ECON 4160, Spring term 2013. Lecture 7

Identification and estimation of SEMs (Part 2)

Ragnar Nymoen

Department of Economics

8 Oct 2013

- References to Davidson and MacKinnon,
  - Ch 8.1-8.5
  - Ch 12.4.15
- For both Lecture 6 and 7

Notation for a structural equation (recap Lect 6) I

- Consider equation # 1 in a SEM as in DM page 522:

$$\mathbf{y}_1 = \mathbf{Z}_1\boldsymbol{\beta}_{11} + \mathbf{Y}_1\boldsymbol{\beta}_{21} + \boldsymbol{\epsilon}_1 \tag{1}$$

- $\mathbf{y}_1$ is $n \times 1$, with observations of the variable that # 1 in the SEM is normalized on.
- $\mathbf{Z}_1$ is $n \times k_{11}$ with observations of the k_{11} included predetermined or exogenous variables.
- $\mathbf{Y}_1$, $n \times k_{12}$ holds the included endogenous explanatory variables.

Notation for a structural equation (recap Lect 6) II

The total number of explanatory variables in the first equation is

$$k_{11} + k_{12} = k_1 \tag{2}$$

For simplicity assume that the structural disturbance is Gaussian white-noise.

$$\boldsymbol{\epsilon}_1 = IN(\mathbf{0}, \sigma_1^2 \mathbf{I}_{nxn})$$

By defining the two partitioned matrices:

$$\mathbf{X}_1 = \begin{pmatrix} \mathbf{Z}_1 & : & \mathbf{Y}_1 \end{pmatrix} \text{ (note change)} \tag{3}$$

$$\boldsymbol{\beta}_1 = \begin{pmatrix} \boldsymbol{\beta}_{11} & : & \boldsymbol{\beta}_{21} \end{pmatrix}' \tag{4}$$

(1) can be written compactly as

$$\mathbf{y}_1 = \mathbf{X}_1 \boldsymbol{\beta}_1 + \boldsymbol{\epsilon}_1 \tag{5}$$

The IV estimator (exact/just identification) I

Define $\mathbf{W}_1$ as

$$\mathbf{W}_1 = \left(\mathbf{Z}_1 \quad : \quad \mathbf{X}_{01} \right) \tag{6}$$

where $\mathbf{X}_{01}$ is $n \times k_{12}$ (number of included endogenous variables minus one), with asymptotic properties:

$$plim(\frac{1}{n}\mathbf{W}_1'\mathbf{X}_1) = \mathbf{S}_{W_1'X_1} \text{ (invertible)} \tag{7}$$

$$plim(\frac{1}{n}\mathbf{W}_1'\epsilon_1) = \mathbf{0} \text{ (independence)} \tag{8}$$

$$plim(\frac{1}{n}\mathbf{W}_1'\mathbf{W}_1) = \mathbf{S}_{W_1'W_1} \text{ (positive definite)} \tag{9}$$

We call $\mathbf{W}_1$ the instrumental variable matrix and find the IV estimator $\hat{\beta}_{1,IV}$ for the first structural equation by solving

The IV estimator (exact/just identification) II

$$\mathbf{W}_1' \left[\mathbf{y}_1 - \mathbf{X}_1 \hat{\beta}_{1,IV} \right] = \mathbf{0} \tag{10}$$

giving

$$\hat{\beta}_{1,IV} = (\mathbf{W}_1'\mathbf{X}_1)^{-1}\mathbf{W}_1'\mathbf{y}_1 \tag{11}$$

