Supplement to "Estimation of Time-invariant Effects in Static Panel Data Models"

by M. Hashem Pesaran, University of Southern California,

Qiankun Zhou, State University of New York (SUNY) at Binghamton

May 2016

This supplement contains three sections. Section A gives the derivation of the modified HT estimator proposed in the paper, and provides a comparison of the modified HT and FEF-IV estimation procedures. Section B includes all the Monte Carlo simulation results discussed in the paper. Section C provides additional simulations for the (unmodified) HT estimation.

A: Modified HT estimators and comparison of FEF-IV and modified HT estimators

Using the same notations as in the main paper, we first note that

$$\mathbf{\Omega}^{-1/2}=\frac{1}{\sigma_{\varepsilon}}\left(\varphi\mathbf{P}_{V}+\mathbf{Q}_{V}\right)\equiv\lambda\mathbf{I}_{NT}+\psi\mathbf{Q}_{V}, \tag{1}$$

where $\mathbf{M}_{T}=\mathbf{I}_{T}-\boldsymbol{\tau}_{T}\left(\boldsymbol{\tau}_{T}^{\prime}\boldsymbol{\tau}_{T}\right)^{-1}\boldsymbol{\tau}_{T}^{\prime}$, $\mathbf{P}_{V}=\mathbf{I}_{N}\otimes\left(\mathbf{I}_{T}-\mathbf{M}_{T}\right)$, $\mathbf{Q}_{V}=\mathbf{I}_{N}\otimes\mathbf{M}_{T}$, $\varphi=\sigma_{\varepsilon}/\sqrt{\sigma_{\varepsilon}^{2}+T\sigma_{\eta}^{2}}$, $\lambda=\varphi/\sigma_{\varepsilon}$ and $\psi=\left(1-\varphi\right)/\sigma_{\varepsilon}$. The (infeasible) modified HT (HTM) estimator is defined by

$$\hat{\boldsymbol{\theta}}_{HTM}=\left(\mathbf{W}^{\prime}\mathbf{\Omega}^{-1/2}\mathbf{P}_{A^{*}}\mathbf{\Omega}^{-1/2}\mathbf{W}\right)^{-1}\left(\mathbf{W}^{\prime}\Omega^{-1/2}\mathbf{P}_{A^{*}}\Omega^{-1/2}\mathbf{y}\right), \tag{2}$$

where $\mathbf{W}=\left[\left(\boldsymbol{\tau}_{N}\otimes\boldsymbol{\tau}_{T}\right),\mathbf{Z}\otimes\boldsymbol{\tau}_{T},\mathbf{X}\right]$, $\mathbf{P}_{A^{*}}=\mathbf{A}^{*}\left(\mathbf{A}^{*\prime}\mathbf{A}^{*}\right)^{-1}\mathbf{A}^{*\prime}$, $\mathbf{A}^{*}=\left[\left(\boldsymbol{\tau}_{N}\otimes\boldsymbol{\tau}_{T}\right),\mathbf{R}\otimes\boldsymbol{\tau}_{T},\mathbf{Q}_{V}\mathbf{X}\right]$, and $\mathbf{R}$ is the $N\times s$ matrix of instrumental variables. The associated variance-covariance matrix is given by

$$Cov\left(\hat{\boldsymbol{\theta}}_{HTM}\right)=\left(\mathbf{W}^{\prime}\mathbf{\Omega}^{-1/2}\mathbf{P}_{A^{*}}\mathbf{\Omega}^{-1/2}\mathbf{W}\right)^{-1}.$$

Since $\mathbf{M}_{T}\boldsymbol{\tau}_{T}=0$, then $\mathbf{Q}_{V}\mathbf{Z}=\left(\mathbf{I}_{N}\otimes\mathbf{M}_{T}\right)\left(\mathbf{Z}\otimes\boldsymbol{\tau}_{T}\right)=\mathbf{0}$, and $\mathbf{P}_{V}\left(\mathbf{Z}\otimes\boldsymbol{\tau}_{T}\right)=\left(\mathbf{I}_{N}\otimes\left(\mathbf{I}_{T}-\mathbf{M}_{T}\right)\right)\left(\mathbf{Z}\otimes\boldsymbol{\tau}_{T}\right)=\left(\mathbf{Z}\otimes\boldsymbol{\tau}_{T}\right)$, and further using (1) it then readily follows that

$$\mathbf{\Omega}^{-1/2}\mathbf{W}==\left[\lambda\left(\boldsymbol{\tau}_{N}\otimes\boldsymbol{\tau}_{T}\right),\lambda\mathbf{Z}\otimes\boldsymbol{\tau}_{T},\mathbf{\Omega}^{-1/2}\mathbf{X}\right]. \tag{3}$$

Also, since $\mathbf{R}^{\prime}\mathbf{Q}_{V}=\mathbf{0}$, then

$$\begin{aligned}\mathbf{A}^{*\prime}\mathbf{A}^{*} &= \begin{pmatrix}\left(\boldsymbol{\tau}_{N}^{\prime}\otimes\boldsymbol{\tau}_{T}^{\prime}\right)\\ \mathbf{R}^{\prime}\otimes\boldsymbol{\tau}_{T}^{\prime}\\ \mathbf{X}^{\prime}\mathbf{Q}_{V}\end{pmatrix}\left(\left(\boldsymbol{\tau}_{N}\otimes\boldsymbol{\tau}_{T}\right),\mathbf{R}\otimes\boldsymbol{\tau}_{T},\mathbf{Q}_{V}\mathbf{X}\right)\\ &\equiv NT\begin{pmatrix}1 & \frac{1}{N}\boldsymbol{\tau}_{N}^{\prime}\mathbf{R} & 0\\ \frac{1}{N}\mathbf{R}^{\prime}\boldsymbol{\tau}_{N} & \frac{1}{N}\mathbf{R}^{\prime}\mathbf{R} & 0\\ 0 & 0 & \frac{1}{NT}\mathbf{X}^{\prime}\mathbf{Q}_{V}\mathbf{X}\end{pmatrix},\end{aligned}$$

and

$$\left(\mathbf{A}^{*\prime}\mathbf{A}^{*}\right)^{-1}=\frac{1}{NT}\begin{pmatrix} 1+\bar{\mathbf{r}}'\mathbf{Q}_{rr,N}^{-1}\bar{\mathbf{r}} & -\bar{\mathbf{r}}'\mathbf{Q}_{rr,N}^{-1} & 0 \\ -\mathbf{Q}_{rr,N}^{-1}\bar{\mathbf{r}} & \mathbf{Q}_{rr,N}^{-1} & 0 \\ 0 & 0 & \mathbf{Q}_{FE,NT}^{-1} \end{pmatrix}, \tag{4}$$

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

Using (3), we have

$$\begin{aligned}\mathbf{A}^{*\prime}\boldsymbol{\Omega}^{-1/2}\mathbf{W} &= \begin{pmatrix} \left(\boldsymbol{\tau}_N'\otimes\boldsymbol{\tau}_T'\right) \\ \mathbf{R}'\otimes\boldsymbol{\tau}_T' \\ \mathbf{X}'\mathbf{Q}_V \end{pmatrix}\left[\lambda\left(\boldsymbol{\tau}_N\otimes\boldsymbol{\tau}_T\right),\lambda\mathbf{Z}\otimes\boldsymbol{\tau}_T,\boldsymbol{\Omega}^{-1/2}\mathbf{X}\right] \\ &= \begin{pmatrix} \lambda NT & \lambda T\left(\boldsymbol{\tau}_N'\mathbf{Z}\right) & \left(\boldsymbol{\tau}_N'\otimes\boldsymbol{\tau}_T'\right)\boldsymbol{\Omega}^{-1/2}\mathbf{X} \\ \lambda T\mathbf{R}'\boldsymbol{\tau}_N & \lambda T\mathbf{R}'\mathbf{Z} & \left(\mathbf{R}'\otimes\boldsymbol{\tau}_T'\right)\boldsymbol{\Omega}^{-1/2}\mathbf{X} \\ \lambda\mathbf{X}'\mathbf{Q}_V\left(\boldsymbol{\tau}_N\otimes\boldsymbol{\tau}_T\right) & \lambda\mathbf{X}'\mathbf{Q}_V\left(\mathbf{Z}\otimes\boldsymbol{\tau}_T\right) & \mathbf{X}'\mathbf{Q}_V\boldsymbol{\Omega}^{-1/2}\mathbf{X} \end{pmatrix},\end{aligned} \tag{5}$$

where

$$\begin{aligned}\left(\boldsymbol{\tau}_N'\otimes\boldsymbol{\tau}_T'\right)\boldsymbol{\Omega}^{-1/2}\mathbf{X} &= \left(\boldsymbol{\tau}_N'\otimes\boldsymbol{\tau}_T'\right)\left(\lambda\mathbf{I}_{NT}+\psi\mathbf{Q}_V\right)\mathbf{X} \\ &= \lambda\left(\boldsymbol{\tau}_N'\otimes\boldsymbol{\tau}_T'\right)\mathbf{X}+\psi\left(\boldsymbol{\tau}_N'\otimes\boldsymbol{\tau}_T'\right)\mathbf{Q}_V\mathbf{X} \\ &= \lambda\left(\boldsymbol{\tau}_N'\otimes\boldsymbol{\tau}_T'\right)\mathbf{X} \\ &= \lambda NT\bar{\mathbf{x}},\end{aligned}$$

