Empirical Banking and Finance
Tutorial 3

Konrad Adler

Institute for Financial Economics and Statistics
University of Bonn

Summer 2020

Tutorial 3

- Overall, good work, again
- Subset: Read the questions carefully
  - Example: test the two requirements!
  - Example II: why could financial development simply be a leading indicator for growth
- 0) Error in my slides:
  - Sorry for the mistake
  - Did not count those parts of the questions
- 1) Discuss the solution
- 2) Small methods add-on to show how the information of several instruments is combined

This Lecture

- Correction
- Tutorial 3
- Combining Instruments

Two Stage Least Squares

 # 3 Overidentifying restrictions test: Are the instruments exogenous?

Mechanics:

Run Two Stage Least Squares, take the residuals $\hat{e}_i$

Regress the residuals $\hat{e}_i$ on all exogenous variables (z and x)

nR^2 of that regression is distributed χ^2_q

q the difference between # of instruments and # of endogenous variables

H_0: all instruments are valid[CORRECT VERSION]

This Lecture

Correction

Tutorial 3

Combining Instruments

Question 1: Motivating the use of IV

a)

b)
   i) Credit to GDP: Research and development is expensive, takes long and needs continuous funding. Firms could finance research using bank credit. If firms are on average successful an increase in credit would precede an increase in technological progress coming from research and development
   ii) Stock Market: Because shares are net present discounted values of future dividends and values the stock market is a leading indicator for economic development. Investors might have information about future economic growth and therefore buy more shares. In that case stock market valuation grows before economic growth follows.

c) An IV would help to address this reverse causality problem by only considering the impact of financial development on growth that is "exogenous".

Question 2: Regression 1: OLS

a)

b)

c) We include the 1960 log gdp because it is 1) correlated with subsequent GDP growth because decreasing returns and 2) because it is probably correlated with 1960 credit to GDP. Leaving it out would therefore bias the coefficient of interest

d)
   i) *private_credit_1960* is positive and statistically significant at 1% level.
   ii) A one sd increase in *private_credit_1960* would increase *gdpgrowth* by 0.44 sd
   iii) 95% Confidence interval: [.023 .068]

Question 3: The instrument

a)

b) i) First requirement: IV has to be correlated with the variable of interest, financial development, this is probably the case, but we can test this

ii) Second requirement: IV should not affect the outcome, *gdpgrowth*, other than through financial development

1) This seems problematic. When a country was forced to adopt/ adopted a legal system this also affected many other institutions as well. Those other institutions might affect gdpgrowth through other channels than the financial system.
2) Example: Former french colonies seem to have inherited worse infrastructure relative to other former colonies. Infrastructure quality affects gdpgrowth directly.

c) See lecture: More credit and larger stock market when investors know they can rely on a legal system to enforce their rights if necessary

Question 3: The instrument (continued)

d) 52% French legal system, 30% UK legal system

e) Somewhat more confident, because at first sight credit to GDP seems to vary quite a bit across legal origins. This indicates that the first stage might work, i.e. legal system can explain the variation in credit to GDP well.

Question 4: IV with one instrument

a) *Important*: include loggdp_1960 in both first and second stage!

b) Both are positive, only the OLS coefficient is statistically significant, but the IV coefficient is more than twice as large as the OLS coefficient

c) Confidence interval: [-0.037 0.27]

d) Exactly identified - *You don't need a test for this!*

e)
   i) First Requirement: we can look at the p-value of the first stage which is 0.058, which is less significant than what the rule of thumb suggests $\rightarrow$ seems that we have a weak instrument
   ii) Second Requirement: Exclusion restriction - can only be tested formally when we have more than 1 instrument
      - What about looking at corr(loggdp_1960, legor_ge)? Can we conclude that the exclusion restriction does not hold when the correlation is positive?
      - No, because even when the exclusion restriction holds the two might be correlated through *private_credit_1960*

Question 5: IV with several instruments

a)

b) Multicollinearity problem in the first stage

c) OLS vs IV: both are positive, IV has smaller t-stat, IV estimate is more than twice as large than OLS

d) Confidence interval [0.015 0.209]

e) Overidentified

Question 5: IV with several instruments (continued)

f) i) First Requirement:

