Comparative Effectiveness Research Methods Training

Module 5: IV Theory & Application

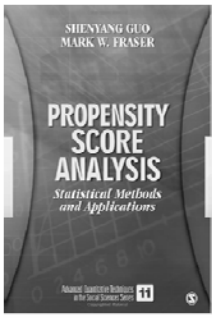
J. Michael Oakes, PhD
Associate Professor
Division of Epidemiology
University of Minnesota
oakes007@umn.edu

CCTS | Center for Clinical & Translational Science

A little (more) about me.

CCTS | Center for Clinical & Translational Science

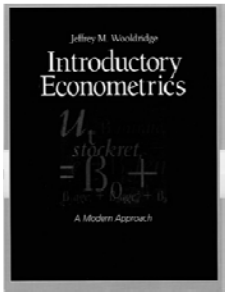
CCTS | Center for Clinical & Translational Science



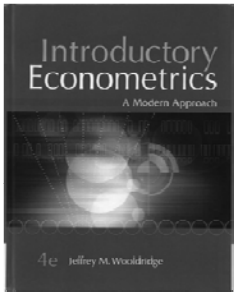
(2010)

(2007)

CCTS | Center for Clinical & Translational Science

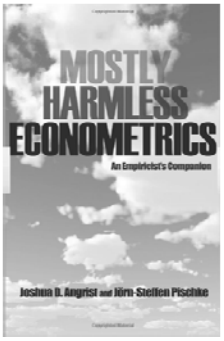


(2000)

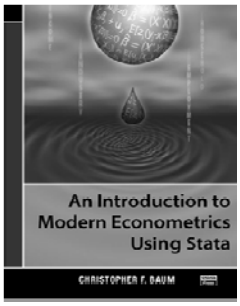


(2009)

CCTS | Center for Clinical & Translational Science



(2008)



(2006)

CCTS | Center for Clinical & Translational Science

Module #5 Outline

1. Review of Core Ideas
2. The Problem
3. IV Model Intuition
4. Examples
5. LATE
6. Worked Example
7. Issues & Assumptions
8. Review
9. Questions

CCTS | Center for Clinical & Translational Science

Caveat lector

CCTS | Center for Clinical & Translational Science

1. Review Core Ideas

CCTS | Center for Clinical & Translational Science

Analysis of Experimental Data

$$Y = \alpha + \beta_1 T + \varepsilon$$

$$\hat{\beta}_1 \Rightarrow \bar{\Delta} = \text{ average causal effect}$$

T is (0,1) treatment indicator which, for large samples, is independent of background characteristics by study design (ie, randomization)

CCTS | Center for Clinical & Translational Science

Absent Randomization

$$Y = \alpha + \beta_1 T + \boldsymbol{\beta}\mathbf{Z} + \varepsilon$$

Covariates, ***Z***, serve to
<u>adjust groups for confounding…</u>

CCTS | Center for Clinical & Translational Science

Absent Randomization

Unless <u>specification of the model, including Z, is perfect</u>, bias results

$$\beta_1 = \bar{\Delta} + \text{BIAS}$$

CCTS | Center for Clinical & Translational Science

2. The Problem

CCTS | Center for Clinical & Translational Science

Compared to other treatments for a hearth attack, does cardiac catheritization reduce mortality?

CCTS | Center for Clinical & Translational Science

If RCT

$$\text{Mortality} = \alpha + \beta_1 T + \varepsilon$$

If No RCT

$$\text{Mortality} = \alpha + \beta_1 T + \boldsymbol{\beta}\mathbf{Z} + \varepsilon$$

CCTS | Center for Clinical & Translational Science

Again, what goes in **Z**?

Patient's age, medical history, time to ED…
Cardiologist's skill, tools available…
Weather, politics, ????

$$\text{Mortality} = \alpha + \beta_1 T + \boldsymbol{\beta}\mathbf{Z} + \varepsilon$$

CCTS | Center for Clinical & Translational Science

Got **Z**?

If you can confidently say you measured (accurately!) <u>all the relevant confounding variables</u> you could claim conditional independence and conduct a propensity score analysis or employ a multiple regression model.

CCTS | Center for Clinical & Translational Science

Really got **Z**?

What of <u>unmeasured or unmeasurable</u> variables?