$\bar{\mathbf{x}}=\frac{1}{NT}\sum_{i,t}\mathbf{x}_{it}$, and

$$\begin{aligned}\left(\mathbf{R}'\otimes\boldsymbol{\tau}_T'\right)\boldsymbol{\Omega}^{-1/2}\mathbf{X} &= \left(\mathbf{R}'\otimes\boldsymbol{\tau}_T'\right)\left(\lambda\mathbf{I}_{NT}+\psi\mathbf{Q}_V\right)\mathbf{X} \\ &= \lambda\left(\mathbf{R}'\otimes\boldsymbol{\tau}_T'\right)\mathbf{X}+\psi\left(\mathbf{R}'\otimes\boldsymbol{\tau}_T'\right)\mathbf{Q}_V\mathbf{X} \\ &= \lambda\left(\mathbf{R}'\otimes\boldsymbol{\tau}_T'\right)\mathbf{X} \\ &= \lambda NT\left(\frac{1}{N}\sum_i\mathbf{r}_i\bar{\mathbf{x}}_i'\right).\end{aligned}$$

Furthermore, $\mathbf{Q}_{r\bar{x}}=\frac{1}{N}\sum_i\mathbf{r}_i\bar{\mathbf{x}}_i'-\bar{\mathbf{r}}\bar{\mathbf{x}}'$, $\mathbf{X}'\mathbf{Q}_V\left(\boldsymbol{\tau}_N\otimes\boldsymbol{\tau}_T\right)=\mathbf{0}$, $\mathbf{X}'\mathbf{Q}_V\left(\mathbf{Z}\otimes\boldsymbol{\tau}_T\right)=\mathbf{0}$, and

$$\begin{aligned}\mathbf{X}'\mathbf{Q}_V\boldsymbol{\Omega}^{-1/2}\mathbf{X} &= \mathbf{X}'\mathbf{Q}_V\left(\lambda\mathbf{I}_{NT}+\psi\mathbf{Q}_V\right)\mathbf{X} \\ &= \left(\lambda+\psi\right)\mathbf{X}'\mathbf{Q}_V\mathbf{X}'=\frac{1}{\sigma_\varepsilon}\mathbf{X}'\mathbf{Q}_V\mathbf{X}'.\end{aligned}$$

Then by using the above results, (5) reduces to

$$\mathbf{A}^{*\prime}\boldsymbol{\Omega}^{-1/2}\mathbf{W}=\lambda NT\begin{pmatrix} 1 & \bar{\mathbf{z}}' & \bar{\mathbf{x}}' \\ \bar{\mathbf{r}} & \frac{1}{N}\sum_i\mathbf{r}_i\mathbf{z}_i' & \frac{1}{N}\sum_i\mathbf{r}_i\bar{\mathbf{x}}_i' \\ 0 & 0 & \frac{1}{\varphi}\mathbf{Q}_{FE,NT} \end{pmatrix}.$$

Hence

$$
\begin{aligned}
& \left(\mathbf{A}^{*\prime}\mathbf{A}^{*}\right)^{-1}\mathbf{A}^{*\prime}\boldsymbol{\Omega}^{-1/2}\mathbf{W} \\
= \; & \lambda \begin{pmatrix} 1+\bar{\mathbf{r}}'\mathbf{Q}_{rr,N}^{-1}\bar{\mathbf{r}} & -\bar{\mathbf{r}}'\mathbf{Q}_{rr,N}^{-1} & 0 \\ -\mathbf{Q}_{rr,N}^{-1}\bar{\mathbf{r}} & \mathbf{Q}_{rr,N}^{-1} & 0 \\ 0 & 0 & \mathbf{Q}_{FE,NT}^{-1} \end{pmatrix} \begin{pmatrix} 1 & \bar{\mathbf{z}}' & \bar{\mathbf{x}}' \\ \bar{\mathbf{r}} & \frac{1}{N}\sum_i \mathbf{r}_i\mathbf{z}_i' & \frac{1}{N}\sum_i \mathbf{r}_i\bar{\mathbf{x}}_i' \\ 0 & 0 & \mathbf{Q}_{FE,NT} \end{pmatrix} \\
= \; & \lambda \begin{pmatrix} 1 & \bar{\mathbf{z}}'+\bar{\mathbf{r}}'\mathbf{Q}_{rr,N}^{-1}\bar{\mathbf{r}}\bar{\mathbf{z}}'-\bar{\mathbf{r}}'\mathbf{Q}_{rr,N}^{-1}\frac{1}{N}\sum_i \mathbf{r}_i\mathbf{z}_i' & \bar{\mathbf{x}}'+\bar{\mathbf{r}}'\mathbf{Q}_{rr,N}^{-1}\bar{\mathbf{r}}\bar{\mathbf{x}}'-\bar{\mathbf{r}}'\mathbf{Q}_{rr,N}^{-1}\frac{1}{N}\sum_i \mathbf{r}_i\bar{\mathbf{x}}_i' \\ 0 & -\mathbf{Q}_{rr,N}^{-1}\bar{\mathbf{r}}\bar{\mathbf{z}}'+\mathbf{Q}_{rr,N}^{-1}\frac{1}{N}\sum_i \mathbf{r}_i\mathbf{z}_i' & -\mathbf{Q}_{rr,N}^{-1}\bar{\mathbf{r}}\bar{\mathbf{x}}'+\mathbf{Q}_{rr,N}^{-1}\frac{1}{N}\sum_i \mathbf{r}_i\bar{\mathbf{x}}_i' \\ 0 & 0 & \frac{1}{\varphi}\mathbf{I}_k \end{pmatrix},
\end{aligned}
$$

where

$$
\begin{aligned}
\bar{\mathbf{z}}'+\bar{\mathbf{r}}'\mathbf{Q}_{rr,N}^{-1}\bar{\mathbf{r}}\bar{\mathbf{z}}'-\bar{\mathbf{r}}'\mathbf{Q}_{rr,N}^{-1}\frac{1}{N}\sum_i \mathbf{r}_i\mathbf{z}_i' &= \bar{\mathbf{z}}'-\bar{\mathbf{r}}'\mathbf{Q}_{rr,N}^{-1}\left(\frac{1}{N}\sum_i \mathbf{r}_i\mathbf{z}_i'-\bar{\mathbf{r}}\bar{\mathbf{z}}'\right) \\
&= \bar{\mathbf{z}}'-\bar{\mathbf{r}}'\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{rz,N},
\end{aligned}
$$

with $\mathbf{Q}_{rz,N}=\frac{1}{N}\sum_i\left(\mathbf{r}_i-\bar{\mathbf{r}}\right)\left(\mathbf{z}_i-\bar{\mathbf{z}}\right)'=\left(\frac{1}{N}\sum_i \mathbf{r}_i\mathbf{z}_i'-\bar{\mathbf{r}}\bar{\mathbf{z}}'\right)$, and

$$
\begin{aligned}
\bar{\mathbf{x}}'+\bar{\mathbf{r}}'\mathbf{Q}_{rr,N}^{-1}\bar{\mathbf{r}}\bar{\mathbf{x}}'-\bar{\mathbf{r}}'\mathbf{Q}_{rr,N}^{-1}\frac{1}{N}\sum_i \mathbf{r}_i\bar{\mathbf{x}}_i' &= \bar{\mathbf{x}}'-\bar{\mathbf{r}}'\mathbf{Q}_{rr,N}^{-1}\left(\frac{1}{N}\sum_i \mathbf{r}_i\bar{\mathbf{x}}_i'-\bar{\mathbf{r}}\bar{\mathbf{x}}'\right) \\
&= \bar{\mathbf{x}}'-\bar{\mathbf{r}}'\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{r\bar{x},N},
\end{aligned}
$$

with $\mathbf{Q}_{r\bar{x},N}=\frac{1}{N}\sum_i\left(\mathbf{r}_i-\bar{\mathbf{r}}\right)\left(\bar{\mathbf{x}}_i-\bar{\mathbf{x}}\right)'=\frac{1}{N}\sum_i \mathbf{r}_i\bar{\mathbf{x}}_i'-\bar{\mathbf{r}}\bar{\mathbf{x}}'$. As a result, we obtain

$$
\left(\mathbf{A}^{*\prime}\mathbf{A}^{*}\right)^{-1}\mathbf{A}^{*\prime}\boldsymbol{\Omega}^{-1/2}\mathbf{W}=\lambda\begin{pmatrix} 1 & \bar{\mathbf{z}}'-\bar{\mathbf{r}}'\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{rz,N} & \bar{\mathbf{x}}'-\bar{\mathbf{r}}'\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{r\bar{x},N} \\ 0 & \mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{rz,N} & \mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{r\bar{x},N} \\ 0 & 0 & \frac{1}{\varphi}\mathbf{I}_k \end{pmatrix},
$$

