar{x},N}\tilde{\boldsymbol{\gamma}} + \mathbf{Q}'_{h\bar{x},N}\tilde{\delta},\\
\mathbf{Q}_{z\bar{y},N} &= \mathbf{Q}_{z\bar{x},N}\tilde{\boldsymbol{\beta}} + \mathbf{Q}_{zz,N}\tilde{\boldsymbol{\gamma}} + \mathbf{Q}'_{hz,N}\tilde{\delta},\\
\mathbf{Q}_{h\bar{y},N} &= \mathbf{Q}_{h\bar{x},N}\tilde{\boldsymbol{\beta}} + \mathbf{Q}_{hz,N}\tilde{\boldsymbol{\gamma}} + \mathbf{Q}_{hh,N}\tilde{\delta}.
\end{aligned}$$

Also, when an intercept is included in the second step of FEVD we have $\mathbf{Q}_{hz,N} = 0 = \mathbf{Q}'_{hz,N}$, (see (43) and (46)), and the normal equations reduce to

$$\mathbf{q}_{p,NT} = \mathbf{Q}_{p,NT}\tilde{\boldsymbol{\beta}} + \mathbf{Q}'_{z\bar{x},N}\tilde{\boldsymbol{\gamma}} + \mathbf{Q}'_{h\bar{x},N}\tilde{\delta}, \tag{48}$$

$$\mathbf{Q}_{z\bar{y},N} = \mathbf{Q}_{z\bar{x},N}\tilde{\boldsymbol{\beta}} + \mathbf{Q}_{zz,N}\tilde{\boldsymbol{\gamma}}, \tag{49}$$

$$\mathbf{Q}_{h\bar{y},N} = \mathbf{Q}_{h\bar{x},N}\tilde{\boldsymbol{\beta}} + \mathbf{Q}_{hh,N}\tilde{\delta}. \tag{50}$$

The FEVD estimator of $\boldsymbol{\gamma}$, namely $\tilde{\boldsymbol{\gamma}}$, can now be obtained using the above system of the equations. The results are summarized in the following proposition.

Proposition 3 *Consider the panel data model (44), and suppose that $\mathbf{Q}_{p,NT}$ and $\mathbf{Q}_{zz,N}$ are non-singular, and $\mathbf{Q}_{hh,N} > 0$. Let*

$$\mathbf{Q}_{NT} = \mathbf{Q}_{p,NT} - \mathbf{Q}'_{z\bar{x},N}\mathbf{Q}^{-1}_{zz,N}\mathbf{Q}_{z\bar{x},N} - \mathbf{Q}'_{h\bar{x},N}\mathbf{Q}^{-1}_{hh,N}\mathbf{Q}_{h\bar{x},N}, \tag{51}$$

and suppose also that $\mathbf{Q}_{NT}$ is non-singular. Then the FEVD estimators proposed by PT and FEF estimators proposed in this paper are identical if an intercept is included in the second step regression of the FEVD procedure, namely $\tilde{\boldsymbol{\gamma}} = \hat{\boldsymbol{\gamma}}$, and $\tilde{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}$. Furthermore $\tilde{\delta}$, the FEVD estimator of δ in the third step of the FEVD procedure, is identically equal to unity.

Proof. See Section A.4 in the Appendix for a proof. ■

Proposition 4 *Suppose that the three-step FEVD estimators (denoted as before by $\tilde{\boldsymbol{\beta}}$, $\tilde{\gamma}$ and $\tilde{\boldsymbol{\delta}}$) are computed without including an intercept in the regression in the second step. In this case we* *continue to have $\tilde{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}$ and $\tilde{\delta} = 1$, but for $\boldsymbol{\gamma}$ we obtain $\tilde{\boldsymbol{\gamma}} = \mathring{\boldsymbol{\gamma}}$, where $\mathring{\boldsymbol{\gamma}}$ is the OLS estimator of the coefficient of $\mathbf{z}_i$ in the OLS regression of $\bar{\hat{u}}_i$ on $\mathbf{z}_i$, without an intercept, and $\mathring{\boldsymbol{\gamma}}$ is biased and inconsistent unless $\alpha E\left[\left(N^{-1}\sum_{i=1}^{N}\mathbf{z}_i\mathbf{z}_i'\right)^{-1}N^{-1}\sum_{i=1}^{N}\mathbf{z}_i\right] = \mathbf{0}$.*

Proof. See Section A.5 in the Appendix for a proof. ■

It is also interesting to compare the covariance of the FEF given by (30), with the one that is obtained when the standard formula for the variance of the pooled OLS estimators is applied to the third step of the FEVD procedure as proposed by PT. Recall that the FEVD estimator of $\boldsymbol{\gamma}$ coincides with $\hat{\boldsymbol{\gamma}}$ if an intercept is included in the second step of the procedure, and pooled OLS applied to the third step will result in a valid inference only if the variance obtained using the FEVD procedure also coincides with $\boldsymbol{\Omega}_{\hat{\gamma}_{FEF}}$. To simplify the comparisons suppose that $\varepsilon_{it} \sim IID(0, \sigma^2)$ for all i and t, and note that in this simple case $\mathbf{V}_{zz}$ and $\boldsymbol{\Omega}_{\hat{\beta}}$ (given by (31) and (14)) reduce to

$$\begin{aligned}
\mathbf{V}_{zz} &= \left(\sigma_\eta^2 + \frac{\sigma^2}{T}\right)\mathbf{Q}_{zz}, \\
\boldsymbol{\Omega}_{\hat{\beta}} &= \frac{\sigma^2}{T}\left[\lim_{N\to\infty}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\mathbf{x}_{it} - \bar{\mathbf{x}}_i)(\mathbf{x}_{it} - \bar{\mathbf{x}}_i)'\right]^{-1} = \frac{\sigma^2}{T}\left(\mathbf{Q}_p - \mathbf{Q}_{\bar{x}\bar{x}}\right)^{-1},
\end{aligned}$$

and we have[6]

$$\begin{aligned}
\boldsymbol{\Omega}_{\hat{\gamma}_{FEF}} &= \mathbf{Q}_{zz}^{-1}\left[\left(\sigma_\eta^2 + \frac{\sigma^2}{T}\right)\mathbf{Q}_{zz} + \mathbf{Q}_{z\bar{x}}\Omega_{\hat{\beta}}\mathbf{Q}_{\bar{x}z}\right]\mathbf{Q}_{zz}^{-1} \\
&= \left(\sigma_\eta^2 + \frac{\sigma^2}{T}\right)\mathbf{Q}_{zz}^{-1} + \frac{\sigma^2}{T}\mathbf{Q}_{zz}^{-1}\mathbf{Q}_{z\bar{x}}\left(\mathbf{Q}_p - \mathbf{Q}_{\bar{x}\bar{x}}\right)^{-1}\mathbf{Q}_{\bar{x}z}\mathbf{Q}_{zz}^{-1}.
\end{aligned}$$

Under the same model specifications the covariance of the FEVD estimator (also scaled by $\sqrt{N}$) is given by

$$\boldsymbol{\Omega}_{\hat{\gamma}_{FEVD}} = \sigma_{\tilde{\varepsilon}}^2\left[\mathbf{Q}_{zz}^{-1} + \mathbf{Q}_{zz}^{-1}\mathbf{Q}_{z\bar{x}}\mathbf{Q}^{-1}\mathbf{Q}_{z\bar{x}}'\mathbf{Q}_{zz}^{-1}\right],$$

where as before $\mathbf{Q} = \mathbf{Q}_p - \mathbf{Q}_{z\bar{x}}'\mathbf{Q}_{zz}^{-1}\mathbf{Q}_{z\bar{x}} - \mathbf{Q}_{h\bar{x}}'\mathbf{Q}_{hh}^{-1}\mathbf{Q}_{h\bar{x}}$,

$$\begin{aligned}
\sigma_{\tilde{\varepsilon}}^2 &= \lim_{N\to\infty}\left[N^{-1}T^{-1}\sum_{t=1}^{T}\sum_{i=1}^{N}\left(y_{it} - \tilde{\alpha} - \mathbf{x}_{it}'\tilde{\boldsymbol{\beta}} - \mathbf{z}_i'\tilde{\boldsymbol{\gamma}} - \tilde{\delta}\hat{h}_i\right)^2\right] \\
&= \lim_{N\to\infty}\left[N^{-1}T^{-1}\sum_{t=1}^{T}\sum_{i=1}^{N}\hat{\tilde{\varepsilon}}_{it}^2\right],
\end{aligned}$$

and

$$\hat{\tilde{\varepsilon}}_{it} = y_{it} - \bar{y} - (\mathbf{x}_{it} - \bar{\mathbf{x}})'\hat{\boldsymbol{\beta}} - (\mathbf{z}_i - \bar{\mathbf{z}})'\hat{\boldsymbol{\gamma}} - \hat{h}_i.$$

[6]Note that since in the present case $\varepsilon_{it}'s$ are serially uncorrelated then $\boldsymbol{\Delta}_{\bar{\xi}} = 0$.

To derive σ_ε^2 note that $\hat{h}_i = \bar{y}_i - \bar{y} - (\bar{\mathbf{x}}_i - \bar{\mathbf{x}})'\hat{\boldsymbol{\beta}} - (\mathbf{z}_i - \bar{\mathbf{z}})'\hat{\boldsymbol{\gamma}}$, and hence

$$\begin{aligned}
\widehat{\bar{\varepsilon}}_{it} &= y_{it} - \bar{y} - (\mathbf{x}_{it} - \bar{\mathbf{x}})' \hat{\boldsymbol{\beta}} - (\mathbf{z}_i - \bar{\mathbf{z}})' \hat{\boldsymbol{\gamma}} - \hat{h}_i \\
&= y_{it} - \bar{y}_i - (\mathbf{x}_{it} - \bar{\mathbf{x}}_i)' \hat{\boldsymbol{\beta}} \\
&= \varepsilon_{it} - \bar{\varepsilon}_i - (\mathbf{x}_{it} - \bar{\mathbf{x}}_i)' \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right).
\end{aligned}$$

Also, in the case where $\varepsilon_{it} \sim IID(0, \sigma^2)$, we have $Cov\left[\varepsilon_{it} - \bar{\varepsilon}_i, (\mathbf{x}_{it} - \bar{\mathbf{x}}_i)' \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)\right] = \mathbf{0}$, and hence

$$E\left(\widehat{\bar{\varepsilon}}_{it}^2\right) = \sigma^2 \left(1 - \frac{1}{T} + \frac{k}{N}\right).$$

A comparison of the expressions derived above for $\boldsymbol{\Omega}_{\hat{\gamma}_{FEF}}$ and $\boldsymbol{\Omega}_{\hat{\gamma}_{FEVD}}$ clearly shows that they differ irrespective of whether T is fixed or $T \to \infty$.

Remark 4 *Plumper and Troeger (2007) argue that the necessity of third step is to correct standard errors of $\mathring{\boldsymbol{\gamma}}$ (Plumper and Troeger (2007) , p129)*, *however, as shown above and in the simulations below, the variance term calculated in the third step of FEVD does not fully correct the bias of the variance estimator in the second step.*[7]

4 FEF-IV estimation of time-invariant effects

The FEF procedure assumes that the time-invariant regressors, $\mathbf{z}_i$, are distributed independently of the individual-specific effects $\eta_i + \bar{\varepsilon}_i$. However, it is relatively straight forward to modify the FEF estimator to allow for possibly endogeneity of the time-invariant regressors, if there exists a sufficient number of valid instruments. In particular, it is possible to derive an IV version of FEF, which we denote by FEF-IV, under the following assumptions:

Assumption P8: There exists the $s \times 1$ vector of instruments $\mathbf{r}_i$ for $\mathbf{z}_i$, $i = 1, 2, \ldots, N$, where $\mathbf{r}_i$ is distributed independently of η_j and $\bar{\varepsilon}_j$ for all i and j, and $s \geq m$, and satisfy the moment condition $E\left\|\mathbf{r}_i - \bar{\mathbf{r}}\right\|^4 < K < \infty$, if it has unbounded support.

Assumption P9: Let $\mathbf{Z} = (\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N)'$, $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)'$. Consider the $s \times m$ matrix $\mathbf{Q}_{rz,N}$, the $s \times k$ matrix $\mathbf{Q}_{rx,N} = N^{-1}\sum_{i=1}^N (\mathbf{r}_i - \bar{\mathbf{r}})(\bar{\mathbf{x}}_i - \bar{\mathbf{x}})'$, and the $s \times s$ matrix $\mathbf{Q}_{rr,N}$ defined by

$$\mathbf{Q}_{rz,N} = N^{-1}\sum_{i=1}^N (\mathbf{r}_i - \bar{\mathbf{r}})(\mathbf{z}_i - \bar{\mathbf{z}})',\ \mathbf{Q}_{r\bar{x},N} = N^{-1}\sum_{i=1}^N (\mathbf{r}_i - \bar{\mathbf{r}})(\bar{\mathbf{x}}_i - \bar{\mathbf{x}})',\ \ \mathbf{Q}_{rr,N} = N^{-1}\sum_{i=1}^N (\mathbf{r}_i - \bar{\mathbf{r}})(\mathbf{r}_i - \bar{\mathbf{r}})', \tag{52}$$

where $\bar{\mathbf{r}} = N^{-1}\sum_{i=1}^N \mathbf{r}_i$. $\mathbf{Q}_{rz,N}$ and $\mathbf{Q}_{rr,N}$ are full rank matrices for all $N > r$, and have finite probability limits as $N \to \infty$ given by $\mathbf{Q}_{rz}$ and $\mathbf{Q}_{rr}$, respectively. Matrices $\mathbf{Q}_{r\bar{x},N}$ and $\mathbf{Q}_{zz,N}$ have finite probability limits given by $\mathbf{Q}_{r\bar{x}}$ and $\mathbf{Q}_{zz}$, respectively, and in cases where $\mathbf{x}_{it}$ and $\mathbf{z}_i$ are

[7] PT state "...only the third stage allows obtaining the correct SE's.", P.129.

stochastic with unbounded supports, then $\lambda_{\min}(\mathbf{Q}_{rr,N}) > 1/K$, for all N, and as $N \rightarrow \infty$, with probability approaching one.

Under the above assumptions and maintaining Assumptions P1-P6 as before, a consistent two-stage estimator of $\boldsymbol{\gamma}$ can be obtained as follows[8]

$$\hat{\boldsymbol{\gamma}}_{FEF-IV} = \left(\mathbf{Q}_{zr,N}\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{zr,N}'\right)^{-1}\left(\mathbf{Q}_{zr,N}\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{r\hat{u},N}\right), \tag{53}$$

where $\mathbf{Q}_{zr,N}$ and $\mathbf{Q}_{rr,N}$ are defined by (52),

$$\mathbf{Q}_{r\hat{u},N} = \frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{r}_i - \bar{\mathbf{r}}\right)\left(\bar{\hat{u}}_i - \bar{\hat{u}}\right),$$

$\bar{\hat{u}} = \frac{1}{N}\sum_{i=1}^{N}\bar{\hat{u}}_i$, $\bar{\hat{u}}_i = \bar{y}_i - \left(\bar{\mathbf{x}}_i - \bar{\mathbf{x}}\right)'\hat{\boldsymbol{\beta}}$, and $\hat{\boldsymbol{\beta}}$ is the FE estimator of $\boldsymbol{\beta}$ from the first stage. It then follows that

$$\sqrt{N}\left(\hat{\boldsymbol{\gamma}}_{FEF-IV} - \boldsymbol{\gamma}\right) = \left(\mathbf{Q}_{zr,N}\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{zr,N}'\right)^{-1}\left(\mathbf{Q}_{zr,N}\mathbf{Q}_{rr,N}^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\mathbf{r}_i - \bar{\mathbf{r}}\right)\zeta_i\right),$$

with $\zeta_i = \eta_i + \bar{\varepsilon}_i - \bar{\mathbf{x}}_i'\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)$, as before (see (16)). Following a similar line of proof as in the case with exogenous $\mathbf{z}_i$, under Assumptions P1-P6 and P8-P9, it can be shown that

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\mathbf{r}_i - \bar{\mathbf{r}}\right)\left(\zeta_i - \bar{\zeta}\right) \rightarrow_d N\left(\mathbf{0}, \sigma_\eta^2\mathbf{Q}_{rr} + \boldsymbol{\Omega}_{\bar{\psi}}\right),$$

where

$$\boldsymbol{\Omega}_{\bar{\psi}} = \lim_{N\rightarrow\infty}\frac{1}{N}\sum_{i=1}^{N}\left[T^{-2}\sum_{t,s=1}^{T}\mathbf{d}_{r,it}\mathbf{d}_{r,is}'E\left(\varepsilon_{it}\varepsilon_{is}\right)\right], \tag{54}$$

where $\mathbf{d}_{r,it} = \left(\mathbf{r}_i - \bar{\mathbf{r}}\right) - \frac{1}{N}\sum_{j=1}^{N}\left(\mathbf{r}_j - \bar{\mathbf{r}}\right)w_{ji,t}$, and $w_{ij,t} = \left(\bar{\mathbf{x}}_i - \bar{\mathbf{x}}\right)'\mathbf{Q}_{FE,NT}^{-1}\left(\mathbf{x}_{jt} - \bar{\mathbf{x}}_j\right)$, as before. Moreover, we note that under Assumption P9

$$\mathbf{Q}_{zr,N}\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{zr,N}' \rightarrow_p \mathbf{Q}_{zr}\mathbf{Q}_{rr}^{-1}\mathbf{Q}_{zr}', \text{ and } \mathbf{Q}_{zr,N}\mathbf{Q}_{rr,N}^{-1} \rightarrow_p \mathbf{Q}_{zr}\mathbf{Q}_{rr}^{-1}.$$

Using the above results and Slutsky's theorem now yields

$$\sqrt{N}\left(\hat{\boldsymbol{\gamma}}_{FEF-IV} - \boldsymbol{\gamma}\right) \rightarrow_d N\left(\mathbf{0}, \boldsymbol{\Omega}_{\hat{\gamma}_{FEF-IV}}\right),$$

where

$$\boldsymbol{\Omega}_{\hat{\gamma}_{FEF-IV}} = \left(\mathbf{Q}_{zr}\mathbf{Q}_{rr}^{-1}\mathbf{Q}_{zr}'\right)^{-1}\mathbf{Q}_{zr}\mathbf{Q}_{rr}^{-1}\left(\sigma_\eta^2\mathbf{Q}_{rr} + \Omega_{\bar{\psi}}\right)\mathbf{Q}_{rr}^{-1}\mathbf{Q}_{zr}'\left(\mathbf{Q}_{zr}\mathbf{Q}_{rr}^{-1}\mathbf{Q}_{zr}'\right)^{-1}. \tag{55}$$

[8] A derivation is available upon request.

