FIN 620 Emp. Methods in Finance

Lecture 5 – Instrumental Variables

Professor Todd Gormley

Announcements

- Rough draft of research proposal due next week...
 - Just 1- to 3-page (single-spaced) sketch of your proposal is fine...
 - Should clearly state your question
 - Should give me idea of where you're going with the identification strategy
 - See grading template on Canvas
 - <u>Upload it to Canvas</u> by noon next week
 I will read and then send brief feedback

Background readings

- Roberts and Whited
 Section 3
- Angrist and Pischke
 Sections 4.1, 4.4, and 4.6
- Wooldridge
 Chapter 5

Greene
Sections 8.2-8.5

Outline for Today

- Quick review of panel regressions
- Discuss IV estimation
 - How does it help?
 - □ What assumptions are needed?
 - □ What are the weaknesses?
- Student presentations of "Panel Data"

Quick Review [Part 1]

- What type of omitted variable does panel data and FE help mitigate, and how?
 - Answer #1 = It can help eliminate omitted variables that don't vary within panel groups
 - Answer #2 = It does this by transforming the data to remove this group-level heterogeneity [or equivalently, directly controls for it using indicator variables as in LSDV]

Quick Review [Part 2]

Why is random effects useless
 [at least in corporate finance settings]?

Quick Review [Part 3]

- What are three limitations of FE?
 - #1 Can't estimate coefficient on variables that don't vary within groups
 - #2 Could amplify any measurement error
 - For this reason, be cautious interpreting zero or small coefficients on possibly mismeasured variables
 - #3 Can't be used in models with lagged values of the dependent variable

Outline for Instrumental Variables

- Motivation and intuition
- Required assumptions
- Implementation and 2SLS
 - Weak instruments problem
 - Multiple IVs and overidentification tests
- Miscellaneous IV issues
- Limitations of IV

Motivating IV [Part 1]

Consider the following estimation

$$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k + u$$

where $\operatorname{cov}(x_1, u) = \dots = \operatorname{cov}(x_{k-1}, u) = 0$ $\operatorname{cov}(x_k, u) \neq 0$

- If we estimate this model, will we get a consistent estimate of β_k?
- When would we get a consistent estimate of the other β's, and is this likely?

Motivation [Part 2]

- Answer #1: No. We will not get a consistent estimate of β_k
- Answer #2: Very unlikely. We will only get consistent estimate of other β if x_k is uncorrelated with all other x
- Instrumental variables provide a *potential* solution to this problem...

Instrumental variables – Intuition

\Box Think of x_k as having 'good' and 'bad' variation

- Good variation is <u>not</u> correlated with *u*
- Bad variation is correlated with *u*
- An IV (let's call it z) is a variable that explains variation in x_k, but doesn't explain y
 - I.e., it only explains the "good" variation in x_k
- Can use the IV to extract the "good" variation and replace x_k with only that component!

Outline for Instrumental Variables

- Motivation and intuition
- Required assumptions
- Implementation and 2SLS
 - Weak instruments problem
 - Multiple IVs and overidentification tests
- Miscellaneous IV issues
- Limitations of IV

Instrumental variables – Formally

- IVs must satisfy two conditions
 - Relevance condition
 - Exclusion condition
 - What are these two conditions?
 - Which is harder to satisfy?
 - Can we test whether they are true?

To illustrate these conditions, let's start with the simplest case, where we have one instrument, z, for the problematic regressor, x_k

Relevance condition [*Part 1*]

How can we test this condition?

The following must be true...

□ In the following model

 $x_{k} = \alpha_{0} + \alpha_{1}x_{1} + \dots + \alpha_{k-1}x_{k-1} + \gamma z + \nu$

z satisfies the relevance condition if $\gamma \neq 0$

- What does this mean in words?
 - Answer: z is relevant to explaining the problematic regressor, x_k, <u>after</u> partialling out the effect of all the other regressors in the original model

Relevance condition [Part 2]

- Easy to test the relevance condition!
 - Just run the regression of x_k on all the other x's and the instrument z to see if z explains x_k
 As we see later, this is what people call the 'first stage' of the IV estimation

Exclusion condition [Part 1]

How can we test this condition?