and

$$
\begin{aligned}
& \mathbf{W}'\boldsymbol{\Omega}^{-1/2}\mathbf{P}_{A^*}\boldsymbol{\Omega}^{-1/2}\mathbf{W} \\
= \; & \lambda^2 NT \begin{pmatrix} 1 & \bar{\mathbf{r}}' & 0 \\ \bar{\mathbf{z}} & \frac{1}{N}\sum_i \mathbf{z}_i\mathbf{r}_i' & 0 \\ \bar{\mathbf{x}} & \frac{1}{N}\sum_i \bar{\mathbf{x}}_i\mathbf{r}_i' & \frac{1}{\varphi}\mathbf{Q}_{FE,NT} \end{pmatrix} \begin{pmatrix} 1 & \bar{\mathbf{z}}'-\bar{\mathbf{r}}'\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{rz,N} & \bar{\mathbf{x}}'-\bar{\mathbf{r}}'\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{r\bar{x},N} \\ 0 & \mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{rz,N} & \mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{r\bar{x},N} \\ 0 & 0 & \frac{1}{\varphi}\mathbf{I}_k \end{pmatrix} \\
= \; & \lambda^2 NT \begin{pmatrix} 1 & \bar{\mathbf{z}}' & \bar{\mathbf{x}}' \\ \bar{\mathbf{z}} & \bar{\mathbf{z}}\bar{\mathbf{z}}'+\mathbf{Q}_{rz,N}'\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{rz,N} & \bar{\mathbf{z}}\bar{\mathbf{x}}'+\mathbf{Q}_{rz,N}'\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{r\bar{x},N} \\ \bar{\mathbf{x}} & \bar{\mathbf{x}}\bar{\mathbf{z}}'+\mathbf{Q}_{r\bar{x},N}'\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{rz,N} & \bar{\mathbf{x}}\bar{\mathbf{x}}'+\mathbf{Q}_{r\bar{x},N}'\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{r\bar{x},N}+\frac{1}{\varphi^2}\mathbf{Q}_{FE,NT} \end{pmatrix} \\
= \; & \lambda^2 NT \begin{pmatrix} 1 & \bar{\mathbf{z}}' & \bar{\mathbf{x}}' \\ \bar{\mathbf{z}} & \bar{\mathbf{z}}\bar{\mathbf{z}}'+\mathbf{F} & \bar{\mathbf{z}}\bar{\mathbf{x}}'+\mathbf{G} \\ \bar{\mathbf{x}} & \bar{\mathbf{x}}\bar{\mathbf{z}}'+\mathbf{G}' & \bar{\mathbf{x}}\bar{\mathbf{x}}'+\mathbf{H}+\frac{1}{\varphi^2}\mathbf{Q}_{FE,NT} \end{pmatrix}, \qquad (6)
\end{aligned}
$$

where

$$\begin{aligned}
\mathbf{F} &= \mathbf{Q}'_{rz,N}\mathbf{Q}^{-1}_{rr,N}\mathbf{Q}_{rz,N}, \mathbf{G} = \mathbf{Q}'_{rz,N}\mathbf{Q}^{-1}_{rr,N}\mathbf{Q}_{r\bar{x},N}, \\
\mathbf{H} &= \mathbf{Q}'_{r\bar{x},N}\mathbf{Q}^{-1}_{rr,N}\mathbf{Q}_{r\bar{x},N}.
\end{aligned}$$

To simplify the derivation, suppose $\bar{\mathbf{z}} = \mathbf{0}, \bar{\mathbf{x}} = \mathbf{0}$, but $\bar{\mathbf{x}}_i \neq \mathbf{0}$ for each i. Then

$$\left(\mathbf{W}'\mathbf{\Omega}^{-1/2}\mathbf{P}_A\mathbf{\Omega}^{-1/2}\mathbf{W}\right)^{-1} = \frac{1}{\lambda^2 NT}\begin{pmatrix} 1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{F} & \mathbf{G} \\ \mathbf{0} & \mathbf{G}' & \mathbf{D} \end{pmatrix}^{-1},$$

where

$$\mathbf{D} = \mathbf{Q}'_{r\bar{x},N}\mathbf{Q}^{-1}_{rr,N}\mathbf{Q}_{r\bar{x},N} + \frac{1}{\varphi^2}\mathbf{Q}_{FE,NT}.$$

Now using the inverse for partitioned matrices, we have

$$Asy.Var\left(\sqrt{NT}\hat{\boldsymbol{\beta}}_{HTM}\right) = \frac{1}{\lambda^2}\lim_{N\to\infty}\left(\mathbf{D} - \mathbf{G}'\mathbf{F}^{-1}\mathbf{G}\right)^{-1},$$

and

$$\begin{aligned}
Asy.Var\left(\sqrt{NT}\hat{\boldsymbol{\gamma}}_{HTM}\right) &= \frac{1}{\lambda^2}\lim_{N\to\infty}\left(\mathbf{F} - \mathbf{G}\mathbf{D}^{-1}\mathbf{G}'\right)^{-1} \\
&= \frac{1}{\lambda^2}\lim_{N\to\infty}\left(\mathbf{F}^{-1} + \mathbf{F}^{-1}\mathbf{G}'\left(\mathbf{D} - \mathbf{G}'\mathbf{F}^{-1}\mathbf{G}\right)^{-1}\mathbf{G}\mathbf{F}^{-1}\right) \\
&= \frac{1}{\lambda^2}\lim_{N\to\infty}\left(\mathbf{F}^{-1} + \mathbf{F}^{-1}\mathbf{G}'\left[Asy.Var\left(\sqrt{NT}\hat{\boldsymbol{\beta}}_{HTM}\right)\right]\mathbf{G}\mathbf{F}^{-1}\right) \\
&= \frac{1}{\lambda^2}\lim_{N\to\infty}\mathbf{F}^{-1} + \frac{1}{\lambda^2}\lim_{N\to\infty}\mathbf{F}^{-1}\mathbf{G}'\left[Asy.Var\left(\sqrt{NT}\hat{\boldsymbol{\beta}}_{HTM}\right)\right]\mathbf{G}\mathbf{F}^{-1},
\end{aligned}$$

or

$$\begin{aligned}
Asy.Var\left(\sqrt{N}\hat{\boldsymbol{\gamma}}_{HTM}\right) &= \frac{1}{\lambda^2 T}\lim_{N\to\infty}\mathbf{F}^{-1} + \lim_{N\to\infty}\mathbf{F}^{-1}\mathbf{G}'\left[Asy.Var\left(\sqrt{N}\hat{\boldsymbol{\beta}}_{HTM}\right)\right]\mathbf{G}\mathbf{F}^{-1} \qquad (7) \\
&= \left(\frac{\sigma^2_\varepsilon}{T} + \sigma^2_\eta\right)\lim_{N\to\infty}\mathbf{F}^{-1} + \lim_{N\to\infty}\mathbf{F}^{-1}\mathbf{G}'\left[Asy.Var\left(\sqrt{N}\hat{\boldsymbol{\beta}}_{HTM}\right)\right]\mathbf{G}\mathbf{F}^{-1}.
\end{aligned}$$

It is easily verified that this expression is the same as $Asy.Var\left(\sqrt{N}\hat{\boldsymbol{\gamma}}_{FEF-IV}\right)$ given by equation (51) of the paper, apart from the choice of formula for $Asy.Var\left(\sqrt{N}\hat{\boldsymbol{\beta}}_{HTM}\right)$.

Now using relevant partitioned inverse of $\mathbf{W}'\mathbf{\Omega}^{-1/2}\mathbf{P}_{A^*}\mathbf{\Omega}^{-1/2}\mathbf{W}$, we obtain

$$Asy.Var\left(\sqrt{N}\hat{\boldsymbol{\beta}}_{HTM}\right) = \left(\frac{\sigma^2_\varepsilon}{T} + \sigma^2_\eta\right)\left\{\begin{array}{c} \mathbf{Q}'_{r\bar{x}}\mathbf{Q}^{-1}_{rr}\mathbf{Q}_{r\bar{x}} + \frac{1}{\varphi^2}\mathbf{Q}_{FE} \\ -\mathbf{Q}'_{r\bar{x}}\mathbf{Q}^{-1}_{rr}\mathbf{Q}_{rz}\left(\mathbf{Q}'_{rz}\mathbf{Q}^{-1}_{rr}\mathbf{Q}_{rz}\right)^{-1}\left(\mathbf{Q}'_{rz}\mathbf{Q}^{-1}_{rr}\mathbf{Q}_{r\bar{x}}\right) \end{array}\right\}^{-1},$$

which is not the same as $Asy.Var\left(\sqrt{N}\hat{\boldsymbol{\beta}}_{FE}\right)$. But in the case where $m = s$ and Q_{rz}^{-1} exists, we have

$$\begin{aligned} Asy.Var\left(\sqrt{N}\hat{\boldsymbol{\beta}}_{HTM}\right) &= \left(\frac{\sigma_\varepsilon^2}{T}+\sigma_\eta^2\right)\left\{\begin{array}{c} \mathbf{Q}_{r\bar{x}}'\mathbf{Q}_{rr}^{-1}\mathbf{Q}_{r\bar{x}}+\frac{1}{\varphi^2}\mathbf{Q}_{FE} \\ -\mathbf{Q}_{r\bar{x}}'\mathbf{Q}_{rr}^{-1}\mathbf{Q}_{rz}\left(\mathbf{Q}_{rz}'\mathbf{Q}_{rr}^{-1}\mathbf{Q}_{rz}\right)^{-1}\left(\mathbf{Q}_{rz}'\mathbf{Q}_{rr}^{-1}\mathbf{Q}_{r\bar{x}}\right) \end{array}\right\}^{-1} \\ &= \left(\frac{\sigma_\varepsilon^2}{T}+\sigma_\eta^2\right)\varphi^2\mathbf{Q}_{FE}^{-1} \\ &= \frac{\sigma_\varepsilon^2}{T}\mathbf{Q}_{FE}^{-1} = Asy.Var\left(\sqrt{N}\hat{\boldsymbol{\beta}}_{FE}\right). \end{aligned}$$

