SIMULTANEOUS EQUATIONS SYSTEMS, II

Econometric Analysis of Cross Section and Panel Data, 2e
MIT Press
Jeffrey M. Wooldridge

1. Cross Equation Restrictions
2. Covariance Restrictions
3. A Subtle Point About Identification
4. SEMs Nonlinear in Endogenous Variables
5. Different Instruments for Different Equations

1. CROSS EQUATION RESTRICTIONS

- Most applications of pure SEMs have restrictions within each equation but not across equation. In the labor supply/wage offer example, when could we justify assume that parameters in the supply and offer functions are related? Probably never.
- But some applications of the statistical model with endogenous explanatory variables are not to SEMs, and in some cases there are cross-equation restrictions.

• Consider a simple example just to show how such restrictions can help with identification:

$$
\begin{aligned}
y_1 &= \gamma_{12} y_2 + \delta_{11} z_1 + \delta_{12} z_2 + \delta_{13} z_3 + u_1 \\
y_2 &= \gamma_{21} y_1 + \delta_{21} z_1 + \delta_{22} z_2 + u_2
\end{aligned}
$$

where, as usual, each z_j is uncorrelated with each u_g. Without further restrictions, the first equation is unidentified and the second is (just) identified if and only if $\delta_{13} \neq 0$.

• Suppose now that $\delta_{12} = \delta_{22}$. Because δ_{22} is identified without this restriction, we can treat $\delta_{22} = \delta_{12}$ as known for determining identification of the second equation.

• Write the first equation as

$$y_1 - \delta_{12}z_2 = \gamma_{12}y_2 + \delta_{11}z_1 + \delta_{13}z_3 + u_1$$

where we can treat the LHS variable as known. This equation is identified now because we can use z_2 as an IV for y_2 provided $\delta_{22} \neq 0$.

• We can even imagine a two-step procedure. First, estimate δ_{22} by IV on the second equation; assume that we can reject $\delta_{22} = 0$. Then, estimate

$$y_1 - \hat{\delta}_{22}z_2 = \gamma_{12}y_2 + \delta_{11}z_1 + \delta_{13}z_3 + error$$

using (z_1, z_2, z_3) as IVs. Would have to account for sampling error in $\hat{\delta}_{22}$ for proper inference.

• Once we know both equations are identified, better to use a system method. Define the vector of unrestricted parameters as

$$\boldsymbol{\beta} = (\gamma_{12}, \delta_{11}, \delta_{12}, \delta_{13}, \gamma_{21}, \delta_{21})'.$$

Then we can write, for a random draw i,

$$\begin{pmatrix} y_{i1} \\ y_{i2} \end{pmatrix} = \begin{pmatrix} y_{i2} & z_{i1} & z_{i2} & z_{i3} & 0 & 0 \\ 0 & 0 & z_{i2} & 0 & y_{i1} & z_{i1} \end{pmatrix} \boldsymbol{\beta} + \begin{pmatrix} u_{i1} \\ u_{i2} \end{pmatrix}$$

or

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

- The instruments are

$$\mathbf{Z}_i = \mathbf{I}_2 \otimes \mathbf{z}_i = \begin{pmatrix} z_{i1} & z_{i2} & z_{i3} & 0 & 0 & 0 \\ 0 & 0 & 0 & z_{i1} & z_{i2} & z_{i3} \end{pmatrix}.$$

- The system is just identified, so system IV is the only estimator.

2. COVARIANCE RESTRICTIONS

- Another source of identifying information that is occasionally used is zero covariance restrictions. (Restrictions on the variances themselves are almost unheard of.)
- As just one simple example, suppose the system is

$$
\begin{aligned}
y_1 &= \gamma_{12} y_2 + \delta_{11} z_1 + \delta_{13} z_3 + u_1 \\
y_2 &= \gamma_{21} y_1 + \delta_{21} z_1 + \delta_{22} z_2 + \delta_{23} z_3 + u_2
\end{aligned}
$$

Without further restrictions, the second equation is not identified and the first equation is just identified if and only if $\delta_{22} \neq 0$; assume this is the case.

- Suppose, however, that

$$Cov(u_1, u_2) = E(u_1 u_2) = 0,$$

so that the 2×2 matrix $\mathbf{\Sigma}$ is diagonal. This assumption serves to identify the second equation.

• Here is a way to see how. Because the first equation is identified, we can treat all of the parameters of the first equation as known for identification of the second equation. But then

$$u_1 = y_1 - \gamma_{12} y_2 - \delta_{11} z_1 - \delta_{13} z_3$$

becomes an observable variable.