- The following must be true...
 - □ In the original model, where

 $y = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k + u$

z satisfies the exclusion condition if cov(z, u)=0

- What does this mean in words?
 - Answer: z is uncorrelated with the disturbance, u...
 i.e., z has no explanatory power with respect to y after conditioning on the other x's;

Exclusion condition [Part 2]

- Trick question! You <u>cannot</u> test the exclusion restriction [Why?]
 - Answer: You can't test it because *u* is <u>unobservable</u>
 - You must find a convincing *economic* argument as to why the exclusion restriction is not violated

Side note – What's wrong with this?

 I've seen some try to use the below argument as support for the exclusion restriction... what's wrong with it?

• Estimate the below regression...

$$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k + \gamma z + u$$

If γ=0, then exclusion restriction likely holds...
 i.e., they argue that z doesn't explain y after conditioning on the other x's

Side note – Answer

If the original regression doesn't give consistent estimates, then neither will this one!

cov(x_k, u)≠0, so the estimates are still biased
 Moreover, if we believe the relevance condition, then the coefficient on z is certainly biased because z is correlated with x_k

What makes a good instrument?

- Bottom line, an instrument must be justified largely on economic arguments
 - Relevance condition can be shown formally, but you should have an economic argument for why
 - Exclusion restriction <u>cannot</u> be tested... you need to provide a convincing economic argument as to why it explains y, but only through its effect on x_k

Outline for Instrumental Variables

- Motivation and intuition
- Required assumptions
- Implementation and 2SLS
 - Weak instruments problem
 - Multiple IVs and overidentification tests
- Miscellaneous IV issues
- Limitations of IV

Implementing IV estimation

- You've found a good IV, now what?
- One can think of the IV estimation as being done in two steps
 - **□** First stage: regress x_k on other x's & z
 - Second stage: take predicted x_k from first stage and use it in original model instead of x_k

This is why we also call IV estimations two stage least squares (2SLS)

First stage of 2SLS

Estimate the following

$$\begin{aligned} x_{k} &= \alpha_{0} + \alpha_{1}x_{1} + \ldots + \alpha_{k-1}x_{k-1} + \gamma z + v \\ & & & & & \\ \\ \text{Problematic regressor} \\ \text{[i.e., } \operatorname{cov}(x_{k}, u) \neq 0] \end{aligned}$$
 All other non-problematic variables that explain y Instrumental variable

Get estimates for the α's and γ
 Calculate predicted values, x̂_k, where

$$\hat{x}_k = \hat{\alpha}_0 + \hat{\alpha}_1 x_1 + \ldots + \hat{\alpha}_{k-1} x_{k-1} + \hat{\gamma} z$$

Second stage of 2SLS

Use predicted values to estimate

$$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_k \hat{x}_k + u$$

Predicted values replace the problematic regressor

 Can be shown (see textbook for math) that this 2SLS estimation yields consistent estimates of all the β when both the relevance and exclusion conditions are satisfied

Intuition behind 2SLS

- Predicted values represent variation in x_k that is 'good' in that it is driven only by factors that are uncorrelated with u
 - Specifically, predicted value is linear function of variables that are uncorrelated with *u*
- Why not just use other *x*'s? Why need *z*?
 - Answer: Can't just use other x's to generate predicted value because then predicted value would be collinear in the second stage

Reduced Form Estimates [Part 1]

 The "reduced form" estimation is when you regress *y* directly onto the instrument, *z*, and other non-problematic *x*'s

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_{k-1} x_{k-1} + \delta z + u$$

□ It is an unbiased and consistent estimate of the effect of z on y (presumably through the channel of z's effect on x_k)

Reduced Form Estimates [Part 2]

• It can be shown that the IV estimate for x_k , $\hat{\beta}_k^{IV}$, is simply given by...