Doctor's fatigue, Docs ability, Unexpected delays, Hospital reimbursement policy, Hospital equipment, ED culture, Patient's choice…????

CCTS | Center for Clinical & Translational Science

$$\text{Mortality} = \alpha + \beta_1 T + \boldsymbol{\beta}\mathbf{Z} + \varepsilon$$

If we don't (or can't!) get everything in Z then there will be a correlation between T and the error term. When this happens the estimated treatment effect is biased.

If no CIA, then

$$E(\varepsilon \mid T) \neq 0$$

$$\beta_1 = \bar{\Delta} + \text{BIAS}$$

CCTS | Center for Clinical & Translational Science

More generally…

$$\text{Y} = \alpha + \beta_1 T + \boldsymbol{\beta}\mathbf{Z} + \varepsilon$$

If, even after adjustment for Z, there is a correlation between the treatment indicator, T, and the error term, we say that the treatment is <u>ENDOGENOUS</u>.

In the jargon of econometrics, T would be considered an <u>endogenous regressor</u>.

This is <u>bad</u> since to estimate valid treatment effects regression models require <u>EXOGENOUS</u> regressors.

CCTS | Center for Clinical & Translational Science

Short of an RCT
what can be done?

CCTS | Center for Clinical & Translational Science

3. IV Model

(Intuition)

CCTS | Center for Clinical & Translational Science

IV models?

The instrumental variables technique has been appreciated for 50+ years. It's a staple of every econometrics class and textbook.

The technique seems to be *de rigueur* in economics.

Not well understood or adopted in epidemiology, where controlled experiments is also difficult. I say this is so because few epidemiologists believe in human choice.

CCTS | Center for Clinical & Translational Science

Simply…

The IV technique fixes the problem of endogenous regressors (ie, omitted variables!)

It uses other exogenous regressors (ie, instrumental variables) to remove the endogenous part of the offending variable, which is then used in a "regular" analysis.

It's helpful to conceive of the approach as a two-step process: (1) fix problem regressor, (2) use fixed regressor in desired model.

CCTS | Center for Clinical & Translational Science

The Good Ol' Two Step

If **Y = a + BT + GZ + e** and we miss some variable in Z that makes T correlated with e (bad news).

IV solution

1. Find an IV (predict T but not Y, unless thru T)
2. **T = a + B(IV) + e** (Regression #1, fix T)
3. Get predicted value of T, called T*
4. Fit **Y = a + BT* + GZ + e** (Regression #2, get better B)

CCTS | Center for Clinical & Translational Science

James Heckman

Schultz Distinguished Service Prof of Economics at U of Chicago, where he has served since 1973. In 2000, he won Nobel Prize in Economic Sciences.

CCTS | Center for Clinical & Translational Science

IV Models

Can overcome problems of

- Omitted variables (often results in endogenous regressor)
- Measurement error in regressors
- Reverse causality

Can not only assess discrete Tx effects, but can also model the less biased effects of other predictors on outcome. Experiments and propensity score models cannot do this.

CCTS | Center for Clinical & Translational Science

Reverse Causation

CCTS | Center for Clinical & Translational Science

Imagine a XS study to estimate the effect of public health spending on incidence of some disease. We might consider equation

Disease Rates = a + B(Spending) + e

We hope that B is less than zero: as spending goes up, disease goes down. But it may be spending goes up because disease has gone up. Disease rates cause spending, which is reverse causation in this context. Regular regression estimates will always be biased no matter how many controls are added.

We need to model the causes of spending in order to estimate effect of spending on disease rates. An IV regression could be used.

CCTS | Center for Clinical & Translational Science

Derived from Prof. Kit Baum

What is an IV?

Intuitively, instruments are variables that move around the probability of treatment but do not affect outcomes except through their effect on treatments.

Put more statistically, instruments are variables that are correlated with the endogenous variable – in this context the treatment indicator – but not correlated with the unobservable in the outcome equation.

CCTS | Center for Clinical & Translational Science

What is an IV?

One can think of the instrumental variable as a device that achieves a pseudorandomization.