For convenience, the above results are summarized in the following theorem.

Theorem 2 *Suppose that Assumptions P1-P6, P8 and P9 hold, and let the FEF-IV estimator be defined as in (53). Then we have*

$$\sqrt{N}\left(\hat{\boldsymbol{\gamma}}_{FEF-IV}-\boldsymbol{\gamma}\right)\rightarrow_{d} N\left(\mathbf{0},\boldsymbol{\Omega}_{\hat{\gamma}_{FEF-IV}}\right),$$

where $\boldsymbol{\Omega}_{\hat{\gamma}_{FEF-IV}}$ is given by (55).

The variance of $\hat{\boldsymbol{\gamma}}_{FEF-IV}$ can now be estimated along similar lines as in Section 3.3. We have

$$\widehat{Var}\left(\hat{\boldsymbol{\gamma}}_{FEF-IV}\right)=N^{-1}\mathbf{H}_{zr,N}\left[\hat{\mathbf{V}}_{rr,N}+\mathbf{Q}_{r\bar{x},N}\left(N\widehat{Var}(\hat{\boldsymbol{\beta}})\right)\mathbf{Q}'_{r\bar{x},N}\right]\mathbf{H}'_{zr,N}, \quad (56)$$

where

$$\mathbf{H}_{zr,N}=\left(\mathbf{Q}_{zr,N}\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}'_{zr,N}\right)^{-1}\mathbf{Q}_{zr,N}\mathbf{Q}_{rr,N}^{-1},$$

$$\mathbf{Q}_{r\bar{x},N}=\frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{r}_i-\overline{\mathbf{r}}\right)\left(\bar{\mathbf{x}}_i-\bar{\mathbf{x}}\right)',$$

$$\hat{\mathbf{V}}_{rr,N}=\frac{1}{N}\sum_{i=1}^{N}(\mathbf{r}_i-\overline{\mathbf{r}})(\mathbf{r}_i-\overline{\mathbf{r}})'\left(\hat{\upsilon}_i-\overline{\hat{\upsilon}}\right)^2,$$

$$\hat{\upsilon}_i-\overline{\hat{\upsilon}}=\bar{y}_i-\bar{y}-\left(\bar{\mathbf{x}}_i-\bar{\mathbf{x}}\right)'\hat{\boldsymbol{\beta}}-\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'\hat{\boldsymbol{\gamma}}_{FEF-IV},$$

and as before $\widehat{Var}(\hat{\boldsymbol{\beta}})$ is given by (37). It can be shown that $\widehat{Var}\left(\sqrt{N}\hat{\boldsymbol{\gamma}}_{FEF-IV}\right)$ is a consistent estimator of $\Omega_{\hat{\gamma}_{FEF-IV}}$ defined by (55) if condition (35) is met. Our simulation results suggest that the above variance estimator performs well even if condition (35) is not satisfied.

5 Hausman and Taylor (1981) estimation procedure

Hausman and Taylor (1981) approach the problem of estimation of the time-invariant effects in the panel data model, (1), by assuming that $\mathbf{x}_{it}$ and $\mathbf{z}_i$ can be partitioned into two parts as $(\mathbf{x}_{1,it},\mathbf{x}_{2,it})$ and $(\mathbf{z}_{1,i},\mathbf{z}_{2,i})$, respectively, such that

$$\begin{aligned} E\left(\mathbf{x}'_{1,it}\eta_i\right) &= \mathbf{0},\ E\left(\mathbf{z}'_{1,i}\eta_i\right)=\mathbf{0}, \\ E\left(\mathbf{x}'_{2,it}\eta_i\right) &\neq \mathbf{0},\ E\left(\mathbf{z}'_{2,i}\eta_i\right)\neq\mathbf{0}. \end{aligned}$$

To compute the HT estimator the panel data model is first written as

$$\mathbf{y}_i=\mathbf{X}_i\boldsymbol{\beta}+\left(\mathbf{z}'_i\boldsymbol{\gamma}+\alpha+\eta_i\right)\boldsymbol{\tau}_T+\boldsymbol{\varepsilon}_i,\ \text{for } i=1,2,...,N, \quad (57)$$

where $\mathbf{X}_i = (\mathbf{x}_{i1}, \ldots, \mathbf{x}_{iT})'$, $\mathbf{y}_i = (y_{i1}, y_{i2}, ..., y_{iT})'$, and $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \varepsilon_{i2}, ..., \varepsilon_{iT})'$. Then the following two-step procedure is used[9]:

Step 1 of HT: As in our approach $\boldsymbol{\beta}$ is estimated by $\hat{\boldsymbol{\beta}}$, the FE estimator, the deviations $\hat{d}_i = \bar{y}_i - \bar{\mathbf{x}}_i'\hat{\boldsymbol{\beta}}$, $i = 1, 2, \ldots, N$, are used to compute the 2SLS (or IV) estimator

$$\hat{\gamma}_{IV} = \left(\mathbf{Z}'\mathbf{P}_A\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{P}_A\hat{\mathbf{d}}, \tag{58}$$

where $\hat{\mathbf{d}} = (\hat{d}_1, \hat{d}_2, ..., \hat{d}_N)'$, $\mathbf{Z} = (\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_N)' = (\mathbf{Z}_1, \mathbf{Z}_2)$, and $\mathbf{P}_A = \mathbf{A}\left(\mathbf{A}'\mathbf{A}\right)^{-1}\mathbf{A}'$ is the orthogonal projection matrix of $\mathbf{A} = \left(\boldsymbol{\tau}_N, \bar{\mathbf{X}}_1, \mathbf{Z}_1\right)$, where $\bar{\mathbf{X}} = (\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2)$, and $\bar{\mathbf{X}} = (\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, ..., \bar{\mathbf{x}}_N)'$, $\bar{\mathbf{x}}_i = (\bar{\mathbf{x}}_{i,1}, \bar{\mathbf{x}}_{i,2})$. Using these initial estimators of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, the error variances σ_ε^2 and σ_η^2 are estimated as

$$\begin{aligned}
\hat{\sigma}_\eta^2 &= s^2 - \hat{\sigma}_\varepsilon^2, \\
\hat{\sigma}_\varepsilon^2 &= \frac{1}{N(T-1)}\sum_{i=1}^{N}\left(\mathbf{y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}_{FE}\right)'\mathbf{M}_T\left(\mathbf{y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}_{FE}\right), \\
s^2 &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(y_{it} - \hat{\mu} - \mathbf{x}_{it}'\hat{\boldsymbol{\beta}}_{FE} - \mathbf{z}_i'\hat{\boldsymbol{\gamma}}_{IV}\right)^2,
\end{aligned}$$

Step 2 of HT : In the second step the N equations in (57) are stacked to obtain

$$\mathbf{y} = \mathbf{W}\boldsymbol{\theta} + (\boldsymbol{\eta} \otimes \boldsymbol{\tau}_T) + \boldsymbol{\varepsilon},$$

where $\mathbf{W} = [(\boldsymbol{\tau}_N \otimes \boldsymbol{\tau}_T), \mathbf{X}, (\mathbf{Z} \otimes \boldsymbol{\tau}_T)]$, $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}', \boldsymbol{\gamma}')'$, $\mathbf{y} = (\mathbf{y}_1', \mathbf{y}_2', ..., \mathbf{y}_N')'$, $\boldsymbol{\eta} = (\eta_1, \eta_2, ..., \eta_N)'$, and $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}_1', \boldsymbol{\varepsilon}_2', ..., \boldsymbol{\varepsilon}_N')'$. Under the assumptions that the errors are cross-sectionally independent, serially uncorrelated and homoskedastic we have

$$\boldsymbol{\Omega} = Var\left[(\boldsymbol{\eta} \otimes \boldsymbol{\tau}_T) + \boldsymbol{\varepsilon}\right] = \sigma_\eta^2\left(\mathbf{I}_N \otimes \boldsymbol{\tau}_T\boldsymbol{\tau}_T'\right) + \sigma_\varepsilon^2\left(\mathbf{I}_N \otimes \mathbf{I}_T\right),$$

which can be written as $\boldsymbol{\Omega} = \left(\sigma_\varepsilon^2 + T\sigma_\eta^2\right)\mathbf{P}_V + \sigma_\varepsilon^2\mathbf{Q}_V$, where $\mathbf{P}_V = \mathbf{I}_N \otimes (\mathbf{I}_T - \mathbf{M}_T)$ and $\mathbf{Q}_V = \mathbf{I}_N \otimes \mathbf{M}_T$. It is now easily verified that $\boldsymbol{\Omega}^{-1/2} = \frac{1}{\sigma_\varepsilon}\left(\varphi\mathbf{P}_V + \mathbf{Q}_V\right)$, where $\varphi = \sigma_\varepsilon/\sqrt{\sigma_\varepsilon^2 + T\sigma_\eta^2}$. Then the transformed model can be written as

$$\boldsymbol{\Omega}^{-1/2}\mathbf{y} = \boldsymbol{\Omega}^{-1/2}\mathbf{W}\boldsymbol{\theta} + \boldsymbol{\Omega}^{-1/2}\left[(\boldsymbol{\eta} \otimes \boldsymbol{\tau}_T) + \boldsymbol{\varepsilon}\right]. \tag{59}$$

To simplify the notations we assume that the first column of $\mathbf{Z}$ is $\boldsymbol{\tau}_N$, and then write the (infeasible) HT estimator as,

$$\hat{\boldsymbol{\theta}}_{HT} = \left(\mathbf{W}'\boldsymbol{\Omega}^{-1/2}\mathbf{P}_A\boldsymbol{\Omega}^{-1/2}\mathbf{W}\right)^{-1}\left(\mathbf{W}'\Omega^{-1/2}\mathbf{P}_A\Omega^{-1/2}\mathbf{y}\right), \tag{60}$$

[9] As noted in the Introduction, HT procedure is further developed and extended in the papers by Amemiya and MaCurdy (1986), Breusch et al. (1989), Im et al. (1999) and Baltagi and Bresson (2012).

where

$$\mathbf{P}_A = \mathbf{A}\left(\mathbf{A}'\mathbf{A}\right)^{-1}\mathbf{A}', \ \mathbf{A} = \left(\boldsymbol{\tau}_N \otimes \boldsymbol{\tau}_T, \mathbf{Q}_V\mathbf{X}, \mathbf{X}^{(1)}, \mathbf{Z}_1 \otimes \boldsymbol{\tau}_T\right),$$

$\mathbf{X}^{(1)} = (\mathbf{X}_{1,1}, \mathbf{X}_{1,2}, ..., \mathbf{X}_{1,N})'$, with $\mathbf{X}'_{1,i} = (\mathbf{x}_{1,i1}, \ldots, \mathbf{x}_{1,iT})$, and $\mathbf{x}_{1,it}$ contains the regressors that are uncorrelated with η_i.[10]

The variance covariance matrix of $\hat{\boldsymbol{\theta}}_{HT}$ in the general case where the fixed effects, η_i, are heteroskedastic and possibly cross-sectionally correlated is given by [11]

$$Var\left(\hat{\boldsymbol{\theta}}_{HT}\right) = \mathbf{Q}^{-1} + \left(\frac{T}{\sigma_\varepsilon^2 + T\sigma_\eta^2}\right)\mathbf{Q}^{-1}\left[\mathbf{W}'\boldsymbol{\Omega}^{-1/2}\mathbf{P}_A\left(\left(\mathbf{V}_\eta - \sigma_\eta^2\mathbf{I}_N\right) \otimes \frac{1}{T}\boldsymbol{\tau}_T\boldsymbol{\tau}_T'\right)\mathbf{P}_A\boldsymbol{\Omega}^{-1/2}\mathbf{W}\right]\mathbf{Q}^{-1}, \tag{61}$$

where $\mathbf{Q} = \mathbf{W}'\boldsymbol{\Omega}^{-1/2}\mathbf{P_A}\boldsymbol{\Omega}^{-1/2}\mathbf{W}$, and $\mathbf{V}_\eta$ represents the covariance matrix of $\boldsymbol{\eta}$. $Var\left(\hat{\boldsymbol{\theta}}_{HT}\right)$ reduces to $\mathbf{Q}^{-1}$ in the standard case where η_i's are assumed to be homoskedastic and cross-sectionally independent, namely when $\mathbf{V}_\eta = \sigma_\eta^2\mathbf{I}_N$. To our knowledge the above general expression for $Var\left(\hat{\boldsymbol{\theta}}_{HT}\right)$ is new.

Remark 5 *The HT approach assumes that the errors η_i and ε_{it} are both homoskedastic, serially uncorrelated and cross-sectionally independent. However, simulations to be reported below suggest that the HT estimator works well even in cases of heteroskedasticity and serially correlated errors, if the orthogonality conditions of HT estimator are satisfied. But it is important to bear in mind that if the orthogonality conditions are not met the HT approach breaks down and must be modified.*

Remark 6 *In the case where the effects of the time-invariant regressors are exactly identified, then the HT estimator of γ, $\hat{\boldsymbol{\gamma}}_{HT}$, is identical to the first stage estimator of $\boldsymbol{\gamma}$, given by (58). See Baltagi and Bresson (2012). It is also easily seen that in such a case, $\hat{\boldsymbol{\gamma}}_{HT}$ is also identical to the FEF-IV estimator.*

6 Monte Carlo Simulation

In order to evaluate the performance of the FEF and the FEF-IV estimators proposed in this paper, we conducted two sets of simulations. One set with exogenous time-invariant regressors and a second set where one of the time-invariant regressors is correlated with the fixed effects. In both sets of experiments the data generating process (DGP) include two time-varying and two time-invariant regressors, and allow for error heteroskedasticity and residual serial correlation.

DGP1: Initially, we consider the following data generating process

$$\begin{aligned} y_{it} &= 1 + \alpha_i + x_{1,it}\beta_1 + x_{2,it}\beta_2 + z_{1i}\gamma_1 + z_{2i}\gamma_2 + \varepsilon_{it}, \\ i &= 1, 2, \ldots, N; t = 1, 2, \ldots, T, \end{aligned}$$

[10] See Amemiya and MaCurdy (1986) and Breusch et al. (1989) for discussion on the choice of instruments for HT estimation.

