ISSN 1440-771X

MONASH University

Australia

Department of Econometrics and Business Statistics

http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers/

Consistent Estimation in Large Heterogeneous Panels with Multifactor Structure Endogeneity

G. Forchini, Bin Jiang and Bin Peng

July 2015

Working Paper 14/15

Consistent Estimation in Large Heterogeneous Panels with Multifactor Structure and Endogeneity

G. Forchini[1]
University of Surrey

Bin Jiang[2]
Monash University

Bin Peng[3]
University of Technology Sydney

8 June 2015

Abstract

The set-up considered by Pesaran (Econometrica, 2006) is extended to allow for endogenous explanatory variables. A class of instrumental variables estimators is studied and it is shown that estimators in this class are consistent and asymptotically normally distributed as both the cross-section and time-series dimensions tend to infinity.

Keywords: Heterogeneous panels, multifactor structure, endogeneity, instrumental variables estimator, common correlated effects estimator, common correlated effects mean group estimator

JEL classification: C33, C36

[1] Address for correspondence: Giovanni Forchini, School of Economics, University of Surrey, Guildford, Surrey GU2 7XH, United Kingdom. E-mail: G.Forchini@surrey.ac.uk. This research was partially supported by Australian Research Council grant DP0985432.

[2] Address for correspondence: Bin Jiang, Department of Econometrics and Business Statistics, Monash University, Caulfield East, VIC 3145, Australia. E-mail: Bin.Jiang@monash.edu.

[3] Address for correspondence: Bin Peng, Economics Discipline Group, University of Technology Sydney, Australia, 2007. E-mail: Bin.Peng@uts.edu.au.

1. Introduction

Pesaran (2006) has introduced a very popular panel data model with cross-sectional heterogeneity exhibited through a factor structure in the errors in which the unobserved factors also affect the observed explanatory variables. The model has been extended in several directions (e.g. Bai (2009), Kapetanios, Pesaran and Yamagata (2011) and Pesaran and Tosetti (2011)) and used to analyse, for instance, health care expenditures (e.g. Baltagi and Moscone (2010) and Hauck and Zhang (2014)), energy policies (e.g. Arouri et al. (2012)) and house prices (e.g. Holly, Pesaran and Yamagata (2010)).

Economic models are often characterised by endogeneity of some of the explanatory variables, and endogeneity usually invalidates estimation methods that do not take it into account. Panel data models with endogenous variables and a factor structure in the errors have been studied by Ahn, Lee and Schmidt (2001), Harding and Lamarche (2011), Ahn, Lee and Schmidt (2013), Robertson and Sarafidis (2015) and Forchini, Jiang and Peng (2015). Ahn, Lee and Schmidt (2001, 2013) and Robertson and Sarafidis (2015) derive and study consistent estimators of the structural parameters in panel data models in which the factors are not regarded as unobservable random variables but as unknown parameters and their results do not extend to the model introduced by Pesaran (2006). Harding and Lamarche (2011) introduce endogeneity in the model of Pesaran (2006) and propose a two steps instrumental variables procedure to estimate it. In the first step, Pesaran (2006)'s model is augmented with the cross-sectional means of the dependent variable and the independent variables affected by the shocks; in the second step, the augmented model is estimated using instrumental variables with instruments given by the cross-sectional means of the instruments, and the dependent variables not affected by the shocks. Forchini, Jiang and Peng (2015) investigate endogeneity in panel data with unobservable factors in the errors, in the explanatory variables and in the instruments when the time-dimension is fixed and have shown that classical estimators such as the two-stage least squares (TSLS) and the limited maximum likelihood estimator are inconsistent when the factor loadings in the errors and the exogenous variables are correlated given the common factors.

In this paper, we suggest a class of instrumental variables (IV) estimators for the model of Pesaran (2006) when some of the explanatory variables are endogenous. These are analog estimators exploiting the compatibility restrictions between the structural equation and reduced form, and extend the IV estimators from a standard set-up to the model of Pesaran (2006). We show that estimators in this class are consistent and asymptotically normal as the cross-section and time-series dimensions tend to infinity.

The rest of the paper is organized as follows. Section 2 describes the model with endogenous explanatory variables, defines its reduced form and discusses their relationship. The assumptions needed to construct an asymptotic theory are stated in Section 3 and Section 4 describes the class of estimators considered. Consistency and asymptotic normality are obtained in Section 5. Section 6 investigates the small sample performance of the proposed estimators and Section 7 concludes. Proofs of all results are in the appendix.

2. Model

Consider the panel data model

(1) $$\underset{(1\times 1)}{y_{1it}} = \underset{(1\times n)}{\lambda_i'}\,\underset{(n\times 1)}{d_t} + \underset{(1\times p)}{\beta_i'}\,\underset{(p\times 1)}{y_{2it}} + \underset{(1\times k_1)}{\theta_i'}\,\underset{(k_1\times 1)}{x_{1it}} + u_{it},$$

where y_{1it} and y_{2it} contain observations on the endogenous variables, d_t is a vector of observed common effects, x_{1it} is a vector of observed unit-specific exogenous regressors. In equation (1) and in the rest of the paper, the brackets appearing under the vectors or matrices the first time they are used denote their dimensions. In order to introduce a multifactor structure we consider the reduced form

(2) $$y_{it} = \begin{pmatrix} y_{1it} \\ y_{2it} \end{pmatrix} = \begin{pmatrix} \underset{(1\times n)}{\alpha_{1i}'} \\ \underset{(p\times n)}{\alpha_{2i}'} \end{pmatrix} d_t + \begin{pmatrix} \underset{(1\times k_1)}{\pi_{11i}'} & \underset{(1\times k_2)}{\pi_{21i}'} \\ \underset{(p\times k_1)}{\Pi_{12i}'} & \underset{(p\times k_2)}{\Pi_{22i}'} \end{pmatrix} \begin{pmatrix} \underset{(k_1\times 1)}{x_{1it}} \\ \underset{(k_2\times 1)}{x_{2it}} \end{pmatrix} + \begin{pmatrix} \underset{(1\times 1)}{e_{1it}} \\ \underset{(p\times 1)}{e_{2it}} \end{pmatrix},$$

where x_{2it} denotes a vector of observed instruments. We assume that the reduced form errors have a multifactor error structure of the form

(3) $$\begin{pmatrix} \underset{(1\times 1)}{e_{1it}} \\ \underset{(p\times 1)}{e_{2it}} \end{pmatrix} = \begin{pmatrix} \underset{(1\times m)}{\gamma_{1i}'} \\ \underset{(p\times m)}{\gamma_{2i}'} \end{pmatrix} \underset{(m\times 1)}{f_t} + \begin{pmatrix} \underset{(1\times 1)}{\varepsilon_{1it}} \\ \underset{(p\times 1)}{\varepsilon_{2it}} \end{pmatrix},$$

in which f_t is a vector of unobserved common effects, γ_{1i} and γ_{2i} are factor loadings and ε_{1it} and ε_{2it} are errors which are independent of f_t, d_t, x_{1it} and x_{2it} for each i and t. Moreover,

(4) $$x_{it} = \begin{pmatrix} x_{1it} \\ x_{2it} \end{pmatrix} = \begin{pmatrix} \underset{(k_1\times n)}{A_{1i}'} \\ \underset{(k_2\times n)}{A_{2i}'} \end{pmatrix} d_t + \begin{pmatrix} \underset{(k_1\times m)}{\Gamma_{1i}'} \\ \underset{(k_2\times m)}{\Gamma_{2i}'} \end{pmatrix} f_t + \begin{pmatrix} \underset{(k_1\times 1)}{v_{1it}} \\ \underset{(k_2\times 1)}{v_{2it}} \end{pmatrix},$$

where A_{1i}, A_{2i}, Γ_{1i} and Γ_{2i} are factor loading matrices and v_{1it} and v_{2it} are the specific components of x_{1it} and x_{2it}. The terms v_{1it} and v_{2it} are independent of d_t and f_t.

Notice that (2), (3) and (4) form a multivariate version of the model of Pesaran (2006). Notice also that the factors affect all exogenous variables including the instruments and that we allow for random coefficients and for the instruments to depend on the common factors.

Compatibility of the structural equation (1) and the reduced form (2) implies that

(5) $$\pi_{21i} = \Pi_{22i}\beta_i$$

(6) $$\alpha_{1i} = \alpha_{2i}\beta_i + \lambda_i$$

(7) $$\pi_{11i} = \Pi_{12i}\beta_i + \theta_i$$

and

(8) $$u_{it} = e_{1it} - \beta' e_{2it} = \left(\gamma_{1i}' - \beta_i'\gamma_{2i}'\right) f_t + \varepsilon_{1it} - \beta_i'\varepsilon_{2it}.$$

The latter expression shows that the structural error u_{it} has a factor structure with factor loadings $\gamma_{1i} - \gamma_{2i}\beta_i$. The relationship in (5) expresses the over-identifying restrictions for each individual structural parameter β_i, and the vector of coefficients β_i is identified if and only if Π_{22i} has rank $p \le k_2$; equation (6) defines the factor loadings associated with the observed common factors in terms of the reduced form factor loadings, and equation (7) defines the vector of parameters γ_i in terms of the reduced form parameters.

3. Assumptions and notation

We make the same assumptions as Pesaran (2006), apart from Assumption 4 which needs to be slightly modified. We use $\|\cdot\|$ to denote the Frobenius norm as in Pesaran (2006).

Assumption 1. The vector of common effects $g_t = \left(d_t', f_t'\right)'$ is covariance stationary with absolute summable autocovariances, and is distributed independently of the unit-specific errors $\left(\varepsilon_{1is}', \varepsilon_{2is}'\right)'$ and $\left(v_{1is}', v_{2is}'\right)'$ for all i, t and s .

Assumption 2. The reduced form errors $\varepsilon_{it} = \left(\varepsilon_{1it}', \varepsilon_{2it}'\right)'$, and $v_{is} = \left(v_{1is}', v_{2is}'\right)'$ are independent for all i , t and s . Moreover for each i , ε_{it} and v_{it} follow linear stationary processes with absolute summable autocovariances, $\varepsilon_{it} = \sum_{l=0}^{\infty} A_{il} \varsigma_{i,t-l}$ and $v_{it} = \sum_{l=0}^{\infty} S_{il} \nu_{i,t-l}$, where $\left(\varsigma_{i,t}', \nu_{i,t}'\right)'$ are $\left(k_1 + k_2 + p + 1\right) \times 1$ vectors of identically, independently distributed random variables with mean zero, identity covariance matrix, and finite fourth order cumulants. In particular they are such that the covariance matrices of $\varepsilon_{it} = \left(\varepsilon_{1it}, \varepsilon_{2it}'\right)'$ and $v_{it} = \left(v_{1it}', v_{2it}'\right)'$ are uniformly bounded.