$\hat{\beta}_{1,IV}$ is clearly a method-of-moments estimator. The only difference from OLS is that $\mathbf{W}_1'$ takes the place of $\mathbf{X}_1'$ in the IV "normal equations", or orthogonality conditions, (10).
This means that, by construction, the IV-residuals

$$\hat{\epsilon}_{\mathbf{IV},\mathbf{1}} = \mathbf{y}_1 - \mathbf{X}_1 \hat{\beta}_{1,IV} \tag{12}$$

are uncorrelated with (all the instuments in) $\mathbf{W}_1$:

$$\mathbf{W}_1' \hat{\epsilon}_{\mathbf{IV},\mathbf{1}} = \mathbf{0}. \tag{13}$$

The IV residual maker I

- If we want, we can define the IV-residual maker as

$$\mathbf{M}_{IV,1} = \left[\mathbf{I} - \mathbf{X}_1(\mathbf{W}_1'\mathbf{X}_1)^{-1}\mathbf{W}_1'\right]$$

(Check that

$$\hat{\epsilon}_{\mathbf{IV},\mathbf{1}} = \mathbf{M}_{IV,1}\mathbf{y}_1)$$

- If $\mathbf{M}_{IV,1}$ is a proper residual maker, regression of $\mathbf{W}_1$ on $\mathbf{W}_1$ should result in zero-residuals.

The IV residual maker II

- Check:

$$\mathbf{M}_{IV,1}\mathbf{W}_1 = \left[\mathbf{I} - \mathbf{X}_1(\mathbf{W}_1'\mathbf{X}_1)^{-1}\mathbf{W}_1'\right]\mathbf{W}_1$$

Use that

$$\mathbf{M}_{IV,1}^{'} = \mathbf{M}_{IV,1}$$

(show for $k_1 = 2$ for example), then

$$\begin{aligned}\mathbf{M}_{IV,1}\mathbf{W}_1 &= \mathbf{M}_{IV,1}^{'}\mathbf{W}_1 \\ &= \left[\mathbf{I} - \mathbf{W}_1'\left\{(\mathbf{W}_1'\mathbf{X}_1)^{-1}\right\}'\mathbf{X}_1^{'}\right]\mathbf{W}_1 \\ &= \left[\mathbf{I} - \mathbf{W}_1'\left\{(\mathbf{W}_1'\mathbf{X}_1)'\right\}^{-1}\mathbf{X}_1^{'}\right]\mathbf{W}_1 \\ &= \left[\mathbf{I} - \mathbf{W}_1'\left\{\mathbf{X}_1'\mathbf{W}_1\right\}^{-1}\mathbf{X}_1^{'}\right]\mathbf{W}_1 \\ &= \mathbf{0}\end{aligned}$$

The IV residual maker III

so that the orthogonality condition (10) can be interpreted as:

$$\mathbf{W}_1^{'}\hat{\epsilon}_{IV,1} = (\mathbf{M}_{IV,1}^{'}\mathbf{W}_1)'\mathbf{y}_1 = \mathbf{0} \tag{14}$$

confirming (13), and showing that the "only" difference compared to OLS is that the set of instruments used to form orthogonality conditions (normal equations) has been changed from "**X**" to "**W**".

Optimal instruments in the overidentified case I

- In the case of *overidentification*, $\mathbf{W}_1$ is $n \times l_1$ where $l_1 > k_1 = k_{11} + k_{12}$, $\mathbf{W}_1'\mathbf{X}_1$ is no longer quadratic.
- There is more that one moment-matrix (based on $\mathbf{W}_1'\mathbf{X}_1$) that are quadratic and invertible
- Each one defines a consistent IV-estimator of $\boldsymbol{\beta}_1$ under the assumptions (7)-(9). Hence we have **over-identification**
- To solve this "luxury problem" we can define another IV matrix $\widehat{\mathbf{W}}_1$ that has dimension $n \times k_1$:

$$\widehat{\mathbf{W}}_1 = \left(\; \mathbf{Z}_1 \quad : \quad \hat{\mathbf{Y}}_1 \; \right), \tag{15}$$

Optimal instruments in the overidentified case II

where $\widehat{\mathbf{Y}}_1$ is $n \times k_{12}$ and is made up of the best linear predictors of the k_{12} endogenous variables included in the first equation:

$$\widehat{\mathbf{Y}}_1 = \begin{pmatrix} \widehat{\mathbf{y}}_2 & \widehat{\mathbf{y}}_3 & \dots \end{pmatrix}_{n \times k_{12}} \tag{16}$$

