Instrumental Variables[1]

STA431 Winter/Spring 2013

[1] See last slide for copyright information.

Overview

Seeking identifiability

We have seen that in simple regression, parameters of a model with measurement error are not identifiable.

$$\begin{aligned} Y_i &= \alpha_1 + \beta_1 X_i + \epsilon_i \\ W_i &= \nu + X_i + e_i, \end{aligned}$$

- For example, X might be income and Y might be credit card debt.
- Include another response variable Y_2, like value of automobile.

Include a second response variable

- Response variable of primary interest is now called $Y_{i,1}$
- The second response variable $Y_{i,2}$ is called an *instrumental variable.*
- It's just a tool.

$$
\begin{array}{rcl}
W_i & = & \nu + X_i + e_i \\
Y_{i,1} & = & \alpha_1 + \beta_1 X_i + \epsilon_{i,1} \\
Y_{i,2} & = & \alpha_2 + \beta_2 X_i + \epsilon_{i,2}
\end{array}
$$

where X_i, e_i, $\epsilon_{i,1}$ and $\epsilon_{i,2}$ are all independent, $Var(X_i) = \phi$, $Var(e_i) = \omega$, $Var(\epsilon_{i,1}) = \psi_1$, $Var(\epsilon_{i,2}) = \psi_2$, $E(X_i) = \mu_x$ and the expected values of all error terms are zero. The regression coefficients α_j and β_j are unknown constants.

Are the parameters identifiable?

$$
\begin{aligned}
W_i &= \nu + X_i + e_i \\
Y_{i,1} &= \alpha_1 + \beta_1 X_i + \epsilon_{i,1} \\
Y_{i,2} &= \alpha_2 + \beta_2 X_i + \epsilon_{i,2}
\end{aligned}
$$

- Assume everything is normal: $\mathbf{D}_i \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.
- $\boldsymbol{\theta} = (\nu, \alpha_1, \alpha_2, \beta_1, \beta_2, \mu_x, \phi, \omega, \psi_1, \psi_2)$: Ten parameters.
- $\boldsymbol{\mu}$ is 3×1.
- $\boldsymbol{\Sigma}$ is 3×3.
- Nine moment structure equations in ten unknowns.

Look at the covariance structure equations

We are pessimistic about the expected values

$$\begin{aligned} W_i &= \nu + X_i + e_i \\ Y_{i,1} &= \alpha_1 + \beta_1 X_i + \epsilon_{i,1} \\ Y_{i,2} &= \alpha_2 + \beta_2 X_i + \epsilon_{i,2} \end{aligned}$$

$$\boldsymbol{\Sigma} = V \left(\begin{array}{c} W_i \\ Y_{i,1} \\ Y_{i,2} \end{array} \right) = \left[\begin{array}{ccc} \phi+\omega & \beta_1\phi & \beta_2\phi \\ & \beta_1^2\phi + \psi_1 & \beta_1\beta_2\phi \\ & & \beta_2^2\phi + \psi_2 \end{array} \right].$$

Six equations in six unknowns

A unique solution is possible but not guaranteed

$$\left[\begin{array}{ccc} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ & \sigma_{22} & \sigma_{23} \\ & & \sigma_{33} \end{array}\right] = \left[\begin{array}{ccc} \phi+\omega & \beta_1\phi & \beta_2\phi \\ & \beta_1^2\phi+\psi_1 & \beta_1\beta_2\phi \\ & & \beta_2^2\phi+\psi_2 \end{array}\right]$$

Identifiability depends on where you are in the parameter space. Consider

- $\beta_1 = 0$ and $\beta_2 = 0$
- $\beta_1 = 0$ and $\beta_2 \neq 0$
- $\beta_1 \neq 0$ and $\beta_2 = 0$
- $\beta_1 \neq 0$ and $\beta_2 \neq 0$

The parameter β_1 is identifiable if $\beta_2 \neq 0$: $\beta_1 = \frac{\sigma_{23}}{\sigma_{13}}$.