Assumption 3. The factor loadings $\left(\gamma_{1i}, \gamma_{2i}, \Gamma_{1i}, \Gamma_{2i}\right)$ are independently and identically distributed across i, and of the specific errors $\varepsilon_{it} = \left(\varepsilon_{1it}, \varepsilon_{2it}'\right)'$ and $v_{it} = \left(v_{1it}', v_{2it}'\right)'$, the common factors $g_t = \left(d_t', f_t'\right)'$ and they have finite means $E\left(\gamma_{1i}, \gamma_{2i}, \Gamma_{1i}, \Gamma_{2i}\right) = \left(\gamma_1, \gamma_2, \Gamma_1, \Gamma_2\right)$ and covariance matrices that are uniformly bounded.

Assumption 4. (a) The reduced form parameter $\Pi_i = \begin{pmatrix} \pi_{11i} & \Pi_{12i} \\ \pi_{21i} & \Pi_{22i} \end{pmatrix}$ follows the random coefficient model $\Pi_i = \Pi + \tilde{\Pi}_i$, where $\Pi = \begin{pmatrix} \pi_{11} & \Pi_{12} \\ \pi_{21} & \Pi_{22} \end{pmatrix}$, $vec\left(\tilde{\Pi}_i\right) \sim iid\left(0, \Omega_{\tilde{\Pi}}\right)$, $\|\Pi\|$ and $\|\Omega_{\tilde{\Pi}}\|$ are uniformly bounded and the deviation $\tilde{\Pi}_i$ is distributed independently of γ_j, Γ_j, ε_{jt}, v_{jt} and g_t for all i, j, t.

(b) The structural parameters β_i and γ_i follow the random coefficient models $\beta_i = \beta + \tilde{\beta}_i$ and $\theta_i = \theta + \tilde{\theta}_i$ where $vec\left(\tilde{\beta}_i\right) \sim iid\left(0, \Omega_{\tilde{\beta}}\right)$, $vec\left(\tilde{\theta}_i\right) \sim iid\left(0, \Omega_{\tilde{\theta}}\right)$, $\|\beta\|$, $\|\theta\|$, $\Omega_{\tilde{\beta}}$ and $\Omega_{\tilde{\theta}}$ are uniformly bounded and the deviations $\tilde{\beta}_i$ and $\tilde{\theta}_i$ are distributed independently of γ_j, Γ_j, ε_{jt}, v_{jt} and g_t for all i, j, t.

Let $y_{it}^{(j)}$ be the j-th component of y_{it}, $j = 1, ..., p+1$, and $z_{it}^{(j)} = \left(y_{it}^{(j)}, x_{it}'\right)'$ and $\bar{z}_{\omega t}^{(j)} = \sum_{l=1}^{N} \omega_l z_{lt}^{(j)}$. Moreover, let D be the $T \times n$ matrix of observations on d_t and $\bar{Z}_{\omega}^{(j)}$ be the $T \times \left(k_1 + k_2 + 1\right)$ matrix of observations on $\bar{z}_{\omega t}^{(j)}$, X_i be the $T \times \left(k_1 + k_2\right)$ matrix of observations on the exogenous variables x_{it}, and $y_i^{(j)}$ be the $T \times \left(p+1\right)$ matrix of observations on the endogenous variables $y_{it}^{(j)}$. Finally let $\bar{H}_{\omega}^{(j)} = \left(D, \bar{Z}_{\omega}^{(j)}\right)$ and $\bar{M}_{\omega}^{(j)} = I_T - \bar{H}_{\omega}^{(j)} \left(\bar{H}_{\omega}^{(j)}{}' \bar{H}_{\omega}^{(j)}\right)^{-} \bar{H}_{\omega}^{(j)}{}'$ where $\left(\bar{H}_{\omega}^{(j)}{}' \bar{H}_{\omega}^{(j)}\right)^{-}$ denotes the generalized inverse of $\bar{H}_{\omega}^{(j)}{}' \bar{H}_{\omega}^{(j)}$.

Similarly, the j-th components of $\varepsilon_{it} = \left(\varepsilon_{1it}, \varepsilon_{2it}'\right)'$ and of $\gamma = \left(\gamma_1, \gamma_2\right)$ are denoted respectively by $\varepsilon_{it}^{(j)}$ and $\gamma^{(j)}$. Finally, $\varepsilon_i^{(j)}$ denotes the $T \times \left(p+1\right)$ matrix of the errors $\varepsilon_{it}^{(j)}$.

Assumption 5. (a) $\frac{1}{T} X_i' \bar{M}_{\omega}^{(j)} X_i$ and $\frac{1}{T} X_i' M_g X_i$ are non-singular for each $j = 1, ..., p+1$, where $M_g = I_T - G\left(G'G\right)^{-} G'$ and G is the $T \times \left(n+m\right)$ matrix of observations on $g_t = \left(d_t', f_t'\right)'$, and $\left(G'G\right)^{-}$ denotes the generalized inverse of $G'G$; (b) $\frac{1}{T} \sum_{i=1}^{N} \omega_i X_i' \bar{M}_{\omega}^{(j)} X_i$ is non-singular for each $j = 1, ..., p+1$ and for scalar weights ω_i satisfying $\omega_i = O\left(\frac{1}{N}\right)$, $\sum_{i=1}^{N} \omega_i = 1$ and $\sum_{i=1}^{N} \left|\omega_i\right| < K$ where K is a finite constant.

Assumptions 1, 2, 3 and 5 are the same assumptions used by Pesaran (2006) and they are fully discussed in that work. Assumption 4 is different from Assumption 4 of Pesaran (2006) because the model we consider involves structural and reduced form parameters for both of which we impose random coefficient structures. We discuss what this entails briefly.

Using the compatibility conditions between structural equation and reduced form (5), (6) and (7), we can write

$$\Pi_i = \begin{bmatrix} \theta_i & \pi_{12i} \\ 0 & \Pi_{22i} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\beta_i & I_p \end{bmatrix}^{-1} = \begin{bmatrix} \theta_i & \pi_{12i} \\ 0 & \Pi_{22i} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \beta_i & I_p \end{bmatrix}$$

Due to Assumption 4, this can be rewritten as

$$\begin{aligned} \Pi + \tilde{\Pi}_i &= \left(\begin{bmatrix} \theta & \pi_{12} \\ 0 & \Pi_{22} \end{bmatrix} + \begin{bmatrix} \tilde{\theta}_i & \tilde{\pi}_{12i} \\ 0 & \tilde{\Pi}_{22i} \end{bmatrix} \right) \left(\begin{bmatrix} 1 & 0 \\ \beta & I_p \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \tilde{\beta}_i & 0 \end{bmatrix} \right) \\ &= \begin{bmatrix} \theta & \pi_{12} \\ 0 & \Pi_{22} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \beta & I_p \end{bmatrix} + \begin{bmatrix} \tilde{\theta}_i + \tilde{\pi}_{12i}\beta + \pi_{12}\tilde{\beta}_i + \tilde{\pi}_{12i}\tilde{\beta}_i & \tilde{\pi}_{12i} \\ \tilde{\Pi}_{22i}\beta + \Pi_{22}\tilde{\beta}_i + \tilde{\Pi}_{22i}\tilde{\beta}_i & \tilde{\Pi}_{22i} \end{bmatrix}. \end{aligned}$$

Thus, by taking expectations of both sides of the equation above we must have

$$\Pi = \begin{bmatrix} \theta & \pi_{12} \\ 0 & \Pi_{22} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \beta & I_p \end{bmatrix}, \tag{9}$$

and

$$\tilde{\Pi}_i = \begin{bmatrix} \tilde{\theta}_i + \tilde{\pi}_{12i}\beta + \pi_{12}\tilde{\beta}_i + \tilde{\pi}_{12i}\tilde{\beta}_i & \tilde{\pi}_{12i} \\ \tilde{\Pi}_{22i}\beta + \Pi_{22}\tilde{\beta}_i + \tilde{\Pi}_{22i}\tilde{\beta}_i & \tilde{\Pi}_{22i} \end{bmatrix}. \tag{10}$$

Notice that the condition $E(\tilde{\Pi}_i) = 0$ holds only if $E(\tilde{\pi}_{12i}\tilde{\beta}_i) = 0$ and $E(\tilde{\Pi}_{22i}\tilde{\beta}_i) = 0$ (cf. Kelejian (1974)). Therefore, Assumption 4, entails that $\tilde{\beta}_i$ is not correlated to $\tilde{\pi}_{12i}$ and $\tilde{\Pi}_{22i}$. Notice also that (9) expresses the standard over-identifying restrictions which in this case apply to the means of the coefficients rather than the coefficients themselves.

Finally, the notation $(N,T) \xrightarrow{j} \infty$ indicates that N and T tend to infinity jointly.

4. Estimators

Provided T is sufficiently large we can estimate the matrix Π_i for each unit i by aggregating the common correlated effects (CCE) estimators for the slope coefficients in the reduced form for each individual endogenous variable:

$$\hat{\Pi}_i = \left[\left(X_i{}'\bar{M}_{\omega}^{(1)} X_i \right)^{-1} X_i{}'\bar{M}_{\omega}^{(1)} y_i^{(1)}, \ldots., \left(X_i{}'\bar{M}_{\omega}^{(p+1)} X_i \right)^{-1} X_i{}'\bar{M}_{\omega}^{(p+1)} y_i^{(p+1)} \right]. \tag{11}$$

One possible choice for estimating Π is the common correlated effects mean group (CCEMG) estimator

$$\hat{\Pi}_{MG} = \left[\frac{1}{N}\sum_{i=1}^{N} \left(X_i{}'\bar{M}_{\omega}^{(1)} X_i \right)^{-1} X_i{}'\bar{M}_{\omega}^{(1)} y_i^{(1)}, \ldots., \frac{1}{N}\sum_{i=1}^{N} \left(X_i{}'\bar{M}_{\omega}^{(p+1)} X_i \right)^{-1} X_i{}'\bar{M}_{\omega}^{(p+1)} y_i^{(p+1)} \right]. \tag{12}$$

An alternative is the common correlated effects pooled (CCEP) estimator:

$$\hat{\Pi}_P = \left[\left(\sum_{i=1}^{N} \omega_i X_i{}'\bar{M}_{\omega}^{(1)} X_i \right)^{-1} \sum_{i=1}^{N} \omega_i X_i{}'\bar{M}_{\omega}^{(1)} y_i^{(1)}, \ldots, \left(\sum_{i=1}^{N} \omega_i X_i{}'\bar{M}_{\omega}^{(p+1)} X_i \right)^{-1} \sum_{i=1}^{N} \omega_i X_i{}'\bar{M}_{\omega}^{(p+1)} y_i^{(p+1)} \right], \tag{13}$$

where we set the aggregating weights equal to the pooling weights (cf. Pesaran (2006, p. 986)). Notice that the columns of (12) and (13) are the CCEMG and CCEP estimators applied to the reduced form for each individual endogenous variables.