- Where does the optimal predictors come from? Since we are looking at a single equation in a system-of-equations, they must come from the reduced form equations for the endogenous variables:

$$\widehat{\mathbf{y}}_j = \mathbf{W}_1 \widehat{\boldsymbol{\pi}}_j, \quad j = 2, ..., k_{12} + 1. \tag{17}$$

where

$$\widehat{\boldsymbol{\pi}}_j = (\mathbf{W}_1' \mathbf{W}_1)^{-1} \mathbf{W}_1' \mathbf{y}_j, \tag{18}$$

Optimal instruments in the overidentified case III

are the OLS estimators of the regression coefficients in the conditional expectation function for each included endogenous variables in the first equation, conditional on the full set of predetermined variables in the system of equations. We can write $\widehat{\mathbf{Y}}_1$ as:

$$\widehat{\mathbf{Y}}_1 = \begin{pmatrix} \mathbf{W}_1(\mathbf{W}_1'\mathbf{W}_1)^{-1}\mathbf{W}_1'\mathbf{y}_2 & \dots & \mathbf{W}_1(\mathbf{W}_1'\mathbf{W}_1)^{-1}\mathbf{W}_1'\mathbf{y}_{(k_{21}+1)} \end{pmatrix},$$

and more compactly:

$$\widehat{\mathbf{Y}}_1 = \mathbf{W}_1(\mathbf{W}_1'\mathbf{W}_1)^{-1}\mathbf{W}_1'\mathbf{Y}_1 = \mathbf{P}_{W_1}\mathbf{Y}_1$$

in terms of the prediction-maker:

$$\mathbf{P}_{W_1} = \mathbf{W}_1(\mathbf{W}_1'\mathbf{W}_1)^{-1}\mathbf{W}_1' \tag{19}$$

The GIV-estimator I

- We define the **Generalized IV estimator** as

$$\hat{\beta}_{1,GIV} = (\widehat{\mathbf{W}}_1'\mathbf{X}_1)^{-1}\widehat{\mathbf{W}}_1'\mathbf{y}_1. \tag{20}$$

with

$$\widehat{\mathbf{W}}_1 = \begin{pmatrix} \mathbf{Z}_1 & : & \widehat{\mathbf{Y}}_1 \end{pmatrix},$$

and

$$\widehat{\mathbf{Y}}_1 = \mathbf{P}_{W_1}\mathbf{Y}_1$$

- $\hat{\beta}_{1,GIV}$ is also known as the 2-stage least squares estimator of β_1 in (5)

GIVE and the 2SLS estimator I

- In s separate note, as a supplement to DM Ch 8 at this point, we show that by carrying through the partitioning of $\boldsymbol{\beta}_1$ and $\mathbf{X}_1$, we can write

$$\hat{\boldsymbol{\beta}}_{1,GIV} = \tag{21}$$

$$\begin{pmatrix} \widehat{\boldsymbol{\beta}}_{11,GIV} \\ \widehat{\boldsymbol{\beta}}_{21,GIV} \end{pmatrix} = \begin{pmatrix} \mathbf{Z}_1'\mathbf{Z}_1 & \mathbf{Z}_1'\mathbf{Y}_1 \\ \widehat{\mathbf{Y}}_1'\mathbf{Z}_1 & \widehat{\mathbf{Y}}_1'\mathbf{Y}_1 \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{Z}_1'\mathbf{y}_1 \\ \widehat{\mathbf{Y}}_1'\mathbf{y}_1 \end{pmatrix}$$

and, by use of prediction maker and residual maker matrices:

$$\begin{pmatrix} \widehat{\boldsymbol{\beta}}_{11,GIV} \\ \widehat{\boldsymbol{\beta}}_{21,GIV} \end{pmatrix} = \begin{pmatrix} \mathbf{Z}_1'\mathbf{Z}_1 & \mathbf{Z}_1'\widehat{\mathbf{Y}}_1 \\ \widehat{\mathbf{Y}}_1'\mathbf{Z}_1 & \widehat{\mathbf{Y}}_1'\widehat{\mathbf{Y}}_1 \end{pmatrix}^{-1} \begin{pmatrix} \widehat{\mathbf{Y}}_1'\mathbf{y}_1 \\ \mathbf{Z}_1'\mathbf{y}_1 \end{pmatrix} \tag{22}$$