[11] See Section A.6 of the Appendix for a derivation.

with $\beta_1 = \beta_2 = 1$ and $\gamma_1 = \gamma_2 = 1$. We generate the fixed effects as $\alpha_i \sim 0.5\left(\chi^2(2) - 2\right)$, for $i = 1, 2, \ldots, N$. For the time varying regressors we consider the following relatively general specifications

$$
\begin{aligned}
x_{1,it} &= 1 + \alpha_i g_{1t} + \omega_{it,1}, \\
x_{2,it} &= 1 + \alpha_i g_{2t} + \omega_{it,2},
\end{aligned}
$$

where the time effects g_{1t} and g_{2t} for $t = 1, 2, \ldots, T$, are generated as $U(0, 2)$ and are then kept fixed across the replications. The stochastic components of the time varying regressors ($\omega_{it,1}$ and $\omega_{it,2}$) are generated as heterogenous $AR(1)$ processes

$$
\omega_{it,j} = \mu_{ij}(1 - \rho_{\omega,ij}) + \rho_{\omega,ij}\omega_{it-1,j} + \sqrt{1 - \rho_{\omega,ij}^2}\epsilon_{\omega it,j} \text{ for } j = 1, 2,
$$

where

$$
\begin{aligned}
\epsilon_{\omega it,j} &\sim IIDN(0, \sigma_{\epsilon i}^2), \text{ for all } i, j \text{ and } t, \\
\sigma_{\epsilon i}^2 &\sim 0.5(1 + 0.5IID\chi^2(2)), \ \omega_{i0,j} \sim IIDN(\mu_i, \sigma_{\epsilon i}^2), \text{ for all } i, j, \\
\rho_{\omega,ij} &\sim IIDU[0, 0.98], \ \mu_{ij} \sim IIDN(0, \sigma_\mu^2), \sigma_\mu^2 = 2, \text{for all } i, j.
\end{aligned}
$$

The time-invariant regressors are generated as

$$
\begin{aligned}
z_{1i} &\sim 1 + N(0, 1), \text{ for } i = 1, 2, \ldots, N, \\
z_{2i} &\sim IU[7, 12], \text{ for } i = 1, 2, \ldots, N,
\end{aligned}
$$

where $IU(7, 12)$ denotes integers uniformly drawn within the range $[7, 12]$.

DGP2: In this case the DGP is the same as DGP1, except that the second time-invariant regressor, z_{2i}, is generated to depend on the fixed effects, α_i, namely

$$
\begin{aligned}
z_{2i} &= r_i + \alpha_i, \text{ for } i = 1, 2, \ldots, N, \\
r_i &\sim IU[7, 12] \text{ for } i = 1, 2, \ldots, N.
\end{aligned} \tag{62}
$$

We use r_i as the instrument for z_{2i} in the FEF-IV estimation procedure.

For each of the above two baseline DGPs, we generate ε_{it} according to

Case 1: Homoskedastic errors:

$$
\varepsilon_{it} \sim IIDN(0, 1), \text{ for } i = 1, 2, \ldots, N; t = 1, 2, \ldots, T.
$$

Case 2: Heteroscedastic errors:

$$
\varepsilon_{it} \sim IIDN(0, \sigma_i^2), \ i = 1, 2, \ldots, N; t = 1, 2, \ldots, T,
$$

where $\sigma_i^2 \sim 0.5(1 + 0.5IID\chi^2(2))$ for all i.

Case 3: Serially correlated and heteroscedastic errors:

$$\varepsilon_{it} = \rho_{\varepsilon i}\varepsilon_{i,t-1} + \sqrt{1 - \rho_{\varepsilon i}^2}v_{it},$$

where

$$\begin{aligned} \varepsilon_{i0} &= 0 \text{ for all } i, \\ v_{it} &\sim IIDN(0, \sigma_{vi}^2), \text{ for all } i \text{ and } t, \\ \sigma_{vi}^2 &\sim 0.5(1 + 0.5IID\chi^2(2)), \\ \rho_{\varepsilon i} &\sim IIDU[0, 0.98], \text{for all } i, \end{aligned}$$

for $t = -49, -48, ..., 0, 1, 2, ..., T$, with $u_{i,-49} = 0$, for all i. The first 50 observations are discarded, and the remaining T observations are used in the experiments.

Remark 7 *The two DGPs are intended to capture the different features of the estimators proposed in this paper and in the literature. DGP1 allows for an arbitrary degree of dependence between the time-varying regressors and the fixed effects, but assumes that the time-invariant regressors are exogenous. This DGP is designed to be applicable to the FEVD and FEF estimators. DGP2 allows one of the time-invariant regressors to be correlated with the fixed effects, and is used to evaluate the small sample properties of FEF-IV in the presence of heteroscedastic and serially correlated errors.*

We use $1,000$ replications for each experiment, and report bias, root mean squared error (RMSE), size and power for different estimators of $\boldsymbol{\gamma}$, namely FEVD with and without intercepts in the second step, and the FEF estimator proposed in this paper for DGP1. We also consider HT and FEF-IV estimators in the case of DGP2 where all the time varying regressors are correlated with the errors.

The results of FEF for DGP1 are summarized in Tables 1-6, and clearly show that the FEF estimator performs well in all experiments, even when the errors are serially correlated and/or heteroscedastic. It has much lower bias and RMSE as compared to the FEVD estimator proposed by PT. However, in accordance with our theoretical findings, the FEF and FEVD estimators become identical when an intercept is included in the second stage of the PT estimation procedure. However, even after this correction the FEVD approach continues to exhibit substantial size distortions due to the use of incorrect standard errors in the third stage of the procedure (see Section 3.4).

The results for DGP2 are summarized in Tables 7-12. The FEF-IV estimator is computed using r_i (defined by (62)) as an instrument for the endogenous time-invariant regressor, z_{2i}. The HT estimator uses time averages of the time-varying regressors, $\bar{x}_{1i}$ and $\bar{x}_{2i}$, as well as z_{i1}, as instruments. The FEF-IV procedure performs well in all cases, irrespective of whether the errors are heteroskedastic and/or serially correlated. In particular the size of the FEF-IV estimator is very close to the 5% nominal value, with the power rising steadily in N. This suggests that the variance estimator for the FEF-IV, (56), is valid in the case of error heteroskedasticity and performs well even the errors are serially correlated (for example, see Tables 11-12). In contrast, the application of the standard HT procedure yields biased estimates and significant size distortions, particularly in the case of γ_2, the coefficient of the endogenous time-invariant regressor , z_{2i}. Perhaps this is not surprising, considering that in these experiments both of the included time-varying regressors are correlated with the fixed effects, and neither cannot be used as valid instruments.[12] It is possible to modify the HT procedure by including r_i as an additional instrument, in which case the HT estimator will become identical to the FEF-IV estimator considering that under DGP2 the parameters of the panel data model are exactly identified.

7 Conclusion

In this paper, we propose the FEF and FEF-IV estimators for panel models with time-invariant regressors. The FEF estimator is computed using a two-step procedure, where in the first step the fixed effects estimators are used to filter the effects of time-varying regressors. In the second step, time averages of the residuals are used in cross-section regressions to estimate the coefficients of time-invariant regressors. We also develop the asymptotic distribution for the FEF, and show that it's unbiased, consistent and asymptotically normally distributed. The FEF estimator is sufficiently robust and allows for cross-sectional heteroskedasticity and serial correlation. An alternative variance estimators of the FEF estimator is also proposed in this paper.

Moreover, when there is correlation between the time-invariant variables and individual effects, we propose the FEF-IV estimator, which can also be calculated by a two step procedure. The first step of FEF-IV is similar to FEF, but in the second step, we use the instrument variable estimation for the time-invariant regressors. We also show that this FEF-IV estimator is consistent and asymptotically normally distributed. Similar to the FEF estimator, the FEF-IV estimator is sufficiently robust and allows for cross-sectional heteroskedasticity and serial correlation. An alternative variance estimator of the FEF-IV estimator is also proposed in this paper. By simulations, we find both the FEF and FEF-IV have better small sample performance in terms of bias and RMSE, and most importantly has the correct size in the presence of correlation of arbitrary degree between the time-varying regressors and the individual effects.

Furthermore, we also contribute to the debate on the FEVD estimator proposed by Plumper and Troeger (2007). We show that the FEVD estimator is exactly the same as our FEF estimator if an intercept is included in the second step of PT's procedure, but the FEVD estimator is inconsistent in general if no intercept is included in the second stage (see equation (5) in PT). Furthermore, even if the FEVD estimator is computed using an intercept in the second stage, it will still lead to misleading inference since contrary to what is claimed by PT, the standard errors computed in the

[12] In implemetation of the HT procedure we use $\mathbf{Q}_V\mathbf{X}$ (the demeaned time varying regressors, see (60)) as instruments.

third stage of the PT procedure are not valid.

Overall, our Monte Carlo simulations suggest that FEF and FEF-IV estimators proposed in this paper perform well in terms of bias, RMSE, size and power. The simulation results also confirm our theoretical derivations showing that in general the FEVD estimator suffers from size distortions. Finally, in cases where none of the time-varying regressors is uncorrelated with the fixed effects, the use of standard HT procedure can lead to bias and significant size distortions. In such cases a modified version of the HT procedure can be considered if there exists a sufficient number of instruments for the endogenous time-invariant regressors. But it is not clear what advantages there might be in following such a modified HT procedure as compared to the FEF-IV estimator proposed in this paper.

Appendix: Mathematical Derivations

Lemma A.1 *Suppose that* $\mathbf{A}$ *is a* $p \times p$ *symmetric where* p *is fixed,* $\lambda_{\min}(\mathbf{A}) \geq 2/K$*, and* $\lambda_{\max}(\mathbf{A}) \leq K/2$*, with* K *being a fixed, non-zero positive constant. Consider now the stochastic matrix* $\hat{\mathbf{A}}_N$*, viewed as an estimator of* $\mathbf{A}$*, such that* $\left\| \hat{\mathbf{A}}_N - \mathbf{A}_N \right\| \to_p 0$*. Then with probability approaching one* $\lambda_{\min}\left(\hat{\mathbf{A}}_N\right) \geq 1/K$ *and* $\lambda_{\max}\left(\hat{\mathbf{A}}_N\right) \leq K$*.*

Source: Lemma A0 in the mathematical supplement to Newey and Windmeijer (2009).

Lemma A.2 *Given the cross-product sample moments defined by (45), (46) and (47), and in view of (42) and (43), we have*

$$\mathbf{Q}_{h\bar{y},N} = \mathbf{Q}_{h\bar{x},N}\hat{\boldsymbol{\beta}} + \mathbf{Q}_{hh,N}, \tag{A.1}$$

$$\mathbf{q}_{p,NT} = \mathbf{Q}_{p,NT}\hat{\boldsymbol{\beta}} + \mathbf{Q}'_{z\bar{x},N}\hat{\boldsymbol{\gamma}} + \mathbf{Q}'_{h\bar{x},N}, \tag{A.2}$$

$$\mathbf{Q}_{z\bar{y},N} = \mathbf{Q}_{z\bar{x},N}\hat{\boldsymbol{\beta}} + \mathbf{Q}_{zz,N}\hat{\boldsymbol{\gamma}}, \tag{A.3}$$

$$\mathbf{Q}_{zz,N}\left(\hat{\boldsymbol{\gamma}} - \tilde{\boldsymbol{\gamma}}\right) + \mathbf{Q}_{z\bar{x},N}\left(\hat{\boldsymbol{\beta}} - \tilde{\boldsymbol{\beta}}\right) = \mathbf{0}, \tag{A.4}$$

$$(\tilde{\delta} - 1)\mathbf{Q}_{hh,N} = \mathbf{Q}_{h\bar{x},N}\left(\hat{\boldsymbol{\beta}} - \tilde{\boldsymbol{\beta}}\right). \tag{A.5}$$

Proof. Using (42) and (43) we first note that

$$\begin{aligned}
\mathbf{Q}_{h\bar{y},N} &= \frac{1}{N}\sum_{i=1}^{N}\left(\hat{h}_i - \bar{\hat{h}}\right)(\bar{y}_i - \bar{y}) \\
&= \frac{1}{N}\sum_{i=1}^{N}\left(\hat{h}_i - \bar{\hat{h}}\right)(\bar{\mathbf{x}}_i - \bar{\mathbf{x}})'\hat{\boldsymbol{\beta}} + \frac{1}{N}\sum_{i=1}^{N}\left(\hat{h}_i - \bar{\hat{h}}\right)(\mathbf{z}_i - \bar{\mathbf{z}})'\hat{\boldsymbol{\gamma}} + \frac{1}{N}\sum_{i=1}^{N}\left(\hat{h}_i - \bar{\hat{h}}\right)^2 \\
&= \mathbf{Q}_{h\bar{x},N}\hat{\boldsymbol{\beta}} + \mathbf{Q}_{hh,N}.
\end{aligned}$$

Similarly

$$\begin{aligned}
\mathbf{q}_{p,NT} &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\mathbf{x}_{it} - \bar{\mathbf{x}})(y_{it} - \bar{y}) \\
&= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\mathbf{x}_{it} - \bar{\mathbf{x}})(\mathbf{x}_{it} - \bar{\mathbf{x}})'\hat{\boldsymbol{\beta}} + \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\mathbf{x}_{it} - \bar{\mathbf{x}})\left(\hat{u}_{it} - \bar{\hat{u}}\right) \\
&= \mathbf{Q}_{p,NT}\hat{\boldsymbol{\beta}} + \mathbf{Q}'_{z\bar{x},N}\hat{\boldsymbol{\gamma}} + \mathbf{Q}'_{h\bar{x},N},
\end{aligned}$$

where

$$\begin{aligned}
\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\mathbf{x}_{it}-\bar{\mathbf{x}}\right)\left(\hat{u}_{it}-\bar{\hat{u}}\right) &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\mathbf{x}_{it}-\bar{\mathbf{x}}_{i}\right)\left(\hat{u}_{it}-\bar{\hat{u}}_{i}\right)+\frac{1}{N}\sum_{i=1}^{N}\left(\bar{\mathbf{x}}_{i}-\bar{\mathbf{x}}\right)\left(\bar{\hat{u}}_{i}-\bar{\hat{u}}\right) \\
&= \frac{1}{N}\sum_{i=1}^{N}\left(\bar{\mathbf{x}}_{i}-\bar{\mathbf{x}}\right)\left(\bar{\hat{u}}_{i}-\bar{\hat{u}}\right) \\
&= \frac{1}{N}\sum_{i=1}^{N}\left(\bar{\mathbf{x}}_{i}-\bar{\mathbf{x}}\right)\left(\mathbf{z}_{i}-\bar{\mathbf{z}}\right)'\hat{\boldsymbol{\gamma}}+\frac{1}{N}\sum_{i=1}^{N}\left(\bar{\mathbf{x}}_{i}-\bar{\mathbf{x}}\right)\left(\hat{h}_{i}-\bar{\hat{h}}\right) \\
&= \mathbf{Q}_{z\bar{x},N}'\hat{\boldsymbol{\gamma}}+\mathbf{Q}_{h\bar{x},N}',
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{Q}_{z\bar{y},N} &= \frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{z}_{i}-\bar{\mathbf{z}}\right)\left(\bar{y}_{i}-\bar{y}\right) \\
&= \frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{z}_{i}-\bar{\mathbf{z}}\right)\left(\bar{\mathbf{x}}_{i}-\bar{\mathbf{x}}\right)'\hat{\boldsymbol{\beta}}+\frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{z}_{i}-\bar{\mathbf{z}}\right)\left(\bar{\hat{u}}_{i}-\bar{\hat{u}}\right) \\
&= \mathbf{Q}_{z\bar{x},N}\hat{\boldsymbol{\beta}}+\mathbf{Q}_{zz,N}\hat{\boldsymbol{\gamma}}.
\end{aligned}$$

Using this result together with (49) now yields

$$\mathbf{Q}_{zz,N}\left(\hat{\boldsymbol{\gamma}}-\tilde{\boldsymbol{\gamma}}\right)+\mathbf{Q}_{z\bar{x},N}\left(\hat{\boldsymbol{\beta}}-\tilde{\boldsymbol{\beta}}\right)=\mathbf{0}.$$

Finally, (A.5) follows immediately using (A.1) and (50). ■

A.1 Proof of Liapunov condition (26)

The Liapunov condition (26), for $\delta=2$ can be written as

$$\lim_{N\to\infty}N^{-2}\sum_{i=1}^{N}E\left\|\bar{\boldsymbol{\xi}}_{i,N}\right\|^{4}=0. \tag{A.6}$$

From (22) we first note that

$$\left\|\bar{\boldsymbol{\xi}}_{i,N}\right\|\leq\frac{1}{T}\sum_{t=1}^{T}\left\|\mathbf{d}_{z,it}\varepsilon_{it}\right\|,$$

and by Holder inequality

$$\sum_{t=1}^{T}\left\|\mathbf{d}_{z,it}\varepsilon_{it}\right\|\leq\left(\sum_{t=1}^{T}\left\|\mathbf{d}_{z,it}\right\|^{4}\right)^{1/4}\left(\sum_{t=1}^{T}\left|\varepsilon_{it}\right|^{4/3}\right)^{3/4},$$

and hence

$$\left\|\bar{\boldsymbol{\xi}}_{i,N}\right\|^{4}\leq\frac{1}{T^{4}}\left(\sum_{t=1}^{T}\left\|\mathbf{d}_{z,it}\right\|^{4}\right)\left(\sum_{t=1}^{T}\left|\varepsilon_{it}\right|^{4/3}\right)^{3}.$$

But under Assumptions P3 and P7, $\mathbf{x}_{it}$ and $\mathbf{z}_{i}$ are distributed independently of ε_{it}, and it follows that

$$E\left\|\bar{\boldsymbol{\xi}}_{i,N}\right\|^{4}\leq\left(T^{-1}\sum_{t=1}^{T}E\left\|\mathbf{d}_{z,it}\right\|^{4}\right)E\left[\left(T^{-1}\sum_{t=1}^{T}\left|\varepsilon_{it}\right|^{4/3}\right)^{3}\right].$$

