ECO375 Tutorial 9
2SLS Applications and Endogeneity Tests

Matt Tudball

University of Toronto Mississauga

November 23, 2017

Hausman Test: Motivation

- Suppose that we want to estimate the simple regression

$$y = \beta_0 + \beta_1 x_1 + u$$

- We know that OLS is BLUE (Gauss-Markov theorem).
- Unless there is a compelling reason, we should use OLS (not IV).
- If we have a suspicious OLS estimator due to potential endogeneity and a plausible 2SLS estimator, we can use a Hausman test to test the consistency of OLS.

Hausman Test

- In general, we can implement a **Hausman test** when we have two estimators (of the entire parameter vector) such that:

	under H_0	under H_1
$\hat{\beta}_1$	Consistent, efficient	Inconsistent
$\hat{\beta}_2$	Consistent, less efficient	Consistent

where $\hat{\beta}_1$ and $\hat{\beta}_2$ are estimators of the parameter vector β.

Hausman Test for IV

- Under the null hypothesis that x_1 is exogenous (i.e. $\text{Cov}(x_1.u) = 0$) then
$$\text{plim}\hat{\beta}_1^{OLS} = \beta_1$$
$$\text{plim}\hat{\beta}_1^{2SLS} = \beta_1$$
that is, $\text{plim}\hat{\beta}_1^{OLS} = \text{plim}\hat{\beta}_1^{2SLS}$.
- Under the alternative hypothesis that x_1 is endogeneous (i.e. $\text{Cov}(x_1.u) \neq 0$) then
$$\text{plim}\hat{\beta}_1^{OLS} \neq \beta_1$$
$$\text{plim}\hat{\beta}_1^{2SLS} = \beta_1$$
- A very natural way to test the null hypothesis (and also, incidentally, whether a violation matters) is to simply compare the two estimators.

Hausman Test for IV

- Consider the test statistic

$$h = \hat{\beta}_1^{OLS} - \hat{\beta}_1^{2SLS}$$

- Under the null hypothesis, a normalised *d* statistic follows a χ^2_g distribution where g is the number of parameters involved in the test.
- To implement this test in Stata using the example from last week (where *nearc*4, *motheduc* and *fatheduc* were instruments for *educ* in a regression of *lwage* on *educ*)

```
reg lwage educ
estimates store ols
ivregress 2sls lwage (educ = nearc4 fatheduc ///
motheduc)
estimates store iv
hausman iv ols, constant sigmamore
```

Review: IV Assumptions

- Suppose that we have valid instruments $z_1, z_2, ..., z_l$ for some endogeneous variable x_1, then each of them must satisfy the following assumptions:

$$\begin{array}{lll} (1) & \text{Cov}(z_j, u) = 0 & \text{exogeneity condition} \\ (2) & \text{Cov}(z_j, x_1) \neq 0 & \text{relevance condition} \end{array}$$

  for $j = 1, ..., l$.

- In most cases, only the second assumption is testable.

Testing the Relevance Condition

- Note that the relevance condition $\text{Cov}(z_j, x_1)$ describes the covariance between two variables that we observe in our data.
- We generally want to test that the instruments are jointly relevant to x_1. We can do this by regressing x_1 on $z_1, z_2, ..., z_l$ and running an F-test of the following form:

$$x_1 = \pi_0 + \pi_1 z_1 + \pi_2 z_2 + ... + \pi_l z_l + \epsilon$$
$$H_0 : \pi_1 = \pi_2 = ... = \pi_l = 0$$

Testing the Relevance Condition: Rule of Thumb

- It is temping to evaluate this F-test using conventional significance levels (5% or 10%).
- However, as is shown in the paper Staiger and Stock (1997), weak instrument bias can be quite large even when the instruments are jointly significant at conventional levels.
- From their work, a general 'rule of thumb' has developed which suggests that the first stage F-statistic should be greater than 10.
- If you are only testing one instrument, recall that the t-statistic is equal to the square of the F-statistic, so the rule of thumb corresponds to a t-statistic of ~ 3.2.

Testing the Relevance Condition: Example

- Ex. Recall the example from last week where we wanted to estimate the model

$$lwage = \beta_0 + \beta_1 educ + u$$

and used *near4c*, *motheduc* and *fatheduc* as instruments for *educ*.