 $\hat{\beta}_{k}^{IV} =$

Reduced form coefficient estimate for zFirst stage coefficient estimate for z

- I.e., if you don't find effect of z on y in reduced form, then IV is unlikely to work
 - IV estimate is just scaled version of reduced form

Practical advice [Part 1]

- Don't state in your paper's intro that you use an IV to resolve an identification problem, unless...
 - □ You also state what the IV you use is
 - And provide a strong economic argument as to why it satisfies the necessary conditions

Don't bury the explanation of your IV! Researchers that do this almost <u>always</u> have a bad IV. If you really have a good IV, you'll be willing to defend it in the intro!

Practical advice [Part 2]

- Don't forget to justify why we should believe the exclusion restriction holds
 - Too many researchers only talk about the relevance condition
 - Exclusion restriction is <u>equally</u> important

Practical Advice [Part 3]

- Do not do two stages on your own!
 - Let the software do it; e.g., in Stata, use the IVREG or XTIVREG (for panel data) commands
- Three ways people will mess up when trying to do 2SLS on their ...
 - #1 Standard errors will be wrong
 - #2 They try using nonlinear models in first stage
 - #3 They will use the fitted values incorrectly

Practical Advice [Part 3-1]

- Why will standard errors be wrong if you try to do 2SLS on your own?
 - Answer: Because the second stage uses 'estimated' values that have their own estimation error. This error needs to be considered when calculating standard errors!

Practical Advice [Part 3-2]

- People will try using predicted values from non-linear model, e.g., Probit or Logit, in a 'second stage' IV regression
 - But <u>only</u> linear OLS in first stage guarantees covariates and fitted values in second stage will be uncorrelated with the error
 - I.e., this approach is **NOT** consistent
 - This is what we call the "forbidden regression"

Practical Advice [Part 3-3]

In models with quadratic terms, e.g.

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + u$$

people often try to calculate one fitted value \hat{x} using one instrument, z, and then plug in \hat{x} and \hat{x}^2 into second stage...

• Seems intuitive, but it is **NOT** consistent!

□ Instead, you should just use z and z^2 as IVs!

Practical Advice [Part 3]

- Bottom line... if you find yourself plugging in fitted values when doing an IV, you are probably doing something wrong!
 - Let the software do it for you; it will prevent you from doing incorrect things

Practical Advice [Part 4]

- All x's that are not problematic, need to be included in the first stage!!!
 - You're <u>not</u> doing 2SLS, and you're <u>not</u> getting consistent estimates if this isn't done
 - □ This includes things like firm and year FE!
- Yet another reason to let statistical software do the 2SLS estimation for you!

Practical Advice [Part 5]

- Always report your first stage results & R²
- There are two good reasons for this...
 [What are they?]
 - Answer #1: It is direct test of relevance condition... i.e., we need to see $\gamma \neq 0$!
 - Answer #2: It helps us determine whether there might be a <u>weak IV problem</u>...
Outline for Instrumental Variables

- Motivation and intuition
- Required assumptions
- Implementation and 2SLS
 - Weak instruments problem
 - Multiple IVs and overidentification tests
- Miscellaneous IV issues
- Limitations of IV

Consistent, but biased

- IV is a consistent, but biased, estimator
 - For any finite number of observations, N, the IV estimates are biased toward the biased OLS estimate
 - But, as N approaches infinity, the IV estimates converge to the true coefficients
- This feature of IV leads to what we call the weak instrument problem...