• Normally, being able to effectively observe u_1 would not help with equation 2. But if we assume u_1 is uncorrelated with u_2, u_1 becomes a potential instrument for y_1. By assumption, u_1 is exogenous to the second equation and clearly it would be correlated with y_1.

$$y_2 = \gamma_{21} y_1 + \delta_{21} z_1 + \delta_{22} z_2 + \delta_{23} z_3 + u_2$$

- Can use a two-step procedure: (1) Estimate the first equation by IV with instruments (z_1, z_2, z_3), get the residuals, $\hat{u}_1$. (2) Estimate the second equation by IV using instruments $(\hat{u}_1, z_1, z_2, z_3)$. In this case, we cannot ignore estimation of the IVs in performing inference.

• A better approach is to use all of the moment conditions at once. Let $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ be the structural parameters in the first and second equations with the restrictions already imposed. Then the moment restrictions we have are

$$\begin{aligned} E[\mathbf{z}' u_1(\boldsymbol{\beta}_1)] &= \mathbf{0} \\ E[\mathbf{z}' u_2(\boldsymbol{\beta}_2)] &= \mathbf{0} \\ E[u_1(\boldsymbol{\beta}_1) u_2(\boldsymbol{\beta}_2)] &= 0, \end{aligned}$$

where the first two sets of restrictions are the standard ones and the third is the zero covariance assumption. (The notation is used to emphasize the moment conditions depend on the parameters.)

- Now, $\boldsymbol{\beta}_1$ has three unknown elements and $\boldsymbol{\beta}_2$ has four, so there are seven unknown parameters. But there are also seven total restrictions.

- Method of moments estimation is now nonlinear in the parameters because of $E[u_1(\boldsymbol{\beta}_1)u_2(\boldsymbol{\beta}_2)] = 0$. We will discuss nonlinear GMM later.

- Again, in good SEM applications, such as supply and demand, it is rare in cross section settings to impose zero covariance restrictions.

- A *recursive system* has a sequential decision making structure:

$$\begin{aligned}
y_1 &= \mathbf{z}\boldsymbol{\delta}_1 + u_1 \\
y_2 &= \gamma_{21} y_1 + \mathbf{z}\boldsymbol{\delta}_2 + u_2 \\
y_3 &= \gamma_{31} y_1 + \gamma_{32} y_2 + \mathbf{z}\boldsymbol{\delta}_3 + u_3 \\
&\vdots \\
y_G &= \gamma_{G1} y_1 + \gamma_{G2} y_2 + \ldots + \gamma_{G,G-1} y_{G-1} + \mathbf{z}\boldsymbol{\delta}_G + u_G
\end{aligned}$$

so that all exogenous variables show up in each equation.

- The first equation is always identified, but that would be the only identified equation without more assumptions. A *fully recursive system* adds the assumption

$$E(u_g u_h) = 0, g \neq h,$$

which adds more than enough conditions to identify all parameters.

- Sequential estimation is possible, starting with OLS on the first equation and using the residuals, $\hat{u}_1$, to instrument for y_1 in the second equation – and so on.

- Joint GMM can be used but moment conditions are nonlinear in parameters.

3. A SUBTLE POINT ABOUT IDENTIFICATION

- In the general SEM written as

$$\begin{aligned} \mathbf{y}\boldsymbol{\gamma}_1 + \mathbf{z}\boldsymbol{\delta}_1 + u_1 &= 0 \\ \mathbf{y}\boldsymbol{\gamma}_2 + \mathbf{z}\boldsymbol{\delta}_2 + u_2 &= 0 \\ &\vdots \\ \mathbf{y}\boldsymbol{\gamma}_G + \mathbf{z}\boldsymbol{\delta}_G + u_G &= 0 \end{aligned}$$

we used the weakest exogeneity requirement on **z**:

$$E(\mathbf{z}'u_g) = \mathbf{0},\ g = 1, \ldots, G.$$

- This means that we can only use $\mathbf{z}$ as the instruments in each equation, not nonlinear functions of $\mathbf{z}$.
- If the entire system is structural, we are often willing to assume correct functional forms in the exogenous variables, which means we assume

$$E(u_g|\mathbf{z}) = 0,\ g = 1,\ldots,G.$$

• Aha! With a zero conditional mean assumption, any functions of $\mathbf{z}$, say z_j^2, $z_j z_k$, $z_j/(1+|z_k|)$, and infinitely other functions are uncorrelated with u_g (under the assumption of finite second moments for u_g and the chosen functions of $\mathbf{z}$).

