Panel Econometrics

Paulo M. M. Rodrigues

February 2022

Today's outline

Instrumental Variables

1. Panel IV estimation
2. Random effects IV estimation
3. Testing regressor exogeneity
4. Up next

Outline

Motivation

Our previous estimators all relied on assumptions that in some cases may be deemed inappropriate.

- RE, FE, and FD all require strict exogeneity of the regressors:

$$\mathrm{E}(u_{it}|\mathbf{x}_{i1}, \ldots, \mathbf{x}_{iT}, c_i) = 0.$$

- POLS and RE require that the regressors be uncorrelated with the unobserved individual effect:

$$\mathrm{E}(\mathbf{x}'_{it} c_i) = \mathbf{0}.$$

Whenever we suspect that strict exogeneity (and perhaps additionally uncorrelatedness with the unobserved individual effect) does not hold ("regressor endogeneity"), we may resort to IV estimation.

Assumptions

We have K regressors and $L \geq K$ instruments $\mathbf{z}_{it}$ for which we assume

$$\mathrm{E}(\mathbf{z}'_{it} u_{it}) = \mathbf{0}, \qquad t = 1, \dots, T,$$

and

$$\mathrm{E}(\mathbf{z}'_{it} c_i) = \mathbf{0}, \qquad t = 1, \dots, T.$$

Note that the second assumption says that the instruments are uncorrelated with the unobserved individual effect which may be strong.

(We also need appropriate rank conditions but like always this is typically satisfied if we use appropriate and strong instruments.)

The basics: Pooled IV estimator

In the *just identified* case $K = L$, we can stack all observations $\mathbf{z}_{it}$ of individual i in a $T \times L$ matrix $\mathbf{Z}_i$, and all observations of all individuals in a $NT \times L$ matrix $\mathbf{Z}$.

A straightforward IV estimator is

$$\hat{\boldsymbol{\beta}}_{PIV} = \left(\mathbf{Z}'\mathbf{X}\right)^{-1}\mathbf{Z}'\mathbf{y} = \left(\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{X}_i\right)^{-1}\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{y}_i$$

Pooled 2SLS estimator

In case of *overidentification* we may use the panel version of the 2SLS estimator.

The P2SLS estimator is correspondingly

$$\begin{aligned}\hat{\boldsymbol{\beta}}_{P2SLS} &= \left(\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{y} \\ &= \left[\left(\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{Z}_i\right)\hat{\mathbf{W}}\left(\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{X}_i\right)\right]^{-1}\left(\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{Z}_i\right)\hat{\mathbf{W}}\left(\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{y}_i\right),\end{aligned}$$

where

$$\hat{\mathbf{W}} = N^{-1}\left(\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{Z}_i\right)^{-1} = N^{-1}(\mathbf{Z}'\mathbf{Z})^{-1}.$$

Asymptotic distribution

The asymptotic distribution of the P2SLS estimator applying the CLT yields

$$N^{-1/2}\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{v}_i \xrightarrow{\text{d}} \text{Normal}(0,\boldsymbol{\Lambda}),$$

where $\boldsymbol{\Lambda} \equiv \mathrm{E}(\mathbf{Z}_i'\mathbf{v}_i\mathbf{v}_i'\mathbf{Z}_i)$. Applying the LLN yields

$$N^{-1}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{Z}_i \xrightarrow{\text{p}} \mathbf{C} \qquad \text{and} \qquad \hat{\mathbf{W}} = N^{-1}\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{Z}_i \xrightarrow{\text{p}} \mathbf{W}.$$

Putting the parts together using Slutsky's theorem yields

$$N^{1/2}(\hat{\boldsymbol{\beta}}_{P2SLS}-\boldsymbol{\beta}) \xrightarrow{\text{d}} \text{Normal}(0,(\mathbf{C}'\mathbf{W}\mathbf{C})^{-1}\mathbf{C}'\mathbf{W}\boldsymbol{\Lambda}\mathbf{W}\mathbf{C}(\mathbf{C}'\mathbf{W}\mathbf{C})^{-1}).$$

Feasible versions

Like always, the unknown population moments can be estimated by sample averages. This yields $\hat{\mathbf{C}}$ and $\hat{\mathbf{W}}$.

What is important in the panel context is that the $\mathbf{v}_i$ are certainly autocorrelated (due to the unobserved individual effect) and probably heteroscedastic.

