Instrumental variables in practice

PSCI 2301: Quantitative Political Science II

Prof. Brenton Kenkel

brenton.kenkel@gmail.com

Vanderbilt University

March 5, 2025

Recap

How can we draw causal inferences when there's unobserved confounding?

One approach — **instrumental variables**

1. Philosophy
    - Find an as-if-random influence on treatment assignment
    - Use it to isolate effect of treatment from confounders
2. Requirements for an instrumental variable
    - Independence: Instrument cannot be confounded (ideally random)
    - First stage: Instrument must affect treatment status (ideally a lot)
    - Exclusion restriction: Instrument only affects outcome through treatment, not directly or through any other channel

Today's agenda

User's guide to instrumental variables

1. Working through AJR data
    - Visual evidence of the relationship
    - Estimating the effect of institutions on growth
2. Practical issues
    - Checking for weak instruments
    - Calculating standard errors
    - When and how to include controls

Analyzing AJR’s data

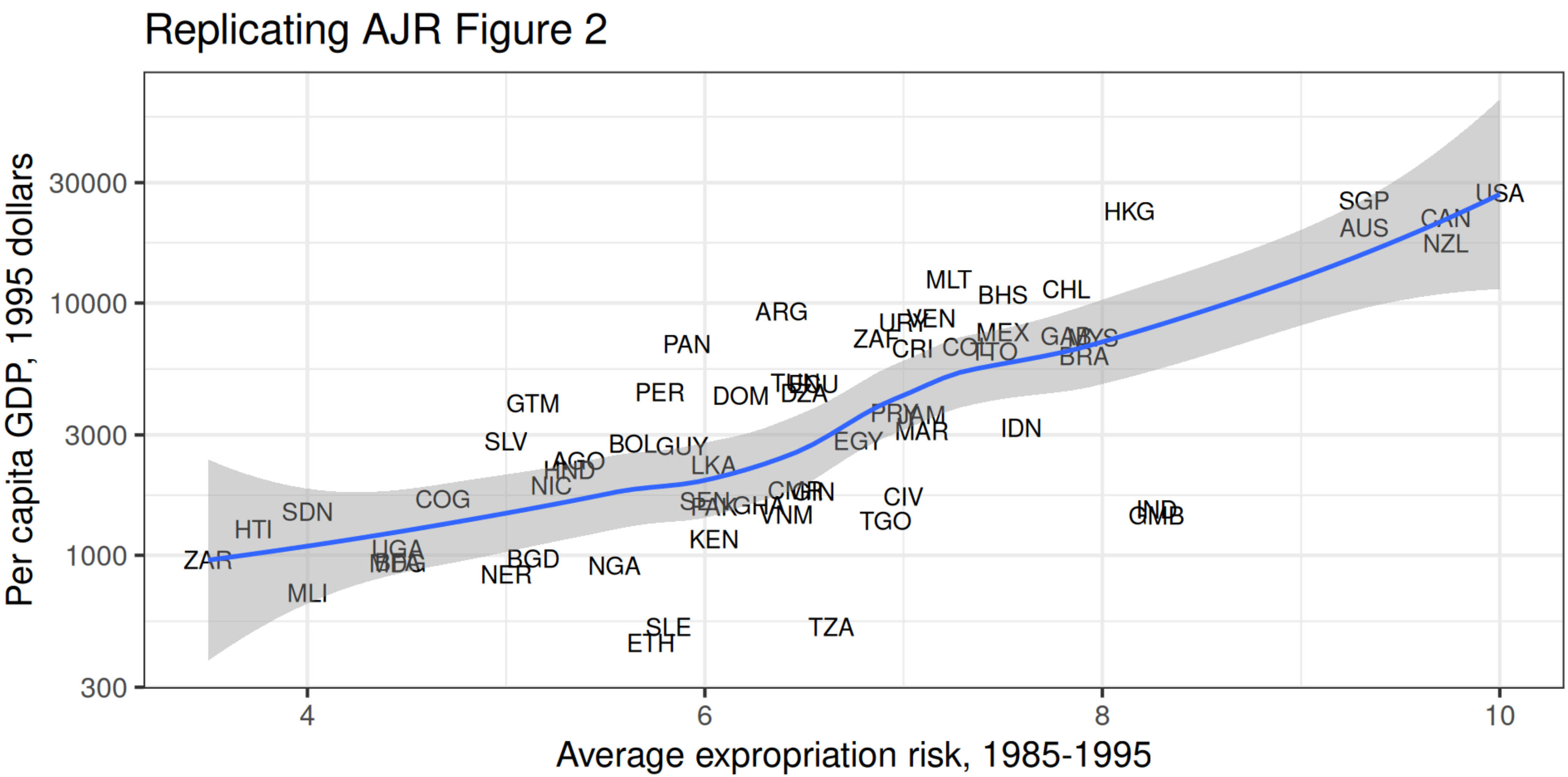
The data

Can obtain from Acemoglu's data archive site

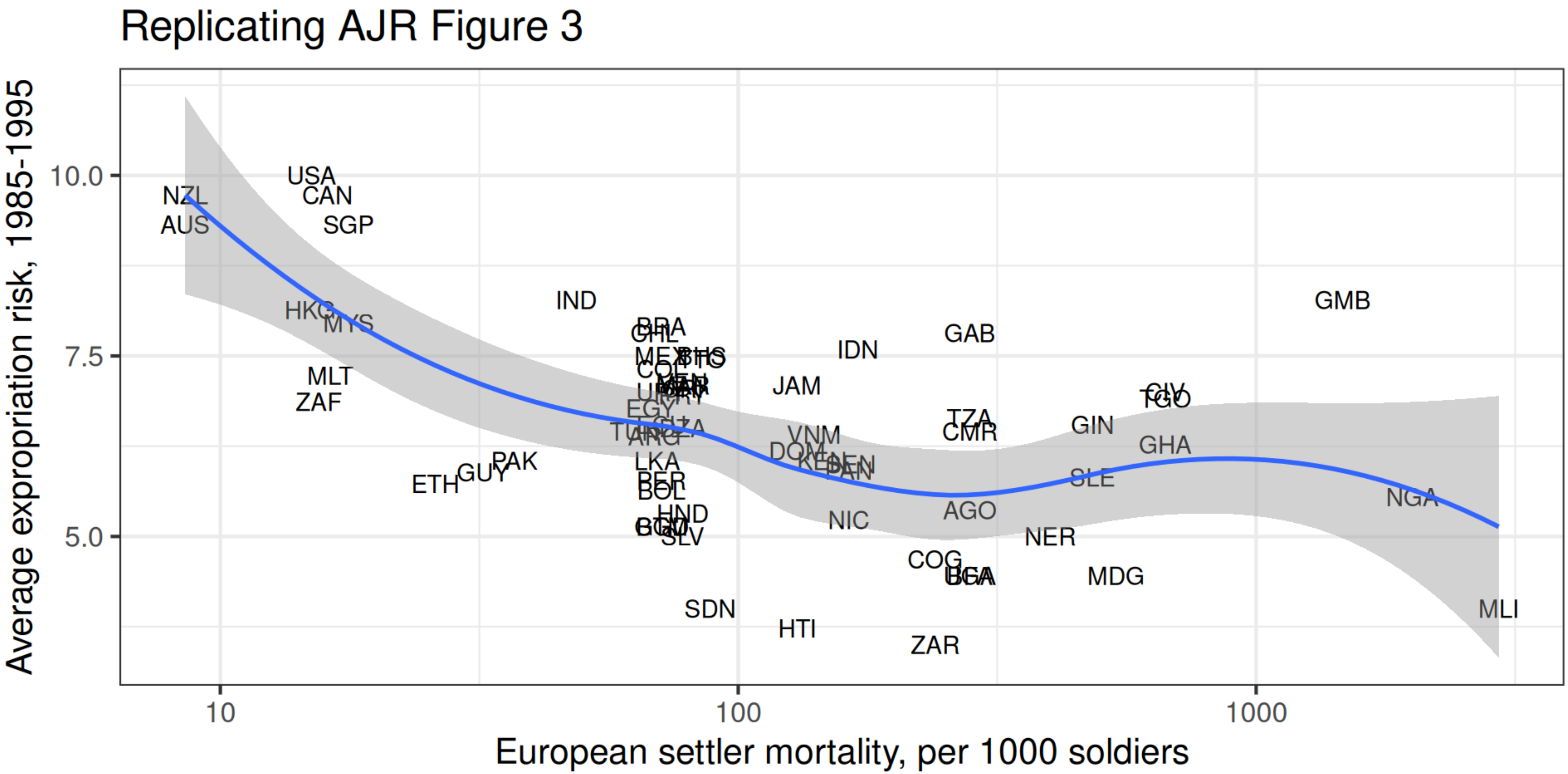
```r
library("haven")  # to read data in proprietary formats
df_ajr <- read_dta("maketable5.dta")
print(df_ajr)
```

```
# A tibble: 163 × 12
  shortnam catho80 muslim80 lat_abst no_cpm80 f_brit f_french avexpr sjlofr logpgp95 logem4 baseco
  <chr>      <dbl>    <dbl>    <dbl>    <dbl>  <dbl>    <dbl>  <dbl>  <dbl>    <dbl>  <dbl>  <dbl>
1 AFG        0       99.3      0.367    0.700      1        0  NA          1    NA       4.54     NA
2 AGO       68.7      0        0.137   11.5        0        0   5.36       1     7.77    5.63      1
3 ARE        0.400   94.9      0.267    4.40       1        0   7.18       0     9.80   NA         NA
4 ARG       91.6      0.200    0.378    5.50       0        0   6.39       1     9.13    4.23      1
5 ARM        0        0        0.444  100          0        0  NA          0     7.68   NA         NA
# i 158 more rows
```