- F-test of whether the four instruments are jointly significant in the first stage.[1]
- H_0: $\beta_{legalorigin1} = 0$ and $\beta_{legalorigin2} = 0$ and $\beta_{legalorigin3} = 0$ and $\beta_{legalorigin3} = 0$
- H_A: H_0 not true, any or all of the coefficients are different from zero
- F-Stat is 3.36, H_0 of all of the instruments being zero is rejected at 5%, but below the rule of thumb of an F-stat of 10
- Still weak instruments

[1] The F-test reported by the standard regression output is testing whether all explanatory variables are zero, not the same thing

Question 5: IV with several instruments (continued)

f) ii) Second Requirement:

- Overidentification test of all instruments
- H_0 all instruments are exogenous
- H_A at least one instrument is not exogenous
- Robust se: $\chi^2(3) = 3.128$
- Without robust se: $\chi^2(3) = 2.896$
- We cannot reject H_0 here, therefore all instruments are exogenous[2]

[2] At least formally, never claim that your instruments are exogenous only based on this test

Question 5: IV with several instruments (continued)

g) Not useful: the test requires that one instrument is truly exogenous, but with the legal origin dummies either all or none are truly endogenous

h) Is *private_credit_1960* really endogenous?

- H_0 *private_credit_1960* is exogenous.
- F-stat of 6.88
- $\chi^2(1) = 3.44$
- F-test: we can reject H_0 and conclude that *private_credit_1960* is endogenous at 5%
- $\chi^2(1)$ test: we cannot reject H_0 and conclude that *private_credit_1960* could be exogenous at 5% .

Question 6 & 7

Answered well

This Lecture

Correction

Tutorial 3

Combining Instruments

Legal origin dummies

Combining several instruments

- This short section tries to give some intuition how the information of several instruments is aggregated
- Two cases
  - 1 dummy instrument
  - several dummy instruments
- In the methods part these two ways of getting IV estimates:
  - Two stage least squares
  - Special case when instruments are dummies $\rightarrow$ next slide

Reference: Angrist (2009) 4.1.3

Instrumental Variables: Special Case

When z_i is a dummy, i.e. UK legal origin, the IV formula simplifies to:

$$\beta_1 = \frac{\mathbb{E}[y_i|z_i = 1] - \mathbb{E}[y_i|z_i = 0]}{\mathbb{E}[\textit{private\_credit\_1960}_i|z_i = 1] - \mathbb{E}[\textit{private\_credit\_1960}_i|z_i = 0]}$$

- Numerator:
  - Compare *gdpgrowth* of countries with and without UK legal orign
  - No matter what *private_credit_1960*
- Denominator:
  - Adjust for differences in *private_credit_1960*

Instrumental Variables: 2SLS

First Stage

$$private\_credit\_1960_i = \beta_0 + \beta_1 legalOrigin + e_i$$

- for simplicity without covariates

Second Stage

$$y_i = \beta_0 + \beta_1 \hat{private\_credit\_1960_i} + e_i$$

Instrumental Variables

How are the two ways related?

- 2SLS vs Special Case (Wald estimator)
- 1 dummy IV
- several dummies IV

1 dummy IV: 2SLS

First Stage

$$\begin{bmatrix} private\_credit\_1960_1 \\ \vdots \\ private\_credit\_1960_N \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 1 \\ 1 & 1 \\ 1 & \underbrace{1}_{\text{French}} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_{French} \end{bmatrix} + \begin{bmatrix} e_1 \\ \vdots \\ e_N \end{bmatrix}$$

How does $\hat{private\_credit\_1960_i}$ look like? Lets look at Stata

1 dummy IV: 2SLS

Second Stage

$$y_i = \beta_0 + .122654 \widehat{private\_credit\_1960}_i + e_i$$

What about the Wald estimator?