But under Assumption P1, $E\left(|\varepsilon_{it}|^4\right) < K$, and since T is finite and for each i, ε_{it} are serially independent, then for some positive finite constant K_1 we have

$$E\left[\left(T^{-1}\sum_{t=1}^{T}|\varepsilon_{it}|^{4/3}\right)^3\right] \leq K_1\ T^{-1}\sum_{t=1}^{T}E\,|\varepsilon_{it}|^4 < K_1K.$$

Hence,

$$N^{-2}\sum_{i=1}^{N}\left\|\boldsymbol{\xi}_{i,N}\right\|^4 < K_1KT^{-1}\sum_{t=1}^{T}\left[N^{-2}\sum_{i=1}^{N}E\left\|\mathbf{d}_{z,it}\right\|^4\right] \tag{A.7}$$

Now recall that

$$\mathbf{d}_{z,it} = (\mathbf{z}_i - \bar{\mathbf{z}}) - \frac{1}{N}\sum_{j=1}^{N}(\mathbf{z}_j - \bar{\mathbf{z}})\,w_{ji,t},$$

where $w_{ji,t} = (\bar{\mathbf{x}}_j - \bar{\mathbf{x}})'\,\mathbf{Q}_{FE,NT}^{-1}(\mathbf{x}_{it} - \bar{\mathbf{x}}_i)$. Then

$$\begin{aligned}\mathbf{d}_{z,it} &= (\mathbf{z}_i - \bar{\mathbf{z}}) - \frac{1}{N}\sum_{j=1}^{N}(\mathbf{z}_j - \bar{\mathbf{z}})(\bar{\mathbf{x}}_j - \bar{\mathbf{x}})'\,\mathbf{Q}_{FE,NT}^{-1}(\mathbf{x}_{it} - \bar{\mathbf{x}}_i) \\ &= \mathbf{z}_i - \bar{\mathbf{z}} - \mathbf{A}_{z\bar{x},N}(\mathbf{x}_{it} - \bar{\mathbf{x}}_i),\end{aligned} \tag{A.8}$$

where $\mathbf{A}_{z\bar{x},N} = \mathbf{Q}_{z\bar{x},N}\mathbf{Q}_{FE,NT}^{-1}$. But[13]

$$\begin{aligned}\left\|\mathbf{A}_{z\bar{x},N}\right\|^2 &= tr\left(\mathbf{Q}_{FE,NT}^{-1}\mathbf{Q}_{z\bar{x},N}'\mathbf{Q}_{z\bar{x},N}\mathbf{Q}_{FE,NT}^{-1}\right) = tr\left(\mathbf{Q}_{FE,NT}^{-2}\mathbf{Q}_{z\bar{x},N}'\mathbf{Q}_{z\bar{x},N}\right) \\ &\leq \lambda_{\max}\left(\mathbf{Q}_{FE,NT}^{-2}\right)tr\left(\mathbf{Q}_{z\bar{x},N}'\mathbf{Q}_{z\bar{x},N}\right) \\ &= \lambda_{\max}^2\left(\mathbf{Q}_{FE,NT}^{-1}\right)\left\|\mathbf{Q}_{z\bar{x},N}\right\|^2 \\ &= \lambda_{\min}^{-2}\left(\mathbf{Q}_{FE,NT}\right)\left\|\mathbf{Q}_{z\bar{x},N}\right\|^2 \\ &= \frac{\left\|\mathbf{Q}_{z\bar{x},N}\right\|^2}{\lambda_{\min}^2\left(\mathbf{Q}_{FE,NT}\right)},\end{aligned}$$

and noting that under Assumption P5, $\lambda_{\min}\left(\mathbf{Q}_{FE,NT}\right) > 1/K$, then

$$\left\|\mathbf{A}_{z\bar{x},N}\right\| \leq K\left\|\mathbf{Q}_{z\bar{x},N}\right\|. \tag{A.9}$$

Also, it is easily seen that

$$\left\|\mathbf{d}_{z,it}\right\|^2 = \mathbf{d}_{z,it}'\mathbf{d}_{z,it} = \left\|\mathbf{z}_i - \bar{\mathbf{z}}\right\|^2 - 2(\mathbf{z}_i - \bar{\mathbf{z}})'\,\mathbf{A}_{z\bar{x},N}(\mathbf{x}_{it} - \bar{\mathbf{x}}_i) + (\mathbf{x}_{it} - \bar{\mathbf{x}}_i)'\,\mathbf{A}_{z\bar{x},N}'\mathbf{A}_{z\bar{x},N}(\mathbf{x}_{it} - \bar{\mathbf{x}}_i),$$

and

$$\begin{aligned}\left\|\mathbf{d}_{z,it}\right\|^4 = {} & \left\|\mathbf{z}_i - \bar{\mathbf{z}}\right\|^4 + 4(\mathbf{z}_i - \bar{\mathbf{z}})'\,\mathbf{A}_{z\bar{x},N}(\mathbf{x}_{it} - \bar{\mathbf{x}}_i)(\mathbf{x}_{it} - \bar{\mathbf{x}}_i)'\,\mathbf{A}_{z\bar{x},N}'(\mathbf{z}_i - \bar{\mathbf{z}}) \\ & + (\mathbf{x}_{it} - \bar{\mathbf{x}}_i)'\,\mathbf{A}_{z\bar{x},N}'\mathbf{A}_{z\bar{x},N}(\mathbf{x}_{it} - \bar{\mathbf{x}}_i)(\mathbf{x}_{it} - \bar{\mathbf{x}}_i)'\,\mathbf{A}_{z\bar{x},N}'\mathbf{A}_{z\bar{x},N}(\mathbf{x}_{it} - \bar{\mathbf{x}}_i) \\ & - 4\left\|\mathbf{z}_i - \bar{\mathbf{z}}\right\|^2(\mathbf{z}_i - \bar{\mathbf{z}})'\,\mathbf{A}_{z\bar{x},N}(\mathbf{x}_{it} - \bar{\mathbf{x}}_i) + 2\left\|\mathbf{z}_i - \bar{\mathbf{z}}\right\|^2(\mathbf{x}_{it} - \bar{\mathbf{x}}_i)'\,\mathbf{A}_{z\bar{x},N}'\mathbf{A}_{z\bar{x},N}(\mathbf{x}_{it} - \bar{\mathbf{x}}_i) \\ & - 4(\mathbf{z}_i - \bar{\mathbf{z}})'\,\mathbf{A}_{z\bar{x},N}(\mathbf{x}_{it} - \bar{\mathbf{x}}_i)(\mathbf{x}_{it} - \bar{\mathbf{x}}_i)'\,\mathbf{A}_{z\bar{x},N}'\mathbf{A}_{z\bar{x},N}(\mathbf{x}_{it} - \bar{\mathbf{x}}_i).\end{aligned}$$

[13] Note that for any $p \times p$ matrices $\mathbf{A}$ and $\mathbf{B}$ such that $\mathbf{A}$ is symmetric and $\mathbf{B}$ positive semi-definite, then $tr(\mathbf{AB}) \leq \lambda_{\max}(\mathbf{A})\,tr(\mathbf{B})$.

Hence

$$\begin{aligned}
\|\mathbf{d}_{z,it}\|^4 \quad \leq \quad & \|\mathbf{z}_i - \bar{\mathbf{z}}\|^4 + 4\|\mathbf{z}_i - \bar{\mathbf{z}}\|^2 \|\mathbf{x}_{it} - \bar{\mathbf{x}}_i\|^2 \left\|\mathbf{A}_{z\bar{x},N}\right\|^2 \\
& + \|\mathbf{x}_{it} - \bar{\mathbf{x}}_i\|^4 \left\|\mathbf{A}_{z\bar{x},N}\right\|^4 + 4\|\mathbf{z}_i - \bar{\mathbf{z}}\|^3 \|\mathbf{x}_{it} - \bar{\mathbf{x}}_i\| \left\|\mathbf{A}_{z\bar{x},N}\right\| \\
& + 2\|\mathbf{z}_i - \bar{\mathbf{z}}\|^2 \|\mathbf{x}_{it} - \bar{\mathbf{x}}_i\|^2 \left\|\mathbf{A}_{z\bar{x},N}\right\|^2 + 4\|\mathbf{z}_i - \bar{\mathbf{z}}\| \|\mathbf{x}_{it} - \bar{\mathbf{x}}_i\|^3 \left\|\mathbf{A}_{z\bar{x},N}\right\|^3,
\end{aligned}$$

and using (A.9) we have

$$\begin{aligned}
N^{-1}\sum_{i=1}^{N}\|\mathbf{d}_{z,it}\|^4 \quad \leq \quad & N^{-1}\sum_{i=1}^{N}\|\mathbf{z}_i - \bar{\mathbf{z}}\|^4 + 4K_1^2\left\|\mathbf{Q}_{z\bar{x},N}\right\|^2 \left[N^{-1}\sum_{i=1}^{N}\|\mathbf{z}_i - \bar{\mathbf{z}}\|^2 \|\mathbf{x}_{it} - \bar{\mathbf{x}}_i\|^2\right] \\
& + K_1^4\left\|\mathbf{Q}_{z\bar{x},N}\right\|^4 \left[N^{-1}\sum_{i=1}^{N}\|\mathbf{x}_{it} - \bar{\mathbf{x}}_i\|^4\right] + 4K_1\left\|\mathbf{Q}_{z\bar{x},N}\right\| \left[N^{-1}\sum_{i=1}^{N}\|\mathbf{z}_i - \bar{\mathbf{z}}\|^3 \|\mathbf{x}_{it} - \bar{\mathbf{x}}_i\|\right] \\
& + 4K_1^3\left\|\mathbf{Q}_{z\bar{x},N}\right\|^3 \left[N^{-1}\sum_{i=1}^{N}\|\mathbf{z}_i - \bar{\mathbf{z}}\| \|\mathbf{x}_{it} - \bar{\mathbf{x}}_i\|^3\right].
\end{aligned}$$

Using this result in (A.7) we now obtain

$$N^{-1}\sum_{i=1}^{N}\left\|\bar{\boldsymbol{\xi}}_{i,N}\right\|^4 \leq K_1 K N^{-1}\sum_{i=1}^{N}\|\mathbf{z}_i - \bar{\mathbf{z}}\|^4 + W_{1N} + W_{2N} + W_{3N} + W_{4N},$$

where

$$\begin{aligned}
W_{1N} &= 4KK_1^3\left\|\mathbf{Q}_{z\bar{x},N}\right\|^2 \left[(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\|\mathbf{z}_i - \bar{\mathbf{z}}\|^2 \|\mathbf{x}_{it} - \bar{\mathbf{x}}_i\|^2\right], \\
W_{2N} &= KK_1^5\left\|\mathbf{Q}_{z\bar{x},N}\right\|^4 \left[(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\|\mathbf{x}_{it} - \bar{\mathbf{x}}_i\|^4\right], \\
W_{3N} &= 4KK_1^2\left\|\mathbf{Q}_{z\bar{x},N}\right\| \left[(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\|\mathbf{z}_i - \bar{\mathbf{z}}\|^3 \|\mathbf{x}_{it} - \bar{\mathbf{x}}_i\|\right], \\
W_{4N} &= 4KK_1^4\left\|\mathbf{Q}_{z\bar{x},N}\right\|^3 \left[N^{-1}\sum_{i=1}^{N}\|\mathbf{z}_i - \bar{\mathbf{z}}\| \|\mathbf{x}_{it} - \bar{\mathbf{x}}_i\|^3\right].
\end{aligned}$$

To investigate the limiting property of $N^{-1}\sum_{i=1}^{N}\left\|\bar{\boldsymbol{\xi}}_{i,N}\right\|^4$, we first note that by Assumption P6 $\left\|\mathbf{Q}_{z\bar{x},N}\right\| \to_p c$, as $N \to \infty$, where c is a finite constant, and by Slutsky's theorem (as $N \to \infty$, for a fixed T) we have

$$\begin{aligned}
W_{1N} &\to_p 4KK_1^3c^2\left[\lim_{N\to\infty}(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}E\left[\|\mathbf{z}_i - \bar{\mathbf{z}}\|^2 \|\mathbf{x}_{it} - \bar{\mathbf{x}}_i\|^2\right]\right], \\
W_{2N} &\to_p KK_1^5c^4\left[\lim_{N\to\infty}(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}E\|\mathbf{x}_{it} - \bar{\mathbf{x}}_i\|^4\right], \\
W_{3N} &\to_p 4KK_1^2c\left[\lim_{N\to\infty}(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}E\left[\|\mathbf{z}_i - \bar{\mathbf{z}}\|^3 \|\mathbf{x}_{it} - \bar{\mathbf{x}}_i\|\right]\right], \\
W_{4N} &\to_p 4KK_1^4c^3\left[\lim_{N\to\infty}N^{-1}\sum_{i=1}^{N}\|\mathbf{z}_i - \bar{\mathbf{z}}\| \|\mathbf{x}_{it} - \bar{\mathbf{x}}_i\|^3\right].
\end{aligned}$$

Hence

$$\begin{aligned}
\lim_{N\to\infty} N^{-1}\sum_{i=1}^{N} E\left\|\bar{\boldsymbol{\xi}}_{i,N}\right\|^4 \quad \leq \quad & K \lim_{N\to\infty} N^{-1}\sum_{i=1}^{N} E\left\|\mathbf{z}_i-\bar{\mathbf{z}}\right\|^4 \\
& +4KK_1^3c^2\left[\lim_{N\to\infty}(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T} E\left[\left\|\mathbf{z}_i-\bar{\mathbf{z}}\right\|^2\left\|\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right\|^2\right]\right] \\
& +KK_1^5c^4\left[\lim_{N\to\infty}(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T} E\left\|\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right\|^4\right] \\
& +4KK_1^2c\left[\lim_{N\to\infty}(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T} E\left[\left\|\mathbf{z}_i-\bar{\mathbf{z}}\right\|^3\left\|\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right\|\right]\right] \\
& +4KK_1^4c^3\left[\lim_{N\to\infty}N^{-1}\sum_{i=1}^{N}\left\|\mathbf{z}_i-\bar{\mathbf{z}}\right\|\left\|\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right\|^3\right].
\end{aligned}$$

Therefore, $N^{-1}\sum_{i=1}^{N} E\left\|\bar{\boldsymbol{\xi}}_{i,N}\right\|^4$ is bounded and converges to a finite limit as $N\to\infty$ (irrespective of whether T is fixed or tends to infinity) if the following conditions hold

$$\lim_{N\to\infty} N^{-1}\sum_{i=1}^{N} E\left\|\mathbf{z}_i-\bar{\mathbf{z}}\right\|^4 < K,$$

$$\begin{aligned}
\lim_{N\to\infty} N^{-1}\sum_{i=1}^{N} E\left\|\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right\|^4 \quad & < \quad K, \text{ when } T \text{ is fixed} \\
\lim_{N \text{ and } T\to\infty}(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T} E\left\|\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right\|^4 \quad & < \quad K
\end{aligned}$$

$$\begin{aligned}
\lim_{N\to\infty} N^{-1}\sum_{i=1}^{N} E\left[\left\|\mathbf{z}_i-\bar{\mathbf{z}}\right\|^2\left\|\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right\|^2\right] \quad & < \quad K, \text{ when } T \text{ is fixed}, \\
\lim_{N \text{ and } T\to\infty}(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T} E\left[\left\|\mathbf{z}_i-\bar{\mathbf{z}}\right\|^2\left\|\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right\|^2\right] \quad & < \quad K,
\end{aligned} \tag{A.10}$$

$$\begin{aligned}
\lim_{N\to\infty} N^{-1}\sum_{i=1}^{N} E\left[\left\|\mathbf{z}_i-\bar{\mathbf{z}}\right\|^3\left\|\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right\|\right] \quad & < \quad K, \text{ when } T \text{ is fixed}, \\
\lim_{N \text{ and } T\to\infty}(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T} E\left[\left\|\mathbf{z}_i-\bar{\mathbf{z}}\right\|^3\left\|\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right\|\right] \quad & < \quad K,
\end{aligned} \tag{A.11}$$

$$\begin{aligned}
\lim_{N\to\infty} N^{-1}\sum_{i=1}^{N} E\left[\left\|\mathbf{z}_i-\bar{\mathbf{z}}\right\|\left\|\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right\|^3\right] \quad & < \quad K, \text{ when } T \text{ is fixed}, \\
\lim_{N \text{ and } T\to\infty}(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T} E\left[\left\|\mathbf{z}_i-\bar{\mathbf{z}}\right\|\left\|\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right\|^3\right] \quad & < \quad K,
\end{aligned} \tag{A.12}$$

The above conditions are clearly satisfied if $\mathbf{x}_{it}$ and $\mathbf{z}_i$ have bounded supports. In the case where $\mathbf{x}_{it}$ and $\mathbf{z}_i$ do not have bounded supports the following moment conditions are sufficient to ensure that $\lim_{N\to\infty} N^{-2}\sum_{i=1}^{N} E\left\|\bar{\boldsymbol{\xi}}_{i,N}\right\|^4 = 0$, as required (applying Cauchy–Schwarz inequality to (A.10) and Holder's Inequality to (A.11) and (A.12)[14]

$$E\left\|\mathbf{z}_i-\bar{\mathbf{z}}\right\|^4<K, \text{ and } E\left\|\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right\|^4<K,$$

for all i,t, N and T. These conditions allow for any degree of dependence between $\mathbf{z}_i$ and $\mathbf{x}_{it}$.