Hence an appropriate estimator of $\boldsymbol{\Lambda}$ is

$$\hat{\boldsymbol{\Lambda}} = N^{-1} \sum_{i=1}^{N} \mathbf{Z}_i' \hat{\mathbf{v}}_i \hat{\mathbf{v}}_i' \mathbf{Z}_i,$$

where $\hat{\mathbf{v}}_i$ are the residuals.

... robustness rules.

Pooled GMM estimator

The P2SLS estimator can be generalized to any appropriate weighting matrix $\hat{\mathbf{W}}$.

This yields a system GMM estimator.

In particular, based on an initial estimator (such as P2SLS), we can estimate the optimal weighting matrix

$$\hat{\mathbf{W}} = \hat{\boldsymbol{\Lambda}}^{-1} = \left(N^{-1} \sum_{i=1}^{N} \mathbf{Z}_i' \hat{\hat{\mathbf{v}}}_i \hat{\hat{\mathbf{v}}}_i' \mathbf{Z}_i \right)^{-1},$$

where $\hat{\hat{\mathbf{v}}}_i$ are the residuals obtained from the initial estimator, and perform optimal GMM.

This estimator is optimal among the GMM estimators that do not impose any structure on $\mathbf{v}_i$ (except for $\mathrm{E}(\mathbf{z}_{it}' c_i) = \mathbf{0}$ and $\mathrm{E}(\mathbf{z}_{it}' u_{it}) = \mathbf{0}$ of course).

Implementation in Stata

Suppose you have one exogenous regressor x1 and one endogenous regressor x2 for which two instruments z1 and z2 are available.

You first have to tell Stata that you have panel data:

```
xtset id year
```

P2SLS with s.e.'s robust to heteroscedasticity and arbitrary correlation within individuals:

```
ivregress 2sls y x1 (x2 = z1 z2), vce(cluster id)
```

Panel GMM with optimal weights and s.e.'s robust to heteroscedasticity and arbitrary correlation within individuals:

```
gmm (y-b1*x1-b2*x2-b0), instruments(x1 z1 z2)
winitial(unadjusted) wmatrix(cluster id) vce(cluster id)
```

Example: Demand for air travel

Example 11.1 taken from Wooldridge's textbook

Question: what is the price elasticity of demand for air travel in the U.S.?

Data set airfare.dta includes annual information for 1,149 U.S. routes for years 1997-2000.

Demand equation:

$$\texttt{lpas}_{it} = \beta_1 \texttt{lfare}_{it} + \beta_2 \texttt{ldist}_{it} + \beta_3 \texttt{ldist}^2_{it} + \texttt{year dummies} + c_{i,1} + u_{it,1}$$

Reduced form equation (first-stage regression) for fare price:

$$\texttt{lfare}_{it} = \gamma_1 \texttt{concen}_{it} + \gamma_2 \texttt{ldist}_{it} + \gamma_3 \texttt{ldist}^2_{it} + \texttt{year dummies} + c_{i,2} + u_{it,2}$$

- $\texttt{lpas}_{it}$: log of average passenger number per day
- $\texttt{lfare}_{it}$: log of average one-way airfare in US dollars
- $\texttt{ldist}_{it}$: log of distance in miles
- $\texttt{ldist}^2_{it}$: squared log of distance
- $\texttt{concen}_{it}$: measure of concentration (market share of largest carrier on this route)

Stata

```
*** load data and set panel ***

use "airfare.dta", clear
xtset id year


*** P2SLS with original data ***

ivregress 2sls lpas (lfare = concen) ldist ldistsq y98 y99
y00, vce(cluster id)


*** GMM yields the same result because just identifying ***
*** Therefore, option onestep is sufficient ***

gmm
(lpas-b1*lfare-b2*ldist-b3*ldistsq-b4*y98-b5*y99-b6*y00-b0),
instruments(concen ldist ldistsq y98 y99 y00)
winitial(unadjusted) wmatrix(cluster id) vce(cluster id)
```