Suppose both $\beta_1 \neq 0$ and $\beta_2 \neq 0$

$$\left[\begin{array}{ccc} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ & \sigma_{22} & \sigma_{23} \\ & & \sigma_{33} \end{array}\right] = \left[\begin{array}{ccc} \phi+\omega & \beta_1\phi & \beta_2\phi \\ & \beta_1^2\phi+\psi_1 & \beta_1\beta_2\phi \\ & & \beta_2^2\phi+\psi_2 \end{array}\right]$$

$$\beta_1 = \frac{\sigma_{23}}{\sigma_{13}}$$
$$\beta_2 = \frac{\sigma_{23}}{\sigma_{12}}$$
$$\phi = \frac{\sigma_{12}\sigma_{13}}{\sigma_{23}}$$

Solve for ω, ψ_1 and ψ_2 by subtraction. Can write

$$\omega = \sigma_{11} - \phi$$
$$\psi_1 = \sigma_{22} - \beta_1^2\phi$$
$$\psi_2 = \sigma_{33} - \beta_2^2\phi$$

Without substituting for parameter that have already been identified. Don't need to give complete explicit solution. This shows it can be done.

What about the expected values?

$$
\begin{array}{rcl}
W_i & = & \nu + X_i + e_i \\
Y_{i,1} & = & \alpha_1 + \beta_1 X_i + \epsilon_{i,1} \\
Y_{i,2} & = & \alpha_2 + \beta_2 X_i + \epsilon_{i,2}
\end{array}
$$

$$
\begin{array}{rcl}
\mu_1 & = & \nu + \mu_x \\
\mu_2 & = & \alpha_1 + \beta_1 \mu_x \\
\mu_3 & = & \alpha_2 + \beta_2 \mu_x
\end{array}
$$

- Three equations in four unknowns, even assuming β_1 and β_2 known.
- Re-parameterize.

Re-parameterize

$$\begin{aligned} \mu_1 &= \nu + \mu_x \\ \mu_2 &= \alpha_1 + \beta_1 \mu_x \\ \mu_3 &= \alpha_2 + \beta_2 \mu_x \end{aligned}$$

- Absorb $\nu, \mu_x, \alpha_1, \alpha_2$ into $\boldsymbol{\mu}$.
- Parameter was $\boldsymbol{\theta} = (\nu, \mu_x, \alpha_1, \alpha_2, \beta_1, \beta_2, \phi, \omega, \psi_1, \psi_2)$
- Now it's $\boldsymbol{\theta} = (\mu_1, \mu_2, \mu_3, \beta_1, \beta_2, \phi, \omega, \psi_1, \psi_2)$
- Dimension of the parameter space is now one less.
- We haven't lost much.

We haven't lost much especially because the model was already re-parameterized

Model is

$$\begin{aligned} W_i &= \nu + X_i + e_i \\ Y_{i,1} &= \alpha_1 + \beta_1 X_i + \epsilon_{i,1} \\ Y_{i,2} &= \alpha_2 + \beta_2 X_i + \epsilon_{i,2} \end{aligned}$$

But of course there is measurement error in Y_1 and Y_2. Recall

$$\begin{aligned} Y &= \alpha + \beta X + \epsilon \\ V &= \nu_0 + Y + e \\ &= \nu_0 + (\alpha + \beta X + \epsilon) + e \\ &= (\nu_0 + \alpha) + \beta X + (\epsilon + e) \\ &= \alpha' + \beta X + \epsilon' \end{aligned}$$

Summary

- Adding the instrumental variable didn't help identify the expected values and intercepts. That's hopeless.
- But we did identify β_1, which is the most interesting parameter.
- Re-parameterizing, absorbed the intercepts and expected values into $\boldsymbol{\mu}$.
- Where β_1 and β_2 are both non-zero, the entire parameter vector is identifiable.
- For maximum likelihood estimation, it helps to have the *entire* parameter vector identifiable at the true parameter value.
- This is definitely a success.