A.2 Proof of Proposition 1

Using (A.8) and noting that $E\left(\varepsilon_{is}\varepsilon_{it}\right)=\gamma_i(s,t)$, $\mathbf{\Omega}_{\bar{\xi},N}$ defined by (28) can be written as

$$\begin{aligned}
\mathbf{\Omega}_{\bar{\xi},N} &= \frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t,s=1}^{T}\mathbf{d}_{z,it}\mathbf{d}_{z,is}'E\left(\varepsilon_{it}\varepsilon_{is}\right)\\
&= \frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t,s=1}^{T}\left[\mathbf{z}_i-\bar{\mathbf{z}}-\mathbf{A}_{z\bar{x}}\left(\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right)\right]\left[\mathbf{z}_i-\bar{\mathbf{z}}-\mathbf{A}_{z\bar{x}}\left(\mathbf{x}_{is}-\bar{\mathbf{x}}_i\right)\right]'\gamma_i(t,s)
\end{aligned}$$

where $\mathbf{A}_{z\bar{x},N}=\mathbf{Q}_{z\bar{x},N}\mathbf{Q}_{FE,NT}^{-1}$. Since

$$\begin{aligned}
&\left[\mathbf{z}_i-\bar{\mathbf{z}}-\mathbf{A}_{z\bar{x},N}\left(\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right)\right]\left[\mathbf{z}_i-\bar{\mathbf{z}}-\mathbf{A}_{z\bar{x},N}\left(\mathbf{x}_{is}-\bar{\mathbf{x}}_i\right)\right]'\\
=\;&\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'-\mathbf{A}_{z\bar{x},N}\left(\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right)\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'-\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)\left(\mathbf{x}_{is}-\bar{\mathbf{x}}_i\right)'\mathbf{A}_{z\bar{x},N}'\\
&+\mathbf{A}_{z\bar{x},N}\left(\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right)\left(\mathbf{x}_{is}-\bar{\mathbf{x}}_i\right)'\mathbf{A}_{z\bar{x},N}',
\end{aligned}$$

then setting $\kappa_{iT}^2=T^{-2}\sum_{t,s=1}^{T}\gamma_i(t,s)$, we obtain (See (18) and (32)).

$$\sigma_\eta^2\mathbf{Q}_{zz}+\mathbf{\Omega}_{\bar{\xi},N}=\mathring{\mathbf{Q}}_{zz,N}+\mathbf{\Delta}_N-\mathbf{\Delta}_{\bar{\xi}N}-\mathbf{\Delta}_{\bar{\xi}N}' \tag{A.13}$$

where

$$\mathring{\mathbf{Q}}_{zz,N}=\frac{1}{N}\sum_{i=1}^{N}\left(\sigma_\eta^2+\kappa_{iT}^2\right)\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)',$$

$$\mathbf{\Delta}_N=\frac{1}{T}\mathbf{A}_{z\bar{x},N}\left[\frac{1}{NT}\sum_{i=1}^{N}\sum_{t,s=1}^{T}\gamma_i(t,s)\left(\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right)\left(\mathbf{x}_{is}-\bar{\mathbf{x}}_i\right)'\right]\mathbf{A}_{z\bar{x},N}', \tag{A.14}$$

and

$$\mathbf{\Delta}_{\bar{\xi},N}=\mathbf{A}_{z\bar{x},N}\left[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t,s=1}^{T}\gamma_i(t,s)\left(\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right)\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'\right]. \tag{A.15}$$

Consider (A.14) and note that using (12) and replacing $\mathbf{A}_{z\bar{x},N}$ by $\mathbf{Q}_{z\bar{x},N}\mathbf{Q}_{FE,NT}^{-1}$, it can be written as

$$\mathbf{\Delta}_N=T^{-1}\mathbf{Q}_{z\bar{x},N}\mathbf{Q}_{FE,NT}^{-1}\mathbf{V}_{FE,NT}\mathbf{Q}_{FE,NT}^{-1}\mathbf{Q}_{z\bar{x},N}',$$

which upon using (11) reduces to

$$\mathbf{\Delta}_N=\mathbf{Q}_{z\bar{x},N}Var(\sqrt{N}\hat{\boldsymbol{\beta}})\mathbf{Q}_{z\bar{x},N}'. \tag{A.16}$$

[14] Note that by Holder's inequality

$$E\left[\left\|\mathbf{z}_i-\bar{\mathbf{z}}\right\|^3\left\|\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right\|\right]\leq\left[E\left(\left\|\mathbf{z}_i-\bar{\mathbf{z}}\right\|^4\right)\right]^{3/4}\left[E\left\|\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right\|^4\right]^{1/4},$$

and

$$E\left[\left\|\mathbf{z}_i-\bar{\mathbf{z}}\right\|\left\|\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right\|^3\right]\leq\left[E\left(\left\|\mathbf{z}_i-\bar{\mathbf{z}}\right\|^4\right)\right]^{1/4}\left[E\left\|\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right\|^4\right]^{3/4}.$$

Similarly, using the expression for $\mathbf{A}_{z\bar{x},N}$ given above, $\Delta_{\bar{\xi},N}$, given by (A.15) can be written as

$$\mathbf{\Delta}_{\bar{\xi},N} = \mathbf{Q}_{z\bar{x},N}\mathbf{Q}_{FE,NT}^{-1}\left[\frac{1}{T^2N}\sum_{i=1}^{N}\sum_{t,s=1}^{T}\gamma_i(t,s)\left(\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right)\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'\right]. \tag{A.17}$$

Substituting (A.16) and (A.17) in (A.13) now yields the desired result.

A.3 Proof of Proposition 2

Consider (30) and the decomposition (33), and note that under Assumptions P1-P5, $\widehat{Var}(\sqrt{N}\hat{\boldsymbol{\beta}})$ defined by (37) tends to $Var\left(\sqrt{N}\hat{\boldsymbol{\beta}}\right)$ for a fixed T and as $N\to\infty$. (See, for example, Arellano (1987)). Consider now $\hat{\mathbf{V}}_{zz,N}$ defined by (38) and note that $\hat{\varsigma}_i-\bar{\hat{\varsigma}}$ defined by (39) can be written as

$$\hat{\varsigma}_i-\bar{\hat{\varsigma}} = (\eta_i-\bar{\eta})+(\bar{u}_i-\bar{u})-(\bar{\mathbf{x}}_i-\bar{\mathbf{x}})'\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)-(\mathbf{z}_i-\bar{\mathbf{z}})'\left(\hat{\boldsymbol{\gamma}}_{FEF}-\boldsymbol{\gamma}\right).$$

Then

$$\hat{\mathbf{V}}_{zz,N} = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{z}_i-\bar{\mathbf{z}})(\mathbf{z}_i-\bar{\mathbf{z}})'\left(\hat{\varsigma}_i-\bar{\hat{\varsigma}}\right)^2 = \mathbf{A}_{1N}+\mathbf{A}_{2N}-\mathbf{A}_{3N},$$

where

$$\begin{aligned}
\mathbf{A}_{1N} &= \frac{1}{N}\sum_{i=1}^{N}(\mathbf{z}_i-\bar{\mathbf{z}})(\mathbf{z}_i-\bar{\mathbf{z}})'\left[(\eta_i-\bar{\eta})+(\bar{\varepsilon}_i-\bar{\varepsilon})\right]^2,\\
\mathbf{A}_{2N} &= \frac{1}{N}\sum_{i=1}^{N}(\mathbf{z}_i-\bar{\mathbf{z}})(\mathbf{z}_i-\bar{\mathbf{z}})'\left[(\bar{\mathbf{x}}_i-\bar{\mathbf{x}})'\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)+(\mathbf{z}_i-\bar{\mathbf{z}})'\left(\hat{\boldsymbol{\gamma}}_{FEF}-\boldsymbol{\gamma}\right)\right]^2\\
\mathbf{A}_{3N} &= \frac{2}{N}\sum_{i=1}^{N}(\mathbf{z}_i-\bar{\mathbf{z}})(\mathbf{z}_i-\bar{\mathbf{z}})'\left[(\eta_i-\bar{\eta})+(\bar{\varepsilon}_i-\bar{\varepsilon})\right]\left[(\bar{\mathbf{x}}_i-\bar{\mathbf{x}})'\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)+(\mathbf{z}_i-\bar{\mathbf{z}})'\left(\hat{\boldsymbol{\gamma}}_{FEF}-\boldsymbol{\gamma}\right)\right]
\end{aligned}$$

Starting with $\mathbf{A}_{1N}$, we have

$$\begin{aligned}
\mathbf{A}_{1N} &= \frac{1}{N}\sum_{i=1}^{N}(\mathbf{z}_i-\bar{\mathbf{z}})(\mathbf{z}_i-\bar{\mathbf{z}})'\left[v_i-(\bar{\eta}+\bar{\varepsilon})\right]^2\\
&= \frac{1}{N}\sum_{i=1}^{N}(\mathbf{z}_i-\bar{\mathbf{z}})(\mathbf{z}_i-\bar{\mathbf{z}})'v_i^2-2\left(\bar{\eta}+\bar{\varepsilon}\right)\frac{1}{N}\sum_{i=1}^{N}(\mathbf{z}_i-\bar{\mathbf{z}})(\mathbf{z}_i-\bar{\mathbf{z}})'v_i\\
&\quad+\frac{(\bar{\eta}+\bar{\varepsilon})^2}{N}\sum_{i=1}^{N}(\mathbf{z}_i-\bar{\mathbf{z}})(\mathbf{z}_i-\bar{\mathbf{z}})'.
\end{aligned}$$

Under Assumptions P1, P2 and P7, $\bar{\eta}+\bar{\varepsilon}=O_p\left(N^{-1/2}\right)+O_p\left(N^{-1/2}T^{-1/2}\right)=o_p(1)$. Also, $N^{-1}\sum_{i=1}^{N}(\mathbf{z}_i-\bar{\mathbf{z}})(\mathbf{z}_i-\bar{\mathbf{z}})'v_i=o_p(1)$, and conditional on $\mathbf{Z}=(\mathbf{z}_1,\mathbf{z}_2,...,\mathbf{z}_N)'$,

$$\frac{1}{N}\sum_{i=1}^{N}(\mathbf{z}_i-\bar{\mathbf{z}})(\mathbf{z}_i-\bar{\mathbf{z}})'\left[v_i^2-E(v_i^2)\right]=o_p(1),$$

as $N\to\infty$. This latter result follows under Assumption P7, $\xi_i=v_i^2-E(v_i^2)$ and $\mathbf{z}_i$ are independently distributed, and v_i are cross-sectionally independent. Under these assumptions

$$E\left[\frac{1}{N}\sum_{i=1}^{N}(\mathbf{z}_i-\bar{\mathbf{z}})(\mathbf{z}_i-\bar{\mathbf{z}})'\left[v_i^2-E(v_i^2)\right]\right]=\frac{1}{N}\sum_{i=1}^{N}E\left[(\mathbf{z}_i-\bar{\mathbf{z}})(\mathbf{z}_i-\bar{\mathbf{z}})'\right]E\left[v_i^2-E(v_i^2)\right]=\mathbf{0},$$

and

$$E\left|\frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'\left[v_i^2-E(v_i^2)\right]\right|\leq\frac{1}{N}\sum_{i=1}^{N}E\left\|\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)\right\|^2 E\left|v_i^2-E(v_i^2)\right|<K<\infty,$$

since by assumption $E\left\|\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)\right\|^2<K$, and $E(v_i^2)<K$. In view of these results it now follows that

$$\mathbf{A}_{1N}=\frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'E\left(v_i^2\right)+o_p(1).$$

Consider now $\mathbf{A}_{2N}$ and note similarly that

$$\begin{aligned}
\mathbf{A}_{2N} &= \frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'\left[\left(\bar{\mathbf{x}}_i-\bar{\mathbf{x}}\right)'\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)+\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'\left(\hat{\boldsymbol{\gamma}}_{FEF}-\boldsymbol{\gamma}\right)\right]^2 \\
&= \frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'\left(\bar{\mathbf{x}}_i-\bar{\mathbf{x}}\right)'\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)'\left(\bar{\mathbf{x}}_i-\bar{\mathbf{x}}\right) \\
&\quad+\frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'\left(\hat{\boldsymbol{\gamma}}_{FEF}-\boldsymbol{\gamma}\right)\left(\hat{\boldsymbol{\gamma}}_{FEF}-\boldsymbol{\gamma}\right)'\left(\mathbf{z}_i-\bar{\mathbf{z}}\right) \\
&\quad+\frac{2}{N}\sum_{i=1}^{N}\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'\left(\bar{\mathbf{x}}_i-\bar{\mathbf{x}}\right)'\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'\left(\hat{\boldsymbol{\gamma}}_{FEF}-\boldsymbol{\gamma}\right),
\end{aligned}$$

and expectations of all the three terms above tend to zero with N. Furthermore

$$\begin{aligned}
&\left\|N^{-1}\sum_{i=1}^{N}\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)\left(\bar{\mathbf{x}}_i-\bar{\mathbf{x}}\right)'\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)'\left(\bar{\mathbf{x}}_i-\bar{\mathbf{x}}\right)\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'\right\| \\
\leq\ &\left\|\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right\|^2\left[N^{-1}\sum_{i=1}^{N}\left\|\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)\left(\bar{\mathbf{x}}_i-\bar{\mathbf{x}}\right)'\right\|^2\right].
\end{aligned}$$

But by Cauchy–Schwarz inequality

$$N^{-1}\sum_{i=1}^{N}E\left[\left\|\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)\left(\bar{\mathbf{x}}_i-\bar{\mathbf{x}}\right)'\right\|^2\right]\leq N^{-1}\sum_{i=1}^{N}\left(E\left\|\mathbf{z}_i-\bar{\mathbf{z}}\right\|^4\right)^{1/2}\left(E\left\|\bar{\mathbf{x}}_i-\bar{\mathbf{x}}\right\|^4\right)^{1/2},$$

and since under Assumption P4 and P7, $E\left\|\mathbf{z}_i-\bar{\mathbf{z}}\right\|^4<K$ and $E\left\|\bar{\mathbf{x}}_i-\bar{\mathbf{x}}\right\|^4<K$, it then follows that $N^{-1}\sum_{i=1}^{N}\left\|\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)\left(\bar{\mathbf{x}}_i-\bar{\mathbf{x}}\right)'\right\|^2$ converges to a finite limit and hence

$$E\left\|N^{-1}\sum_{i=1}^{N}\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)\left(\bar{\mathbf{x}}_i-\bar{\mathbf{x}}\right)'\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)'\left(\bar{\mathbf{x}}_i-\bar{\mathbf{x}}\right)\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'\right\|\leq K\ E\left[\left\|\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right\|^2\right]=O\left(N^{-1}\right).$$

A similar line of argument applies to other terms of $\mathbf{A}_{2,N}$.

Finally, for $\mathbf{A}_{3N}$ we have

$$\mathbf{A}_{3N}=\frac{2}{N}\sum_{i=1}^{N}\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'\left[v_i-(\bar{\eta}+\bar{\varepsilon})\right]\left[\left(\bar{\mathbf{x}}_i-\bar{\mathbf{x}}\right)'\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)+\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'\left(\hat{\boldsymbol{\gamma}}_{FEF}-\boldsymbol{\gamma}\right)\right].$$

Once again noting that $\bar{\eta}+\bar{\varepsilon}=o_p(1)$, $\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}=O_p\left(N^{-1/2}\right)=o_p\left(1\right)$, $\hat{\boldsymbol{\gamma}}_{FEF}-\boldsymbol{\gamma}=O_p\left(N^{-1/2}\right)$, it then follows that

$$\begin{aligned}
\mathbf{A}_{3N} &= \left(\frac{2}{N}\sum_{i=1}^{N}\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'v_i\left(\bar{\mathbf{x}}_i-\bar{\mathbf{x}}\right)'\right)\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right) \\
&\quad-\left(\frac{2}{N}\sum_{i=1}^{N}\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'v_i(\mathbf{z}_i-\bar{\mathbf{z}})'\right)\mathbf{Q}_{zz,N}^{-1}\mathbf{Q}_{z\bar{x},N}\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)+o_p(1)
\end{aligned}$$

so long as

$$E\left\|\frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'v_i\left(\bar{\mathbf{x}}_i-\bar{\mathbf{x}}\right)'\right\|\leq\frac{1}{N}\sum_{i=1}^{N}E\left[\left\|\mathbf{z}_i-\bar{\mathbf{z}}\right\|^2\left\|\bar{\mathbf{x}}_i-\bar{\mathbf{x}}\right\|\right]E\left|v_i\right|<K<\infty,$$

and

$$E\left\|\frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'v_i(\mathbf{z}_i-\bar{\mathbf{z}})'\right\|\leq\frac{1}{N}\sum_{i=1}^{N}E\left\|\mathbf{z}_i-\bar{\mathbf{z}}\right\|^3E\left|v_i\right|<K<\infty.$$

The above conditions are met if $E\left\|\mathbf{z}_i-\bar{\mathbf{z}}\right\|^4<K$ and $E\left\|\bar{\mathbf{x}}_i-\bar{\mathbf{x}}\right\|^2<K$, for all i.