- We can run the first stage regression

$$educ = \pi_0 + \pi_1 near4c + \pi_2 motheduc + \pi_3 fatheduc + \epsilon$$

- The F-statistic from the null hypothesis $H_0 : \pi_1 = \pi_2 = \pi_3 = 0$ is 257.43, which is much larger than 10 and suggests that these three instruments are jointly very relevant.

Testing the Exogeneity Condition

- How can we test whether the exogeneity condition $\text{Cov}(z_j, u) = 0$ holds?
- Since this condition relates to the unobserved error term u it is very difficult to test this condition.
- We usually need to make a case based on intuition or economic theory.
- However, if we assume the validity of one instrument (say, z_1) then we can evaluate a second (say, z_2).
- This is called the **test of over-identifying restrictions**.
  - Use z_1 to estimate the model, recovering an estimate of $\hat{u}_1$.
  - Test whether $\text{Cov}(\hat{u}_1, z_2) = 0$.

Testing the Exogeneity Condition: Steps

1. Retrieve 2SLS residuals from the second stage equation using only z_1:

$$\hat{u}_1 = y - \hat{\beta}_0^{2SLS} - \hat{\beta}_1^{2SLS} x_1$$

2. Regress the residuals on all of the exogenous variables and the instruments

$$\hat{u}_1 = \delta_0 + \delta_1 z_1 + \delta_2 z_2 + \epsilon$$

3. Under the null hypothesis that the instruments are valid

$$R_{uz}^2 \sim \chi_q^2$$

where q is the degree of overidentification; in this case, $q = 1$.

- If the test is rejected, it suggests that at least one of the instruments is invalid.
- Note that this test requires a valid instrument z_1 to begin with. If neither z_1 nor z_2 is valid then the test will not work.

Testing the Exogeneity Condition: Implementation

- Suppose we know that *fatheduc* and *motheduc* are valid instruments and we want to test whether *nearc*4 is also valid.
- We can run this test of overidentifying restrictions in several ways.

```
ivregress 2sls lwage (educ = fatheduc motheduc)
predict u1, residuals
reg u1 fatheduc motheduc nearc4
test fatheduc motheduc nearc4
```

- Stata also has several postestimation commands:

```
ivregress 2sls lwage (educ = fatheduc motheduc ///
nearc4)
estat overid
estat endogenous
```

- The first test will run a test of over-identifying restrictions slightly different from our manual implementation (it will test whether the instruments each produce similar estimates).
- The second test will run a test of exogeneity as described above.

Dealing with Endogeneity in Practice

- Although the test of over-identifying restrictions can be useful, the conditions for it to be informative are rarely met. We must usually rely on economic intuition to justify the exogeneity condition.
- Even though IVs are difficult to justify, the results of some very influential applied economics papers rely on them.
- I will go through the justification behind the clever IVs used in some of these papers.

Natural Experiments

- A **natural experiment** uses variation in explanatory variables generated by plausibly exogenous changes in state laws, government draft mechanisms or other means.
- The main idea behind natural experiments is to find a natural phenomenon which provides conditions similar to that of a controlled experiment (i.e. randomly assigning observations into 'treatment' or 'control' groups).
- Remember that one way of thinking about IVs is as a variable which 'assigns' people into a treatment or control group. In this way, valid IVs produce the conditions of a controlled experiment in the real world.

Example 1: Angrist (1990)

Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records

By JOSHUA D. ANGRIST*

The randomly assigned risk of induction generated by the draft lottery is used to construct estimates of the effect of veteran status on civilian earnings. These estimates are not biased by the fact that certain types of men are more likely than others to service in the military. Social Security administrative records indicate that in the early 1980s, long after their service in Vietnam was ended, the earnings of white veterans were approximately 15 percent less than the earnings of comparable nonveterans. (JEL 824)

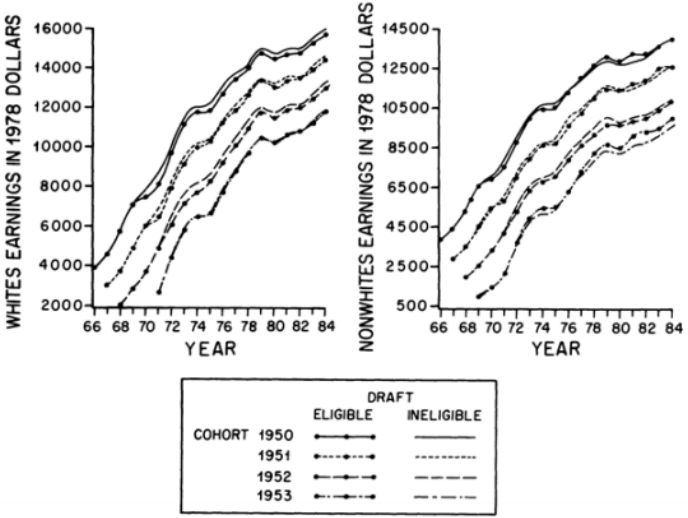
Example 1: Angrist (1990)