• Question: How come we have not solved the identification problem without making any restrictions on the structural parameters?

- Answer: If we assume $E(u_g|\mathbf{z}) = 0$ then the reduced forms satisfy

$$
\begin{aligned}
y_g &= \mathbf{z}\boldsymbol{\pi}_g + v_g \\
E(v_g|\mathbf{z}) &= 0,\ g = 1, 2, \ldots, G,
\end{aligned}
$$

because v_g is a linear function of $u_1, u_2,$ ...,u_G.

- But then for any (vector) function $\mathbf{h}_g(\mathbf{z})$, the linear projection of y_g on $[\mathbf{z}, \mathbf{h}_g(\mathbf{z})]$ is just $\mathbf{z}\boldsymbol{\pi}_g$; it does not depend on $\mathbf{h}_g(\mathbf{z})$.

- Thus, for identification, adding nonlinear functions of the orginal instruments is useless. Our original identification analysis holds under the stronger form of exogeneity.
- $E(u_g|\mathbf{z}) = 0$ can have relevance for efficiency. In particular, if $Var(\mathbf{u}|\mathbf{z})$ is a function of $\mathbf{z}$, then GMM using an efficient weighting matrix and an expanded list of IVs is generally more efficient than 3SLS on the orginal system of GMM on the original system. It is rare in practice to try to exploit such efficiency gains.

4. SEMS NONLINEAR IN ENDOGENOUS VARIABLES

- Consider a version of the labor supply/wage offer example, written in equilibrium:

$$\begin{aligned} h &= \gamma_{12}\log(w) + \gamma_{13}[\log(w)]^2 + \delta_{11}z_1 + u_1 \\ \log(w) &= \gamma_{21}h + \delta_{22}z_2 + u_2 \end{aligned}$$

where z_1 and z_2 are exogenous. This is an *SEM nonlinear in endogenous variables*.

- If $\gamma_{13} = 0$, this would be a linear SEM for the purposes of identification and estimation. It is the presence of the squared term, $[\log(w)]^2$, that makes this a nonlinear model, not the presence of $\log(w)$.

- What about identification of this system? We cannot easily solve for h and $\log(w)$ as functions of $\mathbf{z} = (z_1, z_2)$ and $\mathbf{u} = (u_1, u_2)$. (The quadratic formula is needed, and then it might give a nonsensical answer.)
- But we can count. The second equation passes the order condition because we have an instrument for h, namely, z_1.

- At first glance, the first equation appears to be unidentified: $\log(w)$ and $[\log(w)]^2$ are both endogenous and yet z_2 is the only exogenous variable in the second equation.
- Important: Concluding that the labor supply function is unidentified is too pessimistic. In fact, if z_2 truly appears in the wage offer function, then, except in special cases, we have plenty of instruments for $\log(w)$ and $[\log(w)]^2$.

- Write the system as

$$
\begin{aligned}
y_1 &= \gamma_{12} y_2 + \gamma_{13} y_2^2 + \delta_{11} z_1 + u_1 \\
y_2 &= \gamma_{21} y_1 + \delta_{22} z_2 + u_2
\end{aligned}
$$

where $y_1 = h$ and $y_2 = \log(w)$. This is a nonlinear in endogenous variables because there are only two endogenous variables but we have a nonlinear function of one, y_2, appearing in the system.

- Once nonlinear functions appear, it makes sense to assume

$$E(u_g|\mathbf{z}) = 0$$

for all g ($g = 1, 2$ in this example).

• The zero conditional mean assumption gives us lots of possibilities for instruments in addition to (z_1, z_2). Once we have specified the IVs, estimation is standard (although trying to mimic two stage least squares can lead to trouble!) We can write the two equations as

$$y_1 = \gamma_{12} y_2 + \gamma_{13} y_3 + \delta_{11} z_1 + u_1$$
$$y_2 = \gamma_{21} y_1 + \delta_{22} z_2 + u_2$$

where $y_3 = y_2^2$. The only issue now is that this looks like a two-equation system with three endogenous variables.

• We can use results on optimal instruments to get a start [with the nominal assumption of constant $Var(\mathbf{u}|\mathbf{z})$]. For the first equation, We would have to find $E(y_2|\mathbf{z})$ and $E(y_2^2|\mathbf{z})$, which is very difficult in general.