Considering all the three terms together we now have

$$\hat{\mathbf{V}}_{zz,N}=\frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'E(v_i^2)+o_p(1).$$

We also note that $E(v_i^2)=\omega_{iT}^2$, where ω_{iT}^2 is defined by (18), and hence $\hat{\mathbf{V}}_{zz,N}\rightarrow_p\mathbf{V}_{zz}$, defined by (31) as required.

Finally, since $\mathbf{Q}_{z\bar{x},N}\mathbf{Q}_{FE,NT}^{-1}\rightarrow_p\mathbf{Q}_{z\bar{x}}\mathbf{Q}_{FE,T}^{-1}$, as $N\rightarrow\infty$, which is finite and bounded in N, then for a fixed T, $\Delta_{\bar{\xi},N}$ (defined by (34)) has the same order as $N^{-1}T^{-2}\sum_{i=1}^{N}\sum_{t,s=1}^{T}\gamma_i(t,s)\left(\mathbf{x}_{it}-\bar{\mathbf{x}}_i\right)\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'$, and for a fixed T then $\boldsymbol{\Delta}_{\bar{\xi},N}=o_p(1)$, if condition (35) is met.

A.4 Proof of proposition 3

Rewrite the normal equations of the FEVD procedure, (48)-(50) in the following matrix format

$$\begin{pmatrix}\mathbf{Q}_{p,NT} & \mathbf{Q}'_{z\bar{x},N} & \mathbf{Q}'_{h\bar{x},N}\\ \mathbf{Q}_{z\bar{x},N} & \mathbf{Q}_{zz,N} & 0\\ \mathbf{Q}_{h\bar{x},N} & 0 & \mathbf{Q}_{hh,N}\end{pmatrix}\begin{pmatrix}\tilde{\boldsymbol{\beta}}\\ \tilde{\boldsymbol{\gamma}}\\ \tilde{\delta}\end{pmatrix}=\begin{pmatrix}\mathbf{q}_{p,NT}\\ \mathbf{Q}_{z\bar{y},N}\\ \mathbf{Q}_{h\bar{y},N}\end{pmatrix},$$

and note that the inverse of the LHS coefficient matrix is given by (see Magnus and Neudecker (2007), p12))

$$\begin{pmatrix}\mathbf{Q}_{p,NT} & \mathbf{Q}'_{z\bar{x},N} & \mathbf{Q}'_{h\bar{x},N}\\ \mathbf{Q}_{z\bar{x},N} & \mathbf{Q}_{zz,N} & 0\\ \mathbf{Q}_{h\bar{x},N} & 0 & \mathbf{Q}_{hh,N}\end{pmatrix}^{-1}=$$

$$\begin{pmatrix}\mathbf{Q}_{NT}^{-1} & -\mathbf{Q}_{NT}^{-1}\mathbf{Q}'_{z\bar{x},N}\mathbf{Q}_{zz,N}^{-1} & -\mathbf{Q}_{NT}^{-1}\mathbf{Q}'_{h\bar{x},N}\mathbf{Q}_{hh,N}^{-1}\\ -\mathbf{Q}_{zz,N}^{-1}\mathbf{Q}_{z\bar{x},N}\mathbf{Q}_{NT}^{-1} & \mathbf{Q}_{zz,N}^{-1}+\mathbf{Q}_{zz,N}^{-1}\mathbf{Q}_{z\bar{x},N}\mathbf{Q}_{NT}^{-1}\mathbf{Q}'_{z\bar{x},N}\mathbf{Q}_{zz,N}^{-1} & \mathbf{Q}_{zz,N}^{-1}\mathbf{Q}_{z\bar{x},NT}\mathbf{Q}_{NT}^{-1}\mathbf{Q}'_{h\bar{x},N}\mathbf{Q}_{hh,N}^{-1}\\ -\mathbf{Q}_{hh,N}^{-1}\mathbf{Q}_{h\bar{x},N}\mathbf{Q}^{-1} & \mathbf{Q}_{hh,N}^{-1}\mathbf{Q}_{h\bar{x},N}\mathbf{Q}_{NT}^{-1}\mathbf{Q}'_{z\bar{x},N}\mathbf{Q}_{zz,N}^{-1} & \mathbf{Q}_{hh,N}^{-1}+\mathbf{Q}_{hh,N}^{-1}\mathbf{Q}_{h\bar{x},N}\mathbf{Q}_{NT}^{-1}\mathbf{Q}'_{h\bar{x},N}\mathbf{Q}_{hh,N}^{-1}\end{pmatrix}$$

where $\mathbf{Q}_{NT}$ is given by (51). Hence

$$\tilde{\boldsymbol{\beta}}=\mathbf{Q}_{NT}^{-1}\left(\mathbf{q}_{p,NT}-\mathbf{Q}'_{z\bar{x},N}\mathbf{Q}_{zz,N}^{-1}\mathbf{Q}_{z\bar{y},N}-\mathbf{Q}'_{h\bar{x},N}\mathbf{Q}_{hh,N}^{-1}\mathbf{Q}_{h\bar{y},N}\right).$$

But using the results in the lemma A.2 we note that

$$\begin{aligned}
& \mathbf{q}_{p,NT}-\mathbf{Q}'_{z\bar{x},N}\mathbf{Q}_{zz,N}^{-1}\mathbf{Q}_{z\bar{y},N}-\mathbf{Q}'_{h\bar{x},N}\mathbf{Q}_{hh,N}^{-1}\mathbf{Q}_{h\bar{y},N}\\
= \;& \mathbf{Q}_{p,NT}\hat{\boldsymbol{\beta}}+\mathbf{Q}'_{z\bar{x},N}\hat{\boldsymbol{\gamma}}+\mathbf{Q}'_{h\bar{x},N}-\mathbf{Q}'_{z\bar{x},N}\mathbf{Q}_{zz,N}^{-1}\left(\mathbf{Q}_{z\bar{x},N}\hat{\boldsymbol{\beta}}+\mathbf{Q}_{zz,N}\hat{\boldsymbol{\gamma}}\right)\\
& -\mathbf{Q}'_{h\bar{x},N}\mathbf{Q}_{hh,N}^{-1}\left(\mathbf{Q}_{h\bar{x},N}\hat{\boldsymbol{\beta}}+\mathbf{Q}_{hh,N}\right)\\
= \;& \mathbf{Q}_{p,NT}\hat{\boldsymbol{\beta}}+\mathbf{Q}'_{z\bar{x},N}\hat{\boldsymbol{\gamma}}+\mathbf{Q}'_{h\bar{x},N}-\mathbf{Q}'_{z\bar{x},N}\mathbf{Q}_{zz,N}^{-1}\mathbf{Q}_{z\bar{x},N}\hat{\boldsymbol{\beta}}-\mathbf{Q}'_{z\bar{x},N}\mathbf{Q}_{zz,N}^{-1}\mathbf{Q}_{zz,N}\hat{\boldsymbol{\gamma}}\\
& -\mathbf{Q}'_{h\bar{x},N}\mathbf{Q}_{hh,N}^{-1}\mathbf{Q}_{h\bar{x},N}\hat{\boldsymbol{\beta}}-\mathbf{Q}'_{h\bar{x},N}\mathbf{Q}_{hh,N}^{-1}\mathbf{Q}_{hh,N}\\
= \;& \mathbf{Q}_{p,NT}\hat{\boldsymbol{\beta}}+\mathbf{Q}'_{z\bar{x},N}\mathbf{Q}_{zz,N}^{-1}\left(\mathbf{Q}_{z\bar{y},N}-\mathbf{Q}_{z\bar{x},N}\hat{\boldsymbol{\beta}}\right)+\mathbf{Q}'_{h\bar{x},N}-\mathbf{Q}'_{z\bar{x},N}\mathbf{Q}_{zz,N}^{-1}\mathbf{Q}_{z\bar{x},N}\hat{\boldsymbol{\beta}}\\
& -\mathbf{Q}'_{z\bar{x},N}\mathbf{Q}_{zz,N}^{-1}\mathbf{Q}_{zz,N}\mathbf{Q}_{zz,N}^{-1}\left(\mathbf{Q}_{z\bar{y},N}-\mathbf{Q}_{z\bar{x},N}\hat{\boldsymbol{\beta}}\right)-\mathbf{Q}'_{h\bar{x},N}\mathbf{Q}_{hh,N}^{-1}\mathbf{Q}_{h\bar{x},N}\hat{\boldsymbol{\beta}}\\
& -\mathbf{Q}'_{h\bar{x},N}\mathbf{Q}_{hh,N}^{-1}\mathbf{Q}_{hh,N}\\
= \;& \left(\mathbf{Q}_{p,NT}-\mathbf{Q}'_{z\bar{x},N}\mathbf{Q}_{zz,N}^{-1}\mathbf{Q}_{z\bar{x},N}-\mathbf{Q}'_{h\bar{x},N}\mathbf{Q}_{hh,N}^{-1}\mathbf{Q}_{h\bar{x},N}\right)\hat{\boldsymbol{\beta}}=\mathbf{Q}_{NT}\hat{\boldsymbol{\beta}}.
\end{aligned}$$

Hence, given that $\mathbf{Q}_{NT}$ is non-singular by assumption then $\tilde{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}$. Using this result in (A.4) and (A.5) now establishes that $\tilde{\boldsymbol{\gamma}} = \hat{\boldsymbol{\gamma}}$, and $\tilde{\delta} = 1$, as required.

A.5 Proof of proposition 4

Denote the residuals from the OLS regression of by $\mathring{h}_i$ and note that in this case the FEVD estimators are obtained by application of the pooled OLS procedure to the following regression

$$y_{it} = \tilde{\alpha} + \mathbf{x}_{it}'\tilde{\boldsymbol{\beta}} + \mathbf{z}_i'\tilde{\boldsymbol{\gamma}} + \tilde{\delta}\mathring{h}_i + \tilde{\zeta}_{it},$$

where

$$\mathring{h}_i = \bar{\hat{u}}_i - \mathbf{z}_i'\mathring{\boldsymbol{\gamma}},$$

and $\tilde{\zeta}_{it}$ are the residuals from the pooled OLS regression. Recall also that when an intercept is included in the second step regression we have

$$\hat{h}_i = \bar{\hat{u}}_i - \hat{a}_\gamma - \mathbf{z}_i'\hat{\boldsymbol{\gamma}}.$$

Hence,

$$\mathring{h}_i + \mathbf{z}_i'\mathring{\boldsymbol{\gamma}} = \hat{h}_i + \hat{a}_\gamma + \mathbf{z}_i'\hat{\boldsymbol{\gamma}}$$

Using this result to substitute $\mathring{h}_i$ in terms of $\hat{h}_i$ we obtain

$$y_{it} = \tilde{\alpha} + \mathbf{x}_{it}'\tilde{\boldsymbol{\beta}} + \mathbf{z}_i'\tilde{\boldsymbol{\gamma}} + \tilde{\delta}\left(\hat{h}_i + \hat{a}_\gamma + \mathbf{z}_i'\hat{\boldsymbol{\gamma}} - \mathbf{z}_i'\mathring{\boldsymbol{\gamma}}\right) + \tilde{\zeta}_{it},$$

or

$$y_{it} = \left(\tilde{\alpha} + \tilde{\delta}\hat{a}_\gamma\right) + \mathbf{x}_{it}'\tilde{\boldsymbol{\beta}} + \mathbf{z}_i'\left(\tilde{\boldsymbol{\gamma}} + \hat{\boldsymbol{\gamma}} - \mathring{\boldsymbol{\gamma}}\right) + \tilde{\delta}\hat{h}_i + \tilde{\zeta}_{it}.$$

This is the same regression estimated in the third step of the FEVD procedure when an intercept term is included in the second stage, and the results of proposition 3 applies directly and we must have

$$\tilde{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}},\ \tilde{\delta} = 1,$$

and

$$\left(\tilde{\boldsymbol{\gamma}} + \tilde{\delta}\hat{\boldsymbol{\gamma}} - \tilde{\delta}\mathring{\boldsymbol{\gamma}}\right) = \hat{\boldsymbol{\gamma}},$$

Hence, $\tilde{\boldsymbol{\gamma}} = \mathring{\boldsymbol{\gamma}}$.

To derive the bias of $\mathring{\boldsymbol{\gamma}}$ as an estimator of $\boldsymbol{\gamma}$, note that

$$\begin{aligned}
\mathring{\boldsymbol{\gamma}} &= \left(\sum_{i=1}^{N}\mathbf{z}_i\mathbf{z}_i'\right)^{-1}\sum_{i=1}^{N}\mathbf{z}_i\bar{\hat{u}}_i = \left(\sum_{i=1}^{N}\mathbf{z}_i\mathbf{z}_i'\right)^{-1}\sum_{i=1}^{N}\mathbf{z}_i\left(\bar{y}_i - \bar{\mathbf{x}}_i'\hat{\boldsymbol{\beta}}\right) \\
&= \boldsymbol{\gamma} + \left(\sum_{i=1}^{N}\mathbf{z}_i\mathbf{z}_i'\right)^{-1}\sum_{i=1}^{N}\mathbf{z}_i\left[\alpha + \eta_i + \bar{u}_i - \bar{\mathbf{x}}_i'\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)\right],
\end{aligned}$$

Hence,

$$E\left(\mathring{\boldsymbol{\gamma}}\,|\mathbf{Z}\right) = \boldsymbol{\gamma} + \alpha\left(N^{-1}\sum_{i=1}^{N}\mathbf{z}_i\mathbf{z}_i'\right)^{-1}\bar{\mathbf{z}},$$

and $\mathring{\boldsymbol{\gamma}}$ is an unbiased estimator of $\boldsymbol{\gamma}$, if $\alpha E\left[\left(N^{-1}\sum_{i=1}^{N}\mathbf{z}_i\mathbf{z}_i'\right)^{-1}\bar{\mathbf{z}}\right] = \mathbf{0}$. Note that the bias term does not vanish even for N sufficiently large if $\alpha E(\mathbf{z}_i) \neq \mathbf{0}$, for at least one i.