Therefore, it follows that

$$Asy.Var\left(\sqrt{N}\hat{\boldsymbol{\beta}}_{HTM}\right) = Asy.Var\left(\sqrt{N}\hat{\boldsymbol{\beta}}_{FE}\right), \tag{8}$$

if the errors ε_{it} are homoskedastic and serially uncorrelated, and if $\boldsymbol{\gamma}$ is exactly identified, namely if $m = s$, with $\mathbf{Q}_{rz}$ being nonsingular. The same result holds if

$$\mathbf{Q}_{r\bar{x}} = \lim_{N\to\infty}\frac{1}{N}\sum_i\left(\mathbf{r}_i-\bar{\mathbf{r}}\right)\left(\bar{\mathbf{x}}_i-\bar{\mathbf{x}}\right)' = \mathbf{0},$$

namely, if $\bar{\mathbf{z}} = \mathbf{0}, \bar{\mathbf{x}} = \mathbf{0}$, and $\bar{\mathbf{x}}_i$ and $\mathbf{r}_i$ are uncorrelated.

As a result, by comparing (7) with equation (50) in the paper together with (8), we have

$$Asy.Var\left(\sqrt{N}\hat{\boldsymbol{\gamma}}_{FEF-IV}\right) = Asy.Var\left(\sqrt{N}\hat{\boldsymbol{\gamma}}_{HTM}\right),$$

as required. The above result holds since when $\mathbf{Q}_{r\bar{x}} = \mathbf{0}$, and $\boldsymbol{\gamma}$ is exactly identified, the expression for $\boldsymbol{\Omega}_{\hat{\gamma}_{FEF-IV}}$ simplifies to

$$\boldsymbol{\Omega}_{\hat{\gamma}_{FEF-IV}} = \left(\sigma_\eta^2+\frac{\sigma_\varepsilon^2}{T}\right)\left(\mathbf{Q}_{zr}\mathbf{Q}_{rr}^{-1}\mathbf{Q}_{zr}'\right)^{-1}+\frac{\sigma_\varepsilon^2}{T}\mathbf{Q}_{zr}'^{-1}\mathbf{Q}_{r\bar{x}}\mathbf{Q}_{FE,T}^{-1}\mathbf{Q}_{r\bar{x}}'\mathbf{Q}_{zr}^{-1}.$$

Also, for the modified HT estimator (2), since

$$\boldsymbol{\Omega}^{-1/2}\mathbf{y} = \left(\lambda\mathbf{I}_{NT}+\psi\mathbf{Q}_V\right)\mathbf{y} = \lambda\mathbf{y}+\psi\mathbf{Q}_V\mathbf{y},$$

then

$$\begin{aligned} \mathbf{y}'\boldsymbol{\Omega}^{-1/2}\mathbf{A}^* &= \left(\lambda\mathbf{y}'+\psi\mathbf{y}'\mathbf{Q}_V\right)\left(\left(\boldsymbol{\tau}_N\otimes\boldsymbol{\tau}_T\right),\mathbf{R}\otimes\boldsymbol{\tau}_T,\mathbf{Q}_V\mathbf{X}\right) \\ &= \left(\lambda NT\bar{y},\lambda\mathbf{y}'\left(\mathbf{R}\otimes\boldsymbol{\tau}_T\right),\left(\lambda+\psi\right)\mathbf{y}'\mathbf{Q}_V\mathbf{X}\right) \\ &= \lambda NT\left(\bar{y},\frac{1}{N}\sum_i\bar{y}_i\mathbf{r}_i,\left(1+\frac{\psi}{\lambda}\right)\frac{1}{NT}\mathbf{y}'\mathbf{Q}_V\mathbf{X}\right) \\ &= \lambda NT\left(\bar{y},\frac{1}{N}\sum_i\bar{y}_i\mathbf{r}_i,\frac{1}{\varphi}\frac{1}{NT}\mathbf{y}'\mathbf{Q}_V\mathbf{X}\right). \end{aligned}$$

Hence, using the above derivations, we have

$$\begin{aligned}
&\mathbf{y}'\mathbf{\Omega}^{-1/2}\mathbf{A}^*\left(\mathbf{A}^{*\prime}\mathbf{A}^*\right)^{-1}\mathbf{A}^{*\prime}\mathbf{\Omega}^{-1/2}\mathbf{W} \\
=\;& \lambda^2 NT\left(\bar{y}, \frac{\sum_i y_i\mathbf{r}_i}{N}, \frac{1}{\varphi}\frac{\mathbf{y}'\mathbf{Q}_V\mathbf{X}}{NT}\right)\begin{pmatrix} 1 & \bar{\mathbf{z}}' - \bar{\mathbf{r}}'\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{rz,N} & \bar{\mathbf{x}}' - \bar{\mathbf{r}}'\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{r\bar{x},N} \\ 0 & \mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{rz,N} & \mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{r\bar{x},N} \\ 0 & 0 & \frac{1}{\varphi}\mathbf{I}_k \end{pmatrix} \\
=\;& \lambda^2 NT\begin{pmatrix} \bar{y},\ \bar{y}\bar{\mathbf{z}}' - \bar{y}\bar{\mathbf{r}}'\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{rz,N} + \frac{\sum_i y_i\mathbf{r}_i}{N}\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{rz,N},\ \bar{y}\bar{\mathbf{x}}' - \bar{y}\bar{\mathbf{r}}'\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{r\bar{x},N} \\ +\frac{\sum_i y_i\mathbf{r}_i}{N}\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{r\bar{x},N} + \frac{1}{\varphi^2}\frac{\mathbf{y}'\mathbf{Q}_V\mathbf{X}}{NT} \end{pmatrix} \\
=\;& \lambda^2 NT\left(\bar{y},\ \bar{y}\bar{\mathbf{z}}' + \mathbf{Q}_{\bar{y}r,N}\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{rz,N},\ \bar{y}\bar{\mathbf{x}}' + \mathbf{Q}_{\bar{y}r,N}\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{r\bar{x},N} + \frac{1}{\varphi^2}\frac{\mathbf{y}'\mathbf{Q}_V\mathbf{X}}{NT}\right),
\end{aligned}$$

where

$$\mathbf{Q}_{\bar{y}r,N} = \frac{1}{N}\sum_i\left(\bar{y}_i - \bar{y}\right)\left(\mathbf{r}_i - \bar{\mathbf{r}}\right)'.$$

Hence, under the normalization $\bar{\mathbf{z}} = \mathbf{0}$ and $\bar{\mathbf{x}} = \mathbf{0}$, we have

$$\mathbf{W}'\mathbf{\Omega}^{-1/2}\mathbf{A}^*\left(\mathbf{A}^{*\prime}\mathbf{A}^*\right)^{-1}\mathbf{A}^{*\prime}\mathbf{\Omega}^{-1/2}\mathbf{y} = \lambda^2 NT\begin{pmatrix} \bar{y} \\ \left(\mathbf{Q}_{\bar{y}r,N}\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{rz,N}\right)' \\ \left(\mathbf{Q}_{\bar{y}r,N}\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{r\bar{x},N}\right)' + \frac{1}{\varphi^2}\frac{\mathbf{X}'\mathbf{Q}_V\mathbf{y}}{NT} \end{pmatrix}.$$