- The objective of Angrist (1990) is to evaluate the effect of military service on earnings.
- Sources of endogeneity:
  - Analyses which compare civilian earnings by veteran status may be biased if a non-random group of individuals serves in the military.
  - Those who enlist may have lower human capital than those who do not enlist.
  - Alternatively, military induction may screen out those individuals in worse health.
- Natural experiment:
  - Use the variation in veterans' status caused by the Vietnam-era draft lottery or the World War II draft mechanism.
  - These depended on date of birth.

Example 1: Angrist (1990)

- Y = lifetime earnings
- X = veteran status
- Z = random draft lottery to be evaluated to enter the military during to Vietnam War period.
- Is this a valid instrument?
  - **Relevance condition:** Being selected in the lottery increased the probability of enlisting in the military.
  - **Exogeneity condition:** The lottery was random and so it was independent of other variables (no direct effects on earnings).
- Angrist finds that "long after their service in Vietnam was ended, the earnings of white veterans were approximately 15 percent less than the earnings of comparable non-veterans".

Example 1: Angrist (1990)



Example 2: Acemoglu Johnson Robinson (1990)

The Colonial Origins of Comparative Development: An Empirical Investigation

By Daron Acemoglu, Simon Johnson, and James A. Robinson*

We exploit differences in European mortality rates to estimate the effect of institutions on economic performance. Europeans adopted very different colonization policies in different colonies, with different associated institutions. In places where Europeans faced high mortality rates, they could not settle and were more likely to set up extractive institutions. These institutions persisted to the present. Exploiting differences in European mortality rates as an instrument for current institutions, we estimate large effects of institutions on income per capita. Once the effect of institutions is controlled for, countries in Africa or those closer to the equator do not have lower incomes. (*JEL* O11, P16, P51)

Example 2: Acemoglu Johnson Robinson (1990)

- The authors want to understand the effect of the quality of a country's institutions (as measured by expropriation risk; the risk that the government seizes private assets) on GDP growth.
- Source of endogeneity:
  - There could be reverse causality, in that richer countries can afford better institutions (ex. better education for civil servants and politicians, better trained and equipped police force).
  - There could be omitted variables (ex. civil war, which will reduce GDP growth and also reduce the quality of a country's institutions).
- Natural experiment:
  - Look at former colonies and use early settler morality as an instrument for expropriation risk.

Example 2: Acemoglu Johnson Robinson (1990)

- Why do this?
- The authors say that their theory rests on three premises:
    - There were different types of colonisation policies which created different sets of institutions. The first are called "extractive" colonies (ex. Belgian Congo), which existed to extract resources to the coloniser. The second are called "inclusive" colonies (ex. Canada), which existed to replicate European institutions and become a home for European settlers.
    - The colonisation policy was influenced by the feasibility of settlements. Places with an abundance of tropical diseases such as Sub-Saharan Africa and South America were not suitable locations for "inclusive" colonies.
    - The colonial state and institutions persisted after independence.

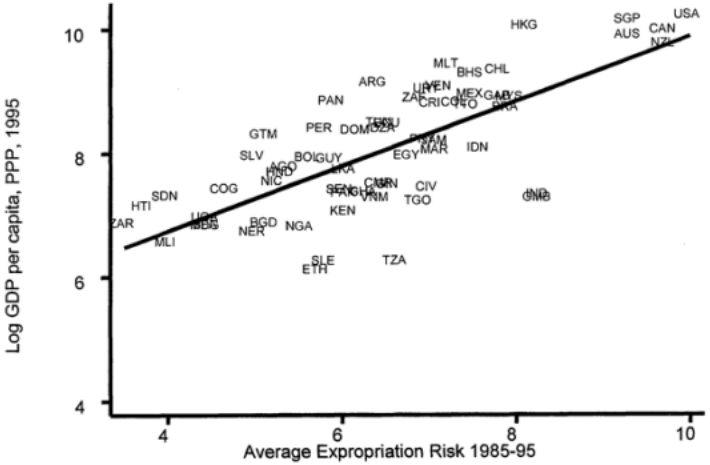


FIGURE 2. OLS RELATIONSHIP BETWEEN EXPROPRIATION RISK AND INCOME

Example 2: Acemoglu Johnson Robinson (1990)