Testing $H_0 : \beta_1 = 0$

The most interesting null hypothesis

- The parameter β_1 is identifiable, so a valid test is possible.
- But the whole parameter *vector* is not identifiable when $\beta_1 = 0$.
- Technical conditions of the likelihood ratio test are not satisfied.
- It becomes quite "interesting."
- Likelihood ratio statistic actually has 2 *df* even though H_0 appears to impose only one restriction on the parameter.
- Too interesting.

It's better with two (or more) instrumental variables.

$$
\begin{array}{rcl}
W_i & = & \nu + X_i + e_i \\
Y_{i,1} & = & \alpha_1 + \beta_1 X_i + \epsilon_{i,1} \\
Y_{i,2} & = & \alpha_2 + \beta_2 X_i + \epsilon_{i,2} \\
Y_{i,3} & = & \alpha_3 + \beta_3 X_i + \epsilon_{i,3},
\end{array}
$$

$$
\boldsymbol{\Sigma} = \left(\begin{array}{cccc}
\phi + \omega & \beta_1 \phi & \beta_2 \phi & \beta_3 \phi \\
 & \beta_1^2 \phi + \psi_1 & \beta_1 \beta_2 \phi & \beta_1 \beta_3 \phi \\
 & & \beta_2^2 \phi + \psi_2 & \beta_2 \beta_3 \phi \\
 & & & \beta_3^2 \phi + \psi_3
\end{array} \right).
$$

With two instrumental variables

- Again, identification of the expected values and intercepts is out of the question.
- So we re-parameterize,
- Absorbing the expected values and intercepts into $\boldsymbol{\mu} = E(\mathbf{D}_i)$
- And look at the covariance structure equations.

Covariance structure equations

$$
V\left(\begin{array}{c} W_i \\ Y_{i,1} \\ Y_{i,2} \\ Y_{i,3} \end{array}\right) = \left(\begin{array}{cccc} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ & \sigma_{22} & \sigma_{23} & \sigma_{24} \\ & & \sigma_{33} & \sigma_{3,4} \\ & & & \sigma_{4,4} \end{array}\right)
$$

$$
= \left(\begin{array}{cccc} \phi+\omega & \beta_1\phi & \beta_2\phi & \beta_3\phi \\ & \beta_1^2\phi+\psi_1 & \beta_1\beta_2\phi & \beta_1\beta_3\phi \\ & & \beta_2^2\phi+\psi_2 & \beta_2\beta_3\phi \\ & & & \beta_3^2\phi+\psi_3 \end{array}\right)
$$

- Ten equations in eight unknowns.
- Unique solution possible but not guaranteed.
- Primary interest is still in β_1.
- Assume $\beta_2 \neq 0$ and $\beta_3 \neq 0$, meaning only that Y_2 and Y_3 are well chosen.

Solve for ϕ

Assuming $\beta_2 \neq 0$ and $\beta_3 \neq 0$

$$\left(\begin{array}{c c c c} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ & \sigma_{22} & \sigma_{23} & \sigma_{24} \\ & & \sigma_{33} & \sigma_{3,4} \\ & & & \sigma_{4,4} \end{array} \right) = \left(\begin{array}{c c c c} \phi+\omega & \beta_1\phi & \beta_2\phi & \beta_3\phi \\ & \beta_1^2\phi+\psi_1 & \beta_1\beta_2\phi & \beta_1\beta_3\phi \\ & & \beta_2^2\phi+\psi_2 & \beta_2\beta_3\phi \\ & & & \beta_3^2\phi+\psi_3 \end{array} \right)$$

$$\frac{\sigma_{13}\sigma_{14}}{\sigma_{34}} = \frac{\beta_2\beta_3\phi^2}{\beta_2\beta_3\phi} = \phi$$

Then all you have to write is

$$
\begin{aligned}
\omega &= \sigma_{11} - \phi \\
\beta_1 &= \frac{\sigma_{12}}{\phi} \\
\beta_2 &= \frac{\sigma_{13}}{\phi} \\
\beta_3 &= \frac{\sigma_{14}}{\phi} \\
\psi_1 &= \sigma_{22} - \beta_1^2\phi \\
\psi_2 &= \sigma_{33} - \beta_2^2\phi \\
\psi_3 &= \sigma_{44} - \beta_3^2\phi
\end{aligned}
$$

Notice again how once we have solved for a model parameter, we may use it to solve for other parameters without explicitly substituting in terms of σ_{ij}.