GIVE and the 2SLS estimator II

- Next, call the reduced form regression that gives rise to $\widehat{\mathbf{Y}}_1$ the **first-stage regression.**
- Then consider the consequences of using OLS on the structural equation (1), but after substitution of $\mathbf{Y}_1$ with $\widehat{\mathbf{Y}}_1$:

$$\begin{aligned} \mathbf{y}_1 &= \mathbf{Z}_1\beta_{11} + \widehat{\mathbf{Y}}_1\beta_{21} + \text{ disturbance} \\ &= \left(\mathbf{Z}_1 \; : \; \widehat{\mathbf{Y}}_1 \right) \begin{pmatrix} \beta_{11} \\ \beta_{21} \end{pmatrix} + \text{disturbance} \end{aligned}$$

GIVE and the 2SLS estimator III

Call the OLS estimator from this second least-square estimation, the **Two-Stage Least Square** estimator: 2SLS:

$$\begin{pmatrix} \widehat{\beta}_{11,2SLS} \\ \widehat{\beta}_{21,2SLS} \end{pmatrix} = \left[\begin{pmatrix} \mathbf{Z}_1 & : & \widehat{\mathbf{Y}}_1 \end{pmatrix}' \begin{pmatrix} \mathbf{Z}_1 & : & \widehat{\mathbf{Y}}_1 \end{pmatrix} \right]^{-1} \begin{pmatrix} \mathbf{Z}_1' \\ \widehat{\mathbf{Y}}_1' \end{pmatrix} \mathbf{y}_1$$

$$= \begin{pmatrix} \mathbf{Z}_1'\mathbf{Z}_1 & \mathbf{Z}_1'\widehat{\mathbf{Y}} \\ \widehat{\mathbf{Y}}_1'\mathbf{Z}_1 & \widehat{\mathbf{Y}}_1'\widehat{\mathbf{Y}}_1 \end{pmatrix}^{-1} \begin{pmatrix} \widehat{\mathbf{Y}}_1'\mathbf{y}_1 \\ \mathbf{Z}_1'\mathbf{y}_1 \end{pmatrix} \qquad (23)$$

which is identical to (22).

- This establishes that in the over-identified case, **GIVE** is identical to the **Two-Stage Least squares** estimator (2SLS)

Quick reference back to the just identified case I

- Since $\mathbf{W}_1$ is $T \times k_1$ (i.e. $l_1 = k_1$), we have:

$$\hat{\boldsymbol{\beta}}_{1,IV} = \hat{\boldsymbol{\beta}}_{1,GIV} = \widehat{\boldsymbol{\beta}}_{1,2SLS}$$

- Moreover, there is also a third estimator called **indirect least squares** ($\hat{\boldsymbol{\beta}}_{1,IV}$) which is the unique and consistent estimator for $\boldsymbol{\beta}_1$ that can be obtained from the reduced form estimators.
- Indirect least squares is cumbersome for all but simple systems-of-equation. But Introductory books usually include an example, for example Bårdsen and Nymoen ch 9.