• But suppose we act as if $\gamma_{13} = 0$. Notice then that each equation in the system would be just identified by our previous identification analysis (assuming $\delta_{11} \neq 0$, $\delta_{22} \neq 0$, which we do).

• We can use the linear version of the system to determine sensible instruments. When $\gamma_{13} = 0$, we can solve for linear reduced forms:

$$y_1 = \mathbf{z}\boldsymbol{\pi}_1 + v_1$$
$$y_2 = \mathbf{z}\boldsymbol{\pi}_2 + v_2$$

and v_1 and v_2 have zero conditional means give $\mathbf{z}$. Therefore, the optimal instrument for y_2 is $\mathbf{z}\boldsymbol{\pi}_2$.

- Further, we can write

$$
\begin{aligned}
y_2^2 &= (\mathbf{z}\boldsymbol{\pi}_2)^2 + 2(\mathbf{z}\boldsymbol{\pi}_2)v_2 + v_2^2 \\
E(y_2^2|\mathbf{z}) &= (\mathbf{z}\boldsymbol{\pi}_2)^2 + 2(\mathbf{z}\boldsymbol{\pi}_2)E(v_2|\mathbf{z}) + E(v_2^2|\mathbf{z}) \\
&= (\mathbf{z}\boldsymbol{\pi}_2)^2 + \tau_2^2
\end{aligned}
$$

because $E(v_2|\mathbf{z}) = 0$ and $E(v_2^2|\mathbf{z})$ is (nominally) assumed to be constant.

• In practice, using $(\mathbf{z}\boldsymbol{\pi}_2)^2$ is restrictive: it is derived under the linear version of the model and homoskedasticity of the structural errors. But it suggests a generally sensible approach: use squares, z_1^2, z_2^2 and the cross product, $z_1 z_2$ in the list of IVs for BOTH equations. So

$$(z_1, z_2, z_1^2, z_2^2, z_1 z_2).$$

In any application, we would add an intercept to both equations and, of course, it would act as its own instrument.

- When $\gamma_{13} \neq 0$, the linear projection of both y_1 and y_2 on $(z_1, z_2, z_1^2, z_2^2, z_1 z_2)$ would generally depend on the nonlinear functions, too.
- We would use the full set of instruments in each equation, whether we use a single equation method or a system method. It is (asymptotically) no less efficient to use the unrestricted vector $(z_1, z_2, z_1^2, z_2^2, z_1 z_2)$ even though the optimal IV for y_2^2 is $(\mathbf{z}\boldsymbol{\pi}_2)^2$.

- Question: If we apply 2SLS to the equation

$$y_1 = \gamma_{12} y_2 + \gamma_{13} y_2^2 + \delta_{11} z_1 + u_1$$

using instruments $(z_1, z_2, z_1^2, z_2^2, z_1 z_2)$, what do the first stage regressions look like?

• Answer: The linear projections underlying the first-stage regressions are

$$\begin{aligned}
y_2 &= \pi_{21}z_1 + \pi_{22}z_2 + \pi_{23}z_1^2 + \pi_{24}z_2^2 + \pi_{25}z_1z_2 + v_2 \\
y_2^2 &= \pi_{31}z_1 + \pi_{32}z_2 + \pi_{33}z_1^2 + \pi_{34}z_2^2 + \pi_{35}z_1z_2 + v_3
\end{aligned}$$

where v_2 and v_3 are uncorrelated with all RHS variables.

• So the correct procedure regresses each of y_2 and y_2^2 on $(z_1, z_2, z_1^2, z_2^2, z_1z_2)$, giving fitted values $\hat{y}_2$ and $\widehat{y_2^2}$.

• The following procedure does *not* work (except in special cases):

(1) Run the regression y_{i2} on z_{i1}, z_{i2} to get the fitted values, $\check{y}_{i2}$.

(2) Run the regression y_{i1} on $\check{y}_{i2}$, $\check{y}_{i2}^2$, z_{i1}.

• Inserting fitted values into nonlinear functions is often called the "forbidden regression."

• A modification *does* work: in step (2), use ($\check{y}_{i2}$, $\check{y}_{i2}^2$, z_{i1}) as IVs, *not* regressors!

• How can we study identification more generally? To get an idea, write an expanded system as

$$
\begin{aligned}
y_1 &= \gamma_{12} y_2 + \gamma_{13} y_3 + \delta_{11} z_1 + u_1 \\
y_2 &= \gamma_{21} y_1 + \delta_{22} z_2 + u_2 \\
y_3 &= \delta_{31} z_1 + \delta_{32} z_2 + \delta_{33} z_1^2 + \delta_{34} z_2^2 + \delta_{35} z_1 z_2 + u_3
\end{aligned}
$$

where $y_3 = y_2^2$.