Weak instruments problem

- A weak instrument is an IV that doesn't explain very much of the variation in the problematic regressor
- Why is this an issue?
 - Small sample bias of estimator is greater when the instrument is weak; *i.e.*, *our estimates*, *which use a finite sample, might be misleading...*
 - □ *t*-stats in finite sample can also be wrong

Weak IV bias can be severe [Part 1]

■ Hahn and Hausman (2005) show that finite sample bias of 2SLS is ≈

$$\frac{j\rho(1-r^2)}{Nr^2}$$

- j = number of IVs [we'll talk about multiple IVs in a second]
- $\Box \ \varrho = \text{correlation between } x_k \text{ and } u$
- \square $r^2 = \mathbb{R}^2$ from first-stage regression
- \square N = sample size

Weak IV bias can be severe [Part 2]

 $\int \frac{j\rho(1-r^2)}{Nr^2}$

More instruments, which we'll talk about later, need not help; they help increase r², but if they are weak (i.e., don't increase r² much), they can still increase finite sample bias A low explanatory power in first stage can result in large bias even if N is large

Detecting weak instruments

- Number of warning flags to watch for...
 - □ Large standard errors in IV estimates
 - You'll get large SEs when covariance between instrument and problematic regressor is low
 - □ Low F statistic from first stage
 - The higher F statistic for excluded IVs, the better
 - Stock, Wright, and Yogo (2002) find that an F statistic above 10 likely means you're okay...

Excluded IVs – Tangent

- Just some terminology...
 - In some ways, can think of all nonproblematic x's as IVs; they all appear in first stage and are used to get predicted values
 - But, when people refer to excluded IVs, they refer to the IVs (i.e., z's) that are excluded from the second stage

Outline for Instrumental Variables

- Motivation and intuition
- Required assumptions
- Implementation and 2SLS
 - Weak instruments problem
 - Multiple IVs and overidentification tests
- Miscellaneous IV issues
- Limitations of IV

More than one problematic regressor

■ Now, consider the following...

$$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k + u$$

where $\operatorname{cov}(x_1, u) = \dots = \operatorname{cov}(x_{k-2}, u) = 0$ $\operatorname{cov}(x_{k-1}, u) \neq 0$ $\operatorname{cov}(x_k, u) \neq 0$

- There are two problematic regressors, x_{k-1} and x_k
- Easy to show that IVs can solve this as well

Multiple IVs [Part 1]

- Just need one IV for each problematic regressor, e.g., z₁ and z₂
- Then, estimate 2SLS in similar way...
 - Regress x_k on all other x's (except x_{k-1}) and both instruments, z_1 and z_2
 - Regress x_{k-1} on all other x's (except x_k) and both instruments, z_1 and z_2
 - Get predicted values, do second stage

Multiple IVs [Part 2]

- Need at least as many IVs as problematic regressors to ensure predicted values are not collinear with the non-problematic x's
 - If # of IVs match # of problematic x's, model is said to be "Just Identified"

"Overidentified" Models

- Can also have models with more IVs than # of problematic regressors
 - E.g., *m* instruments for *h* problematic regressors, where *m* > *h*
 - □ This is what we call an overidentified model
- Can implement 2SLS just as before...

Overidentified model conditions

- Necessary conditions very similar
 - Exclusion restriction = none of the instruments are correlated with *u*

Relevance condition

E.g., you can't just have one IV that is correlated with all the problematic regressors, and all the other IVs are not

- Each first stage (there will be *h* of them) must have at least one IV with non-zero coefficient
 - Of the *m* instruments, there must be at least *b* of them that are partially correlated with problematic regressors [otherwise, model isn't identified]

Benefit of Overidentified Model

- Assuming you satisfy the relevance and exclusion conditions, you will get more asymptotic efficiency with more IVs
 - Intuition: you can extract more 'good' variation from the first stage of the estimation

However, Overidentification Dilemma

- Suppose you are a very clever researcher...
 - You find not just *h* instruments for *h* problematic regressors, you find *m* > *h*
 - First, you should consider yourself very clever [a good instrument is hard to come by]!
 - But why might you not want to use the *m-h* extra instruments?

Answer-Weak instruments

- Again, as we saw earlier, a weak instrument will increase likelihood of finite sample bias and misleading inferences!
 - If have one good IV, not clear you want to add some extra (less good) IVs...