Parameters can be recovered from the covariance matrix

$$
\begin{aligned}
\phi &= \frac{\sigma_{13}\sigma_{14}}{\sigma_{34}} \\
\omega &= \sigma_{11} - \phi \\
\beta_1 &= \frac{\sigma_{12}}{\phi} \\
\beta_2 &= \frac{\sigma_{13}}{\phi}
\end{aligned}
\qquad
\begin{aligned}
\beta_3 &= \frac{\sigma_{14}}{\phi} \\
\psi_1 &= \sigma_{22} - \beta_1^2\phi \\
\psi_2 &= \sigma_{33} - \beta_2^2\phi \\
\psi_3 &= \sigma_{44} - \beta_3^2\phi
\end{aligned}
$$

- Parameter vector is identifiable almost everywhere in the parameter space.
- Everywhere β_2 and β_3 are both non-zero
- $\beta_1 = 0 \Leftrightarrow \sigma_{12} = 0$ presents no problem.

But there is more than one way to recover the parameter values from Σ

$$\left(\begin{array}{cccc} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ & \sigma_{22} & \sigma_{23} & \sigma_{24} \\ & & \sigma_{33} & \sigma_{3,4} \\ & & & \sigma_{4,4} \end{array} \right) = \left(\begin{array}{cccc} \phi+\omega & \beta_1\phi & \beta_2\phi & \beta_3\phi \\ & \beta_1^2\phi+\psi_1 & \beta_1\beta_2\phi & \beta_1\beta_3\phi \\ & & \beta_2^2\phi+\psi_2 & \beta_2\beta_3\phi \\ & & & \beta_3^2\phi+\psi_3 \end{array} \right)$$

$$\begin{aligned} \beta_1 &= \frac{\sigma_{12}\sigma_{34}}{\sigma_{13}\sigma_{14}} \\ \beta_1 &= \frac{\sigma_{23}}{\sigma_{13}} \\ \beta_1 &= \frac{\sigma_{24}}{\sigma_{14}} \end{aligned}$$

Is there a problem?

$$\beta_1 = \frac{\sigma_{12}\sigma_{34}}{\sigma_{13}\sigma_{14}} = \frac{\sigma_{23}}{\sigma_{13}} = \frac{\sigma_{24}}{\sigma_{14}}$$

Does this mean the solution for β_1 is not "unique?"

- No – everything is okay.
- If the parameters can be recovered from the covariances in any way at all, they are identifiable.
- If the model is correct, all the seemingly different ways must be the same.
- That is,
$$\frac{\sigma_{12}\sigma_{34}}{\sigma_{13}\sigma_{14}} = \frac{\sigma_{23}}{\sigma_{13}} \text{ and } \frac{\sigma_{12}\sigma_{34}}{\sigma_{13}\sigma_{14}} = \frac{\sigma_{24}}{\sigma_{14}}$$
- Simplifying a bit,

$$\sigma_{12}\sigma_{34} = \sigma_{14}\sigma_{23} = \sigma_{13}\sigma_{24}$$

Model implies two constraints on the covariance matrix

$$\sigma_{12}\sigma_{34} = \sigma_{14}\sigma_{23} = \sigma_{13}\sigma_{24}$$

- All three products equal $\beta_1\beta_2\beta_3\phi^2$
- True even when some $\beta_j = 0$

$$\left(\begin{array}{cccc} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ & \sigma_{22} & \sigma_{23} & \sigma_{24} \\ & & \sigma_{33} & \sigma_{3,4} \\ & & & \sigma_{4,4} \end{array} \right) = \left(\begin{array}{cccc} \phi+\omega & \beta_1\phi & \beta_2\phi & \beta_3\phi \\ & \beta_1^2\phi+\psi_1 & \beta_1\beta_2\phi & \beta_1\beta_3\phi \\ & & \beta_2^2\phi+\psi_2 & \beta_2\beta_3\phi \\ & & & \beta_3^2\phi+\psi_3 \end{array} \right)$$