Indeed, the actual randomization in an RCT is a special case of IV. Imagine, for example, that one tosses an unbiased coin to assign people to treatment or control groups at random. The outcome of the coin toss, heads or tails, is the IV -- a variable that induces variation in the treatment variable

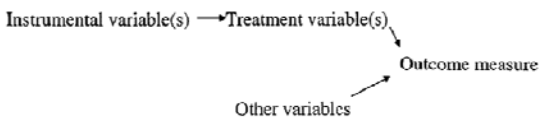


Figure 1 Schematic of Instrumental Variable Estimation.

Newhouse & McClellan 1998

Parts is Parts

- For valid estimate of B in Y= a + BT + e, where T is correlated with e.
- The variation in T can be divided into two parts:
  - Part correlated with e (bad)
  - Part uncorrelated with e (good, just like RCT)
- Want to use the second part, which is what IV does.

What makes a good IV?

Imagine $Y = a + BT + e$ with potential instrument "Z"

Z is a good instrument for endogenous T if Z is

(1) Exogenous: uncorrelated with the errors
(2) Correlated with the endogenous T (predicts T)
(3) Correlated with the Y only through T

Some say "A good instrument should not be correlated with the dependent variable". This is incorrect. Z has to be correlated with Y, otherwise it is useless as an instrument. But it must be correlated with Y only through T.

CCTS | Center for Clinical & Translational Science

What makes a good IV?

Natural Experiments are when mother nature randomizes persons to one condition or another are a useful form of exogenous variation and thus can serve as excellent IVs.

Moreover, the effects of natural experiments should probably be estimated through IV methods.

CCTS | Center for Clinical & Translational Science

IVs are not just another covariate

Assume that Z is both a valid IV and a valid covariate of a True regression model.

If Z is a covariate, the true regression model would be given by $y = X + Z + e$. The estimated model, however, is given by $y = X+e$; Z is excluded from the regression equation (as an IV).

So the estimated error is $e = Z + e$. This error term, e, is obviously correlated with the IV, Z. Therefore, the assumption that a valid IV is also a valid covariate the model leads to a contradiction. Thus, an IV is not a valid covariate in the model.

CCTS | Center for Clinical & Translational Science

4. Examples

Stylized for pedagogy

CCTS | Center for Clinical & Translational Science

I. Catheterization (CATH) after AMI on mortality?

- t-test shows CATH reduced four-year mortality by 37%.
- But no RCT: no exchangeability or balance
- Multiple regression (age, sex, race…) shows 28% mortality reduction due to CATH.
- Use distance to nearest hospital as IV
  - Correlated with CATH but not otherwise to mortality
- A simple IV estimator showed just 7% decrease in mortality.
- IV with more controls (high-volume hospital, rural area, etc) reduce effect of CATH on mortality to 5%.
- CATH matters but by far less than naïve methods suggested
- 37% is far different from 5%

CCTS | Center for Clinical & Translational Science

From McClellan & Newhouse 1998

II. Compulsory school attendance effect income?

Income = a + b_1(educ_attain) + b_2(covariates) + e

- But what of ability and/or motivation? How to measure it?
- It's important because <u>ability affects educational attainment</u> and thus leaving it out of our model creates a correlation between *educ_attain* and *e* and this biases the effect of interest, b1.
- An omitted variable problem.

CCTS | Center for Clinical & Translational Science

From class notes of Prof Haley Fisher citing Angrist & Krueger (1991)

What instrument? Month born!

Child can enter kindergarten if 5 years old by September 1. Thus, children born on August 30 can enter kindergarten but their friend born on Sept 2 cannot. The first child will be young for her grade; the second child will be old for his grade.

Further, most states prevent youth from dropping out of school until they are 16 years old. This creates a natural experiment where children with arbitrary birthdays can drop out when they are in different grades.

Thus, month of birth is correlated to educational attainment (or drop out grade) at 16 but there is no reason to think it is related to income.

CCTS | Center for Clinical & Translational Science

Income = a + b_1(educ_attain) + b_2(covariates) + e

- Education is endogenous (correlated with error due to missing measures of "ability" and stuff) but we will use birth month as instrument to break the correlation and get a better estimate of the effect of education on income.
- Regular regression show educ_attain increases income by 7%
- IV methods
  - Step 1: educ_attain = birth_month + sex + race + hhinc + e
  - Step 2: income = educ_attain* + e
- The IV estimate shows a statistically insignificant effect of education on income!

