ECON 4160, Spring term 2013. Lecture 6

Identification and estimation of SEMs (Part 1)

Ragnar Nymoen

Department of Economics

25 Sep 2013

- References to Davidson and MacKinnon,
  - Ch 8.1-8.5
  - Ch 12.4.15
- For both Lecture 6 and 7 (need more than one to cover all this)

Structural form and reduced form I

- Consider the following two equations for simultaneous equilibrium in a single marked (partial equilibrium)

$$Q_t = \beta_{11} + \beta_{12} P_t + \varepsilon_{dt} \text{ (demand)} \tag{1}$$
$$Q_t = \beta_{21} + \beta_{22} P_t + \varepsilon_{st} \text{ (supply)} \tag{2}$$

- ε_{dt} and ε_{st} are white-noise processes, e.g., $(\varepsilon_{dt}, \varepsilon_{st})' \sim IN(\mathbf{0}, \mathbf{\Sigma})$.
- In this lecture we will not require from the outset that $\mathbf{\Sigma}$ is diagonal: We want to study identification both without imposing any restrictions of the covariance matrix $\mathbf{\Sigma}$, and with such restrictions (typically that $\mathbf{\Sigma}$ is diagonal).

Structural form and reduced form II

- In matrix notation, the model is
$$\begin{pmatrix} 1 & -\beta_{12} \\ 1 & -\beta_{22} \end{pmatrix} \begin{pmatrix} Q_t \\ P_t \end{pmatrix} = \begin{matrix} \beta_{11} + \varepsilon_{dt} \\ \beta_{21} + \varepsilon_{st} \end{matrix}$$
- Since the market is always in equilibrium, we observe not (1) and (2), but a sequence of equilibrium variables $\{\ldots, (Q_{t-1}, P_{t-1}), (Q_t, P_t), (Q_{t+1}, P_{t+1}), \ldots\}$ from the reduced-form of the simultaneous equations model

$$\begin{pmatrix} Q_t \\ P_t \end{pmatrix} = \begin{pmatrix} 1 & -\beta_{12} \\ 1 & -\beta_{22} \end{pmatrix}^{-1} \begin{pmatrix} \beta_{11} + \varepsilon_{dt} \\ \beta_{21} + \varepsilon_{st} \end{pmatrix} \qquad (3)$$

Since

$$\begin{pmatrix} 1 & -\beta_{12} \\ 1 & -\beta_{22} \end{pmatrix}^{-1} = \begin{pmatrix} -\frac{\beta_{22}}{\beta_{12}-\beta_{22}} & \frac{\beta_{12}}{\beta_{12}-\beta_{22}} \\ -\frac{1}{\beta_{12}-\beta_{22}} & \frac{1}{\beta_{12}-\beta_{22}} \end{pmatrix}$$

Structural form and reduced form III

the Reduced Form (RF) becomes:

$$Q_t = \frac{\beta_{12}\beta_{21} - \beta_{11}\beta_{22}}{\beta_{12} - \beta_{22}} + \frac{\beta_{12}\varepsilon_{st} - \beta_{22}\varepsilon_{dt}}{\beta_{12} - \beta_{22}} \quad (4)$$

$$P_t = \frac{\beta_{21} - \beta_{11}}{\beta_{12} - \beta_{22}} + \frac{\varepsilon_{st} - \varepsilon_{dt}}{\beta_{12} - \beta_{22}} \quad (5)$$

- **The identification issue:** From the RF (4)-(5), can we obtain consistent estimators of the parameters of the simultaneous equation model (1)-(2)?
- NOTE: We discuss identification from as a logical property of the theoretical model. This corresponds to the term "Asymptotic Identification" (page 529) used by Davidson and MacKinnon.

Under-identification I

- Given the assumptions we have made, OLS estimation of the RF parameters in (4) and (5) will be consistent MLE estimators.
- Let us therefore assume a **perfect sample,** that gives us knowledge of the *plim*-values of the OLS estimators of $E(Q_t)$ and $E(P_t)$. Denote these *plim*-values by γ_{Q0} and γ_{P0}

$$\gamma_{Q0} = \frac{\beta_{12}\beta_{21} - \beta_{11}\beta_{22}}{\beta_{12} - \beta_{22}} \tag{6}$$

$$\gamma_{P0} = \frac{\beta_{21} - \beta_{11}}{\beta_{12} - \beta_{22}} \tag{7}$$

  - Two equations in **two known RF parameters**, γ_{Q0} and γ_{P0}, and **four unknown structural parameters**.