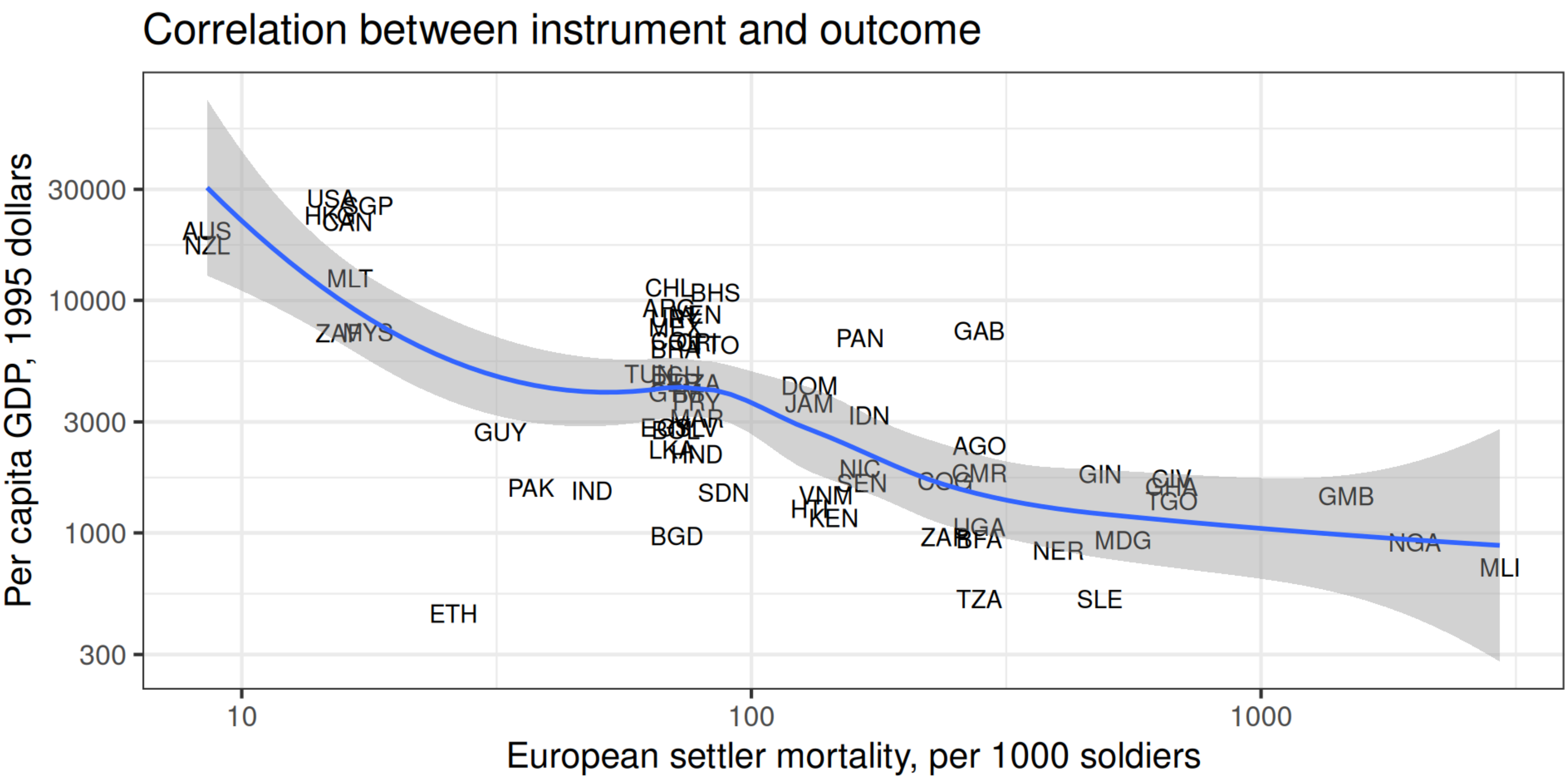
Raw correlation of institutions and development