CCTS | Center for Clinical & Translational Science

III. Cholera in London ~1854

CCTS | Center for Clinical & Translational Science

Recall, Snow hypothesized that cholera was waterborne infection.

But he could not examine water purity and its correlation with the incidence of cholera

Generally, those who drank “bad” water were more likely to be poor, to live in crowded tenements and to live in an environment contaminated in many ways. Snow intuitively appreciated selection bias.

Oh, and he couldn’t see the bacteria in the water anyway. Talk about an unobservable!

CCTS | Center for Clinical & Translational Science

Based on Prof. Kit Baum’s PPT

What could serve as an instrument for the unobservables related to the selection bias?

The water company a household used.

Water company was correlated with exposure to water (good, bad) and cholera but only through the (theorized) biological mechanisms at work. (Snow could not see the bacteria!)

Water company was unlikely to be correlated with other factors influencing cholera (such as the health status of those living in certain neighborhoods) given that the suppliers competed throughout the city.

Snow conceptually employed IV method and determined something in the water caused cholera!

CCTS | Center for Clinical & Translational Science

Although econometricians believe that instrumental variable methods were the product of Sewall Wright’s analysis of agricultural supply and demand in the 1920s, or the work of the Cowles Commission in the 1950s, the method seems to have been discovered by John Snow, <u>the first modern epidemiologist!</u>

CCTS | Center for Clinical & Translational Science

Based on Prof. Kit Baum’s PPT

5. LATE

CCTS | Center for Clinical & Translational Science

Recall, ideally, we'd like the treatment effect for each individual in our study. If we could observe every person and their counterfactual we could just take the average across all persons as an estimate of delta.

$$\tau_i = Y_i(1) - Y_i(0)$$

$$\overline{\tau} \Rightarrow \Delta$$

CCTS | Center for Clinical & Translational Science

But of course we cannot calculate a causal effect for a particular person.

We must move up the population level.

CCTS | Center for Clinical & Translational Science

Recall ATE:

The **average treatment effect** (**ATE**) is the difference in the average of the outcome variable in the treatment group minus the average of the outcome variable in the control group. ATE is the same as the **average causal effect (ACE).**

$$\text{ATE} = \text{ACE} = E[Y(1) - Y(0)]$$

CCTS | Center for Clinical & Translational Science

Recall ATT:

The **average treatment effect on the treated (ATT)** is the mean difference between those actually treated or exposed and their counterfactuals. ATT is the same as the **treatment effect on the treated (TOT)**.

$$\text{ATT} = \text{TOT} = E[Y(1) - Y(0) \mid \text{T=1}]$$

CCTS | Center for Clinical & Translational Science

LATE

Local Average Treatment Effect

The average treatment effect for individuals "who can be induced to change their treatment by a change in the instrument". For example, those who comply with treatment assignment in an RCT.

LATE is the average causal effect of X on Y for "compliers," as opposed to "always takers" or "never takers". LATE does not incorporate those who refuse to be in the "RCT" or will take the "pill" regardless of assignment.

CCTS | Center for Clinical & Translational Science

Who are the “marginal” patients (or “compliers”), whose treatment is effected by the instrument?

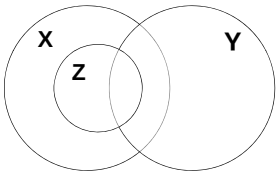
IV estimates treatment effect among these “marginal” patients, which is why the term “local” is used.

LATE rarely a good estimate of the treatment effect in the general population.

IV methods estimate average treatment effects, with the average depending on the instruments.

CCTS | Center for Clinical & Translational Science

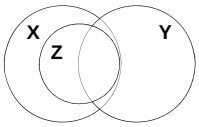
In IV, not all of the available variation in X is used. **Only that portion of X which is “explained” by Z is used to explain Y**



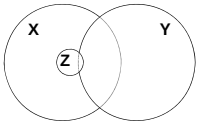
X = Endogenous variable
Y = Response variable
Z = Instrumental variable

CCTS | Center for Clinical & Translational Science

Based on Prof. Robert Apel’s PPT



Best-case scenario: A lot of X is explained by Z, and most of the overlap between X and Y is accounted for



Realistic scenario: Very little of X is explained by Z, or what is explained does not overlap much with Y

CCTS | Center for Clinical & Translational Science

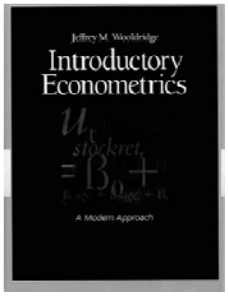
6. Worked Example

CCTS | Center for Clinical & Translational Science

Let's use Stata

What is Stata? Stata is a full-featured statistical programming language for Windows, Macintosh, Unix and Linux. It can be considered a "stat package," like SAS, SPSS, or R.