Under-identification II

- Cannot determine β_{11}, β_{12}, β_{21} and β_{22} from (6) and (7) even if we have perfect knowledge of the reduced form parameters γ_{Q0} and γ_{P0}
- The parameters of neither (1) nor (2) are identified.
- Neither of the equations in the SEM are identified.

Partial identification by imposing restrictions I

- Let us assume completely inelastic supply, $\beta_{22} = 0$. The SEM is now

$$Q_t = \beta_{11} + \beta_{12} P_t + \varepsilon_{dt} \text{ (demand)} \tag{8}$$
$$Q_t = \beta_{21} + \varepsilon_{st} \text{ (supply)} \tag{9}$$

  and keep $(\varepsilon_{dt}, \varepsilon_{st})' \sim IN(\mathbf{0}, \mathbf{\Sigma})$ as before.
- RF for this case:

$$\gamma_{Q0} = \frac{\beta_{12}\beta_{21}}{\beta_{12}} = \beta_{21} \tag{10}$$
$$\gamma_{P0} = \frac{\beta_{21} - \beta_{11}}{\beta_{12}} \tag{11}$$

Partial identification by imposing restrictions II

- Assume again that γ_{Q0} and γ_{P0} are known (the same perfect sample assumption as above).
  - The structural parameter β_{21} is found (determined) from (10), hence the supply equation (9) is identified
  - In (11) we still have "two unknowns" and only one equation: The demand equation (8) is not identified
- This is an example of partial identification (i.e. of one equation namely ((9)) in a SEM.

Full identification by imposing restrictions I

- In addition to $\beta_{22} = 0$ we assume that $\mathbf{\Sigma}$ is diagonal:

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma^2_{d\varepsilon} & 0 \\ 0 & \sigma^2_{s\varepsilon} \end{pmatrix}$$

- Does this lead to (more) identification?

Full identification by imposing restrictions II

- To find the answer, write down all the first and second order parameters (moments) from the reduced form

$$E(Q_t) \equiv \gamma_{Q0} = \beta_{21} \tag{12}$$

$$E(P_t) \equiv \gamma_{P0} = \frac{\beta_{21} - \beta_{11}}{\beta_{12}} \tag{13}$$

$$Var(Q_t) = \frac{\beta_{12}^2 \sigma_{s\varepsilon}^2}{\beta_{12}^2} \tag{14}$$

$$Var(P_t) = \frac{\sigma_{s\varepsilon}^2 - \sigma_{d\varepsilon}^2}{\beta_{12}^2} \tag{15}$$

$$Cov(Q_t, P_t) = \frac{\beta_{12} \sigma_{s\varepsilon}^2}{\beta_{12}^2} \tag{16}$$

Full identification by imposing restrictions III

- We now have a determined equation system for det RF parameters and the structural form parameters: 5 independent equations in the five unknown structural parameters: β_{11}, β_{12}, $\beta_{21}, \sigma^2_{s\varepsilon}$ and $\sigma^2_{d\varepsilon}$.
- Both equations are therefore identified, we call this *exact-identification* or *just-identification*.

Interpreting the exact identified case: recursive structure I

- The implication of $\beta_{21} = 0$ (inelastic supply) and diagonal $\mathbf{\Sigma}$ diagonal is that the structure become **recursive:**
- The expected supply is fixed from period to period, and there is no indirect correlation between ε_{st} and ε_{dt} via $\mathbf{\Sigma}$
- In effect, Q_t is *strictly exogenous* in a regression between P_t and Q_t, and the *plim* of the OLS estimated regression coefficient is

$$\frac{Cov(Q_t, P_t)}{Var(Q_t)} = \frac{\frac{\beta_{12}\sigma^2_{s\varepsilon}}{\beta^2_{12}}}{\frac{\beta^2_{12}\sigma^2_{s\varepsilon}}{\beta^2_{12}}} = \frac{\frac{\sigma^2_{s\varepsilon}}{\beta_{12}}}{\sigma^2_{s\varepsilon}} = \frac{1}{\beta_{12}}$$

Interpreting the exact identified case: recursive structure II

- Hence, the slope parameter $\beta^{'}_{11}$ of the *inverted demand function*

$$P_t = \beta'_{11} + \beta^{'}_{12} Q_t + \varepsilon'_{st}$$

is consistently estimated by OLS, and a consistent estimator of β_{12} is the reciprocal of that OLS estimator $\widehat{\beta^{'}_{12}}$.