Output I

```
. ivregress 2sls lpassen (lfare = concen) ldist ldistsq y98 y99 y00, vce(cluster id)

Instrumental variables (2SLS) regression          Number of obs  =     4596
                                                  Wald chi2(6)   =   168.47
                                                  Prob > chi2    =   0.0000
                                                  R-squared      =        .
                                                  Root MSE       =    .9499

                              (Std. Err. adjusted for 1149 clusters in id)
```

lpassen	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
lfare	-1.776549	.4748196	-3.74	0.000	-2.707178	-.8459196
ldist	-2.498972	.8304964	-3.01	0.003	-4.126715	-.8712285
ldistsq	.2314932	.0704479	3.29	0.001	.0934178	.3695687
y98	.0616171	.0131388	4.69	0.000	.0358656	.0873686
y99	.1241675	.0183135	6.78	0.000	.0882737	.1600613
y00	.2542695	.0457529	5.56	0.000	.1645955	.3439435
_cons	21.21249	3.856458	5.50	0.000	13.65397	28.77101

```
Instrumented:  lfare
Instruments:   ldist ldistsq y98 y99 y00 concen
```

Output II

```
. gmm (lpassen - {b1}*lfare - {b2}*ldist - {b3}*ldistsq - {b4}*y98 - {b5}*y99 -
> {b6}*y00 - {b0}), winitial(unadjusted) onestep instruments(concen ldist ldists
> q y98 y99 y00) vce(cluster id)

Step 1
Iteration 0:   GMM criterion Q(b) =  36.228744
Iteration 1:   GMM criterion Q(b) =  1.234e-15
Iteration 2:   GMM criterion Q(b) =  3.131e-25

GMM estimation

Number of parameters =   7
Number of moments    =   7
Initial weight matrix: Unadjusted                     Number of obs   =      4596

                                  (Std. Err. adjusted for 1149 clusters in id)
------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         /b1 |  -1.776549   .4748196    -3.74   0.000    -2.707178   -.8459196
         /b2 |  -2.498972   .8304968    -3.01   0.003    -4.126715   -.8712278
         /b3 |   .2314932    .070448     3.29   0.001     .0934177    .3695687
         /b4 |   .0616171   .0131388     4.69   0.000     .0358656    .0873686
         /b5 |   .1241675   .0183135     6.78   0.000     .0882737    .1600613
         /b6 |   .2542695   .0457529     5.56   0.000     .1645955    .3439435
         /b0 |   21.21249   3.856459     5.50   0.000     13.65397    28.77101
------------------------------------------------------------------------------
Instruments for equation 1: concen ldist ldistsq y98 y99 y00 _cons
```

Outline

Assumptions

When we accept strict exogeneity of the instruments, we can use a random effects IV estimator. (Should be "more efficient")

We have K regressors and $L \geq K$ instruments $\mathbf{z}_{it}$ for which we assume REIV.1:

$$\text{(a)} \quad \mathrm{E}(u_{it}|\mathbf{z}_{i1}, \ldots, \mathbf{z}_{iT}, c_i) = 0, \qquad t = 1, \ldots, T,$$

and

$$\text{(b)} \quad \mathrm{E}(c_i|\mathbf{z}_{i1}, \ldots, \mathbf{z}_{iT}) = \mathrm{E}(c_i) = 0, \qquad t = 1, \ldots, T.$$

We also need appropriate rank conditions REIV.2: rank $\mathrm{E}(\mathbf{Z}_i'\boldsymbol{\Omega}^{-1}\mathbf{Z}_i) = L$ and rank $\mathrm{E}(\mathbf{Z}_i'\boldsymbol{\Omega}^{-1}\mathbf{X}_i) = K$. (Typically satisfied etc.)

Finally, we may impose the by now well-known error-components structure assumption REIV.3: $\mathrm{E}(\mathbf{u}_i'\mathbf{u}_i|\mathbf{Z}_i, c_i) = \sigma_u^2\mathbf{I}_T$ and $\mathrm{E}(c_i^2|\mathbf{Z}_i) = \sigma_c^2$.

REIV estimator

To estimate the structural equation

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{v}_i$$

by REIV, first transform the equation to

$$\sigma_u\boldsymbol{\Omega}^{-1/2}\mathbf{y}_i = \sigma_u\boldsymbol{\Omega}^{-1/2}\mathbf{X}_i\boldsymbol{\beta} + \sigma_u\boldsymbol{\Omega}^{-1/2}\mathbf{v}_i \qquad \Leftrightarrow \qquad \tilde{\mathbf{y}}_i = \tilde{\mathbf{X}}_i\boldsymbol{\beta} + \tilde{\mathbf{v}}_i.$$