1 dummy IV: Wald estimator

$$\beta_1 = \frac{\mathbb{E}[y_i|z_i = 1] - \mathbb{E}[y_i|z_i = 0]}{\mathbb{E}[\textit{private\_credit\_1960}_i|z_i = 1] - \mathbb{E}[\textit{private\_credit\_1960}_i|z_i = 0]}$$

$$= \frac{0.0095 - 0.0243}{0.1951 - 0.3151}$$

$$= .122654$$

In the special case of 1 dummy variable IV 2SLS and Wald estimator are the same

- Numerator
  - Countries with French legal origin had much lower GDP growth
  - Irrespective of their level of *private_credit_1960*
- Denominator
  - Countries with French legal orign also had lower *private_credit_1960*
  - We have to adjust the difference in GDP growth by that difference

1 dummy IV: Wald estimator

$$\mathbb{E}[\textit{private\_credit\_1960}_i | z_i = 1] - \mathbb{E}[\textit{private\_credit\_1960}_i | z_i = 0]$$

- Denominator
  - Tension between two extreme cases
  - Both terms are equal: weak instrument
  - The two terms are very different:
    - We have to discount the GDP growth difference even more
    - Intuition: countries with French legal origin are *very* different from all other countries, therefore we cannot compare the GDP growth to the others

Legal origin dummies

Several dummies IV

several dummy IV: 2SLS

First Stage

$$\begin{bmatrix} private\_credit\_1960_1 \\ \vdots \\ private\_credit\_1960_N \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & x_1 \\ 1 & 1 & 0 & 0 & x_2 \\ 1 & 0 & 1 & 0 & x_3 \\ 1 & 0 & 1 & 0 & x_4 \\ 1 & 0 & 0 & 1 & x_5 \\ 1 & 0 & 0 & 1 & x_6 \\ 1 & \underbrace{0}_{\text{German}} & \underbrace{0}_{\text{French}} & \underbrace{0}_{\text{Common}} & x_7 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_{German} \\ \beta_{French} \\ \beta_{Common} \end{bmatrix} + \begin{bmatrix} e_1 \\ \vdots \\ e_N \end{bmatrix}$$