Summing up the first example I

- Identification is a logical property of the structural (theory) model. In particular it is **not** a property of the sample.
- Consistent estimation of the RF parameters is necessary for identification of structural parameters.
- Restriction on one or more of the structural parameters (β_{22} above), may lead to at least partial identification
- Restriction on the covariance matrix of the structural disturbances ($\sigma_{sd} = 0$, above) can also increase the degree of identification.
- In the following we shall concentrate on developing an easy-to-use method of checking identification *without* considering any restrictions on the Σ matrix for the disturbances

Summing up the first example II

- This is the method of the Order- and Rank-conditions
- Before presenting the general result, we try to develop our understanding through several steps.

Another partially identified structure I

- Before we restricted (1) and (2), call it *Structure 1*, none of the structural parameters were identified. We can make it more concrete by writing it as

$$\textit{Structure 1}$$
$$Q_t = 15 - P_t + \varepsilon_{dt}$$
$$Q_t = -0.2 + 0.5P_t + \varepsilon_{st}$$

  In *CC4singlemarked1.fl* data is generated for Q_t and P_t according to this structure.

- Intuitively, the lack of identification is due to the fact that there is no *independent variation* in the supply schedule that can help "trace out" the demand curve, and vice versa for the supply curve.

Another partially identified structure II

- When we discuss identification, it is often useful use an "identification table"

	Q	1	P
Demand equation	1	β_{11}	β_{12}
Supply equation	1	β_{21}	β_{22}

- Note that we regard the constant term as an exogenous variable, denoted by the "1" in the table.
- In the detailed *Structure 1* $\beta_{11} = 15$, $\beta_{22} = 0.5$ and so on.

Another partially identified structure III

- Now consider a second structure:

$$\begin{aligned} &\textit{Structure 2} \\ Q_t &= 15 - P_t + 1.5X_{dt} + \varepsilon_{dt} \\ Q_t &= -0.2 + 0.5P_t + \varepsilon_{st} \end{aligned}$$

where X_{dt} is an observable variable, so the identification table is

	Q	1	P	X_d
Demand equation	1	β_{11}	β_{12}	β_{13}
Supply equation	1	β_{21}	β_{22}	0

where the exclusion X_d from the supply equation is market by the "0".

Another partially identified structure IV

- Intuition: the supply equation is now identified, because the demand equation shifts due to changes in X_d, and the supply is unaffected by the X_d shifts
- More formally: We cannot obtain an observationally equivalent supply-equation by taking a linear combination of the two relationships in *Structure 2* (expect by giving zero weight to the demand equation).
- However, such a linear relationship is observationally equivalent to the *Demand equation* which is not identified in *Structure 2*
- *CC4singlemarked2.fl* illustrates *Structure 2*.

What about?

Structure 3

$$Q_t = 15 - P_t + \varepsilon_{dt}$$

$$Q_t = -0.2 + 0.5P_t + 2X_{st} + \varepsilon_{st}$$

	Q	1	P	X_s
Demand equation	1	β_{11}	β_{12}	0
Supply equation	1	β_{21}	β_{22}	β_{23}

- Use CC4singlemarked3.fl for this case

Exact identification I

Structure 4

$$Q_t = 15 - P_t + 1.5X_{dt} + \varepsilon_{dt}$$
$$Q_t = -0.2 + 0.5P_t + 2X_{st} + \varepsilon_{st}$$

	Q	1	P	X_d	X_s
Demand equation	1	β_{11}	β_{12}	β_{13}	0
Supply equation	1	β_{21}	β_{22}	0	β_{24}

- No linear combination can "give back" the original two structural equations, hence both are identified
- Use CC4singlemarked4.fl to illustrate

The order and rank conditions for exact identification

- The following identification rule suggests itself: In a SEM with g equations, any structural equation is identified if that equation excludes $g-1$ variables. The excluded variables can be endogenous or exogenous.
- This is called the **order condition**.
- An equivalent formulation says that the number of *excluded exogenous* variables from the equation must be equal to the number of *included endogenous variables minus one*.
- In fact the order condition is only necessary. The necessary and sufficient condition says that the excluded variables in the equation under inspection must have coefficients that are different from zero in the other equations in the SEM. This is called the **rank condition** for identification.

Overidentification

- We typically distinguish between just (or exact) identification and overidentification.
- **Overidentification** means that we can derive, from the RF parameters, more than one solution for the structural parameters.
- Hence the RF form in this case has more information than we need in order to estimate the structural parameters consistently.
- Overidentification is not an obstacle to estimation of the SEM, as we shall see later, the only problem is how to use the information to give efficient estimation (lowest possible standard errors).