Model implies $\sigma_{12}\sigma_{34} = \sigma_{14}\sigma_{23} = \sigma_{13}\sigma_{24}$

- Parameter is identifiable.
- Ten equations in eight unknowns.
- Call the parameter *over-identifiable.*
- If there were the same number of equations as unknowns, it would be *just identifiable.*
- Model imposes two equality constraints (restrictions) on the covariance matrix: $10 - 8 = 2$
- Sometimes called *over-identifying restrictions.*
- These are the constraints that are tested in the likelihood ratio test for goodness of fit.
- More instrumental variables can't hurt.

Multiple Explanatory Variables

An example with just two explanatory variables (and two instrumental variables)

Independently for $i = 1, \ldots, n$,

$$
\begin{array}{rcl}
W_{i,1} & = & \nu_1 + X_{i,1} + e_{i,1} \\
Y_{i,1} & = & \alpha_1 + \beta_1 X_{i,1} + \epsilon_{i,1} \\
Y_{i,2} & = & \alpha_2 + \beta_2 X_{i,1} + \epsilon_{i,2} \\
W_{i,2} & = & \nu_2 + X_{i,2} + e_{i,2} \\
Y_{i,3} & = & \alpha_3 + \beta_3 X_{i,2} + \epsilon_{i,3} \\
Y_{i,4} & = & \alpha_4 + \beta_4 X_{i,2} + \epsilon_{i,4}
\end{array}
$$

where $E(X_{i,j}) = \mu_j$, $e_{i,j}$ and $\epsilon_{i,j}$ are independent of one another and of $X_{i,j}$, $Var(e_{i,j}) = \omega_j$, $Var(\epsilon_{i,j}) = \psi_j$, and

$$
V \left(\begin{array}{c} X_{i,1} \\ X_{i,1} \end{array} \right) = \left(\begin{array}{cc} \phi_{11} & \phi_{12} \\ \phi_{12} & \phi_{22} \end{array} \right)
$$

As usual, intercepts and expected values can't be recovered individually

- Eight intercepts and expected values of latent variables.
- Six expected values of observable variables.
- Re-parameterize, absorbing them into $\mu_1, \ldots, \mu_6$.
- Estimate with the vector of 6 sample means and set them aside, forever.

Covariance matrix of $(W_{i,1}, Y_{i,1}, Y_{i,2}, W_{i,2}, Y_{i,3}, Y_{i,4})'$

$$[\sigma_{ij}] =$$

$$\left(\begin{array}{cccccc}
\phi_{11}+\omega_1 & \beta_1\phi_{11} & \beta_2\phi_{11} & \phi_{12} & \beta_3\phi_{12} & \beta_4\phi_{12} \\
 & \beta_1^2\phi_{11}+\psi_1 & \beta_1\beta_2\phi_{11} & \beta_1\phi_{12} & \beta_1\beta_3\phi_{12} & \beta_1\beta_4\phi_{12} \\
 & & \beta_2^2\phi_{11}+\psi_2 & \beta_2\phi_{12} & \beta_2\beta_3\phi_{12} & \beta_2\beta_4\phi_{12} \\
 & & & \phi_{22}+\omega_2 & \beta_3\phi_{22} & \beta_4\phi_{22} \\
 & & & & \beta_3^2\phi_{22}+\psi_3 & \beta_3\beta_4\phi_{22} \\
 & & & & & \beta_4^2\phi_{22}+\psi_4
\end{array} \right)$$

$$\boldsymbol{\theta} = (\beta_1, \beta_2, \beta_3, \beta_4, \phi_{11}, \phi_{12}, \phi_{22}, \omega_1, \omega_2, \psi_1, \psi_2, \psi_3, \psi_4)$$

- Does this model pass the test of the parameter count rule?
- Where the parameter vector is identifiable, how many over-identifying restrictions are there?
- How many degrees of freedom in the likelihood ratio test for model fit?