Partition the CCEMG and CCEP estimators of the reduced form parameters conformably to Π_i in equation (2). The panel instrumental variables (IV) estimators based on, respectively, the CCEMG and CCEP estimators of the reduced form parameters are

$$\hat{\beta}_{IV-MG} = \left(\hat{\Pi}_{22MG}{}'\hat{H}\hat{\Pi}_{22MG} \right)^{-1} \left(\hat{\Pi}_{22MG}{}'\hat{H}\hat{\pi}_{21MG} \right) \tag{14}$$

and

$$\hat{\beta}_{IV-P} = \left(\hat{\Pi}_{22P}{}'\hat{H}\hat{\Pi}_{22P} \right)^{-1} \left(\hat{\Pi}_{22P}{}'\hat{H}\hat{\pi}_{21P} \right), \tag{15}$$

where $\hat{H}$ is any positive definite random matrix converging in probability to a positive definite matrix H. Similarly, we can estimate the coefficients of the exogenous variables as

$$\hat{\theta}_{IV-MG} = \hat{\pi}_{11MG} - \hat{\Pi}_{12MG}\hat{\beta}_{IV-MG}, \tag{16}$$

and

(17) $$\hat{\theta}_{IV-P} = \hat{\pi}_{11P} - \hat{\Pi}_{12P}\hat{\beta}_{IV-P} .$$

Estimators analogous to these have been studied by Forchini, Jiang and Peng (2015) for the case where the reduced form parameters are estimated using OLS and T is fixed. They find that the classical panel TSLS and LIML are not consistent if the factor loadings in the errors and the exogenous variables are dependent conditional on the factors as in the case considered here.

It is important to compare the estimators proposed above with the IV estimator proposed by Harding and Lamarche (2011). The structural equation (1) is augmented using the cross-sectional means $\bar{y}_1$, $\bar{y}_2$ and $\bar{X}_1$

$$y_{1i} = y_{2i}\beta + \bar{y}_2\eta_2 + D\lambda_i + X_{1i}\theta_i + \bar{X}_1\eta_3 + \bar{y}_1\eta_1 + \tilde{u}_i \ ,$$

Notice that the endogenous variables are y_{2i} and $\bar{y}_2$. The two stage procedure is implemented as follows:

1) $[y_{2i}, \bar{y}_2]$ is regressed on d_t , $\bar{X}_2$ and the constant to obtain the predicted values $[\hat{y}_{2i}, \hat{\bar{y}}_2] = P_{[D,1,\bar{X}_2]}[y_{2i}, \bar{y}_2]$;

2) $[\hat{y}_{2i}, \hat{\bar{y}}_2]$ is replaced in the augmented equation

$$y_{1i} = P_{[D,1,\bar{X}_2]}y_{2i}\beta + P_{[D,1,\bar{X}_2]}\bar{y}_2\eta_2 + D\lambda_i + X_{1i}\theta + \bar{X}_1\eta_3 + \bar{y}_1\eta_1 + \tilde{\tilde{u}}_i \ ,$$

which is estimated using OLS.

In order to estimate $\begin{pmatrix}\beta\\ \theta\end{pmatrix}$ we premultiply the above equation by $M_{\left[P_{[D,1,\bar{X}_2]}\bar{y}_2, D, \bar{y}_1, \bar{X}_1\right]}$ to obtain

$$M_{\left[P_{[D,1,\bar{X}_2]}\bar{y}_2, D, \bar{y}_1, \bar{X}_1\right]}y_{1i} = M_{\left[P_{[D,1,\bar{X}_2]}\bar{y}_2, D, \bar{y}_1, \bar{X}_1\right]}\left(P_{[D,1,\bar{X}_2]}y_{2i}, X_1\right)\begin{pmatrix}\hat{\beta}\\ \hat{\theta}\end{pmatrix}$$

Adding up over i we have

$$\sum_{i=1}^{N}\left(P_{[D,1,\bar{X}_2]}y_{2i}, X_1\right)' M_{\left[P_{[D,1,\bar{X}_2]}\bar{y}_2, D, \bar{y}_1, \bar{X}_1\right]}y_{1i} = \sum_{i=1}^{N}\left(P_{[D,1,\bar{X}_2]}y_{2i}, X_1\right)' M_{\left[P_{[D,1,\bar{X}_2]}\bar{y}_2, D, \bar{y}_1, \bar{X}_1\right]}\left(P_{[D,1,\bar{X}_2]}y_{2i}, X_1\right)\begin{pmatrix}\hat{\beta}\\ \hat{\theta}\end{pmatrix}$$

or

$$\begin{pmatrix}\hat{\beta}\\ \hat{\theta}\end{pmatrix} = \left(\sum_{i=1}^{N}\left(P_{[D,1,\bar{X}_2]}y_{2i}, X_1\right)' M_{\left[P_{[D,1,\bar{X}_2]}\bar{y}_2, D, \bar{y}_1, \bar{X}_1\right]}\left(P_{[D,1,\bar{X}_2]}y_{2i}, X_1\right)\right)^{-1}\sum_{i=1}^{N}\left(P_{[D,1,\bar{X}_2]}y_{2i}, X_1\right)' M_{\left[P_{[D,1,\bar{X}_2]}\bar{y}_2, D, \bar{y}_1, \bar{X}_1\right]}y_{1i}$$

.

Notice that the IV estimator proposed by Harding and Lamarche (2011) can be thought of as the CCEP estimator of $y_{1i} = P_{[D,1,\bar{X}_2]} y_{2i}\beta + P_{[D,1,\bar{X}_2]} \bar{y}_2\eta_2 + D\lambda_i + X_{1i}\theta + \bar{X}_1\eta_3 + \bar{y}_1\eta_1 + \tilde{\tilde{u}}_{it}$.

Some of the estimators in the class we considered can also be interpreted as two stage procedures. Suppose that $\Pi_i = \Pi$, $\beta_i = \beta$ and $\theta_i = \theta$ as in Harding and Lamarche (2011). Then,

1) The reduced form for y_2 is estimated using (12) or (13) and the predicted value for y_2 is obtained as $\hat{y}_{2it} = \hat{\alpha}_2 ' d_t + \hat{\hat{\Pi}}_{12} ' x_{1it} + \Pi_{22} ' x_{2it}$;
2) The predicted values for y_2 are replaced into (1), which is also augmented as in Pesaran (2006) to eliminate cross sectional dependence, $y_{1i} = D\lambda_i + \hat{y}_{2i}\beta + x_{1i}\theta + \bar{X}_{1\omega}\eta_1 + \bar{y}_{1\omega}\eta_3 + \tilde{\tilde{u}}_i$. The last equation is then estimated using OLS.

It is easy to show that this yields an estimator of the form (14) or (15) with

(18) $$\hat{H} = \sum_{i=1}^{N} X_{2i} ' M_{[D,\bar{X}_2,\bar{y}_1]} X_{2i} - \sum_{i=1}^{N} X_{2i} ' M_{[D,\bar{X}_2,\bar{y}_1]} X_{1i} \left(\sum_{i=1}^{N} X_{1i} ' M_{[D,\bar{X}_2,\bar{y}_1]} X_{1i} \right)^{-1} \sum_{i=1}^{N} X_{1i} ' M_{[D,\bar{X}_2,\bar{y}_1]} X_{2i} .$$

Notice that the first stage of the procedure of Harding and Lamarche (2011) is different from that in our procedure because we take into account the fact that the instruments may also be affected by the common shocks.

5. Main results

First we show that the CCEMG and the CCEP estimators of the reduced form coefficients are consistent and asymptotically normal when both the cross-section and the time dimensions go to infinity. Notice that the reduced form is just a multivariate linear model, so that the results of Pesaran (2006) apply with obvious modification.

Proposition 1. Given Assumptions 1-5,

(a) the CCEMG estimator of the reduced form parameter Π is consistent and asymptotically normal with $\sqrt{N} vec\left(\hat{\Pi}_{MG} - \Pi\right) \rightarrow^{D} N\left(0, \Sigma_{MG}\right)$ as $\left(N,T\right) \xrightarrow{j} \infty$ and the asymptotic covariance matrix Σ_{MG} can be estimated by $\widehat{\Sigma_{MG}} = \frac{1}{N-1} \sum_{i=1}^{N} \left(vec\left(\Pi_i\right) - vec\left(\Pi_{MG}\right) \right) \left(vec\left(\Pi_i\right) - vec\left(\Pi_{MG}\right) \right)'$;

(b) the CCEP estimator of the reduced form parameter Π is consistent and asymptotically normal with $\left(\sum_{i=1}^{N}\omega_i^2\right)^{-\frac{1}{2}} vec\left(\hat{\Pi}_P-\Pi\right)\rightarrow^D N\left(0,\Sigma_P\right)$ as $\left(N,T\right)\overset{j}{\rightarrow}\infty$ and the asymptotic covariance matrix Σ_P can be estimated by $\hat{\widehat{\widehat{\Sigma}}}_P=\left(\sum_{i=1}^{N}\omega_i^2\right)\Phi^{-1}\hat{R}\Phi^{-1}$ with

$$\hat{\Phi}=\begin{bmatrix}\sum_{i=1}^{N}\omega_i\frac{1}{T}X_i'\bar{M}_{\omega}^{(1)}X_i & & 0\\ & \ddots & \\ 0 & & \sum_{i=1}^{N}\omega_i\frac{1}{T}X_i'\bar{M}_{\omega}^{(p+1)}X_i\end{bmatrix},$$

and

$$\hat{R}=\sum_{i=1}^{N}\bar{\omega}_i^2\begin{bmatrix}\frac{1}{T}X_i'\bar{M}_{\omega}^{(1)}X_i & & 0\\ & \ddots & \\ 0 & & \frac{1}{T}X_i'\bar{M}_{\omega}^{(p+1)}X_i\end{bmatrix}\left(vec\left(\hat{\hat{\Pi}}_i\right)-vec\left(\Pi_P\right)\right)\times$$

$$\left(vec\left(\hat{\hat{\Pi}}_i\right)-vec\left(\Pi_P\right)\right)'\begin{bmatrix}\frac{1}{T}X_i'\bar{M}_{\omega}^{(1)}X_i & & 0\\ & \ddots & \\ 0 & & \frac{1}{T}X_i'\bar{M}_{\omega}^{(p+1)}X_i\end{bmatrix},$$

where $\bar{\omega}_i=\omega_i/\sqrt{N^{-1}\sum_{i=1}^{N}\omega_i^2}$.