Order and rank condition, general formulation I

Rank condition

In a SEM with g linear equations, an equation is identified if and only if at least one non-zero $(g-1) \times (g-1)$ determinant is contained in the array of the coefficients which those variables excluded from the equation in question appear in the other equations of the SEM.

Remarks:

- Recall that the rank of a matrix is the order of the largest non-zero determinant that it contains.
- If the rank condition is satisfied, the order condition is automatically satisfied, but not vice versa
- If the order of the non-zero determinant is larger than $(g-1)$ the equation is *overidentified.*

Order and rank condition, general formulation II

- Identities are often part of SEMs. They are counted among the g equation of the model. Identities are identified equations, but the identification of the *other* structural equations should be investigated with the identities taken into account.
- Exclusion restrictions are a special case of linear restrictions on the parameters and an even more general formulation of the order condition is:
  In a SEM with g linear equations, a necessary condition for identification of an equation is that there are $(g-1)\times(g-1)$ linearly independent restrictions on the parameters of the equation.

Order and rank condition, general formulation III

- In PcGive: Identification is always checked prior to estimation, hence it is asymptotic identification that is checked by the program.

Identification of dynamic SEMs

- Predetermined variables count as exogenous variables when we investigate identification
- This is because identification is about obtaining consistent estimators of structural parameters, and as we know, predeterminedness is enough for conistent estimation of RF paramaters, cf. the estimation theory of for VARs!
- Therefore the order and rank conditions apply to dynamic simultaneous equations models

Identification in recursive systems I

- In our first example, we saw that the supply-demand system could be identified if the structure was recursive
- Recursive generally requires two thing:
  1. The matrix of contemporanous coefficients must be (upper or lower) triangular)
  2. The $\boldsymbol{\Sigma}$ matric must be diagonal
- Note from previous lectures, that this is exactly what we achieve if we model the system in terms of conditional and marginal models.

Identification in recursive systems II

- Suppose that the RF of a dynamic SEM for Y_t and X_t is the stationary VAR with white-noise Gaussian:

$$\underbrace{\begin{pmatrix} Y_t \\ X_t \end{pmatrix}}_{\mathbf{y}_t} = \underbrace{\begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix}}_{\mathbf{\Pi}} \underbrace{\begin{pmatrix} Y_{t-1} \\ X_{t-1} \end{pmatrix}}_{\mathbf{y}_{t-1}} + \underbrace{\begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{xt} \end{pmatrix}}_{\boldsymbol{\varepsilon}_t}, \qquad (17)$$

- $Cov(\varepsilon_{yt},\varepsilon_{xt}) \neq 0$ usually. but if we represent the VAR in terms of a conditional ADL and the second line in the VAR

$$Y_t = \phi_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t \qquad (18)$$

$$X_t = \pi_{21} Y_{t-1} + \pi_{22} X_{t-1} + \varepsilon_{xt} \qquad (19)$$

Identification in recursive systems III

we have a model with a recursive structure

$$Y_t - \beta_0 X_t = \phi_1 Y_{t-1} + \beta_1 X_{t-1} + \varepsilon_t \tag{20}$$

$$X_t = \pi_{21} Y_{t-1} + \pi_{22} X_{t-1} + \varepsilon_{xt} \tag{21}$$

since $Cov(\varepsilon_t, \varepsilon_{xt}) = 0$ from the construction of the conditional model

- Because of $Cov(\varepsilon_t, \varepsilon_{xt}) = 0$ OLS on the marginal model and ADL separately give consistent estimators of all the parameters.
- One way to say this, is that, if our parameters of interest are the parameters of the conditional-marginal equation, then that model is the structure, and it is identified.

Notation for a structural equation I

- Without loss of generality, we consider equation # 1 in a SEM
- Apart from the disturbance, adopt the notation in DM page 522:

$$\mathbf{y}_1 = \mathbf{Z}_1\boldsymbol{\beta}_{11} + \mathbf{Y}_1\boldsymbol{\beta}_{21} + \boldsymbol{\epsilon}_1 \tag{22}$$

- $\mathbf{y}_1$ 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 II

The total number of explanatory variable in the first equation is

$$k_{11} + k_{12} = k_1 \tag{23}$$

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

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

Simultaneity bias of OLS estimators I

By defining the two partitioned matrices:

$$\mathbf{X}_1 = \begin{pmatrix} \mathbf{Z}_1 & : & \mathbf{Y}_1 \end{pmatrix} \text{ (!!!)} \tag{24}$$

$$\boldsymbol{\beta}_1 = \begin{pmatrix} \boldsymbol{\beta}_{11} & : & \boldsymbol{\beta}_{21} \end{pmatrix}' \tag{25}$$