Where is the entire parmeter vector identifiable?

$$
\left(\begin{array}{cccccc}
\phi_{11}+\omega_1 & \beta_1\phi_{11} & \beta_2\phi_{11} & \phi_{12} & \beta_3\phi_{12} & \beta_4\phi_{12} \\
 & \beta_1^2\phi_{11}+\psi_1 & \beta_1\beta_2\phi_{11} & \beta_1\phi_{12} & \beta_1\beta_3\phi_{12} & \beta_1\beta_4\phi_{12} \\
 & & \beta_2^2\phi_{11}+\psi_2 & \beta_2\phi_{12} & \beta_2\beta_3\phi_{12} & \beta_2\beta_4\phi_{12} \\
 & & & \phi_{22}+\omega_2 & \beta_3\phi_{22} & \beta_4\phi_{22} \\
 & & & & \beta_3^2\phi_{22}+\psi_3 & \beta_3\beta_4\phi_{22} \\
 & & & & & \beta_4^2\phi_{22}+\psi_4
\end{array}\right)
$$

- What happens if $\beta_1 = \beta_2 = 0$?
- Why is it reasonable to assume $\beta_2 \neq 0$ and $\beta_4 \neq 0$?
- In that case, what else do you need?
- Would any other condition identify the whole parameter vector?

My answer

$$\left(\begin{array}{cccccc}
\phi_{11}+\omega_1 & \beta_1\phi_{11} & \beta_2\phi_{11} & \phi_{12} & \beta_3\phi_{12} & \beta_4\phi_{12} \\
 & \beta_1^2\phi_{11}+\psi_1 & \beta_1\beta_2\phi_{11} & \beta_1\phi_{12} & \beta_1\beta_3\phi_{12} & \beta_1\beta_4\phi_{12} \\
 & & \beta_2^2\phi_{11}+\psi_2 & \beta_2\phi_{12} & \beta_2\beta_3\phi_{12} & \beta_2\beta_4\phi_{12} \\
 & & & \phi_{22}+\omega_2 & \beta_3\phi_{22} & \beta_4\phi_{22} \\
 & & & & \beta_3^2\phi_{22}+\psi_3 & \beta_3\beta_4\phi_{22} \\
 & & & & & \beta_4^2\phi_{22}+\psi_4
\end{array}\right)$$

EITHER

- One of β_1 and β_2 non-zero, and
- One of β_3 and β_4 non-zero, and
- $\phi_{12} \neq 0$

OR, all of $\beta_1, \ldots, \beta_4$ non-zero.

Could we get by with less information?

If we wanted to identify just some interesting parameters?

$$\left(\begin{array}{cccccc}
\phi_{11}+\omega_1 & \beta_1\phi_{11} & \beta_2\phi_{11} & \phi_{12} & \beta_3\phi_{12} & \beta_4\phi_{12} \\
 & \beta_1^2\phi_{11}+\psi_1 & \beta_1\beta_2\phi_{11} & \beta_1\phi_{12} & \beta_1\beta_3\phi_{12} & \beta_1\beta_4\phi_{12} \\
 & & \beta_2^2\phi_{11}+\psi_2 & \beta_2\phi_{12} & \beta_2\beta_3\phi_{12} & \beta_2\beta_4\phi_{12} \\
 & & & \phi_{22}+\omega_2 & \beta_3\phi_{22} & \beta_4\phi_{22} \\
 & & & & \beta_3^2\phi_{22}+\psi_3 & \beta_3\beta_4\phi_{22} \\
 & & & & & \beta_4^2\phi_{22}+\psi_4
\end{array}\right)$$

- Usual rule in Econometrics is at least one instrumental variable for each explanatory variable.
- What if no instrumental variable for X_2?
- What if no response variables at all for X_2?

Observations

- Instrumental variables can solve some of the terrible problems with measurement error in regression.
- General rules like "At least one instrumental variable for each explanatory variable" are useful even if they are over-simplifications.
- Awareness of parameter identifiability is vital in the *planning* of data collection.

- Most observational data sets are collected without the right kind of planning.

Copyright Information