It follows easily from Proposition 1 that:

Proposition 2. Given Assumptions 1-5 and the over-identifying restrictions

(a) the CCEMG estimators of the structural parameters β and θ are consistent as $\left(N,T\right)\overset{j}{\rightarrow}\infty$, and

$$\sqrt{N}\left(\hat{\beta}_{IV-MG}-\beta\right)\rightarrow^D\left((1,-\beta')\otimes\left(\Pi_{22}'H\Pi_{22}\right)^{-1}\left(\Pi_{22}'H\right)\left(0,I_{k_2}\right)\right)N\left(0,\Sigma_{MG}\right)$$

and

$$\sqrt{N}\left(\hat{\theta}_{IV-MG}-\theta\right)\rightarrow^D\left((1,-\beta')\otimes\left(\left(I_{k_1},0\right)-\Pi_{12}\left(\Pi_{22}'H\Pi_{22}\right)^{-1}\left(\Pi_{22}'H\right)\left(0,I_{k_2}\right)\right)\right)N\left(0,\Sigma_{MG}\right).$$

(b) the CCEP estimators of the structural parameters β and θ are consistent as $(N,T)\xrightarrow{j}\infty$, and

$$\sqrt{N}\left(\hat{\beta}_{IV-P}-\beta\right)\to^{D}\left((1,-\beta')\otimes\left(\Pi_{22}'H\Pi_{22}\right)^{-1}\left(\Pi_{22}'H\right)\left(0,I_{k_2}\right)\right)N\left(0,\Sigma_P\right)$$

and

$$\sqrt{N}\left(\hat{\theta}_{IV-P}-\theta\right)\to^{D}\left((1,-\beta')\otimes\left(\left(I_{k_1},0\right)-\Pi_{12}\left(\Pi_{22}'H\Pi_{22}\right)^{-1}\left(\Pi_{22}'H\right)\left(0,I_{k_2}\right)\right)\right)N\left(0,\Sigma_P\right).$$

Therefore, if we have endogenous variables in the model of Pesaran (2006) and instruments are available, it is possible to use the CCEMG and the CCEP estimators of the reduced form parameters to estimate consistently the structural coefficients. These estimators are also asymptotically normal. Notice also that the matrices Σ_{MG} and Σ_P do not have a Kronecker product structures like in the classical instrumental variables model in a cross section. Therefore, it is not possible to choose a value of H which makes the IV estimator asymptotically efficient.

6. Small sample performance

We now study the small sample performance for two of the IV estimators in the class of estimators proposed in the previous sections and compare them with existing estimators. The data generating process is described below. We choose $p=k_1=n=1$, $k_2=2$ and $m=3$.

Factors

The factors are stationary AR(1) processes: $f_t=0.8f_{t-1}+\text{i.i.d. }N\left(0,\Sigma_f\right)$, where Σ_f is an $m\times m$ positive definite matrix generated at the start of the simulations from a Wishart distribution $W_m\left(I_m,m\right)$ for $t=-50,\ldots,-1,0,1,2,...,T$ with $f_{-50}=0$.

Factor loadings

The factor loadings in the errors and those in the exogenous variables are allowed to be correlated. Thus, we assume that

$$vec\left(\gamma_{1i},\gamma_{2i},\Gamma_{1i},\Gamma_{2i}\right)\sim\text{i.i.d.}N\left(i_{(3+k_2)m},\Omega\otimes I_m\right)$$

where

$$\Omega = \begin{pmatrix} \sigma_\gamma^2 I_2 & \rho_{\gamma,\Gamma}\sigma_\gamma\sigma_\Gamma 1_{2\times(1+k_2)} \\ \rho_{\gamma,\Gamma}\sigma_\gamma\sigma_\Gamma 1_{2\times(1+k_2)}' & \sigma_\Gamma^2 I_{1+k_2} \end{pmatrix}.$$

The parameter $-1 < \rho_{\gamma,\Gamma} < 1$ captures the correlation between the factor loadings in the errors and those in the exogenous variables. In the simulations we let $\sigma_\gamma^2 = \sigma_\Gamma^2 = 1$ and $\rho_{\gamma,\Gamma} = 0.2,\ 0.8$.

Exogenous variables

The exogenous variables are

$$\begin{pmatrix} x_{1it} \\ x_{2it} \end{pmatrix} = \begin{pmatrix} A_{1i}' \\ A_{2i}' \end{pmatrix} d_t + \begin{pmatrix} \Gamma_{1i}' \\ \Gamma_{2i}' \end{pmatrix} f_t + \begin{pmatrix} v_{1it} \\ v_{2it} \end{pmatrix}$$

where $d_t = 0.5 d_{t-1} + \text{i.i.d. } N(0, \Sigma_d)$, where Σ_{d_t} is generated at the start of the simulations as $W_n(I_n, n)$ for $t = -50, \ldots, -1, 0, 1, 2, \ldots, T$ with $d_{-50} = 0$. The term $(v_{1it}, v_{2it}')'$ follows a stationary process for each unit i and are independent over i

$$\begin{pmatrix} v_{1it} \\ v_{2it} \end{pmatrix} = 0.5 \begin{pmatrix} v_{1i,t-1} \\ v_{2i,t-1} \end{pmatrix} + \text{i.i.d. } N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{v_1}^2 & \rho_{v_1,v_2}\sigma_{v_1}\sigma_{v_2} 1_{1\times k_2} \\ \rho_{v_1,v_2}\sigma_{v_1}\sigma_{v_2} 1_{1\times k_2}' & \sigma_{v_2}^2 I_{k_2} \end{pmatrix} \right),$$

for $t = -50, \ldots, 0, 1, 2, \ldots, T$ with $\begin{pmatrix} v_{1i,-50} \\ v_{2i,-50} \end{pmatrix} = 0$ and $1_{1\times k_2}$ is a $1 \times k_2$ matrix of ones and $-1 < \rho_{v_1,v_2} < 1$. In the simulations we take $\sigma_{v_1}^2 = \sigma_{v_2}^2 = 1$ and $\rho_{v_1,v_2} = 0.5$. The components of $\begin{pmatrix} A_{1i}' \\ A_{2i}' \end{pmatrix}$ are independently uniformly distributed over the interval $(-1, 2)$.

Error terms

We use the compatibility conditions to generate the structural error u_{it} and the reduced form error e_{2it} as follows

$$\begin{pmatrix} u_{it} \\ e_{2it} \end{pmatrix} = \begin{pmatrix} \tilde{\gamma}_{1i}' \\ \tilde{\gamma}_{2i}' \end{pmatrix} f_t + \begin{pmatrix} \tilde{\varepsilon}_{1it} \\ \tilde{\varepsilon}_{2it} \end{pmatrix},$$

where $\tilde{\gamma}_{2i} = \gamma_{2i}$, $\tilde{\gamma}_{1i}' = \gamma_{1i}' - \beta_i'\gamma_{2i}'$, $\tilde{\varepsilon}_{2it} = \varepsilon_{2it}$ and $\tilde{\varepsilon}_{1it} = \varepsilon_{1it} - \beta_i'\varepsilon_{2it}$. Instead of generating the original quantities γ_{1i}, γ_{2i}, ε_{1it} and ε_{2it} we generate directly $\tilde{\gamma}_{1i}$, $\tilde{\gamma}_{2i}$, $\tilde{\varepsilon}_{1it}$ and $\tilde{\varepsilon}_{2it}$. The idiosyncratic errors $\begin{pmatrix} \tilde{\varepsilon}_{1it} \\ \tilde{\varepsilon}_{2it} \end{pmatrix}$ are generated independently over i as stationary processes

$$\begin{pmatrix} \tilde{\varepsilon}_{1it} \\ \tilde{\varepsilon}_{2it} \end{pmatrix} = 0.5 \begin{pmatrix} \tilde{\varepsilon}_{1i,t-1} \\ \tilde{\varepsilon}_{2i,t-1} \end{pmatrix} + \text{i.i.d. } N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{\tilde{\varepsilon}_1} & \rho_{\tilde{\varepsilon}_1,\tilde{\varepsilon}_2}\sigma_{\tilde{\varepsilon}_1}\sigma_{\tilde{\varepsilon}_2} \\ \rho_{\tilde{\varepsilon}_1,\tilde{\varepsilon}_2}\sigma_{\tilde{\varepsilon}_1}\sigma_{\tilde{\varepsilon}_2} & \sigma^2_{\tilde{\varepsilon}_2} \end{pmatrix} \right)$$

for $t = -50,\ldots,0,1,2,...,T$ and $\begin{pmatrix} \varepsilon_{1i0} \\ \varepsilon_{2i0} \end{pmatrix} = 0$. Notice that the square of the quantity $-1 < \rho_{\tilde{\varepsilon}_1,\tilde{\varepsilon}_2} < 1$ measures the amount of classical endogeneity when no shock is present. We let $\sigma^2_{\tilde{\varepsilon}_1} = \sigma^2_{\tilde{\varepsilon}_2} = 1$ and choose $\rho_{\tilde{\varepsilon}_1,\tilde{\varepsilon}_2} = 0.2,\ 0.8$ to reflect moderate and strong endogeneity respectively.

<u>Reduced form for y_2</u>

The reduced form for y_2 is generated as follows:

$$y_{2it} = \alpha_{2i}'d_t + \left[\Pi_{i12}', \Pi_{i22}'\right] \begin{pmatrix} x_{1it} \\ x_{2it} \end{pmatrix} + e_{2it}$$

where the component of α_{2i} are generated independently from a uniform distribution over the interval $(-1,2)$. The random coefficients satisfies

$$\left[\Pi_{i12}', \Pi_{i22}'\right] \sim \left[0, \sqrt{c}1_{k_2\times1}\right] + N\left(0, I_p \otimes \left(\frac{1}{1+k_2}\Sigma_\Pi\right)\right)$$

where Σ_Π is generated once at the beginning of the simulations as $\Sigma_\Pi \sim W_{1+k_2}\left(1+k_2, I_{1+k_2}\right)$. The parameter c measures the strength of the instruments in the absence of shocks. Moreover, we set $c = 0.8$ so that relatively weak instruments are considered.

<u>Structural equation</u>

Finally, the structural equation is generated as follows

$$y_{1it} = \lambda_i'd_t + \beta_i y_{2it} + x_{1it} + u_{it},$$

where $\beta_i = 1 + i.i.d.N(0,1)$ and λ_i are independently uniformly distributed over the interval $(1,2)$.

We consider two types of estimators in our class corresponding to different choices of $\hat{H}$. For the IV-MG and IV-P estimators, we take $\hat{H}$ equal to identity matrix. For the TSLS-MG and TSLS-P estimators, $\hat{H}$ is defined in (18). We compare these estimators with the OLS (which ignores both common shocks and endogeneity), the CCEMG and the CCEP estimators of Pesaran (2006) (which ignore the classical endogeneity), the panel TSLS of the structural parameters (which does not take the common shocks into account), and the estimator of Harding and Lamarche (2011) denoted by IV-HL. Notice that by neglecting either classical endogeneity or the endogeneity induced by the shocks, the OLS, CCEMG, CCEP and TSLS estimators are not consistent. The IV-HL estimator is also inconsitent when the shocks affect the instruments.