Using this result together with (6) now yields

$$\begin{aligned}
\hat{\boldsymbol{\theta}}_{HTM} =\;& \begin{pmatrix} 1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_{rz,N}'\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{rz,N} & \mathbf{Q}_{rz,N}'\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{r\bar{x},N} \\ \mathbf{0} & \left(\mathbf{Q}_{rz,N}'\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{r\bar{x},N}\right)' & \mathbf{Q}_{r\bar{x},N}'\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{r\bar{x},N} + \frac{1}{\varphi^2}\mathbf{Q}_{FE,NT} \end{pmatrix}^{-1} \\
&\times\begin{pmatrix} \bar{y} \\ \left(\mathbf{Q}_{\bar{y}r,N}\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{rz,N}\right)' \\ \left(\mathbf{Q}_{\bar{y}r,N}\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{r\bar{x},N}\right)' + \frac{1}{\varphi^2}\frac{\mathbf{X}'\mathbf{Q}_V\mathbf{y}}{NT} \end{pmatrix}.
\end{aligned}$$

Therefore, in general, $\hat{\boldsymbol{\beta}}_{HTM}$ and $\hat{\boldsymbol{\gamma}}_{HTM}$ will be different from $\hat{\boldsymbol{\beta}}_{FE}$ and $\hat{\boldsymbol{\gamma}}_{FEF-IV}$. However, $\hat{\boldsymbol{\beta}}_{HTM} = \hat{\boldsymbol{\beta}}_{FE}$, if $\bar{\mathbf{z}} = 0$ and $\bar{\mathbf{x}} = 0$, and $\mathbf{Q}_{r\bar{x},N} = 0$. Under these conditions it follows that

$$\begin{aligned}
\hat{\alpha}_{HTM} &= \bar{y},\ \hat{\boldsymbol{\beta}}_{HTM} = \hat{\boldsymbol{\beta}}_{FE}, \\
\hat{\boldsymbol{\gamma}}_{HTM} &= \left(\mathbf{Q}_{rz,N}'\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{rz,N}\right)^{-1}\left(\mathbf{Q}_{rz,N}'\mathbf{Q}_{rr,N}^{-1}\mathbf{Q}_{r\bar{y},N}\right).
\end{aligned}$$

B. A full set of Monte Carlo simulation results

The DGP considered in the paper is given by

$$\begin{aligned} y_{it} &= 1+\alpha_i+\beta_1 x_{it,1}+\beta_2 x_{it,2}+\gamma_1 z_{i1}+\gamma_2 z_{i2}+\varepsilon_{it}, \\ i &= 1,2,\ldots,N,\ t=1,2,\ldots,T, \end{aligned}$$

with $\beta_1=\beta_2=1$ and $\gamma_1=\gamma_2=1$. An intercept is included, and hence without loss of generality we generate the regressors with zero means. For the time-varying regressors we consider the following relatively general specifications

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

where the time effects g_{1t} and g_{2t} are generated as $U(1,2)$ and are then kept fixed across the replications. Note that

$$\bar{x}_{i,j}=\alpha_i\bar{g}_j+\bar{w}_{i,j},\ \text{for } j=1,2,$$

where $\bar{x}_{i,j}=T^{-1}\sum_{t=1}^{T}x_{it,j}$, $\bar{w}_{ij}=T^{-1}\sum_{t=1}^{T}w_{it,j}$, and $\bar{g}_j=T^{-1}\sum_{t=1}^{T}g_{jt}$. We generate the fixed effects as $\alpha_i\sim 0.5\left(\chi^2(2)-2\right)$, for $i=1,2,\ldots,N$. The stochastic components of the time varying regressors ($w_{it,1}$ and $w_{it,2}$) are generated as heterogenous stationary $AR(1)$ processes

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

where

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

The above DGP allows the individual-specific means, $\bar{x}_{i,j}$, to be non-zero. Note also that the time-varying regressors, $x_{it,j}$, are correlated with the individual effects, α_i, since $\bar{g}_j\neq 0$, for $j=1$ and 2.

The time-invariant regressors, z_{ji}, for $j=1,2$, are generated as

$$\mathbf{z}_i=\begin{pmatrix}1\\1\end{pmatrix}+\mathbf{\Lambda}\bar{\mathbf{w}}_i+\alpha_i\boldsymbol{\phi}+\boldsymbol{\zeta}_i,$$

where $z_i=(z_{i1},z_{i2})'$, $\bar{w}_i=(\bar{w}_{i1},\bar{w}_{i2})'$, and $\zeta_i\sim IIDN\left(\mathbf{0},\mathbf{I}_2\right)$, and

$$\mathbf{\Lambda}=\begin{pmatrix}1&1\\1&1\end{pmatrix}.$$

We consider both cases where the time-invariant regressors, z_i, are exogenous (DGP A) and when they are endogenous (DGP B and C). Under DGP A we set $\boldsymbol{\phi} = (\phi_1, \phi_2)' = \mathbf{0}$, and under DGP B and C we set $\boldsymbol{\phi} = (\phi_1, \phi_2)' = (1, 1)'$. We also assume the 4×1 vector of instruments $\mathbf{r}_i = (r_{i1}, r_{i2}, r_{i3}, r_{i4})'$ are generated as

$$\mathbf{r}_i = \boldsymbol{\Gamma}_\zeta \boldsymbol{\zeta}_i + \boldsymbol{\Gamma}_w \bar{\mathbf{w}}_i + \boldsymbol{\xi}_i, \tag{9}$$

where $\boldsymbol{\xi}_i \sim IIDN\left(\mathbf{0}, \mathbf{I}_4\right)$, with

$$\boldsymbol{\Gamma}_\zeta = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix},$$

and distinguish between DGP B and C by different choices of Γ_w, namely under DGP B we set $\Gamma_w = 0$, and under DGP C we set

$$\boldsymbol{\Gamma}_w = 10 \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Clearly, DGP B and C are relevant only when the time-invariant regressors, z_i, are endogenous. Under DGP B, $Cov(\mathbf{r}_i, \bar{\mathbf{x}}_i) = \mathbf{Q}_{r\bar{x}} = 0$, and under DGP C we have $\mathbf{Q}_{r\bar{x}} = \Gamma_w \neq 0$. Given our theoretical derivations we expect the modified HT and FEF-IV estimators to perform very similarly under DGP B.

For each of the above three DGPs (A, B and C) we consider three different processes for the idiosyncratic errors, ε_{it}:

Case 1: Homoskedastic errors:

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

Case 2: Heteroskedastic errors:

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

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

Case 3: Serially correlated and heteroskedastic 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. We consider the simulation of combinations of $N = 500, 1000, 2000$ and $T = 3, 5, 10$.

A complete set of simulation results are given in Tables B1-B18.

Table B1: Bias, RMSE, size and power of FEF and FEVD estimators for γ_1 in the case of DGP with exogenous time-invariant regressors (DGP A) and homoskedastic and serially uncorrelated errors (Case 1)

N	T	3			5			10		
		FEF	FEVD		FEF	FEVD		FEF	FEVD	
			without	with		without	with		without	with
500	Bias	0.0005	-0.1672	0.0005	-0.0001	-0.1732	-0.0001	0.0012	-0.1829	0.0012
	RMSE	0.0420	0.1830	0.0420	0.0391	0.1893	0.0391	0.0367	0.1972	0.0137
	size	5.7%	91%	44%	5.7%	95%	48%	6.2%	98%	55%
	power	22%	75%	69%	27%	83%	72%	31%	91%	82%
1000	Bias	0.0002	-0.1791	0.0002	0.0012	-0.1855	0.0012	-0.0005	-0.1950	-0.0005
	RMSE	0.0275	0.1864	0.0275	0.0257	0.1927	0.0257	0.0244	0.2016	0.0244
	size	4.7%	99%	40%	4.8%	100%	45%	5.3%	100%	53%
	power	42%	95%	86%	50%	96%	90%	52%	98%	92%
2000	bias	-0.0011	-0.1562	-0.0011	0.0001	-0.1612	0.0001	0.0001	-0.1674	0.00001
	RMSE	0.0203	0.1606	0.0203	0.0194	0.1649	0.0194	0.0181	0.1707	0.0181
	size	5.1%	100%	44%	6.3%	100%	48%	6.2%	100%	57%
	power	68%	97%	95%	76%	99%	98%	80%	100%	99%

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. 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.

Table B2: Bias, RMSE, size and power of FEF and FEVD estimators for γ_2 in the case of DGP with exogenous time-invariant regressors (DGP A) and homoskedastic and serially uncorrelated errors (Case 1)

N	T	3			5			10		
		FEF	FEVD		FEF	FEVD		FEF	FEVD	
			without	with		without	with		without	with
500	Bias	-0.0017	-0.1697	-0.0017	-0.0006	-0.1739	-0.0006	-0.0007	-0.1805	-0.0007
	RMSE	0.0411	0.1854	0.0411	0.0376	0.1892	0.0376	0.0356	0.1946	0.0356
	size	5.7%	92%	43%	4.2%	95%	48%	4.7%	98%	57%
	power	21%	75%	68%	27%	83%	73%	30%	90%	81%
1000	Bias	-0.0009	-0.1836	-0.0009	-0.0002	-0.1877	-0.0002	-0.0002	-0.1963	-0.0002
	RMSE	0.0283	0.1909	0.0283	0.0256	0.1946	0.0256	0.0238	0.2028	0.0238
	size	4.5%	100%	41%	4.1%	100%	45%	4.4%	100%	52%
	power	37%	95%	84%	48%	98%	89%	50%	99%	94%
2000	bias	0.0008	-0.1547	0.0008	0.0003	-0.1600	0.0003	-0.0001	-0.1684	-0.0001
	RMSE	0.0202	0.1589	0.0202	0.0188	0.1637	0.0188	0.0172	0.1716	0.0172
	size	5.5%	100%	41%	6%	100%	47%	4.4%	100%	52%
	power	73%	98%	97%	77%	99%	97%	83%	99%	98%

Notes: 1. Size is calculated under $\gamma_2^{(0)} = 1$, and power under $\gamma_2^{(1)} = 0.95$.