Another equivalent expression of the GIV-estimator I

- Let us again consider over-identification: $l_1 > k_1 = k_{11} + k_{12}$.
- A different method to obtain unique moments is to post-multiply $\mathbf{W}_1$ by the matrix $\mathbf{J}_1$ so that $\mathbf{W}_1\mathbf{J}_1$ is an IV-matrix of dimension $n \times k_1$ and which satisfies (7)-(9).
- $\mathbf{J}_1$ must be $l_1 \times k_1$. If we choose

$$\mathbf{W}_1\mathbf{J}_1 = \mathbf{P}_{W_1}\mathbf{X}_1 \tag{24}$$

as the IV matrix (remember that $\mathbf{P}_{W_1} = \mathbf{W}_1(\mathbf{W}_1'\mathbf{W}_1)^{-1}\mathbf{W}_1'$) we can determine $\mathbf{J}_1$ as

$$\mathbf{J}_1 = (\mathbf{W}_1'\mathbf{W}_1)^{-1}\mathbf{W}_1'\mathbf{X}_1 \tag{25}$$

which has the correct dimension $l_1 \times k_1$.

Another equivalent expression of the GIV-estimator II

- Let us, for the time being denote the estimator that uses $\mathbf{W}_1\mathbf{J}_1$ as IV-matrix by $\hat{\beta}_{1,JGIV}$

$$
\begin{aligned}
\hat{\beta}_{1,JGIV} &= ((\mathbf{W}_1\mathbf{J}_1)'\mathbf{X}_1)^{-1}(\mathbf{W}_1\mathbf{J}_1)'\mathbf{y}_1 \\
&= ((\mathbf{P}_{W_1}\mathbf{X}_1)'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{P}_{W_1}\mathbf{y}_1 \\
&= (\mathbf{X}_1'\mathbf{P}_{W_1}\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{P}_{W_1}\mathbf{y}_1
\end{aligned}
$$

which is the expression for the GIV estimator in equation (8.29) in DM (p 321).

Another equivalent expression of the GIV-estimator III

- But we have already

$$\hat{\beta}_{1,GIV} = (\widehat{\mathbf{W}}_1'\mathbf{X}_1)^{-1}\widehat{\mathbf{W}}_1'\mathbf{y}_1. \tag{26}$$

with

$$\widehat{\mathbf{W}}_1 = \left(\mathbf{Z}_1 \quad : \quad \widehat{\mathbf{Y}}_1 \right),$$

$$\widehat{\mathbf{Y}}_1 = \mathbf{P}_{W_1}\mathbf{Y}_1$$

from before. Is the an internal inconsistency? Can $\hat{\beta}_{1,GIV} \neq \hat{\beta}_{1,JGIV}$?

Another equivalent expression of the GIV-estimator IV

- The answer is "no", since

$$\widehat{\mathbf{W}}_1 = \left(\ \mathbf{Z}_1 \ \ : \ \ \widehat{\mathbf{Y}}_1\ \right) = \left(\ \mathbf{P}_{W_1}\mathbf{Z}_1 \ \ : \ \ \mathbf{P}_{W_1}\widehat{\mathbf{Y}}_1\ \right) = \mathbf{P}_{W_1}\mathbf{X}_1$$

we can replace $\widehat{\mathbf{W}}_1$ in $\hat{\beta}_{1,GIV}$ by $\mathbf{P}_{W_1}\mathbf{X}_1$ and obtain

$$\begin{aligned}\hat{\beta}_{1,GIV} &= (\widehat{\mathbf{W}}_1'\mathbf{X}_1)^{-1}\widehat{\mathbf{W}}_1'\mathbf{y}_1 \\ &= (\mathbf{X}_1'\mathbf{P}_{W_1}\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{P}_{W_1}\mathbf{y}_1 \\ &= \hat{\beta}_{1,JGIV}\end{aligned}$$

IV-criterion function I

- The projection matrix $\mathbf{P}_{W_1}$ is also central in the *IV-criterion function* $Q(\boldsymbol{\beta}_1, \mathbf{y}_1)$ which DM defines on page 321:

$$Q(\boldsymbol{\beta}_1, \mathbf{y}_1) = (\mathbf{y}_1 - \mathbf{X}_1\boldsymbol{\beta}_1)'\mathbf{P}_{W_1}(\mathbf{y}_1 - \mathbf{X}_1\boldsymbol{\beta}_1). \qquad (27)$$