The outcomes of the simulation experiments based on 10000 replications are summarized in Tables 1 to 4. The tables capture different strengths of correlation between the factor loadings in the errors and in the exogenous variables and different levels of endogeneity. The IV-MG, IV-P, TSLS-MG and TSLS-P have very small bias independently of the degree of endogeneity and correlation between the factor loadings. The IV-HL appears to have small bias when the correlation between the factor loadings is small but can have very large bias when this is not the case. This is not unexpected since the IV-HL implicitly assumes that the instruments do not depend on the common factors. The OLS, TSLS, CCEMG and CCEP estimators can be seriously biased as one would expect since they have not been devised for this particular model.

The performance of the IV-MG, IV-P, TSLS-MG and TSLS-P estimators in terms of MSE is not as impressive as the one in terms of bias. This is partly due to the choice of the matrix $\hat{H}$ or to the slow convergence towards the asymptotic limits. The IV-HL estimator seems to perform well in terms of MSE when the sample size and the correlation among the factors are small. The OLS, CCEMG and CCEP estimators may have very low MSE. It often happens in structural equations models that an estimator like the OLS estimator that does not take into account simultaneity and is thus inconsistent has very low MSE due to the fact that it converges quickly as the sample size grows to infinity (and thus has low variability) and the bias is small for certain combinations of the parameters. We suspect this is what happens here.

In order to see what the distributions of the various estimators look like, we plot the densities of the difference between the estimator and the true parameter β in Figures 1 to 4 for $N = T = 100$ in the set-up described above. The densities of the IV-MG and TSLS-MG estimators are almost identical in all situations. The same applies to the densities of the IV-P and TSLS-P estimators. They are correctly centred around zero. The densities of the other estimators are shifted to the right and their relative positions change depending on the situations considered.

7. Conclusions

The paper has extended the model of Pesaran (2006) to allow for endogenous explanatory variables and for the instruments influenced by the common shocks. We have exploited the compatibility conditions between structural equation and reduced form to construct a class of estimators analogous to the classical IV estimator in cross-sectional structural equations models. Since the reduced form parameters can be estimated consistently using the CCEMG and the CCEP estimators, the structural parameters can also be consistently estimated. Moreover, since the estimators in our class have asymptotically normal distribution, tests and confidence intervals can be easily constructed. A simulation experiment, suggests that in small samples the IV estimators proposed perform reasonably well especially in terms of bias.

Appendix: proofs

Proof of Proposition 1

(a) Notice that

$$\sqrt{N}vec\left(\hat{\Pi}_{MG}-\Pi\right)=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}vec\left(\tilde{\Pi}_i\right)+\begin{bmatrix}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\frac{1}{T}X_i'\bar{M}_{\omega}^{(1)}X_i\right)^{-1}\left(\frac{1}{T}X_i'\bar{M}_{\omega}^{(1)}\varepsilon_i^{(1)}\right)\\ \dots\\ \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\frac{1}{T}X_i'\bar{M}_{\omega}^{(p+1)}X_i\right)^{-1}\left(\frac{1}{T}X_i'\bar{M}_{\omega}^{(p+1)}\varepsilon_i^{(p+1)}\right)\end{bmatrix}.$$

$$+\begin{bmatrix}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\frac{1}{T}X_i'\bar{M}_{\omega}^{(1)}X_i\right)^{-1}\left(\frac{1}{T}X_i'\bar{M}_{\omega}^{(1)}F\gamma_i^{(1)}\right)\\ \dots\\ \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\frac{1}{T}X_i'\bar{M}_{\omega}^{(p+1)}X_i\right)^{-1}\left(\frac{1}{T}X_i'\bar{M}_{\omega}^{(p+1)}F\gamma_i^{(p+1)}\right)\end{bmatrix}.$$

Either equation (56) or the equation just above (60) of Pesaran (2006) can be obtained for each component of the matrices above. Thus, as $(N,T)\overset{j}{\rightarrow}\infty$, $\sqrt{N}vec\left(\hat{\Pi}_{MG}-\Pi\right)\rightarrow N\left(0,\Sigma_{MG}\right)$ and the asymptotic covariance matrix Σ_{MG} can be estimated using $\hat{\Sigma}_{MG}$ specified in Proposition 1.

(b) Similarly, we can write

$$vec\left(\hat{\Pi}_P-\Pi\right)=\begin{bmatrix}\left(\sum_{i=1}^{N}\omega_i\frac{1}{T}X_i'\bar{M}_{\omega}^{(1)}X_i\right)^{-1} & & 0\\ & \ddots & \\ 0 & & \left(\sum_{i=1}^{N}\omega_i\frac{1}{T}X_i'\bar{M}_{\omega}^{(p+1)}X_i\right)^{-1}\end{bmatrix}\times\left(\begin{bmatrix}\sum_{i=1}^{N}\omega_i\frac{1}{T}X_i'\bar{M}_{\omega}^{(1)}X_i\tilde{\Pi}_i^{(1)}\\ \dots\\ \sum_{i=1}^{N}\omega_i\frac{1}{T}X_i'\bar{M}_{\omega}^{(p+1)}X_i\tilde{\Pi}_i^{(p+1)}\end{bmatrix}+\right.$$

$$\left.\begin{bmatrix}\sum_{i=1}^{N}\omega_i\frac{1}{T}X_i'\bar{M}_{\omega}^{(1)}F\gamma_i^{(1)}\\ \dots\\ \sum_{i=1}^{N}\omega_i\frac{1}{T}X_i'\bar{M}_{\omega}^{(p+1)}F\gamma_i^{(p+1)}\end{bmatrix}+\begin{bmatrix}\sum_{i=1}^{N}\omega_i\frac{1}{T}X_i'\bar{M}_{\omega}^{(1)}\varepsilon_i^{(1)}\\ \dots\\ \sum_{i=1}^{N}\omega_i\frac{1}{T}X_i'\bar{M}_{\omega}^{(p+1)}\varepsilon_i^{(p+1)}\end{bmatrix}\right)$$

Proceeding as in the proof of Theorem 3 of Pesaran (2006), we can conclude that every component of $vec\left(\hat{\Pi}_P-\Pi\right)$ satisfies (B.5) of Pesaran (2006), so that asymptotic normality follows by taking $(N,T)\overset{j}{\rightarrow}\infty$.□

Proof of Proposition 2.

(a) For the CCEMG estimator, notice that $\hat{\beta}_{IV-MG} = \left(\hat{\Pi}_{22MG}'\hat{H}\Pi_{22MG}\right)^{-1}\left(\Pi_{22MG}'\hat{H}\hat{\pi}_{21MG}\right)$ is a continuous function of $\hat{\Pi}_{22MG}$, $\hat{\pi}_{12MG}$ and $\hat{H}$. By construction $\hat{H} \to^{P} H$ and by Proposition 1, $\hat{\Pi}_{22MG} \to^{P} \Pi_{22}$ and $\hat{\pi}_{21MG} \to^{P} \pi_{21} = \Pi_{22}\beta$, thus $\hat{\beta}_{IV-MG} \to^{P} \left(\Pi_{22}'H\Pi_{22}\right)^{-1}\left(\Pi_{22}'H\Pi_{22}\beta\right) = \beta$. It follows from Proposition 1 and consistency of $\hat{\beta}_{IV-MG}$ that

$$\hat{\theta}_{IV-MG} = \hat{\pi}_{11MG} - \hat{\Pi}_{12MG}\hat{\beta}_{IV-MG} \to^{P} \pi_{11} - \Pi_{12}\beta = \theta .$$

Notice that if the over-identifying restrictions hold,

$$\begin{aligned}\sqrt{N}\left(\hat{\beta}_{IV-MG}-\beta\right) &= \left(\hat{\Pi}_{22MG}'\hat{H}\Pi_{22MG}\right)^{-1}\left(\Pi_{22MG}'H\right)\sqrt{N}\left(\hat{\pi}_{21MG},\Pi_{22MG}\right)\begin{pmatrix}1\\-\beta\end{pmatrix}\\ &= \left(\hat{\Pi}_{22MG}'\hat{H}\Pi_{22MG}\right)^{-1}\left(\Pi_{22MG}'H\right)\sqrt{N}\left(\left(\hat{\pi}_{21MG},\Pi_{22MG}\right)-\left(\pi_{21},\Pi_{22}\right)\right)\begin{pmatrix}1\\-\beta\end{pmatrix},\end{aligned} \tag{19}$$

so that

$$\sqrt{N}\left(\hat{\beta}_{IV-MG}-\beta\right) = \left((1,-\beta')\otimes\left(\hat{\Pi}_{22MG}'\hat{H}\Pi_{22MG}\right)^{-1}\left(\Pi_{22MG}'H\right)\left(0,I_{k_2}\right)\right)\sqrt{N}\left(vec\left(\hat{\Pi}_{MG}\right)-vec(\Pi)\right)$$

and the first part of (a) follows. For the estimator of the exogenous variables, using the over-identifying restrictions and (19) we can write

$$\begin{aligned}\sqrt{N}\left(\hat{\theta}_{IV-MG}-\theta\right) &= \sqrt{N}\left(\left(\hat{\pi}_{11MG},\hat{\Pi}_{12MG}\right)-\left(\pi_{11},\Pi_{12}\right)\right)\begin{pmatrix}1\\-\beta\end{pmatrix} - \Pi_{12MG}\sqrt{N}\left(\hat{\beta}_{IV-MG}-\beta\right)\\ &= \left(\left(I_{k_1},0\right)-\hat{\Pi}_{12MG}\left(\hat{\Pi}_{22MG}'\hat{H}\Pi_{22MG}\right)^{-1}\left(\Pi_{22MG}'H\right)\left(0,I_{k_2}\right)\right)\sqrt{N}\left(\hat{\Pi}_{MG}-\Pi\right)\begin{pmatrix}1\\-\beta\end{pmatrix}.\end{aligned}$$

Noticing that

$$\sqrt{N}\left(\hat{\theta}_{IV-MG}-\theta\right) = \left((1,-\beta')\otimes\left(\left(I_{k_1},0\right)-\hat{\Pi}_{12MG}\left(\hat{\Pi}_{22MG}'\hat{H}\Pi_{22MG}\right)^{-1}\left(\Pi_{22MG}'H\right)\left(0,I_{k_2}\right)\right)\right)\sqrt{N}vec\left(\Pi_{MG}-\Pi\right)$$

the result follows.

(b) The same argument applies to the CCEP estimators of the coefficients of both the endogenous and the exogenous variables. □

References

Ahn, SC, Lee, YH and Schmidt, P 2001, 'GMM Estimation of Linear Panel Data Models with Time-Varying Individual Effects', *Journal of Econometrics*, vol 102, pp. 219-255.