First stage: Settler mortality and institutions



Reduced form: Settler mortality and development



Instrumental variables “by hand”

```r
# Effect of settler mortality on institutions
fit_first <- lm(avexpr ~ logem4, data = df_ajr)
tidy(fit_first)
```

```
# A tibble: 2 × 5
  term        estimate std.error statistic  p.value
  <chr>          <dbl>     <dbl>     <dbl>    <dbl>
1 (Intercept)    9.34      0.611      15.3  1.02e-22
2 logem4        -0.607     0.127      -4.79 1.08e- 5
```

```r
# Effect of settler mortality on development
fit_reduced <- lm(logpgp95 ~ logem4, data = df_ajr)
tidy(fit_reduced)
```

```
# A tibble: 2 × 5
  term        estimate std.error statistic  p.value
  <chr>          <dbl>     <dbl>     <dbl>    <dbl>
1 (Intercept)   10.7      0.367       29.2  5.80e-38
2 logem4        -0.573    0.0762      -7.52 2.66e-10
```

Instrumental variables “by hand”

```r
coef_first <- coef(fit_first)["logem4"]
coef_reduced <- coef(fit_reduced)["logem4"]
coef_reduced / coef_first
```

```
   logem4
0.9442794
```

	Base sample (1)	Base sample (2)
Average protection against expropriation risk 1985–1995	0.94	1.00
	(0.16)	(0.22)
Latitude		−0.65
		(1.34)