  note the close relationship to the sum of squared residuals.
- Minimization of $Q(\boldsymbol{\beta}_1, \mathbf{y}_1)$ wrt $\boldsymbol{\beta}_1$ gives the 1 oc:

$$\mathbf{X}_1'\mathbf{P}_{W_1}(\mathbf{y}_1 - \mathbf{X}_1\hat{\boldsymbol{\beta}}_{1,GIV}) = \mathbf{0} \qquad (28)$$

  which gives $\hat{\boldsymbol{\beta}}_{1,GIV}$ as solution

$$\hat{\boldsymbol{\beta}}_{1,GIV} = (\mathbf{X}_1'\mathbf{P}_{W_1}\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{P}_{W_1}\mathbf{y}_1$$

IV-criterion function II

- Note that from (24)

$$\mathbf{X}_1'\mathbf{P}_{W_1} = \mathbf{J}_1'\mathbf{W}_1'$$

we can write (28) as:

$$\mathbf{J}_1'\mathbf{W}_1'(\mathbf{y}_1 - \mathbf{X}_1\hat{\beta}_{1,GIV}) = \mathbf{0} \tag{29}$$

We can then represent the just identified case by setting $\mathbf{J}_1$ to a non-singular $k_1 \times k_1$ matrix. Pre-multiplication in (29) by $(\mathbf{J}_1')^{-1}$ gives

$$\mathbf{W}_1'(\mathbf{y}_1 - \mathbf{X}_1\hat{\beta}_{1,GIV}) = \mathbf{0}$$

confirming that in the just identified case, $\hat{\beta}_{1,GIV} = \hat{\beta}_{1,IV}$.

IV-criterion function III

- In the just identified case the minimized value of the IV-criterion function is zero:

$$
\begin{aligned}
Q(\hat{\beta}_{1,IV}, \mathbf{y}_1) &= (\mathbf{y}_1 - \mathbf{X}_1\hat{\beta}_{1,IV})'\mathbf{P}_{W_1}(\mathbf{y}_1 - \mathbf{X}_1\hat{\beta}_{1,IV}) \\
&= \hat{\epsilon}'_{\mathbf{IV,1}} \left[\mathbf{W}_1(\mathbf{W}_1'\mathbf{W}_1)^{-1}\mathbf{W}_1'\right] \hat{\epsilon}_{\mathbf{IV,1}} \\
&= \hat{\epsilon}'_{\mathbf{IV,1}}\mathbf{W}_1(\mathbf{W}_1'\mathbf{W}_1)^{-1} \left[\mathbf{W}_1'\hat{\epsilon}_{\mathbf{IV,1}}\right] = \mathbf{0} \qquad (30)
\end{aligned}
$$

from (10).

Sargan test and J-test I

- When $l_1 - k_1 \geq 1$ the validity of the over-identifying instruments can be tested.
- DM Ch 8.6
- Intuition: If the instruments are valid, they should have no significant explanatory power in an (auxiliary) regression that has the GIV-residuals as regressand.
- Simplest case. Assume that there is one endogenous explanatory variable in the first structural equation.
- The regression of $\hat{\boldsymbol{\epsilon}}_{\mathbf{GIV},\mathbf{1}}$ on the instruments and a constant produces an R^2 that we can call R^2_{GIVres}.

Sargan test and J-test II

- The test called *Specification test* when IV-estimation Pc-Give single equation is:

$$Specification\ test = nR^2_{GIVres} \underset{a}{\sim} \chi^2(l_1 - k_1) \qquad (31)$$

under the H_0 of valid instrumental variables.