A.6 Covariance matrix of the HT estimator in the case where the fixed effects are heteroskedastic and cross sectionally correlated

Starting with (60), and using (57) we have

$$\begin{aligned}\hat{\boldsymbol{\theta}}_{HT} &= \mathbf{Q}^{-1}\left[\mathbf{W}'\boldsymbol{\Omega}^{-1/2}\mathbf{P}_A\boldsymbol{\Omega}^{-1/2}\mathbf{y}\right] \\ &= \boldsymbol{\theta}+\mathbf{Q}^{-1}\left[\mathbf{W}'\boldsymbol{\Omega}^{-1/2}\mathbf{P}_A\boldsymbol{\Omega}^{-1/2}\mathbf{u}\right],\end{aligned}$$

where $\mathbf{Q}=\mathbf{W}'\boldsymbol{\Omega}^{-1/2}\mathbf{P}_A\boldsymbol{\Omega}^{-1/2}\mathbf{W}$, $\mathbf{W}=[(\boldsymbol{\tau}_N\otimes\boldsymbol{\tau}_T),\mathbf{X},(\mathbf{Z}\otimes\boldsymbol{\tau}_T)]$, and $\mathbf{u}=(\boldsymbol{\eta}\otimes\boldsymbol{\tau}_T)+\boldsymbol{\varepsilon}$. Hence, conditional on $\mathbf{W}$ we have

$$Var\left(\hat{\boldsymbol{\theta}}_{HT}\right)=\mathbf{Q}^{-1}\left[\mathbf{W}'\boldsymbol{\Omega}^{-1/2}\mathbf{P}_A\Omega^{-1/2}Var\left(\mathbf{u}\right)\boldsymbol{\Omega}^{-1/2}\mathbf{P}_A\boldsymbol{\Omega}^{-1/2}\mathbf{W}\right]\mathbf{Q}^{-1}. \tag{A.18}$$

where

$$Var\left(\mathbf{u}\right)=Var\left((\boldsymbol{\eta}\otimes\boldsymbol{\tau}_T)+\boldsymbol{\varepsilon}\right)=\mathbf{V}_\eta\otimes\boldsymbol{\tau}_T\boldsymbol{\tau}_T'+\left(\mathbf{I}_N\otimes\mathbf{I}_T\right)\sigma_\varepsilon^2,$$

and $\mathbf{V}_\eta=E(\boldsymbol{\eta}\boldsymbol{\eta}')$. Recalling that $\boldsymbol{\Omega}^{-1/2}=\frac{1}{\sigma_\varepsilon}\left[\mathbf{I}_N\otimes\mathbf{M}_T+\varphi\mathbf{I}_N\otimes(\mathbf{I}_T-\mathbf{M}_T)\right]$ with $\varphi=\sigma_\varepsilon/\sqrt{\sigma_\varepsilon^2+T\sigma_\eta^2}$, then we have

$$\begin{aligned}&Var\left(\mathbf{u}\right)=\boldsymbol{\Omega}^{-1/2}Var\left((\boldsymbol{\eta}\otimes\boldsymbol{\tau}_T)+\boldsymbol{\varepsilon}\right)\boldsymbol{\Omega}^{-1/2}\\ =\;&\frac{\varphi^2}{\sigma_\varepsilon^2}\left(\mathbf{V}_\eta\otimes\boldsymbol{\tau}_T\boldsymbol{\tau}_T'\right)+\mathbf{I}_N\otimes\mathbf{M}_T+\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2+T\sigma_\eta^2}\mathbf{I}_N\otimes\left(\mathbf{I}_T-\mathbf{M}_T\right)\\ =\;&\frac{T}{\sigma_\varepsilon^2+T\sigma_\eta^2}\left(\mathbf{V}_\eta\otimes\frac{1}{T}\boldsymbol{\tau}_T\boldsymbol{\tau}_T'\right)+\mathbf{I}_N\otimes\mathbf{M}_T+\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2+T\sigma_\eta^2}\mathbf{I}_N\otimes\left(\mathbf{I}_T-\mathbf{M}_T\right)\\ =\;&\mathbf{I}_N\otimes\mathbf{I}_T-\left(1-\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2+T\sigma_\eta^2}\right)\mathbf{I}_N\otimes\frac{1}{T}\boldsymbol{\tau}_T\boldsymbol{\tau}_T'+\frac{T}{\sigma_\varepsilon^2+T\sigma_\eta^2}\left(\mathbf{V}_\eta\otimes\frac{1}{T}\boldsymbol{\tau}_T\boldsymbol{\tau}_T'\right)\\ =\;&\mathbf{I}_N\otimes\mathbf{I}_T+\frac{T}{\sigma_\varepsilon^2+T\sigma_\eta^2}\left(\left(\mathbf{V}_\eta-\sigma_\eta^2\mathbf{I}_N\right)\otimes\frac{1}{T}\boldsymbol{\tau}_T\boldsymbol{\tau}_T'\right).\end{aligned}$$

Using this result in (A.18) and after some algebra we obtain (conditional on $\mathbf{W}$)

$$\begin{aligned}Var\left(\hat{\boldsymbol{\theta}}_{HT}\right) &= \mathbf{Q}^{-1}\left[\mathbf{W}'\boldsymbol{\Omega}^{-1/2}\mathbf{P}_A\left(\mathbf{I}_N\otimes\mathbf{I}_T+\frac{T}{\sigma_\varepsilon^2+T\sigma_\eta^2}\left(\left(\mathbf{V}_\eta-\sigma_\eta^2\mathbf{I}_N\right)\otimes\frac{1}{T}\boldsymbol{\tau}_T\boldsymbol{\tau}_T'\right)\right)\mathbf{P}_A\boldsymbol{\Omega}^{-1/2}\mathbf{W}\right]\mathbf{Q}^{-1}\\ &= \mathbf{Q}^{-1}+\left(\frac{T}{\sigma_\varepsilon^2+T\sigma_\eta^2}\right)\mathbf{Q}^{-1}\left[\mathbf{W}'\Omega^{-1/2}\mathbf{P}_A\left(\left(\mathbf{V}_\eta-\sigma_\eta^2\mathbf{I}_N\right)\otimes\frac{1}{T}\boldsymbol{\tau}_T\boldsymbol{\tau}_T'\right)\mathbf{P}_A\boldsymbol{\Omega}^{-1/2}\mathbf{W}\right]\mathbf{Q}^{-1}.\end{aligned}$$

References

Amemiya, T. and T. MaCurdy (1986). Instrumental-variable estimation of an error-component model. *Econometrica 54*, 869–880.

Arellano, M. (1987). Computing robust standard errors for within-groups estimators. *Oxford Bulletin of Economics and Statistics 49*, 431–434.

Baltagi, B. and G. Bresson (2012). A robust hausman-taylor estimator. In *Badi H. Baltagi, R. Carter Hill, Whitney K. Newey, Halbert L. White (ed.) Essays in Honor of Jerry Hausman (Advances in Econometrics, Volume 29), pp.175-214.* Emerald Group Publishing Limited.

Beck, N. (2011). Of fixed-effects and time-invariant variables. *Political Analysis 19*, 119–122.

Breusch, T., G. Mizon, and P. Schmidt (1989). Efficient estimation using panel data. *Econometrica 57*, 695–700.

Breusch, T., M. Ward, H. Nguyen, and T. Kompas (2011a). Fevd: Just iv or just mistaken? *Political Analysis 19*, 165–169.

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

Davidson, J. (1994). *Stochastic Limit Theory.* Oxford University Press, Oxford.

Greene, W. (2011a). Fixed effects vector decomposition: A magical solution to the problem of time-invariant variables in fixed effects models? *Political Analysis 19*, 135–146.

Greene, W. (2011b). Reply to rejoinder by plumper and troeger. *Political Analysis 19*, 170–172.

Hausman, B. Y. J. A. and W. Taylor (1981). Panel data and unobservable individual effects. *Econometrica 49*, 1377–1398.

Im, K., S. Ahn, P. Schmidt, and J. Wooldridge (1999). Efficient estimation of panel data models with strictly exogenous explanatory variables. *Journal of Econometrics 93*, 177–201.

Magnus, J. R. and H. Neudecker (2007). *Matrix Differential Calculus with Applications in Statistics and Econometrics.* Wiley, New York.

Newey, W. K. and F. Windmeijer (2009). Generalized method of moments with many weak moment conditions. *Econometrica 77*, 687–719.

Plumper, T. and V. E. Troeger (2007). Efficient estimation of time-invariant and rarely changing variables in finite sample panel analyses with unit fixed effects. *Political Analysis 15*, 124–139.

Plumper, T. and V. E. Troeger (2011). Fixed-effects vector decomposition: Properties, reliability, and instruments. *Political Analysis 19*, 147–164.

Table 1: RMSE, size and power of FEF and FEVD estimators for γ_1 in Case 1 of DGP1

N	T	3			5			10		
		FEF	FEVD		FEF	FEVD		FEF	FEVD	
			without	with		without	with		without	with
100	estimate	0.9981	0.5556	0.9981	1.0006	0.5476	1.0006	1.0029	0.5525	1.0029
	bias	-0.0019	-0.4444	-0.0019	0.0006	-0.4524	0.0006	0.0029	-0.4475	0.0029
	rmse	0.1191	0.5010	0.1191	0.1147	0.5095	0.1147	0.1061	0.5024	0.1061
	size	3.7%	41%	12%	6.1%	44%	35%	6%	72%	58%
	power	5.9%	47%	5.9%	8%	51%	22%	8.4%	57%	61%
500	estimate	0.9996	0.5538	0.9996	1.0018	0.5532	1.0018	0.9985	0.5552	0.9985
	bias	-0.0004	-0.4462	-0.0004	0.0018	-0.4468	0.0018	-0.0015	-0.4448	-0.0015
	rmse	0.0519	0.4577	0.0519	0.0478	0.4581	0.0478	0.0476	0.4570	0.0476
	size	5%	100%	43%	4.9%	100%	47%	5%	100%	58%
	power	20%	97%	61%	19%	100%	67%	20%	100%	76%
1000	estimate	0.9985	0.5499	0.9985	1.0003	0.5521	1.0003	1.0003	0.5524	1.0003
	bias	-0.0015	-0.4501	-0.0015	0.0003	-0.4479	0.0003	0.0003	-0.4476	0.0003
	rmse	0.0360	0.4559	0.0360	0.0351	0.4536	0.0351	0.0328	0.4537	0.0328
	size	4.1%	100%	42%	5.4%	100%	48%	5.5%	100%	58%
	power	25%	100%	72%	33%	100%	79%	34%	100%	86%
2000	estimate	1.0001	0.5505	1.0001	1.0002	0.5485	1.0002	0.9997	0.5514	0.9997
	bias	0.0001	-0.4495	0.0001	0.0002	-0.4515	0.0002	-0.0003	-0.4496	-0.0003
	rmse	0.0260	0.4525	0.0260	0.0245	0.4543	0.0245	0.0229	0.4515	0.0229
	size	5.3%	100%	41%	4.9%	100%	47%	4.7%	100%	54%
	power	52%	100%	87%	52%	100%	92%	57%	100%	98%

Notes: 1. Size is calculated under $\gamma_1^{(0)} = 1$, and power under $\gamma_1^{(1)} = 0.95$.

2. For FEVD estimators, "with" refers to the FEVD estimator when an intercept is included in the second step, and "without" refers to the case where the FEVD estimator is computed without an intercept.

3. The number of replication is set at $R = 1000$, and the 95% confidence interval for size 5% is [3.6%, 6.4%].

4. The FEF estimator and its variance are computed using (4) and (36). The FEVD estimator and its variance are computed using the three step procedure described in Section 3.4.

Table 2: Bias, RMSE, size and power of FEF and FEVD estimators for γ_2 in Case 1 of DGP1

N	T	3			5			10		
		FEF	FEVD		FEF	FEVD		FEF	FEVD	
			without	with		without	with		without	with
100	estimate	0.9948	-0.0124	0.9948	1.0010	-0.0118	1.0010	1.0021	-0.0117	1.0021
	bias	-0.0052	-1.0124	-0.0052	0.0010	-1.0118	0.0010	0.0021	-1.0117	0.0021
	rmse	0.0622	1.0127	0.0622	0.0568	1.0121	0.0568	0.0529	1.0120	0.0529
	size	4.9%	45%	26%	6.2%	66%	47%	5.5%	95%	56%
	power	13%	55%	17%	18%	63%	59%	19%	93%	72%
500	estimate	1.0017	-0.0114	1.0017	0.9992	-0.0117	0.9992	0.9995	-0.0118	0.9995
	bias	0.0017	-1.0114	0.0017	-0.0008	-1.0117	-0.0008	-0.0005	-1.0118	-0.0005
	rmse	0.0266	1.0115	0.0266	0.0237	1.0118	0.0237	0.0229	1.0118	0.0229
	size	5.2%	100%	43%	4.2%	100%	46%	5.6%	100%	55%
	power	52%	100%	88%	52%	100%	91%	57%	100%	94%
1000	estimate	0.9990	-0.0111	0.9990	1.0009	-0.0112	1.0009	1.0002	-0.0115	1.0002
	bias	-0.0010	-1.0111	-0.0010	0.0009	-1.0113	0.0009	0.0002	-1.0115	0.0002
	rmse	0.0185	1.0112	0.0185	0.0173	1.0113	0.0173	0.0161	1.0115	0.0161
	size	5.7%	100%	41%	6.1%	100%	44%	4.2%	100%	57%
	power	78%	100%	96%	85%	100%	99%	87%	100%	99%
2000	estimate	1.0000	-0.0113	1.0000	1.0001	-0.0111	1.0001	1.0005	-0.0113	1.0005
	bias	0.0000	-1.0113	0.0000	0.0001	-1.0111	0.0001	0.0005	-1.0113	0.0005
	rmse	0.0128	1.0113	0.0128	0.0124	1.0111	0.0124	0.0121	1.0113	0.0121
	size	4.6%	100%	40%	4.9%	100%	48%	5.5%	100%	59%
	power	97%	100%	100%	98%	100%	100%	99%	100%	100%

Notes: Size is calculated under $\gamma_2^{(0)} = 1$, and power under $\gamma_2^{(1)} = 0.95$. See also the notes to Table 1.

Table 3: Bias, RMSE, size and power of FEF and FEVD estimators for γ_1 in Case 2 of DGP1

N	T	3			5			10		
		FEF	FEVD		FEF	FEVD		FEF	FEVD	
			without	with		without	with		without	with
100	estimate	0.9960	0.5512	0.9960	1.0041	0.5552	1.0041	1.0016	0.5482	1.0016
	bias	-0.0040	-0.4488	-0.0040	0.0041	-0.4448	0.0041	0.0016	-0.4518	0.0016
	rmse	0.1212	0.5062	0.1212	0.1175	0.5028	0.1175	0.1033	0.5006	0.1033
	size	3.8%	13%	41%	5.6%	36%	50%	4.7%	72%	55%
	power	6.3%	6.3%	46%	9.3%	23%	54%	8%	57%	60%
500	estimate	1.0019	0.5534	1.0019	1.0014	0.5492	1.0014	1.0011	0.5552	1.0011
	bias	0.0019	-0.4466	0.0019	0.0014	-0.4508	0.0014	0.0011	-0.4448	0.0011
	rmse	0.0528	0.4587	0.0528	0.0497	0.4623	0.0497	0.0477	0.4564	0.0477
	size	5.2%	100%	44%	5.1%	100%	48%	5.7%	100%	58%
	power	19%	97%	61%	19%	100%	67%	20%	100%	75%
1000	estimate	1.0005	0.5509	1.0005	1.0003	0.5501	1.0003	1.0010	0.5521	1.0010
	bias	0.0005	-0.4491	0.0005	0.0003	-0.4499	0.0003	0.0010	-0.4479	0.0010
	rmse	0.0379	0.4552	0.0379	0.0343	0.4560	0.0343	0.0345	0.4539	0.0345
	size	5.9%	100%	45%	4.7%	100%	47%	6.5%	100%	59%
	power	30%	100%	72%	32%	100%	78%	33%	100%	85%
2000	estimate	1.0007	0.5512	1.0007	0.9990	0.5488	0.9990	0.9988	0.5487	0.9988
	bias	0.0007	-0.4488	0.0007	-0.0010	-0.4512	-0.0010	-0.0012	-0.4512	-0.0012
	rmse	0.0260	0.4519	0.0260	0.0242	0.4541	0.0242	0.0231	0.4541	0.0231
	size	5%	100%	43%	4.1%	100%	46%	4.9%	100%	58%
	power	52%	100%	89%	51%	100%	92%	55%	100%	95%

Note: See the notes to Table 1.

Table 4: Bias, RMSE, size and power of FEF and FEVD estimators for γ_2 in Case 2 of DGP1

N	T	3			5			10		
		FEF	FEVD		FEF	FEVD		FEF	FEVD	
			without	with		without	with		without	with
100	estimate	0.9966	-0.0118	0.9966	0.9964	-0.0123	0.9964	0.9998	-0.0115	0.9998
	bias	-0.0034	-1.0118	-0.0034	-0.0036	-1.0123	-0.0036	-0.0002	-1.0115	-0.0002
	rmse	0.0624	1.0121	0.0624	0.0557	1.0126	0.0557	0.0544	1.0118	0.0544
	size	4.8%	29%	45%	4.3%	70%	46%	5.9%	97%	56%
	power	14%	19%	55%	15%	60%	59%	18%	93%	72%
500	estimate	1.0000	-0.0116	1.0000	0.9999	-0.0112	0.9999	0.9997	-0.0118	0.9997
	bias	0.0000	-1.0116	0.0000	-0.0001	-1.0112	-0.0001	-0.0003	-1.0118	-0.0003
	rmse	0.0260	1.0116	0.0260	0.0242	1.0113	0.0242	0.0238	1.0118	0.0238
	size	5.1%	100%	42%	5%	100%	47%	5.4%	100%	60%
	power	49%	100%	88%	54%	100%	92%	57%	100%	95%
1000	estimate	1.0006	-0.0112	1.0006	0.9995	-0.0113	0.9995	0.9997	-0.0115	0.9997
	bias	0.0006	-1.0112	0.0006	-0.0005	-1.0113	-0.0005	-0.0003	-1.0115	-0.0003
	rmse	0.0183	1.0112	0.0183	0.0176	1.0113	0.0176	0.0156	1.0115	0.0156
	size	4.3%	100%	45%	5.4%	100%	46%	4.3%	100%	55%
	power	79%	100%	98%	82%	100%	98%	86%	100%	100%
2000	estimate	1.0002	-0.0114	1.0002	1.0002	-0.0111	1.0002	1.0004	-0.0111	1.0004
	bias	0.0002	-1.0114	0.0002	0.0002	-1.0111	0.0002	0.0004	-1.0111	0.0004
	rmse	0.0126	1.0114	0.0126	0.0128	1.0111	0.0128	0.0119	1.0111	0.0119
	size	4.8%	100%	41%	5.5%	100%	50%	6%	100%	59%
	power	98%	100%	100%	98%	100%	100%	99%	100%	100%

Notes: Size is calculated under $\gamma_2^{(0)} = 1$, and power under $\gamma_2^{(1)} = 0.95$. See also the notes to Table 1.