Using transformed instruments $\tilde{\mathbf{Z}}_i = \sigma_u\boldsymbol{\Omega}^{-1/2}\mathbf{Z}_i$ and stacking all observations $i = 1, \ldots, N$ yields the estimator

$$\begin{aligned}\hat{\boldsymbol{\beta}}_{REIV} &= \left(\tilde{\mathbf{X}}'\tilde{\mathbf{Z}}(\tilde{\mathbf{Z}}'\tilde{\mathbf{Z}})^{-1}\tilde{\mathbf{Z}}'\tilde{\mathbf{X}}\right)^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{Z}}(\tilde{\mathbf{Z}}'\tilde{\mathbf{Z}})^{-1}\tilde{\mathbf{Z}}'\tilde{\mathbf{y}} \\ &= \left\{[\mathbf{X}'(\mathbf{I}_N \otimes \boldsymbol{\Omega}^{-1})\mathbf{Z}][\mathbf{Z}'(\mathbf{I}_N \otimes \boldsymbol{\Omega}^{-1})\mathbf{Z}][\mathbf{Z}'(\mathbf{I}_N \otimes \boldsymbol{\Omega}^{-1})\mathbf{Z}]\right\}^{-1} \\ &\quad \times [\mathbf{X}'(\mathbf{I}_N \otimes \boldsymbol{\Omega}^{-1})\mathbf{Z}][\mathbf{Z}'(\mathbf{I}_N \otimes \boldsymbol{\Omega}^{-1})\mathbf{Z}][\mathbf{Z}'(\mathbf{I}_N \otimes \boldsymbol{\Omega}^{-1})\mathbf{y}].\end{aligned}$$

Obviously, to make the REIV estimator feasible, we have to estimate $\boldsymbol{\Omega}$ in a preliminary step (just like for plain vanilla RE).

Quasi de-meaning

Since $\mathbf{\Omega}$ is the same as for the standard RE estimator, we can analogously obtain the transformations by quasi de-meaning:

$$\begin{aligned}\tilde{\mathbf{y}}_i &\equiv \sigma_u \mathbf{\Omega}^{-1/2}\mathbf{y}_i = \mathbf{y}_i - (1-\phi)\bar{\mathbf{y}}_i\\ \tilde{\mathbf{X}}_i &\equiv \sigma_u \mathbf{\Omega}^{-1/2}\mathbf{X}_i = \mathbf{X}_i - (1-\phi)\bar{\mathbf{X}}_i\\ \tilde{\mathbf{Z}}_i &\equiv \sigma_u \mathbf{\Omega}^{-1/2}\mathbf{Z}_i = \mathbf{Z}_i - (1-\phi)\bar{\mathbf{Z}}_i\end{aligned}$$

where

$$\phi^2 = \frac{\sigma_u^2}{T\sigma_c^2 + \sigma_u^2}.$$

The variance components can be estimated the same way as for RE:

- Wooldridge's approach is to apply P2SLS in the first step and then use the P2SLS residuals to estimate σ_c^2 and σ_u^2.
- A Swamy-Arora type approach would use IV fixed effects and IV between estimation.

Outline

Robustness vs. efficiency

Suppose we have an error-components model

$$y_{it} = \mathbf{x}_{1,it}\boldsymbol{\beta}_1 + \mathbf{x}_{2,it}\boldsymbol{\beta}_2 + \mathbf{x}_{3,it}\boldsymbol{\beta}_3 + v_{it}$$

where

- $\mathbf{x}_{1,it}$ is surely exogenous and used as its own instrument,
- $\mathbf{x}_{2,it}$ is surely endogenous and instrumented by $\mathbf{z}_{2,it}$, and
- $\mathbf{x}_{3,it}$ is a vector of J_3 regressors for which it is unclear.

Taking $\mathbf{x}_{3,it}$ as exogenous leads us to the instrument set $[\mathbf{x}_{1,it}, \mathbf{z}_{2,it}, \mathbf{x}_{3,it}]$, where we assume that $\mathbf{z}_{2,it}$ is sufficient to identify the parameters of $\mathbf{x}_{2,it}$.

Let us denote the instrument set that is certainly exogenous as $\mathbf{z}_{it} = [\mathbf{x}_{1,it}, \mathbf{z}_{2,it}]$.

How can we test the null of exogeneity, $H_0 : \mathrm{E}(\mathbf{x}'_{3,it} v_{it}) = \mathbf{0}$?