2. See also the notes 2-3 of Table B1.

Table B3: Bias, RMSE, size and power of FEF and FEVD estimators for γ_1 in the case of DGP with exogenous time-invariant regressors (DGP A) and heteroskedastic and serially uncorrelated errors (Case 2)

N	T	3			5			10		
		FEF	FEVD		FEF	FEVD		FEF	FEVD	
			without	with		without	with		without	with
500	Bias	0.0006	-0.1665	0.0006	0.0006	-0.1698	0.0006	0.0018	-0.1818	0.0018
	RMSE	0.0433	0.1829	0.0433	0.0375	0.1851	0.0375	0.0169	0.1946	0.0169
	size	6.5%	95%	44%	5.1%	97%	46%	41.%	99%	54%
	power	23%	85%	69%	27%	88%	74%	33%	95%	84%
1000	Bias	-0.0004	-0.1813	-0.0004	-0.0014	-0.1885	-0.0014	0.0006	-0.1975	0.0006
	RMSE	0.0282	0.1891	0.0282	0.0261	0.1954	0.0261	0.0256	0.2041	0.0256
	size	4.3%	100%	42%	4.6%	100%	47%	6%	100%	55%
	power	39%	97%	84%	47%	100%	87%	54%	100%	92%
2000	bias	0.0003	-0.1540	0.0003	0.0009	-0.1591	0.0009	-0.0012	-0.1697	-0.0012
	RMSE	0.0197	0.1584	0.0197	0.0184	0.1628	0.0184	0.0175	0.1732	0.0175
	size	4.8%	100%	42%	5.1%	100%	45%	4.8%	100%	52%
	power	71%	100%	97%	78%	100%	98%	80%	100%	99%

Notes: see notes 1-3 of Table B1.

Table B4: Bias, RMSE, size and power of FEF and FEVD estimators for γ_2 in the case of DGP with exogenous time-invariant regressors (DGP A) and heteroskedastic and serially uncorrelated errors (Case 2)

N	T	3			5			10		
		FEF	FEVD		FEF	FEVD		FEF	FEVD	
			without	with		without	with		without	with
500	Bias	0.0003	-0.1684	0.0003	-0.0008	-0.1756	-0.0008	-0.0006	-0.1803	-0.0006
	RMSE	0.0419	0.1844	0.0419	0.0373	0.1904	0.0373	0.0344	0.1930	0.0344
	size	5.7%	95%	43%	5.2%	97%	46%	42.%	99%	54%
	power	24%	84%	69%	26%	90%	73%	29%	94%	81%
1000	Bias	-0.0003	-0.1824	-0.0003	0.0007	-0.1852	0.0007	-0.0008	-0.1942	-0.0008
	RMSE	0.0285	0.1903	0.0285	0.0261	0.1920	0.0261	0.0247	0.2010	0.0247
	size	5.1%	100%	42%	5.4%	100%	44%	4.8%	100%	54%
	power	40%	97%	83%	47%	100%	90%	51%	99%	94%
2000	bias	-0.0003	-0.1554	-0.0003	-0.0012	-0.1624	-0.0012	0.0003	-0.1660	0.0003
	RMSE	0.0196	0.1596	0.0196	0.0192	0.1662	0.0192	0.0176	0.1694	0.0176
	size	4.4%	100%	43%	5.7%	100%	48%	5.2%	100%	53%
	power	71%	100%	96%	75%	100%	97%	82%	100%	99%

Notes: see notes 1-2 of Table B2.

Table B5: Bias, RMSE, size and power of FEF and FEVD estimators for γ_1 in the case of DGP with exogenous time-invariant regressors (DGP A) and serially correlated errors (Case 3)

N	T	3			5			10		
		FEF	FEVD		FEF	FEVD		FEF	FEVD	
			without	with		without	with		without	with
500	Bias	0.0016	-0.1668	0.0016	-0.0003	-0.1721	-0.0003	-0.0008	-0.1850	-0.0008
	RMSE	0.0419	0.1836	0.0419	0.0400	0.1874	0.0400	0.0371	0.1998	0.0371
	size	3.7%	91%	58%	4.2%	94%	60%	4.6%	98%	64%
	power	22%	76%	78%	23%	84%	80%	26%	92%	87%
1000	Bias	-0.0024	-0.1861	-0.0024	-0.0013	-0.1888	-0.0013	-0.0002	-0.1977	-0.0002
	RMSE	0.0299	0.1936	0.0299	0.0288	0.1961	0.0288	0.0282	0.2049	0.0282
	size	5.4%	100%	56%	5.5%	100%	58%	5.9%	100%	64%
	power	32%	96%	87%	39%	98%	89%	46%	100%	92%
2000	bias	-0.0004	-0.1548	-0.0004	0.0007	-0.1607	0.0007	0.0005	-0.1658	0.0005
	RMSE	0.0223	0.1590	0.0223	0.0210	0.1646	0.0210	0.0189	0.1694	0.0189
	size	5.6%	100%	60%	6%	100%	62%	3.4%	100%	66%
	power	62%	98%	96%	69%	100%	98%	74%	100%	100%

Notes: see notes 1-3 of Table B1.

Table B6: Bias, RMSE, size and power of FEF and FEVD estimators for γ_2 in the case of DGP with exogenous time-invariant regressors (DGP A) and serial correlated errors (Case 3)

N	T	3			5			10		
		FEF	FEVD		FEF	FEVD		FEF	FEVD	
			without	with		without	with		without	with
500	Bias	-0.0012	-0.1694	-0.0012	-0.0007	-0.1761	-0.0007	0.0001	-0.1809	0.0001
	RMSE	0.0425	0.1861	0.0425	0.0405	0.1921	0.0405	0.0387	0.1960	0.0387
	size	4.4%	91%	57%	4.5%	95%	60%	5.3%	97%	65%
	power	19%	78%	76%	23%	85%	80%	27%	91%	83%
1000	Bias	0.0024	-0.1781	0.0024	0.0003	-0.1859	0.0003	0.0001	-0.1956	0.0001
	RMSE	0.0291	0.1860	0.0291	0.0292	0.1933	0.0292	0.0283	0.2030	0.0283
	size	4.1%	100%	55%	4.9%	99%	59%	6.3%	100%	64%
	power	40%	94%	90%	40%	97%	91%	43%	99%	93%
2000	bias	0.0009	-0.1548	0.0009	-0.0003	-0.1613	-0.0003	-0.0003	-0.1671	-0.0003
	RMSE	0.0220	0.1588	0.0220	0.0209	0.1654	0.0209	0.0196	0.1707	0.0196
	size	4.3%	100%	61%	5.7%	100%	59%	4.6%	100%	65%
	power	65%	98%	96%	68%	99%	97%	74%	100%	100%

Notes: see notes 1-2 of Table B2.

Table B7: Bias, RMSE, size and power of FEF-IV and HTM estimators for γ_1 in the DGP with endogenous time-invariant regressors (DGP B) and homoskedastic and serially uncorrelated errors (Case 1)

N	T	3		5		10	
		FEF-IV	HTM	FEF-IV	HTM	FEF-IV	HTM
500	Bias	0.0030	0.0028	0.0030	0.0029	0.0052	0.0051
	RMSE	0.0689	0.0683	0.0647	0.0644	0.0592	0.0591
	size	3.7%	3.7%	4.6%	4.2%	4.1%	4%
	power	18%	17%	16%	17%	20%	19%
1000	Bias	0.0005	0.0005	0.0022	0.0022	0.0021	0.0021
	RMSE	0.0455	0.0453	0.0438	0.0437	0.0408	0.0408
	size	4.2%	4%	5%	4.6%	4%	4.3%
	power	23%	23%	28%	28%	29%	29%
2000	Bias	0.0000	0.0000	0.0014	0.0014	0.0015	0.0015
	RMSE	0.0316	0.0315	0.0295	0.0294	0.0257	0.0257
	size	4.4%	4.2%	3.4%	3.3%	6.2%	5.6%
	power	39%	39%	43%	43%	45%	45%

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. "FEF-IV" refers to the FEF-IV estimation, "HTM" refers to the modified HT estimation.