- Y = log GDP per capita
- X = quality of a country's institutions (average expropriation risk)
- Z = log of settler mortality
  - "We use data on the mortality rates of soldiers, bishops,and sailors stationed in the colonies between the seventeenth and nineteenth centuries ....These give a good indication of the mortality rates faced by settlers. Europeans were well informed about these mortality rates at the time, even though they did not know how to control the diseases that caused these high mortality rates."
- Is this a valid instrument? Discuss.
- The authors find that the quality of a country's institutions are a significant determinant of GDP growth (and that the OLS estimates seem to be downward biased).

Example 3: Acemogly Johnson Robinson (1990)

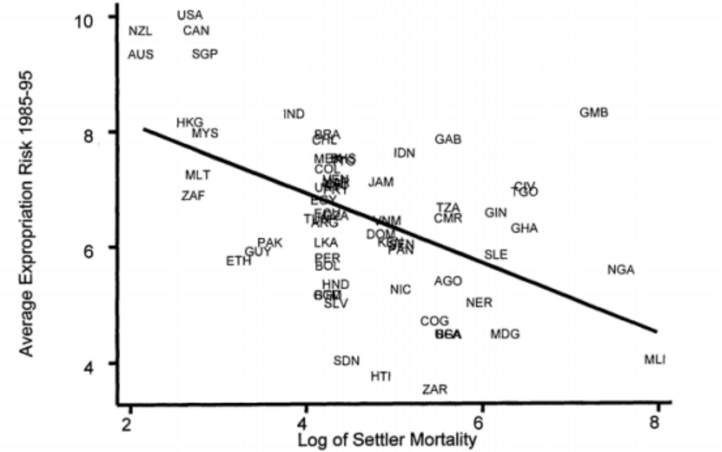


FIGURE 3. FIRST-STAGE RELATIONSHIP BETWEEN SETTLER MORTALITY AND EXPROPRIATION RISK

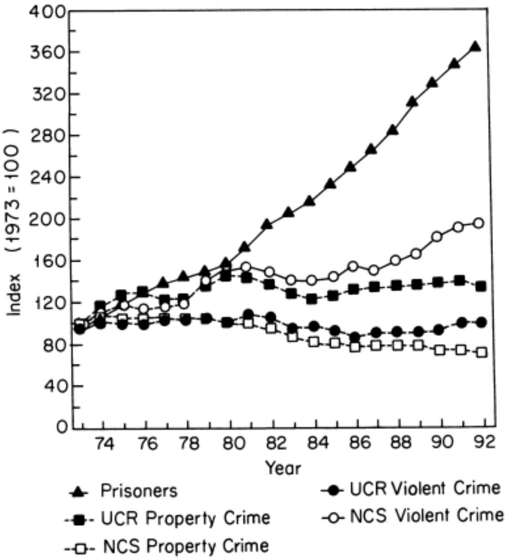
Example 3: Levitt (1996)

THE EFFECT OF PRISON POPULATION SIZE ON CRIME RATES: EVIDENCE FROM PRISON OVERCROWDING LITIGATION*

STEVEN D. LEVITT

Simultaneity between prisoner populations and crime rates makes it difficult to isolate the causal effect of changes in prison populations on crime. To break that simultaneity, this paper uses prison overcrowding litigation in a state as an instrument for changes in the prison population. The resulting elasticities are two to three times greater than those of previous studies. A one-prisoner reduction is associated with an increase of fifteen Index I crimes per year. While calculations of the costs of crime are inherently uncertain, it appears that the social benefits associated with crime reduction equal or exceed the social costs of incarceration for the marginal prisoner.

Example 3: Levitt (1996)



Example 3: Levitt (1996)

- The author wants to understand the effect of prison population on crime rates at the state level.
- Source of endogeneity:
  - There is very clear reverse causality here. States with higher crime rates are going to have a higher prison population.
- Natural experiment:
  - Use state-level prison overcrowding litigation as an instrument for prison population.
  - Levitt argues that prison overcrowding litigation will be related to crime rates only through its impact on prison populations, making the exclusion of litigation status from the crime equation valid.