• The third equation is a "reduced form," but it is useful to add it as part of the system to emphasize that $y_3 = y_2^2$ comes with its own instruments.

• This is now a $G = 3$ equation system. If we use rank analysis on the first equation, we have

$$\mathbf{R}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

where the parameter vector is

$$\boldsymbol{\beta}_1 = (-1, \gamma_{12}, \gamma_{13}, \delta_{11}, \delta_{12}, \delta_{13}, \delta_{14}, \delta_{15})'$$

- When we apply $\mathbf{R}_1$ to this expanded system, we get

$$\mathbf{R}_1\mathbf{B} = \begin{pmatrix} 0 & \delta_{22} & \delta_{32} \\ 0 & 0 & \delta_{33} \\ 0 & 0 & \delta_{34} \\ 0 & 0 & \delta_{35} \end{pmatrix},$$

and the key for identification is $\delta_{22} \neq 0$. We also need at least one of δ_{33}, δ_{34}, and δ_{35} different from zero, but that follows except by fluke if either z_1 or z_2 appears somewhere in the system.

• Simple approach that does not require us to expand the system in an essentially arbitrary way:

1. In the *original* system, label all nonredundant nonlinear functions as endogenous variables; $y_3 = y_2^2$ in the previous example.

2. Apply the rank condition to the this system *without* increasing the number of equations. If the equation satisfies the rank condition, it is generally identified.

• Now when we define $\mathbf{R}_1$, we do *not* include additional parameters on any new nonlinear functions of exogenous variables. In the previous example

$$
\begin{aligned}
y_1 &= \gamma_{12} y_2 + \gamma_{13} y_3 + \delta_{11} z_1 + u_1 \\
y_2 &= \gamma_{21} y_1 + \delta_{22} z_2 + u_2
\end{aligned}
$$

$$\boldsymbol{\beta}_1 = (-1, \gamma_{12}, \gamma_{13}, \delta_{11}, \delta_{12})',$$

$$\mathbf{R}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

- Then

$$\mathbf{R}_1\mathbf{B} = (0, \delta_{22})$$

and we need this matrix to have rank $G - 1 = 2 - 1 = 1$. (Note that we are back to $G = 2$, the original number of equations in the system.) The new rank condition holds if and only if $\delta_{22} \neq 0$.

- Technically, the rank condition in nonlinear models is not necessary for identification, but we should think of it as such.

• Suppose we have

$$
\begin{aligned}
y_1 &= \gamma_{12} y_2 + \gamma_{13} y_2^2 + \delta_{11} z_1 + \delta_{12} z_2 + u_1 \\
y_2 &= \gamma_{21} y_1 + \delta_{21} z_1 + u_2
\end{aligned}
$$

The first equation fails the modified rank condition because if fails the order condition. However, suppose $\gamma_{13} \neq 0$ and $\gamma_{21} \neq 0$. Then $E(y_2|\mathbf{z})$ is generally a nonlinear function of $\mathbf{z}$ (as is, of course, $E(y_2^2|\mathbf{z})$). Therefore, it is likely that $(z_1^2, z_2^2, z_1 z_2)$ can be used to identify the first equation.

- The problem is that we are getting identification off of a nonlinearity: if $\gamma_{13} = 0$, the first equation would not be identified. It is called a *poorly identified equation*. Achieving identification in this way is rarely convincing (cannot test $\gamma_{13} = 0$ because the equation is not identified under the null).

- EXERCISE: Consider

$$
\begin{aligned}
y_1 &= \gamma_{12} y_2 + \delta_{11} z_1 + u_1 \\
y_2 &= \gamma_{12} \exp(y_1) + \delta_{22} z_2 + u_2
\end{aligned}
$$

(for example, $y_1 = \log(wage)$, $y_2 = hours$). Propose a single-equation IV estimator for the second equation.

Estimation

- Once an equation is known to be identified, can use a standard single-equation method. But do not try to mimic 2SLS by plugging in fitted values into nonlinear functions in a second-stage OLS regression!
- Choosing instruments is challenging. If the first equation is

$$y_1 = \mathbf{q}_1(\mathbf{y},\mathbf{z})\boldsymbol{\beta}_1 + u_1,$$

choose functions of $\mathbf{z}$ to approximate $E[\mathbf{q}_1(\mathbf{y},\mathbf{z})|\mathbf{z}]$. Case-by-case basis.