Ahn, SC, Lee, YH and Schmidt, P 2013, 'Panel Data Models with Multiple Time-Varying Individual Effects', *Journal of econometrics*, vol 174, pp. 1-14.

Arouri, M, Youssef, A, M'henni, H and Rault, C 2012, 'Energy consumption, economic growth and CO 2 emissions in Middle East and North African countries', *Energy Policy*, pp. 342–349.

Bai, J 2009, 'Panel Data Models With Interactive Fixed Effects', *Econometrica*, vol 77, pp. 1229–1279.

Baltagi, B and Moscone, F 2010, 'Health care expenditure and income in the OECD reconsidered: Evidence from panel data', *Economic Modelling*, p. Economic Modelling.

Forchini, G, Jiang, B and Peng, B 2015, 'Common Shocks in Panels with Endogenous Regressors', *Mimeo*.

Harding, M and Lamarche, C 2011, 'Least squares estimation of a panel data model with multifactor error structure and endogenous covariates', *Economics Letters*, vol 111, pp. 197-199.

Hauck, K and Zhang, X 2014, 'Heterogeneity in the Effect of Common Shocks on Health Care Expenditure Growth: A Bayesian Model Averaging on Regressions with Multifactor Error Structure', *Mimeo*.

Holly, S, Pesaran, MH and Yamagata, T 2010, 'A spatio-temporal model of house prices in the USA', *Journal of Econometrics*, vol 158, pp. 160–173.

Kapetanios, G, Pesaran, MH and Yamagata, T 2011, 'Panels with non-stationary multifactor error structures', *Journal of Econometrics*, vol 160, pp. 326–348.

Kelejian, HH 1974, 'Random Parameters in a Simultaneous Equation Framework: Identification and Estimation', *Econometrica*, p. 517528.

Pesaran, MH 2006, 'Estimation and Inference in Large Heterogeneous Panels with Multifactor Error Structure', *Econometrica*, pp. 967-1012.

Pesaran, MH and Tosetti, E 2011, 'Large panels with common factors and spatial correlation', *Journal of Econometrics*, vol 161, pp. 182–202.

Robertson, D and Sarafidis, V 2015, 'IV Estimation of Panels with Factor Residuals', *Journal of Econometrics*, vol 185, pp. 526-541.

Tables and figures

	T	N	IV-MG	IV-P	TSLS-MG	TSLS-P	OLS	CCEMG	CCEP	TSLS	IV-HL
BIAS	25	25	-0.0192	-0.0263	0.0058	0.0036	0.0965	0.0610	0.0262	0.1527	0.0654
		50	-0.0375	-0.0329	-0.0264	-0.0212	0.0942	0.0638	0.0151	0.1465	0.0683
		75	-0.0148	-0.0165	-0.0037	-0.0015	0.0963	0.0730	0.0238	0.1722	0.0733
		100	-0.0114	-0.0096	-0.0029	-0.0017	0.0893	0.0567	0.0198	0.1456	0.0711
	50	25	-0.0233	-0.0070	0.0056	0.0174	0.0936	0.0601	0.0237	0.1509	0.0848
		50	-0.0285	-0.0273	-0.0173	-0.0150	0.0957	0.0530	0.0203	0.1557	0.0688
		75	-0.0065	-0.0011	0.0024	0.0087	0.0970	0.0718	0.0238	0.1656	0.0776
		100	-0.0048	-0.0090	0.0001	-0.0034	0.0938	0.0646	0.0197	0.1594	0.0718
	75	25	-0.0268	-0.0164	-0.0018	0.0067	0.0989	0.0718	0.0304	0.1707	0.0754
		50	-0.0170	-0.0154	-0.0068	-0.0045	0.0934	0.0619	0.0178	0.1550	0.0729
		75	-0.0132	-0.0111	-0.0070	-0.0039	0.0974	0.0630	0.0214	0.1619	0.0767
		100	-0.0026	-0.0003	0.0049	0.0050	0.0967	0.0680	0.0216	0.1610	0.0749
	100	25	-0.0057	-0.0084	0.0176	0.0152	0.1031	0.0766	0.0293	0.1709	0.0798
		50	-0.0018	-0.0029	0.0104	0.0096	0.0976	0.0676	0.0216	0.1627	0.0796
		75	-0.0081	-0.0145	0.0001	-0.0051	0.0957	0.0600	0.0181	0.1600	0.0729
		100	0.0000	-0.0022	0.0055	0.0030	0.0966	0.0648	0.0205	0.1655	0.0803
MSE	25	25	0.7605	0.6375	0.8013	0.6692	0.1618	0.3538	0.2473	0.4410	0.2794
		50	0.5241	0.4295	0.5511	0.4539	0.1343	0.2423	0.1898	0.3250	0.1950
		75	0.4058	0.3357	0.4236	0.3533	0.1227	0.2047	0.1442	0.2842	0.1550
		100	0.3686	0.2997	0.3846	0.3150	0.1131	0.1812	0.1353	0.2456	0.1463
	50	25	0.4633	0.4091	0.4798	0.4242	0.1368	0.2354	0.1720	0.3209	0.2182
		50	0.3154	0.2752	0.3232	0.2793	0.1199	0.1666	0.1294	0.2446	0.1561
		75	0.2543	0.2249	0.2610	0.2290	0.1135	0.1494	0.1071	0.2269	0.1361
		100	0.2217	0.1902	0.2288	0.1987	0.1067	0.1311	0.0878	0.2103	0.1202
	75	25	0.3719	0.3246	0.3808	0.3411	0.1298	0.1927	0.1472	0.2947	0.1970
		50	0.2550	0.2264	0.2586	0.2315	0.1102	0.1393	0.1035	0.2256	0.1437
		75	0.2100	0.1851	0.2135	0.1871	0.1093	0.1190	0.0856	0.2092	0.1231
		100	0.1812	0.1593	0.1866	0.1658	0.1063	0.1127	0.0763	0.1972	0.1129
	100	25	0.3038	0.2734	0.3121	0.2834	0.1284	0.1751	0.1359	0.2728	0.1888
		50	0.2152	0.1991	0.2184	0.2045	0.1134	0.1300	0.0946	0.2167	0.1391
		75	0.1768	0.1569	0.1808	0.1605	0.1058	0.1089	0.0792	0.1960	0.1146
		100	0.1472	0.1340	0.1484	0.1343	0.1041	0.1019	0.0687	0.1908	0.1120

Table 1. Bias and MSE for the IV-MG, IV-P, TSLS-MG, TSLS-P, OLS, CCEMG, CCEP, TSLS and IV-HL estimators when both classical and factor endogeneity are small: $c = 0.8$, $\rho_{\tilde{\varepsilon}_1,\tilde{\varepsilon}_2} = 0.2$, $\rho_{\gamma,\Gamma} = 0.2$.

	T	N	IV-MG	IV-P	TSLS-MG	TSLS-P	OLS	CCEMG	CCEP	TSLS	IV-HL
BIAS	25	25	-0.0319	-0.0286	-0.0086	-0.0082	0.0802	0.0647	0.0289	0.2107	0.2383
		50	-0.0319	-0.0395	-0.0242	-0.0254	0.0720	0.0768	0.0287	0.2082	0.2300
		75	-0.0101	-0.0031	0.0010	0.0063	0.0768	0.0746	0.0231	0.2261	0.2360
		100	-0.0051	-0.0003	0.0043	0.0059	0.0668	0.0641	0.0233	0.2098	0.2287
	50	25	-0.0345	-0.0278	-0.0036	0.0008	0.0707	0.0601	0.0173	0.2156	0.2400
		50	-0.0333	-0.0305	-0.0200	-0.0154	0.0761	0.0611	0.0245	0.2106	0.2404
		75	-0.0015	0.0011	0.0070	0.0097	0.0753	0.0750	0.0228	0.2215	0.2413
		100	-0.0096	-0.0063	-0.0047	-0.0013	0.0729	0.0660	0.0206	0.2139	0.2372
	75	25	-0.0117	-0.0164	0.0163	0.0145	0.0754	0.0768	0.0285	0.2294	0.2420
		50	-0.0187	-0.0236	-0.0069	-0.0100	0.0729	0.0653	0.0183	0.2177	0.2382
		75	-0.0161	-0.0151	-0.0074	-0.0056	0.0744	0.0659	0.0216	0.2164	0.2394
		100	-0.0012	0.0021	0.0071	0.0088	0.0739	0.0737	0.0218	0.2230	0.2400
	100	25	-0.0052	-0.0142	0.0204	0.0126	0.0808	0.0833	0.0302	0.2279	0.2471
		50	-0.0084	-0.0077	0.0060	0.0059	0.0751	0.0704	0.0215	0.2223	0.2468
		75	-0.0115	-0.0176	-0.0026	-0.0061	0.0743	0.0638	0.0198	0.2164	0.2395
		100	-0.0041	-0.0022	0.0034	0.0027	0.0743	0.0705	0.0203	0.2235	0.2420
MSE	25	25	0.8013	0.6713	0.8473	0.6871	0.1686	0.3688	0.2632	0.4706	0.3760
		50	0.5331	0.4498	0.5719	0.4730	0.1268	0.2599	0.2099	0.3609	0.3034
		75	0.4378	0.3675	0.4619	0.3854	0.1176	0.2223	0.1538	0.3278	0.2825
		100	0.3875	0.3130	0.4072	0.3281	0.1019	0.1911	0.1445	0.2919	0.2676
	50	25	0.4726	0.4245	0.4864	0.4387	0.1297	0.2381	0.1743	0.3536	0.3254
		50	0.3338	0.2990	0.3410	0.3032	0.1116	0.1783	0.1359	0.2858	0.2874
		75	0.2663	0.2458	0.2747	0.2525	0.1000	0.1552	0.1106	0.2700	0.2701
		100	0.2222	0.1959	0.2275	0.2037	0.0933	0.1341	0.0943	0.2545	0.2607
	75	25	0.3713	0.3270	0.3772	0.3356	0.1197	0.1996	0.1544	0.3242	0.3361
		50	0.2630	0.2393	0.2691	0.2435	0.0995	0.1472	0.1081	0.2686	0.2734
		75	0.2166	0.1951	0.2203	0.1986	0.0927	0.1239	0.0902	0.2535	0.2624
		100	0.1871	0.1691	0.1917	0.1747	0.0892	0.1205	0.0801	0.2507	0.2584
	100	25	0.3048	0.2883	0.3145	0.2936	0.1158	0.1807	0.1407	0.3026	0.3208
		50	0.2246	0.2112	0.2302	0.2165	0.0986	0.1327	0.0966	0.2645	0.2791
		75	0.1816	0.1704	0.1837	0.1717	0.0900	0.1137	0.0803	0.2439	0.2604
		100	0.1528	0.1423	0.1551	0.1428	0.0856	0.1076	0.0685	0.2435	0.2571

Table 2. Bias and MSE for the IV-MG, IV-P, TSLS-MG, TSLS-P, OLS, CCEMG, CCEP, TSLS and IV-HL estimators when classical endogeneity is small but factor endogeneity is large: $c = 0.8$, $\rho_{\tilde{\varepsilon}_1,\tilde{\varepsilon}_2} = 0.2$, $\rho_{\gamma,\Gamma} = 0.8$.