Example 3: Levitt (1996)

- Y = log of crime rate
- X = prison population
- Z = prison-overcrowding lawsuits (10 variables)
  - Overcrowding litigation and incarceration rate changes are negatively correlated
  - "Over the past 30 years, prisoners' rights groups have brought numerous civil suits alleging unconstitutional conditions in prisons. In twelve states the entire state prison system either is currently or has formerly been under court order concerning overcrowding."
- Is this a valid instrument? Discuss.
- Levitt finds that a 1 prisoner reduction in the prison population results in an increase in fifteen Index I crimes (property and violent crimes) per year.
- This is evidence against reducing the prison population.

In-Class Exercise 1

THE

QUARTERLY JOURNAL OF ECONOMICS

Vol. CVI November 1991 Issue 4

DOES COMPULSORY SCHOOL ATTENDANCE AFFECT SCHOOLING AND EARNINGS?*

JOSHUA D. ANGRIST AND ALAN B. KRUEGER

We establish that season of birth is related to educational attainment because of school start age policy and compulsory school attendance laws. Individuals born in the beginning of the year start school at an older age, and can therefore drop out after completing less schooling than individuals born near the end of the year. Roughly 25 percent of potential dropouts remain in school because of compulsory schooling laws. We estimate the impact of compulsory schooling on earnings by using quarter of birth as an instrument for education. The instrumental variables estimate of the return to education is close to the ordinary least squares estimate, suggesting that there is little bias in conventional estimates.

In-Class Exercise 1

Please form into groups of 2 or 3 and write up your answers (written answers and do-file). Send them to my email (matthew.tudball@mail.utoronto.ca) when you have finished.

This is going to be an in-depth replication exercise. Download the dataset QOB1980.dta from my website (matthewtudball.com). This dataset comes from the very influential paper Angrist and Krueger (1991) which contains data on weekly wages (*log_weekly_wage*), completed years of education (*education*), year of birth (*year_of_birth*) and quarter of birth (*quarter_of_birth*) from individuals between 1930 - 1939 .

- This paper centres around the question of the causal effect of education on earnings, which is plagued with endogeneity problems.
- The authors propose a novel instrument for an individual's level of education: quarter of birth.

In-Class Exercise 1

- The idea exploits the variation induced by compulsory schooling laws in the US: most states require students to enter school in the calendar year they turn 6. In addition, these laws require students to remain in school at least until their 16th birthday.
- For instance, a student born in January starts school at 6 (and 8 months), and at her 16th birthday she will have 9 years of completed schooling. A student born in December starts school at 5 (and 8 months), and when she turns 16 she will have 10 years of schooling.
- Then, depending on the date of birth, students will be in different grades when they reach the legal drop-out age.

In-Class Exercise 1

1. We will first discuss the relevance condition. Start by replicating Figure 1 on page 5 which shows the relationship between completed years of education and quarter of birth.
   - Hint: You will first want to generate a year-quarter interaction. You can do this using
     ```
     egen qy = group(year_of_birth quarter_of_birth)
     ```
   - Hint: You will then want to collapse the data by *qy*. You can do this using
     ```
     collapse (mean) log_weekly_wage education, by(qy)
     ```
     You will now have a dataset of average weekly wages and average completed years of education by each year-quarter group.
2. Reload the raw data and test the relevance condition by regressing *education* on each of these year-quarter dummies and performing an F-test. Are the instruments jointly significant?
   - Hint: You can run a regression on dummies by using
     ```
     reg education i.yq
     ```

In-Class Exercise 1

3. Let's now discuss the exogeneity condition. Within your group discuss whether you think the authors' justification for their instrument is valid. Can you think of any violations which might invalidate it? Write down your answer.
4. Let's run the regression. Replicate Column (2) in Table V of page 22. Read carefully what variables that table is including in the regression. You should get the same coefficient on education as the one in that table.
5. Notice that our model is over-identified (i.e. we have more instruments than endogenous variables). Perform Stata's in-built test of over-identifying restrictions described on Slide 12. What do you conclude?
6. Now perform the Hausman test for the exogeneity of *education* as described on Slide 5. What do you conclude?