- For systems, easiest to add reduced form linear projections for the extra nonlinear functions of endogenous varaibles.

EXAMPLE (with intercept explicitly included):

$$
\begin{aligned}
y_1 &= \gamma_{12} y_2 + \gamma_{13} y_2^2 + \gamma_{14} z_1 y_2 + \delta_{10} + \delta_{11} z_1 + \delta_{12} z_2 + u_1 \\
y_2 &= \gamma_{21} y_1 + \delta_{20} + \delta_{22} z_2 + \delta_{23} z_3 + u_2
\end{aligned}
$$

- First equation is generally identified if $\delta_{23} \neq 0$. If we estimate just this equation, choose IVs, say

$$
(1, z_1, z_2, z_3, z_1^2, z_2^2, z_3^2, z_1 z_2, z_1 z_3, z_2 z_3).
$$

- After defining the RHS variables and the instruments, just list each.

- In Stata:

```
. gen y2sq = y2^2
. gen z1y2 = z1*y2
. gen z1sq = z1^2
. gen z2sq = z2^2
. gen z3sq = z3^2
. gen z1z2 = z1*z2
. gen z1z3 = z1*z3
. gen z2z3 = z2*z3

. ivreg y1 z1 z2 (y2 y2sq z1y2 = z1 z2 z3 z1sq z2sq z3sq z1z2 z1z3 z2z3)
```

The above command implements the correct version of 2SLS.

• For 3SLS, would specific the full system. But there must be heteroskedasticity in at least the reduced forms!

```
reg3 (y1 y2 y2sq z1y2 z1 z2) (y2 y1 z2 z3)
     (y2sq z1 z2 z3 z1sq z2sq z3sq z1z2 z1z3 z2z3)
     (z1y2 z1 z2 z3 z1sq z2sq z3sq z1z2 z1z3 z2z3)
```

• Because the last two equations are just identified, the estimates of the first two equations are as if we estimate the first two equations by 3SLS.

• Would not hurt to do efficient GMM to allow system heteroskedasticity throughout.