	T	N	IV-MG	IV-P	TSLS-MG	TSLS-P	OLS	CCEMG	CCEP	TSLS	IV-HL
BIAS	25	25	-0.0145	-0.0232	0.0088	0.0085	0.1036	0.2465	0.0926	0.1513	0.0662
		50	-0.0342	-0.0315	-0.0213	-0.0210	0.1014	0.2523	0.0781	0.1482	0.0694
		75	-0.0122	-0.0145	-0.0006	0.0006	0.1040	0.2535	0.0843	0.1742	0.0759
		100	-0.0109	-0.0093	-0.0013	-0.0020	0.0958	0.2487	0.0820	0.1465	0.0718
	50	25	-0.0204	-0.0080	0.0100	0.0187	0.1005	0.2408	0.0886	0.1513	0.0865
		50	-0.0267	-0.0250	-0.0137	-0.0110	0.1028	0.2461	0.0834	0.1573	0.0701
		75	-0.0065	-0.0009	0.0035	0.0105	0.1038	0.2637	0.0875	0.1657	0.0776
		100	-0.0037	-0.0072	0.0017	-0.0010	0.1006	0.2563	0.0810	0.1604	0.0727
	75	25	-0.0205	-0.0112	0.0078	0.0141	0.1063	0.2615	0.0960	0.1726	0.0744
		50	-0.0172	-0.0155	-0.0051	-0.0025	0.1005	0.2553	0.0813	0.1550	0.0736
		75	-0.0115	-0.0099	-0.0044	-0.0029	0.1038	0.2571	0.0841	0.1615	0.0766
		100	-0.0011	0.0004	0.0069	0.0066	0.1028	0.2588	0.0817	0.1613	0.0747
	100	25	-0.0027	-0.0048	0.0221	0.0208	0.1098	0.2635	0.0931	0.1729	0.0805
		50	0.0009	-0.0017	0.0150	0.0132	0.1040	0.2592	0.0860	0.1625	0.0806
		75	-0.0063	-0.0127	0.0023	-0.0024	0.1020	0.2538	0.0809	0.1603	0.0728
		100	0.0001	-0.0016	0.0059	0.0035	0.1030	0.2555	0.0812	0.1663	0.0814
MSE	25	25	0.7624	0.6375	0.7911	0.6652	0.1660	0.4261	0.2677	0.4379	0.2801
		50	0.5208	0.4249	0.5507	0.4474	0.1392	0.3462	0.2065	0.3263	0.1951
		75	0.4056	0.3378	0.4214	0.3541	0.1289	0.3202	0.1693	0.2873	0.1567
		100	0.3694	0.2996	0.3863	0.3147	0.1181	0.3048	0.1610	0.2467	0.1476
	50	25	0.4593	0.4035	0.4759	0.4208	0.1417	0.3338	0.1956	0.3194	0.2178
		50	0.3154	0.2758	0.3233	0.2782	0.1255	0.2958	0.1572	0.2455	0.1569
		75	0.2550	0.2255	0.2611	0.2301	0.1193	0.2963	0.1409	0.2279	0.1366
		100	0.2229	0.1912	0.2300	0.1998	0.1126	0.2825	0.1234	0.2110	0.1206
	75	25	0.3728	0.3253	0.3820	0.3422	0.1359	0.3198	0.1781	0.2977	0.1956
		50	0.2560	0.2276	0.2599	0.2323	0.1165	0.2878	0.1337	0.2243	0.1442
		75	0.2095	0.1852	0.2130	0.1873	0.1151	0.2793	0.1220	0.2086	0.1235
		100	0.1822	0.1612	0.1867	0.1673	0.1120	0.2777	0.1147	0.1976	0.1128
	100	25	0.3025	0.2731	0.3112	0.2830	0.1341	0.3109	0.1658	0.2716	0.1890
		50	0.2150	0.1983	0.2188	0.2046	0.1189	0.2859	0.1310	0.2168	0.1401
		75	0.1762	0.1561	0.1803	0.1599	0.1116	0.2725	0.1165	0.1957	0.1147
		100	0.1476	0.1342	0.1486	0.1345	0.1100	0.2707	0.1098	0.1917	0.1129

Table 3. Bias and MSE for the IV-MG, IV-P, TSLS-MG, TSLS-P, OLS, CCEMG, CCEP, TSLS and IV-HL estimators when classical endogeneity is large but factor endogeneity is small: $c = 0.8$, $\rho_{\tilde{\varepsilon}_1,\tilde{\varepsilon}_2} = 0.8$, $\rho_{\gamma,\Gamma} = 0.2$.

	T	N	IV-MG	IV-P	TSLS-MG	TSLS-P	OLS	CCEMG	CCEP	TSLS	IV-HL
BIAS	25	25	-0.0346	-0.0252	-0.0055	-0.0034	0.0867	0.2607	0.1008	0.2070	0.2362
		50	-0.0319	-0.0384	-0.0197	-0.0203	0.0799	0.2715	0.0931	0.2081	0.2322
		75	-0.0083	-0.0037	0.0046	0.0068	0.0841	0.2671	0.0867	0.2254	0.2361
		100	-0.0129	-0.0047	-0.0049	-0.0026	0.0746	0.2707	0.0923	0.2070	0.2287
	50	25	-0.0301	-0.0246	0.0015	0.0064	0.0792	0.2603	0.0886	0.2178	0.2429
		50	-0.0355	-0.0322	-0.0200	-0.0151	0.0830	0.2672	0.0928	0.2105	0.2398
		75	-0.0024	0.0010	0.0076	0.0108	0.0829	0.2782	0.0896	0.2216	0.2411
		100	-0.0084	-0.0045	-0.0024	0.0015	0.0800	0.2713	0.0859	0.2157	0.2386
	75	25	-0.0105	-0.0163	0.0201	0.0161	0.0838	0.2783	0.1014	0.2288	0.2418
		50	-0.0201	-0.0234	-0.0069	-0.0094	0.0795	0.2717	0.0862	0.2188	0.2378
		75	-0.0140	-0.0134	-0.0046	-0.0039	0.0813	0.2716	0.0888	0.2169	0.2395
		100	-0.0003	0.0033	0.0084	0.0098	0.0806	0.2799	0.0890	0.2229	0.2399
	100	25	-0.0029	-0.0128	0.0249	0.0156	0.0888	0.2812	0.1006	0.2271	0.2502
		50	-0.0063	-0.0063	0.0103	0.0098	0.0820	0.2750	0.0899	0.2233	0.2469
		75	-0.0109	-0.0162	-0.0017	-0.0046	0.0819	0.2696	0.0871	0.2167	0.2404
		100	-0.0043	-0.0025	0.0044	0.0035	0.0817	0.2740	0.0876	0.2227	0.2421
MSE	25	25	0.7993	0.6734	0.8435	0.6914	0.1736	0.4482	0.2864	0.4712	0.3746
		50	0.5369	0.4525	0.5725	0.4703	0.1310	0.3689	0.2280	0.3595	0.3055
		75	0.4394	0.3697	0.4589	0.3865	0.1226	0.3412	0.1763	0.3262	0.2825
		100	0.3859	0.3111	0.4037	0.3275	0.1072	0.3261	0.1736	0.2894	0.2672
	50	25	0.4786	0.4302	0.4948	0.4448	0.1348	0.3510	0.1980	0.3548	0.3287
		50	0.3338	0.2985	0.3396	0.3027	0.1168	0.3190	0.1673	0.2865	0.2870
		75	0.2634	0.2446	0.2714	0.2522	0.1059	0.3113	0.1447	0.2683	0.2695
		100	0.2233	0.1963	0.2286	0.2044	0.0990	0.2977	0.1300	0.2559	0.2623
	75	25	0.3740	0.3311	0.3800	0.3402	0.1253	0.3369	0.1867	0.3242	0.3322
		50	0.2645	0.2396	0.2714	0.2438	0.1047	0.3058	0.1403	0.2690	0.2725
		75	0.2176	0.1963	0.2215	0.2001	0.0983	0.2934	0.1291	0.2542	0.2626
		100	0.1874	0.1695	0.1916	0.1752	0.0948	0.2982	0.1234	0.2504	0.2582
	100	25	0.3056	0.2891	0.3160	0.2947	0.1216	0.3272	0.1752	0.3020	0.3248
		50	0.2255	0.2118	0.2310	0.2168	0.1039	0.3006	0.1344	0.2657	0.2794
		75	0.1812	0.1691	0.1830	0.1707	0.0966	0.2876	0.1206	0.2442	0.2612
		100	0.1523	0.1422	0.1544	0.1426	0.0920	0.2882	0.1140	0.2428	0.2571

Table 4. Bias and MSE for the IV-MG, IV-P, TSLS-MG, TSLS-P, OLS, CCEMG, CCEP, TSLS and IV-HL estimators when both classical and factor endogeneity are large: $c = 0.8$, $\rho_{\tilde{\varepsilon}_1,\tilde{\varepsilon}_2} = 0.8$, $\rho_{\gamma,\Gamma} = 0.8$.