Practical Advice – Overidentified IV

- Helpful to always show results using "just identified" model with your best IVs
 - □ It is least likely to suffer small sample bias
 - In fact, the just identified model is medianunbiased making weak instruments critique less of a concern

Overidentification "Tests" [Part 1]

- When model is overidentified, you can supposedly "test" the quality of your IVs
- The logic of the tests is as follows...
 - If all IVs are valid, then we can get consistent estimates using any subset of the IVs
 - So, compare IV estimates from different subsets; if find they are similar, this suggests the IVs okay

Overidentification "Tests" [Part 2]

- But I see the following <u>all</u> the time...
 - Researcher has overidentified IV model
 - <u>All</u> the IVs are highly questionable in that they lack convincing economic arguments
 - But authors argue that because their model passes some "overidentification test" that the IVs must be okay
- What is wrong with this logic?

Overidentification "Tests" [Part 3]

- **Answer** = All the IVs could be junk!
 - The "test" implicitly assumes that some subset of instruments is valid
 - This may not be the case!
- To reiterate my earlier point...
 - There is no test to prove an IV is valid! Can only motivate that the IV satisfies exclusion restriction using <u>economic</u> theory

"Informal" checks – Tangent

- It is useful, however, to try some
 "informal" checks on validity of IV
 - E.g., One could show the IV is uncorrelated with other non-problematic regressors or with *y* that pre-dates the instrument
 - Could help bolster economic argument that IV isn't related to outcome *y* for other reasons
 - But don't do this for your actual outcome, *y*, why?
 Answer = It would suggest a weak IV (at best)

Outline for Instrumental Variables

- Motivation and intuition
- Required assumptions
- Implementation and 2SLS
 - Weak instruments problem
 - Multiple IVs and overidentification tests
- Miscellaneous IV issues
- Limitations of IV

Miscelleneous IV issues

- IVs with interactions
- Constructing additional IVs
- Using lagged y or lagged x as IVs
- Using group average of x as IV for x
- Using IV with FE
- Using IV with measurement error

IVs with interactions

Suppose you want to estimate

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + u$$

where $cov(x_1, u) = 0$ $cov(x_2, u) \neq 0$

 Now, both x₂ and x₁x₂ are problematic
 Suppose you can only find one IV, *z*. Is there a way to get consistent estimates? IVs with interactions [Part 2]

- Answer = Yes! In this case, one can
 construct other instruments from the one IV
 - Use z as IV for x_2
 - $\Box \text{ Use } x_1 z \text{ as IV for } x_1 x_2$
- Same <u>economic argument</u> used to support z as IV for x_2 will carry through to using $x_1 z$ as IV for $x_1 x_2$

Constructing additional IV

■ Now, suppose you want to estimate

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u$$



Suppose you can only find one IV, z, and you think z is correlated with both x₂ and x₃...
 Can you use z and z² as IVs?

Constructing additional IV [Part 2]

- Answer = Technically, yes. But probably not advisable...
 - Absent an economic reason for why z² is correlated with either x₂ or x₃ after partialling out z, it's probably not a good IV
 - Even if it satisfies the relevance condition, it might be a 'weak' instrument, which can be problematic [as seen earlier]

Lagged instruments

- It has become common in CF to use lagged variables as instruments
- This usually takes two forms
 - Instrumenting for a lagged y in dynamic panel model with FE using a lagged lagged y
 - Instrumenting for problematic x or lagged y using lagged version of the same x

Example where lagged IVs are used

 As noted last week, we cannot estimate models with both a lagged dep. var. and unobserved FE

$$y_{i,t} = \alpha + \rho y_{i,t-1} + \beta x_{i,t} + f_i + u_{i,t}, \quad |\rho| < 1$$

- The lagged y independent variable will be correlated with the error, n
- One proposed solution is to use lagged values of y as IV for problematic $y_{i,t-1}$

Using lagged y as IV in panel models

- Specifically, papers propose using first differences combined with lagged values, like y_{i,t-2}, as instrument for y_{i,t-1}
 - *Could* work in theory, ...
 - Lagged *y* will likely satisfy relevance criteria
 - But exclusion restriction requires lagged values of y to be <u>uncorrelated</u> with differenced residual, $u_{i,t} u_{i,t-1}$

Is this plausible in corporate finance?