CCTS | Center for Clinical & Translational Science



http://fmwww.bc.edu/gstat/examples/wooldridge/wooldridge15.html

CCTS | Center for Clinical & Translational Science

```
. use http://fmwww.bc.edu/ec-p/data/wooldridge/MROZ

. d

Contains data from http://fmwww.bc.edu/ec-p/data/wooldridge/MROZ.dta
  obs:           753
 vars:            22                          10 Jan 2000 16:54
 size:        72,288 (99.9% of memory free)
-------------------------------------------------------------------------------
              storage  display     value
variable name   type   format      label      variable label
-------------------------------------------------------------------------------
inlf            float  %9.0g
hours           float  %9.0g
kidslt6         float  %9.0g
kidsge6         float  %9.0g
age             float  %9.0g
educ            float  %9.0g
wage            float  %9.0g
repwage         float  %9.0g
hushrs          float  %9.0g
husage          float  %9.0g
huseduc         float  %9.0g
huswage         float  %9.0g
faminc          float  %9.0g
mtr             float  %9.0g
motheduc        float  %9.0g
fatheduc        float  %9.0g
unem            float  %9.0g
city            float  %9.0g
exper           float  %9.0g
nwifeinc        float  %9.0g
lwage           float  %9.0g
expersq         float  %9.0g
-------------------------------------------------------------------------------
Sorted by:
```

CCTS | Center for Clinical & Translational Science

Example 15.1: Estimating the Return to Education for Married Women

$$\log(wage) = \beta_0 + \beta_1 educ + \upsilon$$

CCTS | Center for Clinical & Translational Science

Example 15.1: Estimating the Return to Education for Married Women

$$\log(wage) = \beta_0 + \beta_1 educ + \upsilon$$

The variable "ability" is missing

CCTS | Center for Clinical & Translational Science

First, the <u>incorrect</u> regular regression model

```
. reg lwage educ

      Source |       SS       df       MS              Number of obs =     428
-------------+------------------------------           F(  1,   426) =   56.93
       Model |  26.3264237     1  26.3264237           Prob > F      =  0.0000
    Residual |  197.001028   426  .462443727           R-squared     =  0.1179
-------------+------------------------------           Adj R-squared =  0.1158
       Total |  223.327451   427  .523015108           Root MSE      =  .68003

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .1086487   .0143998     7.55   0.000     .0803451    .1369523
       _cons |  -.1851969   .1852759    -1.00   0.318    -.5492674    .1788735
------------------------------------------------------------------------------
```

CCTS | Center for Clinical & Translational Science

```
. reg lwage educ

      Source |       SS       df       MS
-------------+------------------------------
       Model |  26.3264237     1  26.3264237
    Residual |  197.001028   426  .462443727
-------------+------------------------------
       Total |  223.327451   427  .523015108

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|
-------------+----------------------------------------------------------------
        educ |   .1086487   .0143998     7.55   0.000
       _cons |  -.1851969   .1852759    -1.00   0.318
```

CCTS | Center for Clinical & Translational Science

Now, the IV regression using father's education as an IV for the unmeasured variable "ability"

```
. ivreg lwage (educ = fatheduc )

Instrumental variables (2SLS) regression

      Source |       SS       df       MS              Number of obs =     428
-------------+------------------------------           F(  1,   426) =    2.84
       Model |  20.8673618     1  20.8673618           Prob > F      =  0.0929
    Residual |  202.460089   426  .475258426           R-squared     =  0.0934
-------------+------------------------------           Adj R-squared =  0.0913
       Total |  223.327451   427  .523015108           Root MSE      =  .68939

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .0591735   .0351418     1.68   0.093    -.0098994    .1282463
       _cons |   .4411035   .4461018     0.99   0.323    -.4357311    1.317938
------------------------------------------------------------------------------
Instrumented:  educ
Instruments:   fatheduc
------------------------------------------------------------------------------
```