several dummy IV: 2SLS

Second Stage

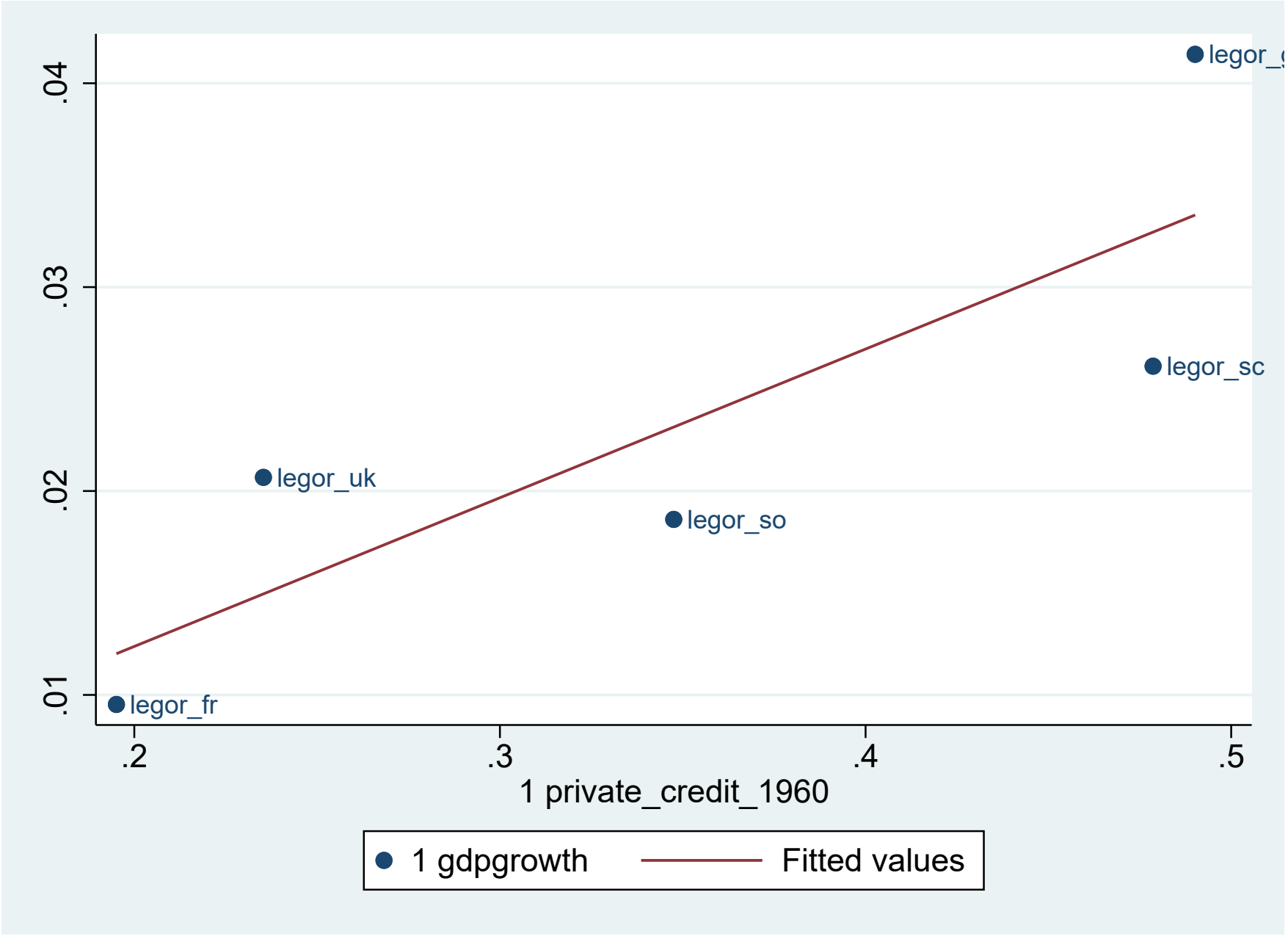
$$y_i = \beta_0 + .0803088 \widehat{private\_credit\_1960}_i + e_i$$

- Combine Wald estimators?
- Wald estimator for each instrument separately?

several dummy IV: Combine Wald estimators

- The 2SLS can be constructed by aggregating IV estimates
- IV for each legal origin category separately
- How does this work?
  - i) Compute $\mathbb{E}[y_i|z_i = 1]$
  - ii) Compute $\mathbb{E}[FinDev_i|z_i = 1]$
  - iii) Do this for all five z_i
  - iv) Run a regression the computed averages of gdp growth and financial development

several dummy IV: Combine Wald estimators



several dummy IV: Combine Wald estimators

- The slope, after reweighting, gives exactly the 2SLS estimate
- Implications
  - We don't need "micro" data to estimate IV
  - Group averages by instrument are enough
- How well does the line fit?
  - Test of whether the parameter of interest should be constant or not
- Take-away
  - 2SLS with several dummy instruments is just a linear combination of each separate IV estimate

several dummy IV: Separate Wald estimators

Wald estimator

$$\beta_1 = \frac{\mathbb{E}[y_i|z_i = 1] - \mathbb{E}[y_i|z_i = 0]}{\mathbb{E}[\mathrm{R}_i|z_i = 1] - \mathbb{E}[\mathrm{R}_i|z_i = 0]}$$

- The assumptions about the instruments should hold for each of them
- The results using individual instruments should be similar to the aggregated one

several dummy IV: Separate Wald estimators

legal origin	wald estimator
legor_fr	.12
legor_uk	-.23
legor_so	.02
legor_ge	.10
legor_sc	.04
2SLS	0.08

- The assumptions about the instruments should hold for each of them
- The results using individual instruments should be similar to the aggregated one

Combining instruments

Take-aways

- 1 dummy IV, no covariates
    - 2SLS = Wald estimator
- several dummy IVs, no covariates
    - 2SLS = linear combination of Wald type estimators
- Practical implications
    - IV can be used having only aggregate data
    - Assess the "fit" across instruments