Table B8: Bias, RMSE, size and power of FEF-IV and HTM estimators for γ_2 in the DGP with endogenous time-invariant regressors (DGP B) and homoskedastic and serially uncorrelated errors (Case 1)

	T	3		5		10	
N		FEF-IV	HTM	FEF-IV	HTM	FEF-IV	HTM
500	Bias	-0.0022	-0.0025	0.0024	0.0023	0.0022	0.0022
	RMSE	0.0653	0.0647	0.0623	0.0621	0.0599	0.0598
	size	4%	3.1%	4.4%	4%	4.1%	3.8%
	power	14%	14%	17%	17%	18%	17%
1000	Bias	0.0000	-0.0001	0.0014	0.0014	-0.0003	-0.0003
	RMSE	0.0459	0.0456	0.0442	0.0441	0.0404	0.0404
	size	4.5%	4.6%	5.1%	4.4%	3.8%	3.7%
	power	23%	23%	28%	28%	25%	25%
2000	Bias	0.0000	0.0000	-0.0002	-0.0002	0.0.041	0.0041
	RMSE	0.0318	0.0317	0.0292	0.0291	0.0224	0.0223
	size	5.1%	4.8%	3.2%	3.4%	3.2%	2.8%
	power	38%	37%	41%	41%	43%	43%

Notes: 1. Size is calculated under $\gamma_2^{(0)} = 1$, and power under $\gamma_2^{(1)} = 0.95$.

2. See notes 2-3 of Table B7.

Table B9: Bias, RMSE, size and power of FEF-IV and HTM estimators for γ_1 in the DGP with endogenous time-invariant regressors (DGP B) and heteroskedastic and serially uncorrelated errors (Case 2)

	T	3		5		10	
N		FEF-IV	HTM	FEF-IV	HTM	FEF-IV	HTM
500	Bias	0.0026	0.0024	0.0038	0.0037	0.0015	0.0014
	RMSE	0.0647	0.0642	0.0617	0.0615	0.0560	0.0559
	size	2.7%	2.6%	4%	3.7%	3.3%	3.3%
	power	16%	15%	17%	16%	17%	16%
1000	Bias	0.0007	0.0007	0.0023	0.0023	0.0014	0.0014
	RMSE	0.0460	0.0458	0.0435	0.0433	0.0406	0.0406
	size	4.2%	4.2%	4.4%	4.9%	5.3%	4.5%
	power	23%	23%	27%	27%	28%	27%
2000	Bias	0.0007	0.0007	0.0025	0.0025	0.0016	0.0016
	RMSE	0.0326	0.0325	0.0305	0.0304	0.0287	0.0287
	size	5.4%	5.4%	5%	5.4%	5.8%	5.2%
	power	39%	38%	44%	45%	46%	46%

Notes: see notes 1-3 of Table B7.

Table B10: Bias, RMSE, size and power of FEF-IV and HTM estimators for γ_2 in the DGP with endogenous time-invariant regressors (DGP B) and heteroskedastic and serially uncorrelated errors (Case 2)

	T	3		5		10	
N		FEF-IV	HTM	FEF-IV	HTM	FEF-IV	HTM
500	Bias	0.0037	0.0035	0.0049	0.0049	0.0033	0.0032
	RMSE	0.0660	0.0654	0.0655	0.0652	0.0590	0.0589
	size	3.8%	3.8%	3.7%	3.3%	4.2%	4.1%
	power	16%	16%	19%	18%	19%	18%
1000	Bias	-0.0001	-0.0002	0.0053	0.0053	0.0032	0.0032
	RMSE	0.0465	0.0463	0.0434	0.0432	0.0412	0.0412
	size	4.1%	4.2%	4.3%	4.3%	5.1%	5.3%
	power	23%	23%	28%	28%	30%	30%
2000	Bias	0.0013	0.0012	0.0002	0.0002	0.0004	0.0004
	RMSE	0.0307	0.0306	0.0298	0.0297	0.0281	0.0280
	size	3.9%	3.9%	4.2%	4.2%	5.2%	5.8%
	power	38%	38%	41%	40%	42%	42%

Notes: see notes 1-2 of Table B8.

Table B11: Bias, RMSE, size and power of FEF-IV and HTM estimators for γ_1 in the DGP with endogenous time-invariant regressors (DGP B) and serially correlated errors (Case 3)

	T	3		5		10	
N		FEF-IV	HTM	FEF-IV	HTM	FEF-IV	HTM
500	Bias	0.0032	0.0031	0.0021	0.0021	0.0053	0.0053
	RMSE	0.0691	0.0689	0.0693	0.0692	0.0639	0.0639
	size	2.6%	2.4%	4.1%	3.5%	4%	4.5%
	power	13%	13%	15%	15%	17%	16%
1000	Bias	0.0017	0.0017	0.0011	0.0011	0.0014	0.0014
	RMSE	0.0507	0.0506	0.0460	0.0460	0.0444	0.0444
	size	4.5%	4.4%	3.3%	3%	3.2%	3.6%
	power	21%	21%	20%	20%	24%	24%
2000	Bias	0.0015	0.0015	0.0004	0.0004	0.0003	0.0003
	RMSE	0.0349	0.0348	0.0332	0.0332	0.0311	0.0310
	size	4.3%	4.2%	3.8%	3.7%	3.8%	4.2%
	power	35%	35%	35%	35%	37%	37%

Notes: see notes 1-3 of Table B7.

Table B12: Bias, RMSE, size and power of FEF-IV and HTM estimators for γ_2 in the DGP with endogenous time-invariant regressors (DGP B) and serially correlated errors (Case 3)

	T	3		5		10	
N		FEF-IV	HTM	FEF-IV	HTM	FEF-IV	HTM
500	Bias	0.0038	0.0038	0.0025	0.0024	0.0056	0.0056
	RMSE	0.0710	0.0707	0.0694	0.0693	0.0686	0.0685
	size	4.2%	4.1%	4%	4%	3.9%	4.1%
	power	14%	14%	15%	14%	17%	17%
1000	Bias	0.0014	0.0014	0.0013	0.0012	0.0018	0.0017
	RMSE	0.0510	0.0509	0.0480	0.0479	0.0471	0.0471
	size	4.1%	3.9%	4.5%	4.5%	4.9%	4.7%
	power	21%	21%	23%	22%	25%	25%
2000	Bias	-0.0002	-0.0002	0.0017	0.0017	0.0008	0.0008
	RMSE	0.0356	0.0356	0.0343	0.0342	0.0318	0.0318
	size	4.2%	4.4%	4.9%	4.9%	5.8%	6.4%
	power	32%	32%	36%	36%	38%	38%

Notes: see notes 1-2 of Table B8.

Table B13: Bias, RMSE, size and power of FEF-IV and HTM estimators for γ_1 in the DGP with endogenous time-invariant regressors (DGP C) and homoskedastic and serially uncorrelated errors (Case 1)

	T	3		5		10	
N		FEF-IV	HTM	FEF-IV	HTM	FEF-IV	HTM
500	Bias	0.0032	0.0024	-0.0003	-0.0006	0.0016	0.0015
	RMSE	0.0490	0.0486	0.0434	0.0433	0.0419	0.0418
	size	5.9%	5.9%	4.6%	4.7%	5.4%	5.1%
	power	21%	20%	22%	20%	24%	23%
1000	Bias	0.0005	0.0000	0.0000	-0.0002	0.0013	0.0012
	RMSE	0.0335	0.0334	0.0315	0.0315	0.0288	0.0287
	size	5%	5.2%	5.3%	5%	4.8%	4.6%
	power	32%	32%	38%	36%	41%	42%
2000	Bias	-0.0001	-0.0003	0.0010	0.0009	-0.0019	-0.0020
	RMSE	0.0235	0.0235	0.0206	0.0205	0.0184	0.0184
	size	4.9%	4.4%	3.3%	3.3%	4.4%	4.6%
	power	59%	59%	66%	66%	67%	67%

Notes: see notes 1-3 of Table B7.

Table B14: Bias, RMSE, size and power of FEF-IV and HTM estimators for γ_2 in the DGP with endogenous time-invariant regressors (DGP C) and homoskedastic and serially uncorrelated errors (Case 1)

	T	3		5		10	
N		FEF-IV	HTM	FEF-IV	HTM	FEF-IV	HTM
500	Bias	-0.0027	-0.0036	-0.0003	-0.0007	-0.0019	-0.0020
	RMSE	0.0496	0.0492	0.0438	0.0436	0.0414	0.0414
	size	6.5%	5.6%	4.8%	4.7%	5.1%	5.8%
	power	20%	18%	22%	20%	20%	19%
1000	Bias	0.0001	-0.0005	-0.0005	-0.0007	-0.00011	-0.0012
	RMSE	0.0331	0.0329	0.0312	0.0311	0.0288	0.0286
	size	6.1%	5.9%	4.9%	4.8%	4.9%	4.8%
	power	32%	31%	37%	36%	38%	38%
2000	Bias	-0.0002	-0.0004	-0.0007	-0.0008	0.0009	0.0008
	RMSE	0.0235	0.0234	0.0208	0.0208	0.0179	0.0179
	size	4.9%	4.8%	3.6%	3.7%	4.8%	4%
	power	57%	56%	63%	63%	65%	65%

Notes: see notes 1-2 of Table B8.