CCTS | Center for Clinical & Translational Science

```
. ivreg lwage (educ = fatheduc )

Instrumental variables (2SLS) regression

      Source |       SS       df       MS
-------------+------------------------------
       Model |  20.8673618     1  20.8673618
    Residual |  202.460089   426  .475258426
-------------+------------------------------
       Total |  223.327451   427  .523015108
```

CCTS | Center for Clinical & Translational Science

```
. ivreg lwage (educ = fatheduc )

Instrumental variables (2SLS) regression

      Source |       SS       df       MS
-------------+------------------------------
       Model |  20.8673618     1  20.8673618
    Residual |  202.460089   426  .475258426
-------------+------------------------------
       Total |  223.327451   427  .523015108

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|
-------------+----------------------------------------------------------------
        educ |   .0591735   .0351418     1.68   0.093
        cons |   .4411035   .4461018     0.99   0.323
------------------------------------------------------------------------------
Instrumented:  educ
Instruments:   fatheduc
------------------------------------------------------------------------------
```

CCTS | Center for Clinical & Translational Science

IV regression using father's & mother's education as an IV for the unmeasured variable "ability", and add experience and experience-squared

```
. ivreg lwage (educ = motheduc fatheduc) exper expersq

Instrumental variables (2SLS) regression

      Source |       SS       df       MS              Number of obs =     428
-------------+------------------------------           F(  3,   424) =    8.14
       Model |  30.3074295     3  10.1024765           Prob > F      =  0.0000
    Residual |  193.020022   424   .455359             R-squared     =  0.1357
-------------+------------------------------           Adj R-squared =  0.1296
       Total |  223.327451   427  .523015108           Root MSE      =  .67471

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .0613966   .0314367     1.95   0.051     .0003945    .1231878
       exper |   .0443704   .0114375     3.79   0.000     .0217679    .0669729
     expersq |   -.000899   .0004017    -2.24   0.026    -.0016885   -.0001094
       _cons |   .0481003   .4001281     0.12   0.904    -.7187745    .8149751
------------------------------------------------------------------------------
Instrumented:  educ
Instruments:   exper expersq motheduc fatheduc
------------------------------------------------------------------------------
```

CCTS | Center for Clinical & Translational Science

```
. ivreg lwage (educ = motheduc fatheduc) exper expersq
Instrumental variables (2SLS) regression

      Source |       SS       df       MS              Number of obs =     428
-------------+------------------------------           F(  3,   424) =
       Model |  30.3074295     3  10.1024765           Prob > F      =
    Residual |  193.020022   424    .4552359           R-squared     =
-------------+------------------------------           Adj R-squared =
       Total |  223.327451   427  .523015108           Root MSE      =

------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf.
------------------------------------------------------------------------
```

CCTS | Center for Clinical & Translational Science

```
------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|
-------------+----------------------------------------------------------
        educ |   .0613966   .0314367     1.95   0.051
       exper |   .0441704   .0134325     3.29   0.001
     expersq |   -.000899   .0004017    -2.24   0.026
       _cons |   .0481001   .4001281     0.12   0.904
------------------------------------------------------------------------
Instrumented:  educ
Instruments:   exper expersq motheduc fatheduc
------------------------------------------------------------------------
```

CCTS | Center for Clinical & Translational Science

Add some diagnostics, by using "ivreg2"

You need to install ivreg2 and it's friend ranktest.