- Show in class by simple PCM example
- As noted by DM on p 338 this is rightly called a **Sargan test** , because of the contribution of Denis Sargan who worked on IV-estimation theory during the 1950 and 1960.
- In terms of computation Sargan test can be computed from the IV criterion function:
  - From (30) $Q(\hat{\boldsymbol{\beta}}_{1,IV}, \mathbf{y}_1) = 0$, the IV-residuals are orthogonal to all instrumental variables

Sargan test and J-test III

- In the over-identified case, $Q(\hat{\boldsymbol{\beta}}_{1,GIV}, \mathbf{y}_1) > 0$ since the GIV-residuals are uncorrelated with the optimal instruments, but not each individual.
- Therefore, a test of the validity of the over-identification can be based on $Q(\hat{\boldsymbol{\beta}}_{1,GIV}, \mathbf{y}_1) - Q(\hat{\boldsymbol{\beta}}_{1,IV}, \mathbf{y}_1)$:

$$\textit{Specification test} = \frac{Q(\hat{\boldsymbol{\beta}}_{1,GIV}, \mathbf{y}_1)}{\hat{\sigma}_1^2} \underset{a}{\sim} \chi^2(l_1 - k_1) \qquad (32)$$

where $\hat{\sigma}_1^2$ is the usual consistent estimator for σ_1^2.

- The two ways of computing the Sargan *Specification test* are numerically-identical.

Sargan test and J-test IV

- Many (papers) and software-packages report "J-tests" or "Hansen test" for instrument validity. This test uses the F-statistic from the auxiliary regression instead of R^2_{GIVres}:

$$J - test = l_1 F^2_{GIVres} \underset{a}{\sim} \chi^2(l_1 - k_1)$$

Asymptotic properties of GIV I

- As noted at the end of Lecture 6, with reference to the asymptotic nature of the IV requirements (7)-(9) it is not surprising that the known properties of the IV and GIVE estimators are asymptotic.
- In the case of overidentification, and maintaining,

$$\epsilon_1 = IN(\mathbf{0}, \sigma_1^2 \mathbf{I}_{nxn}) \tag{33}$$

$\hat{\boldsymbol{\beta}}_{1,GIV}$ and $\hat{\boldsymbol{\beta}}_{1,2SLS}$ are consistent and asymptotically efficient. The use of optimal instruments based on the reduced form of the system-of-equations (and (33)) are the main drivers behind this result.

Asymptotic properties of GIV II

- If (33) cannot be maintained, or if we want to take into account correlation between ϵ_1 and other disturbances in the SEM, other estimation methods are more efficient: GMM and 3SLS for example.
- Lecture 6: asymptotic normality of "t-ratios" etc, and the form of the variance estimator.

Finite sample properties and weak instruments I

- $\widehat{\mathbf{Y}}_1$is asymptotically uncorrelated with ϵ_1, but in a finite sample there is always some correlation so that $E(\hat{\beta}_{1,GIV}) \neq \beta_1$.
- See Ch 8.4 in DM for discussion and analysis
- Weak instruments: The instruments are poorly correlated with the included endogenous variables in the structural equation.
- When instrumentst are weak, the asymptotic distribution reflect the true finite sample distribution poorly. The asymptotic inference theory loses its relevance.
- The practical issue is then how poor the instruments can be before we must conclude that the IV estimation and inference becomes unreliable.

Finite sample properties and weak instruments II

- For $k_{12} = 1$ (only one endogenous explanatory variable) there is a simple test:
  - Regress the endogenous variables on the l_1 instrumental variables.
  - A rule-of-thumb test due to Stock and Watson is that if the $F - statistic$ of the test of joint significance > 10 , weak instruments is not a problem
- In general, the danger of weak instruments means that the system, the reduced form, the VAR should be evaluated in its own merit (do we have a well specified statistical system that has predictive power for the variables) before moving to estimation of the structural equation.

IV estimation of simultaneous equation models and looking ahead I

- We have focused on the "first equation" in SEM
- Obviously: Can apply the same 2-SLS method to **all** the identified equations of a SEM.
- However, if we are interested in more than one identified structural equation, there is really no reason to use limited information methods (based on uncorrelated disturbances).
- We will therefore turn to **Full Information Maximum Likelihood** (FIML) and 3SLS in lecture 8
- First in Lecture 7, will talk briefly cover an important development of "single equation IV" called **Generalized Method of Moments** (GMM, Ch 8 in DM) and show examples of IV estimation and test of forward-looking models.