Aside: Interpreting regressions with logs

Y specification	X specification	Interpretation
Y	X	1 unit increase in X $\rightsquigarrow$ β unit increase in Y
Y	ln X	1% increase in X $\rightsquigarrow$ $\beta/100$ unit increase in Y
ln Y	X	1 unit increase in X $\rightsquigarrow$ $100(e^{\beta}-1)$ percent change in Y
ln Y	ln X	1% increase in X $\rightsquigarrow$ β percent change in Y

1. Reduced form regression: ln(GDP per capita) ~ ln(mortality)
    - Coefficient estimate: $\hat{\beta} = -0.573$
    - 1% mortality increase $\rightsquigarrow$ 0.573% decrease in GDP per capita
2. IV estimate: ln(GDP per capita) ~ expropriation risk index
    - Coefficient estimate: $\hat{\beta} = 0.944$
    - $e^{0.944} \approx 2.57$, use `exp()` function in R
    - 1 unit risk increase $\rightsquigarrow$ 157% increase in GDP per capita

Some questions at this point

Is this instrument strong enough that we can rely on it?

- Typical rule: F-statistic of first-stage regression should be 10 or higher
- (don't worry if that sounds like gobbledygook at this point)

Is there enough evidence against zero causal effect?

- We need standard errors to calculate hypothesis tests
- How do we calculate them?

`ivreg()` from the `AER` package solves both of these issues at once

Using ivreg()

```r
library("AER")
fit_iv <- ivreg(logpgp95 ~ avexpr | logem4, data = df_ajr)
summary(fit_iv, diagnostics = TRUE)
```

```

Call:
ivreg(formula = logpgp95 ~ avexpr | logem4, data = df_ajr)

Residuals:
     Min       1Q   Median       3Q      Max 
-2.44903 -0.56242  0.07311  0.69564  1.71752 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   1.9097     1.0267   1.860   0.0676 .  
avexpr        0.9443     0.1565   6.033  9.8e-08 ***

Diagnostic tests:
                 df1 df2 statistic  p-value    
Weak instruments   1  62     22.95 1.08e-05 ***
```

Instrumental variables versus ordinary regression

Regression estimate, ignoring confounding:

```r
fit_ols <- lm(logpgp95 ~ avexpr, data = df_ajr)
tidy(fit_ols)
```

```
# A tibble: 2 × 5
  term        estimate std.error statistic  p.value
  <chr>          <dbl>     <dbl>     <dbl>    <dbl>
1 (Intercept)    4.66     0.409      11.4  7.19e-17
2 avexpr         0.522    0.0612      8.53 4.72e-12
```

IV standard errors typically much larger

```r
tidy(fit_iv)
```

```
# A tibble: 2 × 5
  term        estimate std.error statistic      p.value
  <chr>          <dbl>     <dbl>     <dbl>        <dbl>
1 (Intercept)    1.91      1.03       1.86 0.0676
2 avexpr         0.944     0.157      6.03 0.0000000980
```

Including controls with instruments

A confounded instrument

Possible worry: Settler mortality doesn't satisfy independence condition

Potential confounding effect of geography

- Tropical location $\rightsquigarrow$ higher settler mortality
- Tropical location $\rightsquigarrow$ present-day growth differences

Luckily, geography is easily measurable

But need to use two stage least squares for estimation

Two stage least squares requirements

Basic ingredients

- Outcome of interest Y_i
- Treatment variable D_i
- Instrument Z_i
- Observed confounders $X_{i1}, \ldots, X_{iK}$

Now we assume conditional independence of the instrument:

- Instrument "assignment" as-if random among observations with same X's
  - → e.g., no confounders for settler mortality in countries at same latitude
- Still allowing for unobserved confounding in *treatment* assignment

Two stage least squares methodology

1. First stage regression
    - Run regression of the form `treatment ~ instrument + confounders`
    - Save predicted values from that regression, `pred_treatment`
    - These represent the as-if random *component* of treatment assignment
2. Final regression
    - Run regression of the form `outcome ~ pred_treatment + confounders`
    - Coefficient on `pred_treatment` = 2SLS estimate of treatment effect