In Stata type

. ssc install ivreg2
. ssc install ranktest

CCTS | Center for Clinical & Translational Science

```
. ivreg2 lwage (educ = motheduc fatheduc) exper expersq

IV (2SLS) estimation

Estimates efficient for homoskedasticity only
Statistics consistent for homoskedasticity only

                                                      Number of obs =     428
                                                      F(  3,   424) =    8.24
                                                      Prob > F      =  0.0000
Total (centered) SS     =  223.3274513                Centered R2   =  0.1357
Total (uncentered) SS   =   829.594813                Uncentered R2 =  0.7671
Residual SS             =  193.0200218                Root MSE      =  .6716

       lwage |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .061966   .0317895     1.95   0.050     .0000704    .1227228
       exper |  .0441704   .0133696     3.30   0.001     .0179665    .0703743
     expersq |  -.000899   .0003998    -2.25   0.025    -.0016826   -.0001154
       _cons |  .0481003    .398453     0.12   0.904    -.7328512    .8290538
-------------+----------------------------------------------------------------
Underidentification test (Anderson canon. corr. LM statistic):          88.840
                                                   Chi-sq(2) P-val =    0.0000
-------------------------------------------------------------------------------
Weak identification test (Cragg-Donald Wald F statistic):               55.400
Stock-Yogo weak ID test critical values: 10% maximal IV size            19.93
                                         15% maximal IV size            11.59
                                         20% maximal IV size             8.75
                                         25% maximal IV size             7.25
Source:  Stock-Yogo (2005).  Reproduced by permission.
-------------------------------------------------------------------------------
Sargan statistic (overidentification test of all instruments):           0.378
                                                 Chi-sq(1) P-val =       0.5386
-------------------------------------------------------------------------------
Instrumented:         educ
Included instruments: exper expersq
Excluded instruments: motheduc fatheduc
```

CCTS | Center for Clinical & Translational Science


```
. ivreg2 lwage (educ = motheduc fatheduc) exper expersq

IV (2SLS) estimation

Estimates efficient for homoskedasticity only
Statistics consistent for homoskedasticity only

                                                      Number of o
                                                      F(  3,   4
                                                      Prob > F
Total (centered) SS     =  223.3274513                Centered R2
Total (uncentered) SS   =   829.594813                Uncentered
Residual SS             =  193.0200218                Root MSE

       lwage |      Coef.   Std. Err.      z    P>|z|     [95% co
```

CCTS | Center for Clinical & Translational Science


```
Underidentification test (Anderson canon. corr. LM statistic):          88.840
                                                   Chi-sq(2) P-val =    0.0000
-------------------------------------------------------------------------------
Weak ide[illegible]tion test (Cragg-Donald Wald F statistic):           55.400
Stock-Y[illegible]o test critical values: 10% maximal IV size            19.93
                                          15% maximal IV size            11.59
                                          20% maximal IV size             8.75
                                          25% maximal IV size             7.25
Source: S[illegible]Yogo (2005).  Reproduced by permission.
-------------------------------------------------------------------------------
Sargan st[illegible]ic (overidentification test of all instruments):     0.378
                                                 Chi-sq(1) P-val =       0.5386
-------------------------------------------------------------------------------
```

Test of whether IVs are correlated with endogenous variable (ie, treatment).

Small p-value indicates good correlation.

CCTS | Center for Clinical & Translational Science

```
underidentification test (Anderson canon. corr. LM statistic):          88.840
                                                   Chi-sq(2) P-val =    0.0000
------------------------------------------------------------------------------
weak identification test (Cragg-Donald Wald F statistic):               55.400
Stock-Yogo weak ID test critical values: 10% maximal IV size             19.93
                                         15% maximal IV size             11.59
                                         20% maximal IV size              8.75
                                         25% maximal IV size              7.25
Source: Stock-Yogo (2005).  Reproduced by permission.
------------------------------------------------------------------------------
Sargan statistic (overidentification test of all instruments):           0.378
                                                   Chi-sq(1) P-val =    0.5386
------------------------------------------------------------------------------
```

A test of how biased the IV estimates are relative to OLS model: in other words, how strong is your IV?

Some dispute on interpretation in diff applications.

CCTS | Center for Clinical & Translational Science


```
underidentification test (Anderson canon. corr. LM statistic):          88.840
                                                   Chi-sq(2) P-val =    0.0000
------------------------------------------------------------------------------
weak identification test (Cragg-Donald Wald F statistic):               55.400
Stock-Yogo weak ID test critical values: 10% maximal IV size             19.93
                                         15% maximal IV size             11.59
                                         20% maximal IV size              8.75
                                         25% maximal IV size              7.25
Source: Stock-Yogo (2005).  Reproduced by permission.
------------------------------------------------------------------------------
Sargan statistic (overidentification test of all instruments):           0.378
                                                   Chi-sq(1) P-val =    0.5386
------------------------------------------------------------------------------
```


Test of whether IVs are valid instruments (ie, not correlated with error term).

Small p-value is bad.