Table B15: Bias, RMSE, size and power of FEF-IV and HTM estimators for γ_1 in the DGP with endogenous time-invariant regressors (DGP C) and heteroskedastic and serially uncorrelated errors (Case 2)

	T	3		5		10	
N		FEF-IV	HTM	FEF-IV	HTM	FEF-IV	HTM
500	Bias	0.0006	-0.0002	-0.0012	-0.0015	-0.0015	-0.0016
	RMSE	0.0449	0.0446	0.0437	0.0436	0.0408	0.0407
	size	4.1%	3.6%	5.2%	4.6%	4.6%	4.2%
	power	18%	16%	21%	20%	23%	22%
1000	Bias	0.0009	0.0004	-0.0016	-0.0018	-0.0011	-0.0011
	RMSE	0.0327	0.0325	0.0313	0.0313	0.0295	0.0294
	size	4.9%	4.6%	5.2%	4.9%	4%	4.1%
	power	34%	34%	37%	37%	41%	41%
2000	Bias	-0.0001	-0.0004	0.0009	0.0008	0.0009	0.0009
	RMSE	0.0233	0.0232	0.0221	0.0221	0.0204	0.0204
	size	4.6%	4.4%	5.7%	5.4%	5%	4.6%
	power	59%	58%	64%	62%	68%	67%

Notes: see notes 1-3 of Table B7.

Table B16: Bias, RMSE, size and power of FEF-IV and HTM estimators for γ_2 in the DGP with endogenous time-invariant regressors (DGP C) and heteroskedastic and serially uncorrelated errors (Case 2)

	T	3		5		10	
N		FEF-IV	HTM	FEF-IV	HTM	FEF-IV	HTM
500	Bias	0.0009	0.0000	0.0003	0.0000	0.0014	0.0013
	RMSE	0.0460	0.0456	0.0430	0.0429	0.0404	0.0403
	size	3.6%	3.9%	5.2%	4.5%	4.3%	4.4%
	power	18%	16%	21%	20%	24%	23%
1000	Bias	-0.0002	-0.0007	0.0013	0.0011	0.0010	0.0009
	RMSE	0.0326	0.0325	0.0312	0.0312	0.0289	0.0289
	size	3.8%	4.1%	5.7%	5.3%	4%	4.2%
	power	32%	30%	38%	37%	42%	42%
2000	Bias	0.0003	0.0000	-0.0013	-0.0014	-0.0006	-0.0007
	RMSE	0.0233	0.0232	0.0224	0.0224	0.0216	0.0215
	size	4.6%	4.5%	5.9%	6.2%	5%	4.4%
	power	57%	56%	62%	61%	66%	65%

Notes: see notes 1-2 of Table B8.

Table B17: Bias, RMSE, size and power of FEF-IV and HTM estimators for γ_1 in the DGP with endogenous time-invariant regressors (DGP C) and serially correlated errors (Case 3)

	T	3		5		10	
N		FEF-IV	HTM	FEF-IV	HTM	FEF-IV	HTM
500	Bias	0.0000	-0.0004	0.0001	-0.0001	-0.0001	-0.0002
	RMSE	0.0498	0.0497	0.0484	0.0483	0.0448	0.0448
	size	4.3%	4.6%	4.4%	4.7%	3.4%	3.8%
	power	17%	16%	18%	18%	21%	21%
1000	Bias	-0.0004	-0.0006	-0.0004	-0.0005	0.0002	0.0002
	RMSE	0.0362	0.0362	0.0350	0.0350	0.0317	0.0317
	size	4.5%	4.3%	6.3%	6.3%	4.5%	4.4%
	power	30%	29%	30%	30%	34%	34%
2000	Bias	0.0006	0.0005	-0.0009	-0.0009	-0.0007	-0.0007
	RMSE	0.0253	0.0252	0.0245	0.0245	0.0231	0.0231
	size	4.9%	5%	5.3%	5.1%	4.2%	4.2%
	power	51%	51%	53%	52%	55%	55%

Notes: see notes 1-3 of Table B7.

Table B18: Bias, RMSE, size and power of FEF-IV and HTM estimators for γ_2 in the DGP with endogenous time-invariant regressors (DGP C) and serially correlated errors (Case 3)

	T	3		5		10	
N		FEF-IV	HTM	FEF-IV	HTM	FEF-IV	HTM
500	Bias	0.0000	-0.0003	0.0003	0.0001	0.0000	-0.0001
	RMSE	0.0496	0.0495	0.0485	0.0484	0.0458	0.0458
	size	5%	4.9%	5.4%	5.4%	5.2%	4.7%
	power	17%	16%	19%	18%	20%	20%
1000	Bias	0.0006	0.0003	0.0001	0.0000	0.0009	0.0009
	RMSE	0.0365	0.0365	0.0348	0.0348	0.0318	0.0318
	size	5.1%	5.1%	6.4%	6.2%	4.3%	4.3%
	power	30%	29%	30%	30%	35%	35%
2000	Bias	-0.0011	-0.0012	0.0004	0.0004	0.0004	0.0004
	RMSE	0.0247	0.0247	0.0247	0.0247	0.232	0.0232
	size	4.2%	4.5%	5.4%	5.4%	4.6%	5.4%
	power	50%	50%	55%	54%	57%	56%

Notes: see notes 1-2 of Table B8.

C. Simulation results for the (unmodified) HT estimator

In this section, we provide simulation results for the (unmodified) HT estimator when there are in fact no valid exogenous time-varying variables which can be used as instruments (as required by HT) for the endogenous time-invariant regressors. We closely follow the Monte Carlo design of the paper and generate y_{it} as

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

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 $i=1,2,\ldots,N$. Both time-varying regressors, $x_{1,it}$ and $x_{2,it}$ are generated to be correlated with the fixed effects:

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

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. It is clear that this DGP does not meet one of the requirements of the HT procedure, which assumes that one or more time varying regressors are uncorrelated with α_i. The stochastic components of the time varying regressors ($\omega_{it,1}$ and $\omega_{it,2}$) are generated as heterogenous $AR(1)$ processes

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

where

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

The time-invariant regressors are generated as

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

where $IU(7,12)$ denotes integers uniformly drawn within the range $[7,12]$. Note that the second time-invariant regressor, z_{2i}, is generated to depend on the fixed effects, α_i, to deal with this endogeneity we use r_i as the instrument for z_{2i} in the FEF-IV estimation procedure. 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: Heteroskedastic 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 heteroskedastic 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.

In computing the HT estimator, we use time averages of the time-varying regressors, namely $\bar{x}_{1i}$ and $\bar{x}_{2i}$, as well as z_{i1}, as instruments. The simulation results are summarized in Tables C1-C6. As can be seen, the unmodified HT estimator which uses invalid instruments, $\bar{x}_{1i}$ and $\bar{x}_{2i}$, is biased and shows substantial size distortions.

Table C1: Bias, RMSE, size and power of FEF-IV and HT estimators for γ_1 in the DGP of endogenous time-invariant regressors and homoskedastic and serially uncorrelated errors (Case 1)

	T	3		5		10	
N		FEF-IV	HT	FEF-IV	HT	FEF-IV	HT
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%
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%
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 and HT use (r_i) and $(\bar{x}_{1i},\ \bar{x}_{2i})$ as the instruments for z_{2i}, respectively.

Table C2: Bias, RMSE, size and power of FEF-IV and HT estimators for γ_2 in the DGP of endogenous time-invariant regressors and homoskedastic and serially uncorrelated errors (Case 1)

	T	3		5		10	
N		FEF-IV	HT	FEF-IV	HT	FEF-IV	HT
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%
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%
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 $\gamma_2^{(1)} = 0.95$.

2. See Notes 2-3 of Table C1.

Table C3: Bias, RMSE, size and power of FEF-IV and HT estimators for γ_1 in the DGP of endogenous time-invariant regressors and heteroskedastic and serially uncorrelated errors (Case 2)

	T	3		5		10	
N		FEF-IV	HT	FEF-IV	HT	FEF-IV	HT
500	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	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	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 1-3 to Table C1.

Table C4: Bias, RMSE, size and power of FEF-IV and HT estimators for γ_2 in the DGP of endogenous time-invariant regressors and heteroskedastic and serially uncorrelated errors (Case 2)

	T	3		5		10	
N		FEF-IV	HT	FEF-IV	HT	FEF-IV	HT
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%
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%
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 1-2 to Table C2.

Table C5: Bias, RMSE, size and power of FEF-IV and HT estimators for γ_1 in the DGP with endogenous time-invariant regressors and serially correlated errors (Case 3)

	T	3		5		10	
N		FEF-IV	HT	FEF-IV	HT	FEF-IV	HT
500	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	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	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 1-3 to Table C1.

Table C6: Bias, RMSE, size and power of FEF-IV and HT estimators for γ_2 in the DGP with endogenous time-invariant regressors and serially correlated errors (Case 3)

	T	3		5		10	
N		FEF-IV	HT	FEF-IV	HT	FEF-IV	HT
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%
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%
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 1-2 to Table C2.