A simple regression-based test

Partial out the effect of the instruments $\mathbf{z}_{it}$ that are certainly exogenous,

$$\mathbf{x}_{3,it} = \mathbf{z}_{it}\mathbf{\Pi} + \mathbf{w}_{it}.$$

Endogeneity arises if $\mathbf{w}_{it}$ (the part of $\mathbf{x}_{3,it}$ that is not due to exogenous variation) correlates with the structural disturbance v_{it}. Hence

$$H_0 : \mathrm{E}(\mathbf{w}'_{it} v_{it}) = \mathbf{0}.$$

If we observed $\mathbf{w}_{it}$, we could estimate the augmented equation

$$y_{it} = \mathbf{x}_{1,it}\boldsymbol{\beta}_1 + \mathbf{x}_{2,it}\boldsymbol{\beta}_2 + \mathbf{x}_{3,it}\boldsymbol{\beta}_3 + \mathbf{w}_{it}\boldsymbol{\gamma} + \varepsilon_{it}$$

by REIV with IVs $[\mathbf{x}_{1,it}, \mathbf{z}_{2,it}, \mathbf{x}_{3,it}, \mathbf{w}_{it}]$, and test $H_0 : \boldsymbol{\gamma} = \mathbf{0}$ (Wald test).

Since we do not observe $\mathbf{w}_{it}$, we estimate the reduced-form equation(s), compute the residuals $\hat{\mathbf{w}}_{it}$ and plug them into the augmented model.

The detailed procedure

1. For each single regressor in $\mathbf{x}_{3,it} = [\mathbf{x}_{3,it}^{(1)}, \ldots, \mathbf{x}_{3,it}^{(J_3)}]$ estimate by pooled OLS or standard RE

$$\mathbf{x}_{3,it}^{(j)} = \mathbf{z}_{it}\boldsymbol{\pi}^{(j)} + \mathbf{w}_{it}^{(j)}, \qquad j = 1, \ldots, J_3.$$

2. Collect $\hat{\mathbf{w}}_{it} = [\hat{\mathbf{w}}_{it}^{(1)}, \ldots, \hat{\mathbf{w}}_{it}^{(J_3)}]$, construct the instrument set $\mathbf{z}_{it} = [\mathbf{x}_{1,it}, \mathbf{z}_{2,it}, \mathbf{x}_{3,it}, \hat{\mathbf{w}}_{it}]$ and estimate by REIV the augmented equation

$$y_{it} = \mathbf{x}_{1,it}\boldsymbol{\beta}_1 + \mathbf{x}_{2,it}\boldsymbol{\beta}_2 + \mathbf{x}_{3,it}\boldsymbol{\beta}_3 + \hat{\mathbf{w}}_{it}\boldsymbol{\gamma} + \varepsilon_{it}.$$

3. Compute the Wald statistic W for $H_0 : \boldsymbol{\gamma} = \mathbf{0}$ and compare it with a critical value of the $\chi^2(J_3)$ distribution. To make the test robust to arbitrary within-individual correlation, compute a robust variance estimator in step (3).

Test of overidentifying restrictions

If $L > K$ we impose more orthogonality restrictions $\mathrm{E}(\tilde{\mathbf{Z}}_i' \tilde{\mathbf{v}}_i) = \mathbf{0}$ than necessary. Hence, we have more instruments that needed: "excess instruments".

A test which can be made robust to arbitrary within-individual correlation is proposed by Wooldridge. This is what we consider in the following.

The idea is straightforward: we check whether we are allowed to leave out the "excess instruments" from the structural equation.

A J test

Suppose we have an error-components model with $K = K_1 + K_2$ regressors

$$y_{it} = \mathbf{x}_{1,it}\boldsymbol{\beta}_1 + \mathbf{x}_{2,it}\boldsymbol{\beta}_2 + v_{it}$$

where

- $\mathbf{x}_{1,it}$ contains K_1 exogenous regressors used as their own instruments,
- $\mathbf{x}_{2,it}$ contains K_2 endogenous regressors instrumented by
- $\mathbf{z}_{2,it} = [\mathbf{g}_{it}, \mathbf{h}_{it}]$ which contains K_2 instruments $\mathbf{g}_{it}$ and L_2 instruments $\mathbf{h}_{it}$.