Lagged y values as instruments?

Probably not...

- Lagged values of *y* will be correlated with changes in errors if errors are serially correlated
- This is common in corporate finance, suggesting this approach is <u>not</u> helpful

[See Holtz-Eakin, Newey, and Rosen (1988), Arellano and Bond (1991), Blundell and Bond (1998) for more details on these type of IV strategies]

Lagged x values as instruments? [Part 1]

- Another approach is to make assumptions about how $x_{i,t}$ is correlated with $u_{i,t}$
 - Idea behind relevance condition is x is persistent and predictive of future x or future y [depends on what you're trying to instrument]
 - And exclusion restriction is satisfied if we assume $x_{i,t}$ is uncorrelated with future shocks, *u*

Lagged x values as instruments? [Part 2]

- Just not clear how plausible this is...
 - Again, serial correlation in *u* (which is very common in CF) all but guarantees the IV is invalid
 - An <u>economic argument</u> is generally lacking,
 [and for this reason, I'm very skeptical of these strategies]

[See Arellano and Bond (1991), Arellano and Bover (1995) for more details on these type of IV strategies]

Using group averages as IVs [Part 1]

Will often see the following...

$$y_{i,j} = \alpha + \beta x_{i,j} + u_{i,j}$$

□ $y_{i,j}$ is outcome for observation *i* (e.g., firm) in group *j* (e.g., industry)

□ Researcher worries that $cov(x, u) \neq 0$

• So, they use group average, $\overline{x}_{-i,j}$, as IV

$$\overline{x}_{-i,j} = \frac{1}{J-1} \sum_{\substack{i \in j \\ k \neq i}} x_{k,j}$$

J is # of observations in the group

Using group averages as IVs [Part 2]

- They say...
 - "group average of x is likely correlated with own x" i.e., relevance condition holds
 - "*but* group average doesn't directly affect y"
 i.e., exclusion restriction holds
- Anyone see a problem?

Using group averages as IVs [Part 3]

Answer =

- Relevance condition implicitly assumes some common group-level heterogeneity, *f_j*, that is correlated with *x_{ij}*
- But if model has f_j (i.e., group fixed effect), then $\overline{x}_{-i,j}$ must violate exclusion restriction!
- This is a bad IV [see Gormley and Matsa (2014) for more details]

?
Other Miscellaneous IVs

- As noted last class, IVs can also be useful in panel estimations
 - #1 Can help identify effect of variables that don't vary within groups [which we can't estimate directly in FE model]
 - #2 Can help with measurement error

#1 – IV and FE models [Part 1]

- Use the following three steps to identify variables that don't vary within groups...
 - **#1** Estimate the FE model
 - #2 Take group-averaged residuals, regress them onto variable(s), x', that don't vary in groups (*i.e.*, the variables you couldn't estimate in FE model)
 - Why is this second step (on its own) problematic?
 - Answer: because unobserved heterogeneity (which is still collinear with x') will still be in error (because it partly explains group-average residuals)

#1 – IV and FE models [Part 2]

- Solution in second step is to use IV!
 - #3 Use covariates that *do* vary in group (from first step) as instruments in second step
 - Which x's from first step are valid IVs?
 - Answer = those that don't co-vary with unobserved heterogeneity but do co-vary with variables that don't vary within groups [again, economic argument needed here]
 - See Hausman and Taylor (1981) for details
 Done in Stata using XTHTAYLOR

#2 – IV and measurement error [Part 1]

- As discussed last week, measurement error can be a problem in FE models
- IVs provide a potential solutions
 - □ Pretty simple idea...
 - □ Find *z* correlated to mismeasured variable, but not correlated with *u*; use IV

#2 – IV and measurement error [Part 2]

- But easier said then done!
 - Identifying a valid instrument requires researcher to understand exact source of measurement error
 - This is because the disturbance, *u*, will include the measurement error; hence, how can you make an economic argument that *z* is uncorrelated with it if you don't understand the measurement error?