- Labor Supply Example: Add $lwage^2$, $age \bullet lwage$ to the equation.

```
. ivreg hours educ nwifeinc age kidslt6 kidsge6 (lwage lwagesq agelwage =
   exper expersq educsq agesq ageeduc ageexper educexper nwifeincsq), robust

Instrumental variables (2SLS) regression               Number of obs =     428

------------------------------------------------------------------------------
             |               Robust
       hours |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lwage |  -118.8399   1450.236    -0.08   0.935    -2969.485    2731.805
     lwagesq |  -515.1494    347.704    -1.48   0.139    -1198.611     168.312
    agelwage |   52.92906   25.35424     2.09   0.037     3.091703    102.7664
        educ |  -87.16321   57.42158    -1.52   0.130    -200.0335    25.70704
    nwifeinc |  -7.329748   4.456165    -1.64   0.101    -16.08897    1.429477
         age |  -72.57273   33.17705    -2.19   0.029    -137.7869   -7.358534
     kidslt6 |   -147.618   186.3906    -0.79   0.429    -513.9951    218.7591
     kidsge6 |  -79.28874   45.54577    -1.74   0.082    -168.8154    10.23793
       _cons |   4198.667    1742.77     2.41   0.016     773.0052    7624.329
------------------------------------------------------------------------------
Instrumented:  lwage lwagesq agelwage
Instruments:   educ nwifeinc age kidslt6 kidsge6 exper expersq educsq agesq
               ageeduc ageexper educexper nwifeincsq
------------------------------------------------------------------------------
```

```
. * The lwage coefficient looks strange because it is the slope when age = 0.

. sum age if lwage != .

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         age |       428    41.97196    7.721084         30         60

. gen age42lwage = (age - 42)*lwage
(325 missing values generated)
```

```
. ivreg hours educ nwifeinc age kidslt6 kidsge6 (lwage lwagesq age42lwage
   = exper expersq educsq agesq ageeduc ageexper educexper nwifeincsq), robust

------------------------------------------------------------------------------
             |               Robust
       hours |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lwage |   2104.181   858.9774     2.45   0.015     415.7388    3792.622
     lwagesq |  -515.1495    347.704    -1.48   0.139    -1198.611     168.312
  age42lwage |   52.92906   25.35424     2.09   0.037     3.091702    102.7664
        educ |  -87.16321   57.42158    -1.52   0.130    -200.0335    25.70705
    nwifeinc |  -7.329747   4.456165    -1.64   0.101    -16.08897    1.429477
         age |  -72.57273   33.17705    -2.19   0.029    -137.7869   -7.358533
     kidslt6 |   -147.618   186.3906    -0.79   0.429    -513.9951    218.7591
     kidsge6 |  -79.28874   45.54577    -1.74   0.082    -168.8154    10.23793
       _cons |   4198.667    1742.77     2.41   0.016      773.005    7624.328
------------------------------------------------------------------------------
Instrumented:  lwage lwagesq age42lwage
Instruments:   educ nwifeinc age kidslt6 kidsge6 exper expersq educsq agesq
               ageeduc ageexper educexper nwifeincsq
------------------------------------------------------------------------------
```

```
. * Drop the quadratic; the turning point is pretty far out, anyway:

. di 2104.18/(2*515.15)
2.0422984

. count if lwage > 2 & lwage != .
   45

. * Only 45 observations have lwage past the turning point.
```

```
. ivreg hours educ nwifeinc age kidslt6 kidsge6 (lwage age42lwage =
   exper expersq educsq agesq ageeduc ageexper educexper nwifeincsq), robust

------------------------------------------------------------------------------
             |               Robust
       hours |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lwage |   1192.105   393.4895     3.03   0.003     418.6511    1965.559
  age42lwage |   57.21932   22.80467     2.51   0.012     12.39381    102.0448
        educ |  -137.7452   46.94343    -2.93   0.004    -230.0186    -45.4719
    nwifeinc |  -8.016031   4.569481    -1.75   0.080    -16.99793    .9658697
         age |   -78.6189   30.10629    -2.61   0.009    -137.7967   -19.44112
     kidslt6 |  -220.7256   162.3955    -1.36   0.175    -539.9348     98.4837
     kidsge6 |  -91.58226   45.43596    -2.02   0.044    -180.8925   -2.272062
       _cons |   5218.437   1308.244     3.99   0.000     2646.915    7789.959
------------------------------------------------------------------------------
Instrumented:  lwage age42lwage
Instruments:   educ nwifeinc age kidslt6 kidsge6 exper expersq educsq agesq
               ageeduc ageexper educexper nwifeincsq
------------------------------------------------------------------------------
```

```
. * Now allow slope to depend on young children:

. gen kidslt6lwage = kidslt6*lwage
. gen kidslt6age = kidslt6*age
. gen kidslt6educ = kidslt6*educ
. gen kidslt6exper = kidslt6*exper
```

```
. ivreg hours educ nwifeinc age kidslt6 kidsge6 (lwage age42lwage kidslt6lwage
   = exper expersq educsq agesq ageeduc ageexper ageexpersq
   educexper nwifeincsq  kidslt6age kidslt6educ kidslt6exper ), robust

------------------------------------------------------------------------------
             |               Robust
       hours |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lwage |   1131.537   344.7533     3.28   0.001     453.8754    1809.198
  age42lwage |   58.61914    25.2722     2.32   0.021      8.94304    108.2952
kidslt6lwage |   93.45248    694.575     0.13   0.893    -1271.833    1458.738
        educ |  -132.9736   42.82474    -3.11   0.002    -217.1517   -48.79551
    nwifeinc |  -7.827732   4.473806    -1.75   0.081    -16.62163    .9661676
         age |  -80.09144   32.32569    -2.48   0.014    -143.6322    -16.5507
     kidslt6 |  -333.7642   839.6618    -0.40   0.691    -1984.238     1316.71
     kidsge6 |  -94.00267   44.31191    -2.12   0.034     -181.104   -6.901314
       _cons |    5291.99   1427.653     3.71   0.000     2485.736    8098.244
------------------------------------------------------------------------------
Instrumented:  lwage age42lwage kidslt6lwage
Instruments:   educ nwifeinc age kidslt6 kidsge6 exper expersq educsq agesq
               ageeduc ageexper ageexpersq educexper nwifeincsq kidslt6age
               kidslt6educ kidslt6exper
------------------------------------------------------------------------------

. * Not much interactive effect with young children.

. * Now system estimation for the age-lwage interactive model. Should have
. * standard errors robust to system heteroskedasticity.
```

```
. reg3 (hours lwage age42lwage educ nwifeinc age kidslt6 kidsge6)
   (lwage educ exper expersq) (age42lwage educ nwifeinc age kidslt6 kidsge6
   exper expersq educsq agesq ageeduc ageexper ageexpersq educexper nwifeincsq)

Three-stage least-squares regression
----------------------------------------------------------------------


------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
hours        |
       lwage |   1619.417    280.716     5.77   0.000     1069.224     2169.61
  age42lwage |   41.47732    19.0728     2.17   0.030     4.095314    78.85933
        educ |  -192.5874    37.4431    -5.14   0.000    -265.9746   -119.2003
    nwifeinc |  -2.539116   3.995325    -0.64   0.525    -10.36981    5.291577
         age |  -64.94568   23.24058    -2.79   0.005    -110.4964   -19.39498
     kidslt6 |  -258.2699   106.2244    -2.43   0.015     -466.466    -50.0738
     kidsge6 |  -86.67123   33.38772    -2.60   0.009    -152.1099    -21.2325
       _cons |   4729.413   1042.356     4.54   0.000     2686.433    6772.393
-------------+----------------------------------------------------------------
lwage        |
        educ |   .1077954    .014065     7.66   0.000     .0802286    .1353623
       exper |   .0396293   .0101335     3.91   0.000      .019768    .0594906
     expersq |  -.0007131   .0002934    -2.43   0.015    -.0012881    -.000138
       _cons |  -.5236813   .1925401    -2.72   0.007    -.9010531   -.1463096
-------------+----------------------------------------------------------------
```

```
age42lwage   |
        educ |  -2.513084   1.110369    -2.26   0.024    -4.689368   -.3368002
    nwifeinc |   .0353513   .0589488     0.60   0.549    -.0801862    .1508888
         age |  -.3865226    .475865    -0.81   0.417    -1.319201    .5461558
     kidslt6 |   .2503764   .6957768     0.36   0.719    -1.113321    1.614074
     kidsge6 |   .1566072   .2251101     0.70   0.487    -.2846004    .5978149
       exper |  -1.433462   .5798369    -2.47   0.013    -2.569922   -.2970031
     expersq |   .0536639   .0213064     2.52   0.012     .0119041    .0954237
      educsq |  -.0719416   .0304118    -2.37   0.018    -.1315477   -.0123356
       agesq |   .0031755   .0049889     0.64   0.524    -.0066025    .0129535
     ageeduc |   .0952531   .0164616     5.79   0.000     .0629889    .1275173
    ageexper |    .017147   .0123111     1.39   0.164    -.0069823    .0412763
  ageexpersq |  -.0007993   .0004201    -1.90   0.057    -.0016227    .0000241
   educexper |   .0151728   .0149373     1.02   0.310    -.0141037    .0444494
  nwifeincsq |  -.0002933   .0008569    -0.34   0.732    -.0019729    .0013862
       _cons |   5.541071   13.86928     0.40   0.690    -21.64222    32.72436
------------------------------------------------------------------------------
Endogenous variables:  hours lwage age42lwage
Exogenous variables:   educ nwifeinc age kidslt6 kidsge6 exper expersq educsq
     agesq ageeduc ageexper ageexpersq educexper nwifeincsq
------------------------------------------------------------------------------
```