This yields the instrument set

$$\mathbf{z}_{it} = [\mathbf{x}_{1,it}, \mathbf{z}_{2,it}] = [\mathbf{x}_{1,it}, \mathbf{g}_{it}, \mathbf{h}_{it}],$$

which is of size $L = K_1 + K_2 + L_2 = K + L_2 > K$.

Note that it does not matter how $\mathbf{z}_{2,it}$ is split into $\mathbf{g}_{it}$ and $\mathbf{h}_{it}$ (except that intercept and time dummies are always part of $\mathbf{x}_{1,it}$ and thus cannot be part of $\mathbf{h}_{it}$).

The detailed procedure I

1. Apply REIV with instruments $\mathbf{z}_{it} = [\mathbf{x}_{1,it}, \mathbf{g}_{it}, \mathbf{h}_{it}]$ to estimate the structural equation

$$y_{it} = \mathbf{x}_{1,it}\boldsymbol{\beta}_1 + \mathbf{x}_{2,it}\boldsymbol{\beta}_2 + v_{it}.$$

   This yields $\hat{\phi}$ and quasi-demeaned residuals $\hat{\tilde{v}}_{it}$ and allows to construct $\tilde{\mathbf{x}}_{it}$ and $\tilde{\mathbf{z}}_{it}$.

2. For each single regressor in $\tilde{\mathbf{x}}_{2,it} = [\tilde{\mathbf{x}}_{2,it}^{(1)}, \ldots, \tilde{\mathbf{x}}_{2,it}^{(K_2)}]$ estimate by pooled OLS

$$\tilde{\mathbf{x}}_{2,it}^{(j)} = \tilde{\mathbf{z}}_{it}\boldsymbol{\pi}^{(j)} + \tilde{\mathbf{w}}_{it}^{(j)}, \qquad j = 1, \ldots, K_2.$$

   Store the fitted values $\hat{\tilde{\mathbf{x}}}_{2,it} = [\hat{\tilde{\mathbf{x}}}_{2,it}^{(1)}, \ldots, \hat{\tilde{\mathbf{x}}}_{2,it}^{(K_2)}]$. By construction, they include that part of the endogenous regressors $\tilde{\mathbf{x}}_{2,it}$ that is exogenous.

The detailed procedure II

3. For each single regressor in $\tilde{\mathbf{h}}_{it} = [\tilde{\mathbf{h}}_{it}^{(1)}, \ldots, \tilde{\mathbf{h}}_{it}^{(L_2)}]$ estimate by pooled OLS

$$\tilde{\mathbf{h}}_{it}^{(j)} = \tilde{\mathbf{x}}_{1,it}\boldsymbol{\psi}_1^{(j)} + \hat{\tilde{\mathbf{x}}}_{2,it}\boldsymbol{\psi}_2^{(j)} + \tilde{\mathbf{r}}_{it}^{(j)}, \qquad j = 1, \ldots, L_2.$$

   Store the residuals $\hat{\tilde{\mathbf{r}}}_{it} = [\hat{\tilde{\mathbf{r}}}_{it}^{(1)}, \ldots, \hat{\tilde{\mathbf{r}}}_{it}^{(L_2)}]$. By construction, they include that part of $\tilde{\mathbf{h}}_{it}$ that is orthogonal to the exogenous regressors $\tilde{\mathbf{x}}_{1,it}$ and the exogenous part of the endogenous regressors $\tilde{\mathbf{x}}_{2,it}$.

4. Estimate by pooled OLS the auxiliary equation

$$\hat{\tilde{v}}_{it} = \hat{\tilde{\mathbf{r}}}_{it}\boldsymbol{\delta} + \varepsilon_{it}.$$

The detailed procedure III

5. Compute the Wald statistic W for $H_0 : \boldsymbol{\delta} = \mathbf{0}$ and compare it with a critical value of the $\chi^2(L_2)$ distribution. To make the test robust to arbitrary within-individual correlation, compute a robust variance estimator in step (4).
6. The test effectively tells us whether the "excess instruments" $\tilde{\mathbf{h}}_{it}$, from which we have partialled out all exogenous regressor influence, can explain the lhs variable of our structural equation, y_{it}, from which we have partialled out all regressor influence. If not, we do not reject H_0 and it is a good idea to use $\tilde{\mathbf{h}}_{it}$ as instruments. If yes, we reject H_0 and should consider different instruments.

Outline

Coming up

More instrumental variables