[See Biorn (2000) and Almeida, Campello, and Galvao (RFS 2010) for examples of this strategy]

Outline for Instrumental Variables

- Motivation and intuition
- Required assumptions
- Implementation and 2SLS
 - Weak instruments problem
 - Multiple IVs and overidentification tests
- Miscellaneous IV issues
- Limitations of IV

Limitations of IV

- There are two main limitations to discuss
 - Finding a good instrument is hard; even the seemingly best IVs can have problems
 - External validity can be a concern

Subtle violations of exclusion restriction

- Even the seemingly best IVs can violate the exclusion restriction
 - Roberts and Whited (pg. 31, 2011) provide a good example of this in description of Bennedsen et al. (2007) paper
 - Whatever group is discussing this paper next week should look... ③

Bennedsen et al. (2007) example [Part 1]

- Paper studies effect of family CEO succession on firm performance
 - IVs for family CEO succession using gender of first-born child
 - Families where the first child was a boy are more likely to have a family CEO succession
 - Obviously, gender of first-born is totally random; seems like a great IV...

Any guesses as to what might be wrong?

Bennedsen et al. (2007) example [Part 2]

- Problem is that first-born gender may be correlated with disturbance *u*
 - Girl-first families may only turnover firm to a daughter when she is very talented
 - Therefore, effect of family CEO turnover might depend on gender of first born
 - I.e., gender of first born is correlated with *u* because it includes interaction between problematic *x* and the instrument, *z*!

External vs. Internal validity

- External validity is another concern of IV [and other identification strategies]
 - Internal validity is when the estimation strategy successfully uncovers a causal effect
 - External validity is when those estimates are predictive of outcomes in <u>other</u> scenarios
 - IV (done correctly) gives us internal validity
 - But it doesn't necessarily give us external validity

External validity [Part 1]

- Issue is that IV estimates only tell us about subsample where the instrument is predictive
 - Remember, you're only making use of variation in x driven by z
 - So, we aren't learning effect of x for observations where z doesn't explain x!
- It's a version of LATE (local average treatment effect) and affects interpretation

External validity [Part 2]

- Again, consider Bennedsen et al (2007)
 - Gender of first born may only predict likelihood of family turnover in certain firms...
 - I.e., family firms where CEO thinks females (including daughters) are less suitable for leadership positions
 - Thus, we only learn about effect of family succession for these firms
 - Why might this matter?

External validity [Part 3]

- Answer: These firms might be different in other dimensions, which limits the external validity of our findings
 - □ E.g., Could be that these are poorly run firms...
 - If so, then we only identify effect for such poorly run firms using the IV
 - And effect of family succession in well-run firms might be quite different...

External validity [Part 4]

- Possible test for external validity problems
 - Size of residual from first stage tells us something about importance of IV for certain observations
 - Large residual means IV didn't explain much
 - Small residual means it did
 - Compare characteristics (i.e., other x's) of observations of groups with small and large residuals to make sure they don't differ much

Summary of Today [Part 1]

- IV estimation is one possible way to overcome identification challenges
- A good IV needs to satisfy two conditions
 Relevance condition
 Exclusion condition
- Exclusion condition cannot be tested; must use economic argument to support it

Summary of Today [Part 2]

- IV estimations have their limits
 - **Really hard to find good IV**
 - Weak instruments can be a problem, particularly when you have more IVs than problematic regressors
 - External validity can be a concern

In First Half of Next Class

- Natural experiments [Part 1]
 - □ How do they help with identification?
 - What assumptions are necessary to make causal inferences?
 - What are their limitations?
- Related readings... see syllabus

Assign papers for next week...

Gormley (JFI 2010)

Foreign bank entry and credit access

Bennedsen, et al. (QJE 2007)

□ CEO family succession and performance

- Giroud, et al (RFS 2012)
 - Debt overhang and performance

Break Time

- Let's take our 10-minute break
- We'll do presentations when we get back