CCTS | Center for Clinical & Translational Science

7. Issues & Assumptions

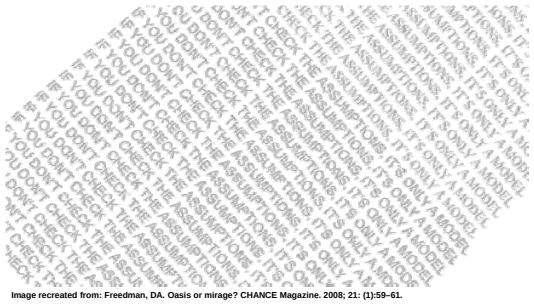


Image recreated from: Freedman, DA. Oasis or mirage? CHANCE Magazine. 2008; 21: (1):59–61.

CCTS | Center for Clinical & Translational Science

The Good & Bad

Good: IV methods can account for unmeasured factors correlated with the outcome

Good: IV methods also permit analysis of other factors that explain outcome besides treatment itself

Bad: IVs are hard to find and defend

Bad: IV standard errors tend to be large, especially when corr(x,z) is very small, which can lead to type II errors

CCTS | Center for Clinical & Translational Science

Must have valid IV

Consistent instrumental variables (IV) estimation requires instruments which are valid, which is to say, uncorrelated with the error term in the regression equation. In practice, however, this condition is unlikely to be satisfied.

Moreover, this assumption is virtually impossible to test since the relevant error term is not directly observable.

Consequently, the validity of IV-based parameter inference largely rests on a statistical assumption which is both suspect and often untestable.

CCTS | Center for Clinical & Translational Science

Cannot have weak IV

IV estimates using instruments which are weak (i.e., only weakly correlated with the endogenous variables) are known to yield unreliable parameter inference even when these instruments are valid, in the sense of being asymptotically uncorrelated with the model error term.

CCTS | Center for Clinical & Translational Science

Small Sample Properties

IV estimators can behave badly in small samples.

CCTS | Center for Clinical & Translational Science

Good Practice with IV Models

- TELL A STORY about why a particular IV is a “good instrument”
- Something to consider when thinking about whether a particular IV is “good”
  - Does the IV, for all intents and purposes, randomize the endogenous regressor?

CCTS | Center for Clinical & Translational Science

Based on Prof. Robert Apel’s PPT

IV > RCT?

In principle, RCTs can be designed to answer any health or policy question. In practice, however, random assignment experiments have important limitations.

- Not everything of interest can or should be randomized
- Participation varies and attrition is a big problem
- RCTs black-box mechanisms

- IV techniques do not require randomization
- Varying participation rates can be studied and explained
- Mechanisms of effects can be investigated

CCTS | Center for Clinical & Translational Science

8. Review

I. Review of Core Ideas
   a. Causal inference
   b. Counterfactuals
   c. Experimental Ideal

II. Problem
   a. Omitted variables
   b. Endogeneity

III. IV Model Intuition
   a. IV is correlated with Tx but not error term
   b. Mother nature randomizes

IV. Examples
   I. Catheritization
   II. Compulsory Schooling
   III. Cholera

V. LATE
   a. Effect on "compliers"
   b. Depends on IV

VI. Worked Example in Stata

IV. Assumptions & Issues
   a. Valid IV
   b. Hard to find IV
   c. Lots of data
   d. IV > RCT

CCTS | Center for Clinical & Translational Science

9. Questions

1. What problem(s) do IV methods address?
2. Name a few IVs that have been successfully used
3. How does IV differ from RCTs and Propensity Score methods?
4. Name two important limitation of IV methods

CCTS | Center for Clinical & Translational Science

References & Resources

1. Newhouse, JP, McClellan, M. Econometrics in outcomes research: The use of instrumental variables. Annual Review of Public Health. 1998; 19:17–34.
2. Frerichs RR. John Snow Site. http://www.ph.ucla.edu/epi/snow/highressnowmap.html. Updated October 4, 2004. Accessed December 14, 2011.
3. Slide 26 uses a picture of James Heckman from http://www.flickr.com/photos/institutoayrtonsenna/ (Creative Commons)
4. Freedman, DA. Oasis or mirage? CHANCE Magazine. 2008; 21: (1):59–61.

CCTS | Center for Clinical & Translational Science