5. DIFFERENT INSTRUMENTS FOR DIFFERENT EQUATIONS

- We discussed earlier how traditional 3SLS maintains that a variable exogenous in any equation is exogenous in all equations. Can easily relax this by using GMM approach.
- Identification is generally more difficult, but in practice, it is usually straightforward.

$$
\begin{aligned}
hours &= \gamma_{12} lwage + \delta_{10} + \delta_{11} educ + \delta_{12} othinc + \beta_{13} kids + u_1 \\
lwage &= \gamma_{23} educ + \delta_{20} + \delta_{21} exper + \delta_{22} exper^2 + u_2 \\
educ &= \delta_{30} + \delta_{31} exper + \delta_{32} exper^2 + \delta_{33} othinc + \delta_{34} kids \\
&\quad + \delta_{35} motheduc + \delta_{36} fatheduc + u_3
\end{aligned}
$$

• Assume *educ* is uncorrelated with u_1 but might be correlated with u_2. The available IVs for the first equation are

$$(1, educ, othinc, kids, exper, exper^2, motheduc, fatheduc)$$

The IVs for the second and third equations are

$$(1, othinc, kids, exper, exper^2, motheduc, fatheduc)$$

• No difficulty provided we can specify the IV matrix to reflect these choices and then use GMM (whether 3SLS or more general weighting matrix).