(22) can be written compactly as

$$\mathbf{y}_1 = \mathbf{X}_1 \boldsymbol{\beta}_1 + \boldsymbol{\epsilon}_1 \tag{26}$$

which is the notation in Ch 8.3 in DM

Simultaneity bias of OLS estimators II

- (26) looks like an ordinary regression, **but of course we know better**: Since it is the first structural equation in a SEM (a system!) we have

$$plim(\frac{1}{n}\mathbf{X}_1'\epsilon_1) \neq \mathbf{0} \tag{27}$$

since $\mathbf{X}_1$ includes k_{12} endogenous explanatory variables.

- Hence the OLS estimator $\hat{\boldsymbol{\beta}}_1$ will be inconsistent by simultaneous equations bias.
- All the estimators will be affected, not only $\hat{\boldsymbol{\beta}}_{21}$.

Instrumental variables matrix I

- Since (22) is assumed to be exactly identified, it is logically consistent to assume that the SEM defines a matrix $\mathbf{W}_1$ with the properties

$$plim(\frac{1}{n}\mathbf{W}_1'\mathbf{X}_1) = \mathbf{S}_{W_1'X_1} \text{ (invertible)} \tag{28}$$

$$plim(\frac{1}{n}\mathbf{W}_1'\epsilon_1) = \mathbf{0} \text{ (independence)} \tag{29}$$

$$plim(\frac{1}{n}\mathbf{W}_1'\mathbf{W}_1) = \mathbf{S}_{W_1'W_1} \text{ (positive definite)} \tag{30}$$

- Since (28) required invertibility, the number of columns in $\mathbf{W}_1$ must k_1.

Instrumental variables matrix II

- The k_{11} predetermined variables included eq # 1 are of course in $\mathbf{W}_1$. In addition we need k_{12} instrumental variables for the included endogenous explanatory variables in eq # 1.
- The k_{12} instrumental variables must be "taken from" the predetermined variables in the SEM that are excluded from the first equation.
- But if structural equation # 1 is **just-identified**, the number of excluded predetermined variables is exactly equal to the number of included endogenous variables in the equation minus one. Hence we can write $\mathbf{W}_1$ as

$$\mathbf{W}_1 = \left(\mathbf{Z}_1 \;\; : \;\; \mathbf{X}_{01} \right) \tag{31}$$

where $\mathbf{X}_{01}$ is $n \times k_{11}$(number of included endogenous variables minus one).

Instrumental variables matrix III

- We see that defined in this way, $\mathbf{W}_1$ will satisfy both
  - **instrument relevance** (28), and
  - **instrument validity** in (29)
- In the absence of perfect collinearity among instruments, also (30).

The IV estimator I

If we apply the method of moments to the structural equation $\mathbf{y}_1 = \mathbf{X}_1\beta_1 + \epsilon_\mathbf{1}$, but with $\mathbf{W}_1$ instead of $\mathbf{X}_1$, we obtain the IV-estimator

$$\hat{\beta}_{1,IV} = (\mathbf{W}_1'\mathbf{X}_1)^{-1}\mathbf{W}_1'\mathbf{y}_1 \tag{32}$$

which an introductory course usually presents for the case of a single endogenous X.

Among the results for the IV-estimator that we know are:

- $\hat{\beta}_{1,IV}$ is a consistent estimator but it is biased in finite sample
- Asymptotic inference based on "t-ratios" and Chi-squared statistics for joint hypotheses are valid

The IV estimator II

- The estimated standard error of $\hat{\beta}_{1,IV}$ can be considerably larger than $\hat{\beta}_{1,OLS}$ if the instruments are **weak**, meaning that they are almost independent from the endogenous variables that they act as instrumenst for.
- Building on the asymptotic theory in Ch 6, DM show that the asymptotic covariance matrix of the vector $(\hat{\beta}_{1,IV} - \beta_1)$ is

$$Var(\hat{\beta}_{1,IV} - \beta_1) = \sigma_1^2 \mathbf{S}_{W_1'X_1}^{-1} \mathbf{S}_{W_1'W_1} \mathbf{S}_{W_1'X_1}^{-1} \qquad (33)$$

see (8.17)

$$Var(\hat{\beta}_{1,IV} - \beta_1) = \sigma_1^2 plim(n^{-1}\mathbf{X}_1'\mathbf{P}_{W_1}\mathbf{X}_1)^{-1} \qquad (34)$$

where:

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

The IV estimator III

our old friend the *prediction maker*.

- What role does it take here?