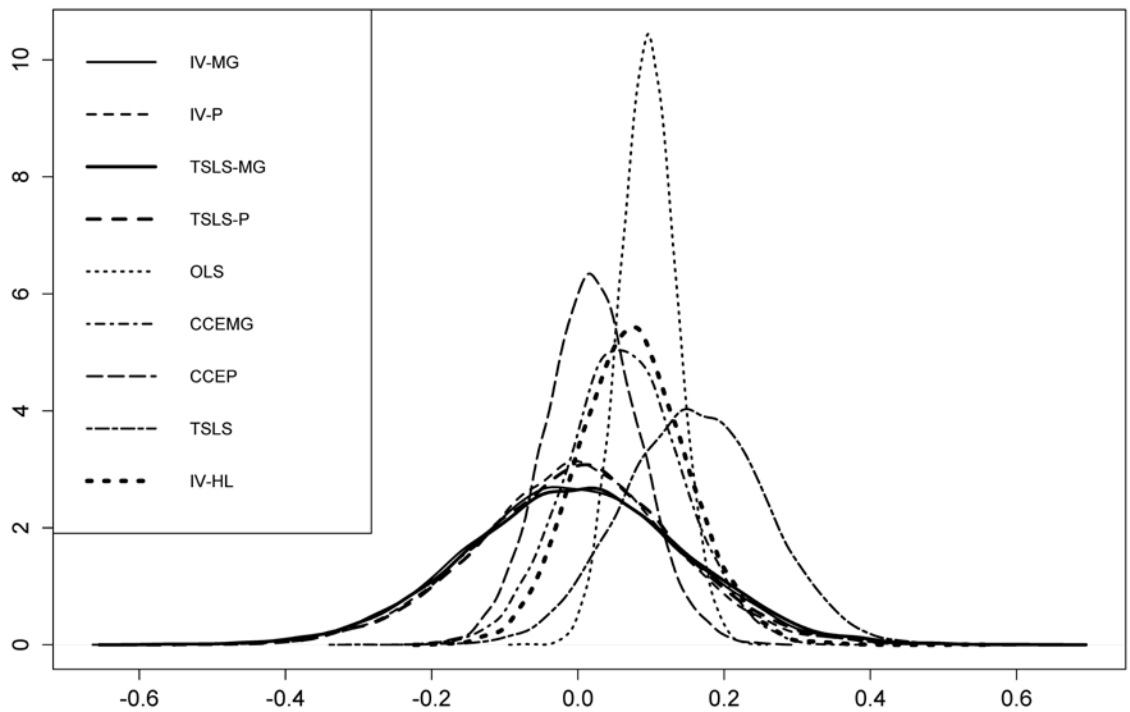

Figure 1: Densities of the IV-MG, IV-P, TSLS-MG, TSLS-P, OLS, CCEMG, CCEP, TSLS and IV-HL estimators when both classical and factor endogeneity are small: $c = 0.8$, $\rho_{\tilde{\varepsilon}_1,\tilde{\varepsilon}_2} = 0.2$, $\rho_{\gamma,\Gamma} = 0.2$, $N = T = 100$ based on 10000 replications.

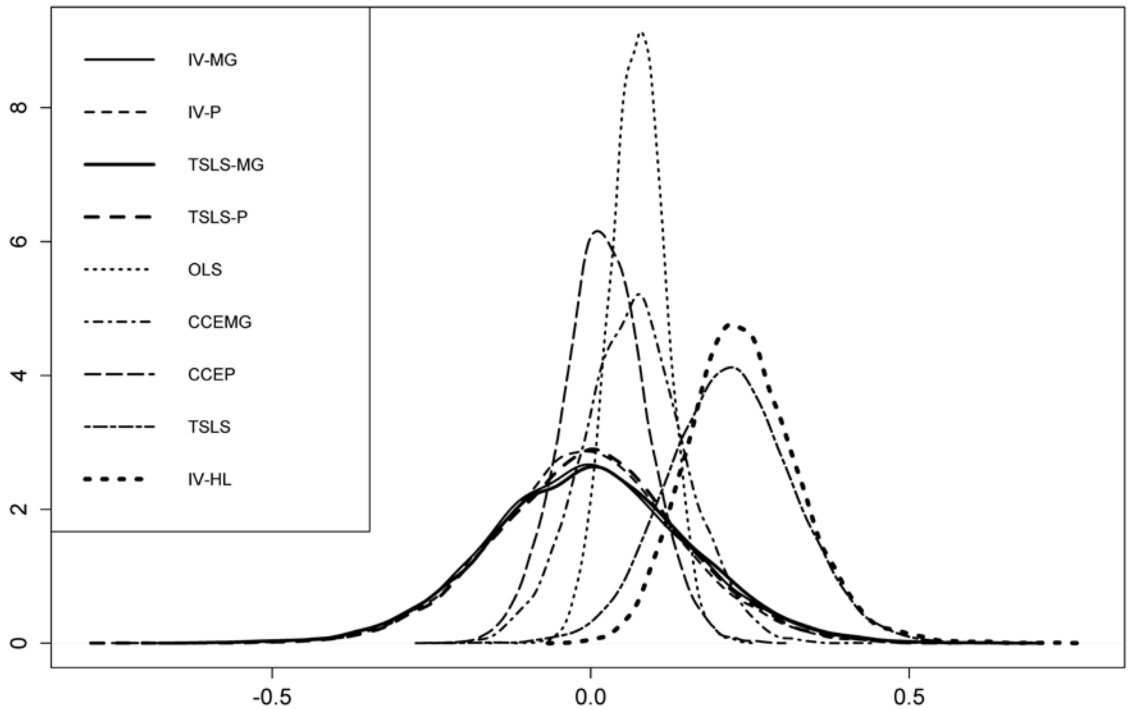

Figure 2: Densities of IV-MG, IV-P, TSLS-MG, TSLS-P, OLS, CCEMG, CCEP, TSLS and IV-HL estimators when classical endogeneity is small but factor endogeneity is large: $c = 0.8$, $\rho_{\tilde{\varepsilon}_1,\tilde{\varepsilon}_2} = 0.2$, $\rho_{\gamma,\Gamma} = 0.8$, $N = T = 100$ based on 10000 replications.

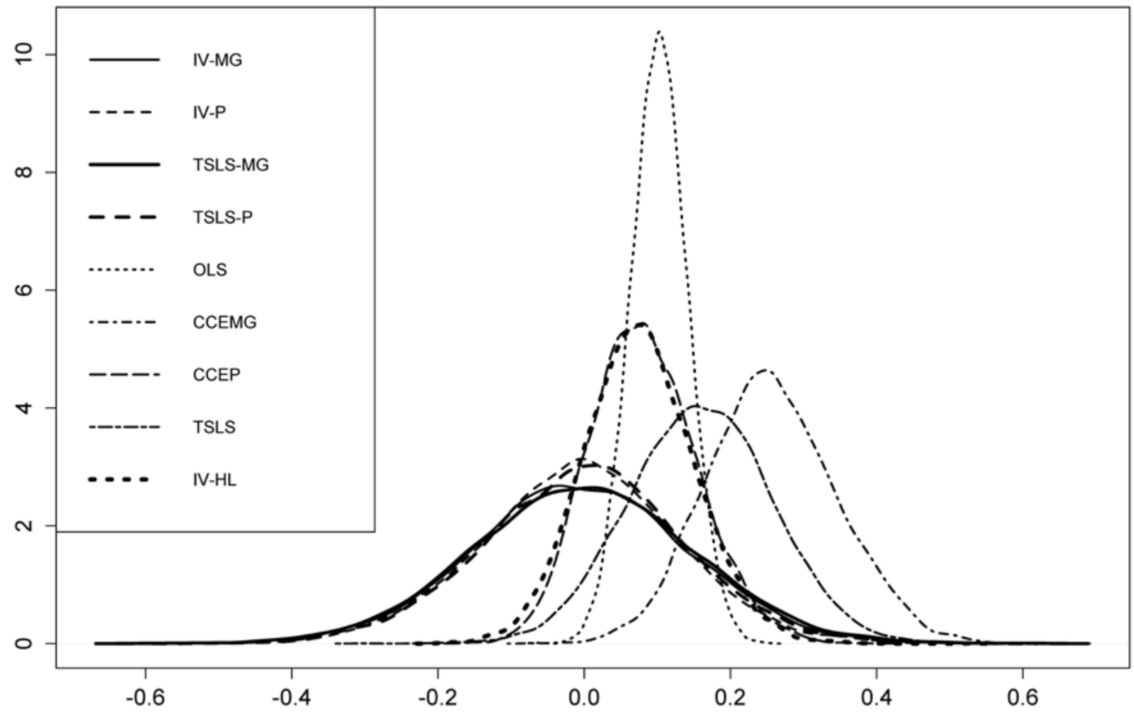

Figure 3: Densities of the IV-MG, IV-P, TSLS-MG, TSLS-P, OLS, CCEMG, CCEP, TSLS and IV-HL estimators when classical endogeneity is large but factor endogeneity is small: $c = 0.8$, $\rho_{\tilde{\varepsilon}_1,\tilde{\varepsilon}_2} = 0.8$, $\rho_{\gamma,\Gamma} = 0.2$, $N = T = 100$ based on 10000 replications.

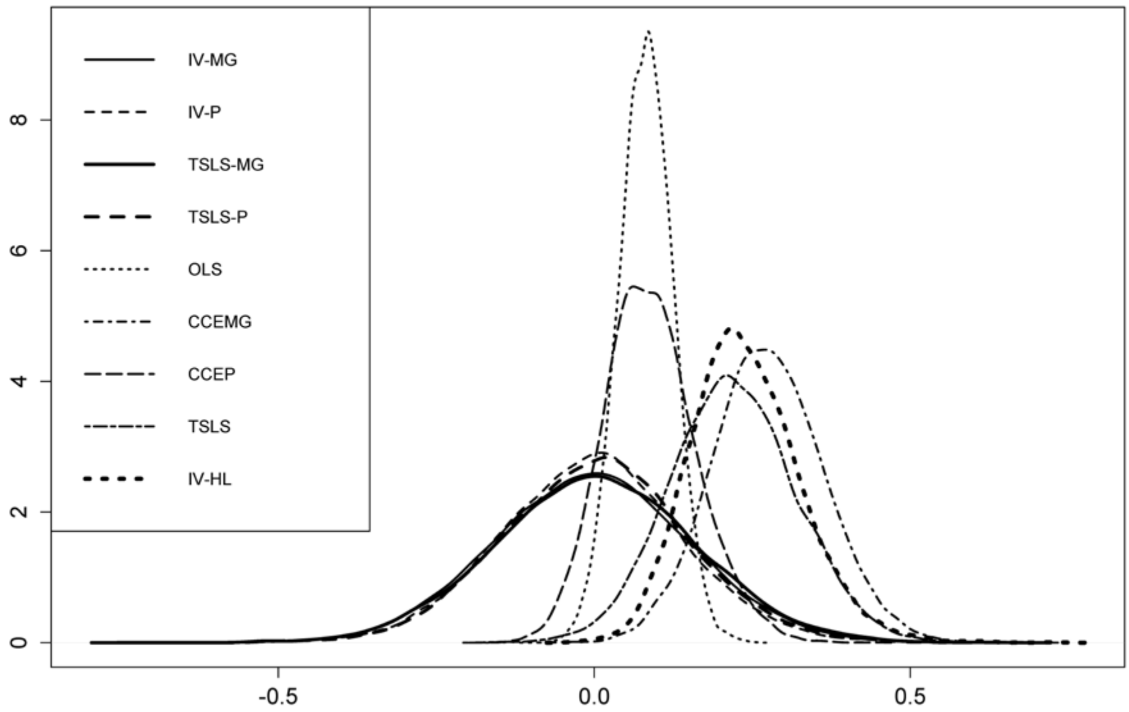

Figure 4; Densities of the IV-MG, IV-P, TSLS-MG, TSLS-P, OLS, CCEMG, CCEP, TSLS and IV-HL estimators when both classical and factor endogeneity are large: $c = 0.8$, $\rho_{\tilde{\varepsilon}_1,\tilde{\varepsilon}_2} = 0.8$, $\rho_{\gamma,\Gamma} = 0.8$, $N = T = 100$ based on 10000 replications.