Without confounders, 2SLS yields same answer

```r
df_ajr_aug <- augment(fit_first, newdata = df_ajr)
print(df_ajr_aug)
```

```
# A tibble: 64 × 16
  shortnam catho80 muslim80 lat_abst no_cpm80 f_brit f_french avexpr sjlofr logpgp95 logem4 baseco
  <chr>      <dbl>    <dbl>    <dbl>    <dbl>  <dbl>    <dbl>  <dbl>  <dbl>    <dbl>  <dbl>  <dbl>
1 AGO       68.7      0        0.137    11.5       0        0   5.36      1     7.77   5.63      1
2 ARG       91.6      0.200    0.378     5.50      0        0   6.39      1     9.13   4.23      1
3 AUS       29.6      0.200    0.300    46.7       1        0   9.32      0     9.90   2.15      1
4 BFA        9       43        0.144    46.4       0        1   4.45      1     6.85   5.63      1
5 BGD        0.200   85.9      0.267    13.7       1        0   5.14      0     6.88   4.27      1
# i 59 more rows
# i 4 more variables: pgp95 <dbl>, em4 <dbl>, .fitted <dbl>, .resid <dbl>
```

```r
fit_2sls <- lm(logpgp95 ~ .fitted, data = df_ajr_aug)
tidy(fit_2sls)
```

```
# A tibble: 2 × 5
  term        estimate std.error statistic  p.value
  <chr>          <dbl>     <dbl>     <dbl>    <dbl>
1 (Intercept)    1.91      0.823      2.32 2.37e- 2
2 .fitted        0.944     0.126      7.52 2.66e-10
```

2SLS with confounders

```r
# Run first stage regression
fit_first_lat <- lm(avexpr ~ logem4 + lat_abst, dat

# Extract predicted values
df_ajr_aug <- augment(fit_first_lat, newdata = df_a

# Run final regression
fit_2sls_lat <- lm(logpgp95 ~ .fitted + lat_abst, d
tidy(fit_2sls_lat)
```

```
# A tibble: 3 × 5
  term      estimate std.error statistic p.value
  <chr>        <dbl>     <dbl>     <dbl>   <dbl>
1 (Inte…       1.69      0.965     1.75  8.45e-2
2 .fitt…       0.996     0.165     6.02  1.08e-7
3 lat_a…      -0.647     0.996    -0.650 5.18e-1
```

	Base sample (1)	Base sample (2)
Average protection against expropriation risk 1985–1995	0.94 (0.16)	1.00 (0.22)
Latitude		−0.65 (1.34)

Getting the standard errors right

```
fit_iv_lat <- ivreg(logpgp95 ~ avexpr + lat_abst | logem4 + lat_abst, data = df_ajr)
summary(fit_iv_lat, diagnostics = TRUE)
```

```
Call:
ivreg(formula = logpgp95 ~ avexpr + lat_abst | logem4 + lat_abst,
    data = df_ajr)

Residuals:
    Min      1Q  Median      3Q     Max 
-2.5611 -0.6557  0.0732  0.7572  1.8803 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   1.6918     1.2930   1.308    0.196    
avexpr        0.9957     0.2217   4.492 3.21e-05 ***
lat_abst     -0.6472     1.3351  -0.485    0.630    

Diagnostic tests:
```

Wrapping up

What we did today

Using instrumental variables in practice

1. Look at the data to get a gut check
2. Fit model using `ivreg()`
    - Check for weak instruments statistic >10
    - Don't trust "by hand" standard errors
3. Include controls if instrument is confounded

After spring break

Assignments

- Problem Set 4 to be posted today, due **Wednesday, March 19**
- Final project proposals to be graded by end of this week
- Problem Set 3 to be graded (+ answer key posted) over the break

Topic for the week after spring break — **regression discontinuity**

1. Read Hall's "What Happens When Extremists Win Primaries?"
2. Read chapter 4 of *Mastering 'Metrics*