Table 5: Bias, RMSE, size and power of FEF and FEVD estimators for γ_1 in Case 3 of DGP1

N	T	3			5			10		
		FEF	FEVD		FEF	FEVD		FEF	FEVD	
			without	with		without	with		without	with
100	estimate	1.0050	0.5585	1.0050	0.9978	0.5421	0.9978	0.9989	0.5576	0.9989
	bias	0.0050	-0.4415	0.0050	-0.0022	-0.4579	-0.0022	-0.0011	-0.4424	-0.0011
	rmse	0.1310	0.5007	0.1310	0.1294	0.5112	0.1294	0.1136	0.5054	0.1136
	size	5.4%	31%	54%	6.3%	64%	60%	5.2%	82%	63%
	power	7.7%	21%	60%	9.1%	49%	62%	7.3%	75%	67%
500	estimate	0.9984	0.5476	0.9984	0.9975	0.5469	0.9975	0.9981	0.5447	0.9981
	bias	-0.0016	-0.4524	-0.0016	-0.0025	-0.4531	-0.0025	-0.0019	-0.4653	-0.0019
	rmse	0.0574	0.4650	0.0574	0.0547	0.4652	0.0547	0.0520	0.4667	0.0520
	size	5.3%	100%	59%	5.3%	100%	60%	5.2%	100%	64%
	power	15%	99%	70%	15%	100%	72%	16%	100%	79%
1000	estimate	1.0004	0.5533	1.0004	1.0020	0.5479	1.0020	1.0004	0.5517	1.0004
	bias	0.0004	-0.4467	0.0004	0.0020	-0.4521	0.0020	0.0004	-0.4483	0.0004
	rmse	0.0406	0.4525	0.0406	0.0380	0.4590	0.0380	0.0359	0.4542	0.0359
	size	4.5%	100%	59%	4.9%	100%	61%	4.9%	100%	67%
	power	25%	100%	79%	28%	100%	83%	29%	100%	86%
2000	estimate	1.0007	0.5495	1.0007	1.0012	0.5520	1.0012	0.9998	0.5516	0.9998
	bias	0.0007	-0.4505	0.0007	0.0012	-0.4480	0.0012	-0.0002	-0.4484	-0.0002
	rmse	0.0279	0.4537	0.0279	0.0267	0.4510	0.0267	0.0269	0.4513	0.0269
	size	5.3%	100%	58%	4.5%	100%	60%	5.2%	100%	69%
	power	43%	100%	89%	46%	100%	93%	51%	100%	93%

See the notes to Table 1.

Table 6: Bias, RMSE, size and power of FEF and FEVD estimators for γ_2 in Case 3 of DGP1

N	T	3			5			10		
		FEF	FEVD		FEF	FEVD		FEF	FEVD	
			without	with		without	with		without	with
100	estimate	0.9992	-0.0124	0.9992	0.9981	-0.0111	0.9981	0.9964	-0.0125	0.9964
	bias	-0.0008	-1.0124	-0.0008	-0.0019	-1.0111	-0.0019	-0.0036	-1.0125	-0.0036
	rmse	0.0649	1.0127	0.0649	0.0627	1.0114	0.0627	0.0584	1.0129	0.0584
	size	6.8%	72%	55%	6.2%	97%	57%	5.9%	100%	65%
	power	12%	63%	68%	13%	95%	69%	13%	100%	73%
500	estimate	0.9996	-0.0109	0.9996	1.0001	-0.0109	1.0001	1.0010	-0.0105	1.0010
	bias	-0.0004	-1.0109	-0.0004	0.0001	-1.0109	0.0001	0.0010	-1.0105	0.0010
	rmse	0.0289	1.0010	0.0289	0.0274	1.0109	0.0274	0.0260	1.0106	0.0260
	size	5%	100%	60%	5%	100%	59%	4.9%	100%	66%
	power	41%	100%	90%	46%	100%	92%	51%	100%	96%
1000	estimate	0.9996	-0.0116	0.9996	1.0002	-0.0112	1.0002	1.0014	-0.0114	1.0014
	bias	-0.0004	-1.0116	-0.0004	0.0002	-1.0112	0.0002	0.0014	-1.0114	0.0014
	rmse	0.0193	1.0116	0.0193	0.0186	1.0113	0.0186	0.0183	1.0114	0.0183
	size	4.8%	100%	55%	4.3%	100%	58%	5.5%	100%	67%
	power	72%	100%	98%	75%	100%	98%	80%	100%	99%
2000	estimate	1.0005	-0.0111	1.0005	1.0007	-0.0113	1.0007	0.9996	-0.0114	0.9996
	bias	0.0005	-1.0111	0.0005	0.0007	-1.0113	0.0007	-0.0004	-1.0114	-0.0004
	rmse	0.0140	1.0111	0.0140	0.0140	1.0113	0.0140	0.0129	1.0115	0.0129
	size	2.8%	100%	59%	6.1%	100%	61%	4.7%	100%	66%
	power	95%	100%	100%	96%	100%	100%	97%	100%	100%

Notes: Size is calculated under $\gamma_2^{(0)} = 1$, and power under $\gamma_2^{(1)} = 0.95$. See also the notes to Table 1.

Table 7: Bias, RMSE, size and power of FEF-IV and HT estimators for γ_1 in Case 1 of DGP2

	T	3		5		10	
N		FEF-IV	HT	FEF-IV	HT	FEF-IV	HT
	estimate	0.9978	1.0025	0.9982	0.9900	0.9988	1.0030
100	bias	-0.0022	0.0025	-0.0018	-0.0100	-0.0012	0.0030
	rmse	0.1193	0.2248	0.1125	0.2614	0.1052	0.2230
	size	3.9%	1.3%	4.9%	1.3%	5.8%	2.3%
	power	6.5%	1.8%	7.4%	2.4%	8.3%	2.5%
	estimate	0.9993	0.9972	1.0011	1.0017	0.9997	1.0002
500	bias	-0.0007	-0.0028	0.0011	0.0017	-0.0003	0.0002
	rmse	0.0518	0.0921	0.0497	0.0939	0.0454	0.0914
	size	5.6%	3.7%	4.9%	4.7%	4.7%	4%
	power	15%	7.1%	19%	8%	19%	7.9%
	estimate	0.9979	1.0001	0.9972	1.0026	0.9996	0.9955
1000	bias	-0.0021	0.0001	-0.0028	0.0026	-0.0004	-0.0045
	rmse	0.0353	0.0683	0.0334	0.0639	0.0338	0.0651
	size	4.3%	4.9%	4%	3.8%	5.9%	4.9%
	power	25%	12%	27%	12%	33%	11%
	estimate	1.0001	1.0009	1.0002	1.0026	1.0005	0.9980
2000	bias	0.0001	0.0009	0.0002	0.0026	0.0005	-0.0020
	rmse	0.0260	0.0481	0.0245	0.0455	0.0237	0.0457
	size	5.3%	5.3%	5.1%	4.3%	4.7%	5.4%
	power	52%	18%	53%	20%	56%	18%

Notes: 1. Size is calculated under $\gamma_1^{(0)} = 1$, and power under $\gamma_1^{(1)} = 0.95$.

2. The number of replication is set at $R = 1000$, and the 95% confidence interval for size 5% is [3.6%, 6.4%].

3. The FEF-IV estimator and its variance are computed using (53) and (56), with r_i, defined by (62) as the instrument. The HT estimator and its variance are computed using (60) and (61), with time averages of the time-varying regressors, $\bar{x}_{1i}$ and $\bar{x}_{2i}$, and z_{i1} as instruments. In computing the variance of the HT estimator we set $\mathbf{V}_\eta = \sigma_\eta^2 \mathbf{I}_N$ in (61), as assumed under the standard HT procedure.

Table 8: Bias, RMSE, size and power of FEF-IV and HT estimators for γ_2 in Case 1 of DGP2

	T	3		5		10	
N		FEF-IV	HT	FEF-IV	HT	FEF-IV	HT
	estimate	0.9969	1.9548	0.9999	1.9619	0.9979	1.9768
100	bias	-0.0031	0.9548	-0.0001	0.9619	-0.0021	0.9768
	rmse	0.0624	1.0705	0.0572	1.1176	0.0527	1.0589
	size	4.2%	64%	5.8%	89%	5.5%	88%
	power	18%	69%	18%	92%	23%	89%
	estimate	1.0012	1.9850	0.9988	2.0002	0.9995	1.9958
500	bias	0.0012	0.9850	-0.0012	1.0002	-0.0005	0.9958
	rmse	0.0266	1.0023	0.0238	1.0107	0.0232	1.0058
	size	5.2%	100%	4.1%	100%	5.2%	100%
	power	53%	100%	51%	100%	57%	100%
	estimate	0.9996	1.9955	0.9993	1.9972	0.9998	1.9991
1000	bias	-0.0004	0.9955	-0.0007	0.9972	-0.0002	0.9991
	rmse	0.0178	1.0039	0.0177	1.0024	0.0168	1.0038
	size	4.6%	100%	5.2%	100%	5.1%	100%
	power	78%	100%	80%	100%	84%	100%
	estimate	0.9999	1.9913	1.0000	2.0004	0.9999	1.9999
2000	bias	-0.0001	0.9913	0.0000	1.0004	-0.0001	0.9999
	rmse	0.0128	0.9955	0.0124	1.0029	0.0117	1.0024
	size	4.6%	100%	5.2%	100%	4.7%	100%
	power	96%	100%	98%	100%	98%	100%

Notes: 1. Size is calculated under $\gamma_2^{(0)} = 1$, and power under $\gamma_2^{(1)} = 0.95$. Also see the notes to Table 7.

Table 9: Bias, RMSE, size and power of FEF-IV and HT estimators for γ_1 in Case 2 of DGP2

	T	3		5		10	
N		FEF-IV	HT	FEF-IV	HT	FEF-IV	HT
100	estimate	1.0019	1.0081	1.0027	0.9963	1.0026	1.0029
	bias	0.0019	0.0081	0.0027	-0.0037	0.0026	0.0029
	rmse	0.1165	0.2196	0.1161	0.2443	0.1032	0.2521
	size	4.6%	1.5%	4.2%	2.5%	5%	1.8%
	power	7.4%	1.9%	8.8%	3.6%	8%	2.7%
500	estimate	0.9994	1.0020	1.0015	1.0019	0.9971	0.9935
	bias	-0.0006	0.0020	0.0015	0.0019	-0.0029	-0.0065
	rmse	0.0529	0.0937	0.0497	0.0919	0.0460	0.0905
	size	5.3%	3.4%	5.1%	3.8%	4%	4.2%
	power	19%	7%	19%	8.3%	18%	7.1%
1000	estimate	1.0010	0.9971	0.9989	1.0004	1.0011	0.9988
	bias	0.0010	-0.0029	-0.0011	0.0004	0.0011	-0.0012
	rmse	0.0355	0.0670	0.0344	0.0681	0.0346	0.0660
	size	3.8%	4.4%	5.4%	5.1%	6.2%	5.2%
	power	30%	12%	30%	13%	34%	11%
2000	estimate	1.0006	1.0011	0.9990	0.9999	0.9999	1.0018
	bias	0.0006	0.0011	-0.0010	-0.0001	-0.0001	0.0018
	rmse	0.0259	0.0477	0.0242	0.0453	0.0237	0.0447
	size	5.2%	5.7%	4.1%	4.7%	5.4%	4.1%
	power	51%	19%	51%	19%	55%	20%

See the notes to Table 7.

Table 10: Bias, RMSE, size and power of FEF-IV and HT estimators for γ_2 in Case 2 of DGP2

	T	3		5		10	
N		FEF-IV	HT	FEF-IV	HT	FEF-IV	HT
	estimate	0.9959	1.9536	0.9938	2.0047	0.9972	2.0584
100	bias	-0.0041	0.9536	-0.0062	1.0047	-0.0028	1.0584
	rmse	0.0583	1.0909	0.0584	1.1266	0.0550	1.6905
	size	3.5%	60%	4.8%	86%	5.9%	87%
	power	15%	67%	18%	88%	21%	89%
	estimate	0.9994	1.9874	0.9994	1.9973	0.9993	1.9959
500	bias	-0.0006	0.9874	-0.0006	0.9937	-0.0007	0.9959
	rmse	0.0260	1.0060	0.0243	1.0046	0.0240	1.0063
	size	4.8%	100%	5.2%	100%	5.9%	100%
	power	49%	100%	54%	100%	58%	100%
	estimate	1.0000	1.9974	0.9987	1.9963	0.9996	2.0018
1000	bias	0.0000	0.9974	-0.0013	0.9963	-0.0004	1.0018
	rmse	0.0186	1.0058	0.0182	1.0019	0.0157	1.0069
	size	5.8%	100%	7%	100%	3.8%	100%
	power	77%	100%	80%	100%	84%	100%
	estimate	1.0001	1.9980	1.0000	1.9981	1.0001	1.9993
2000	bias	0.0001	0.9980	0.0000	0.9981	0.0001	0.9993
	rmse	0.0126	1.0022	0.0128	1.0008	0.0117	1.0018
	size	4.9%	100%	5.2%	100%	4%	100%
	power	97%	100%	97%	100%	99%	100%

See the notes to Table 7 and 8.

Table 11: Bias, RMSE, size and power of FEF-IV and HT estimators for γ_1 in Case 3 of DGP2

N	T	3		5		10	
		FEF-IV	HT	FEF-IV	HT	FEF-IV	HT
100	estimate	1.0022	1.0064	0.9978	1.0063	0.9935	1.0012
	bias	0.0022	0.0064	-0.0022	0.0063	-0.0065	0.0012
	rmse	0.1336	0.2313	0.1332	0.2634	0.1212	0.2257
	size	5.5%	2.1%	6.8%	2.3%	7.1%	2.4%
	power	6.8%	2.6%	8.6%	2.8%	7.9%	3.3%
500	estimate	1.0006	0.9974	0.9987	1.0007	0.9993	1.0039
	bias	0.0006	-0.0026	-0.0013	0.0007	-0.0007	0.0039
	rmse	0.0576	0.0964	0.0543	0.0951	0.0526	0.0965
	size	4.3%	3.3%	5.1%	4.1%	5.7%	5.5%
	power	15%	6.8%	15%	7.2%	18%	9.8%
1000	estimate	0.9996	1.0070	0.9991	1.0012	0.9990	1.0012
	bias	-0.0004	0.0070	-0.0009	0.0012	-0.0010	0.0012
	rmse	0.0396	0.0666	0.0386	0.0644	0.0349	0.0660
	size	4.4%	3.7%	5.2%	3.2%	3.8%	4.9%
	power	22%	12%	25%	11%	26%	11%
2000	estimate	1.0003	1.0020	1.0001	1.0002	0.9998	1.0000
	bias	0.0003	0.0020	0.0001	0.0002	-0.0002	0.0000
	rmse	0.0291	0.0489	0.0266	0.0483	0.0256	0.0456
	size	6.3%	5.2%	4.7%	5.6%	4.7%	3.8%
	power	42%	19%	45%	20%	48%	20%

See the notes to Table 7.

Table 12: Bias, RMSE, size and power of FEF-IV and HT estimators for γ_2 in Case 3 of DGP2

	T	3		5		10	
N		FEF-IV	HT	FEF-IV	HT	FEF-IV	HT
	estimate	0.9977	1.9641	0.9962	2.0007	0.9958	2.0528
100	bias	-0.0023	0.9641	-0.0038	1.0007	-0.0042	1.0528
	rmse	0.0692	1.0573	0.0663	1.1191	0.0596	1.8064
	size	7%	80%	6.9%	90%	5.1%	87%
	power	15%	84%	16%	92%	18%	88%
	estimate	0.9989	2.0003	0.9996	1.9918	1.0008	1.9959
500	bias	-0.0011	1.0003	-0.0004	0.9918	0.0008	0.9959
	rmse	0.0290	1.0139	0.0281	1.0022	0.0259	1.0068
	size	4.5%	100%	6.1%	100%	5.4%	100%
	power	42%	100%	46%	100%	51%	100%
	estimate	1.0007	1.9979	0.9991	2.0047	0.9996	2.0000
1000	bias	0.0007	0.9979	-0.0009	1.0047	-0.0004	1.0000
	rmse	0.0198	1.0050	0.0187	1.0101	0.0177	1.0051
	size	4.9%	100%	4.3%	100%	4.3%	100%
	power	72%	100%	71%	100%	77%	100%
	estimate	0.9994	2.0029	1.0006	1.9992	0.9997	1.9969
2000	bias	-0.0006	1.0029	0.0006	0.9992	-0.0003	0.9969
	rmse	0.0141	1.0063	0.0137	1.0019	0.0129	0.9994
	size	5.2%	100%	5.7%	100%	4.6%	100%
	power	93%	100%	95%	100%	96%	100%

See the notes to Tables 7 and 8.