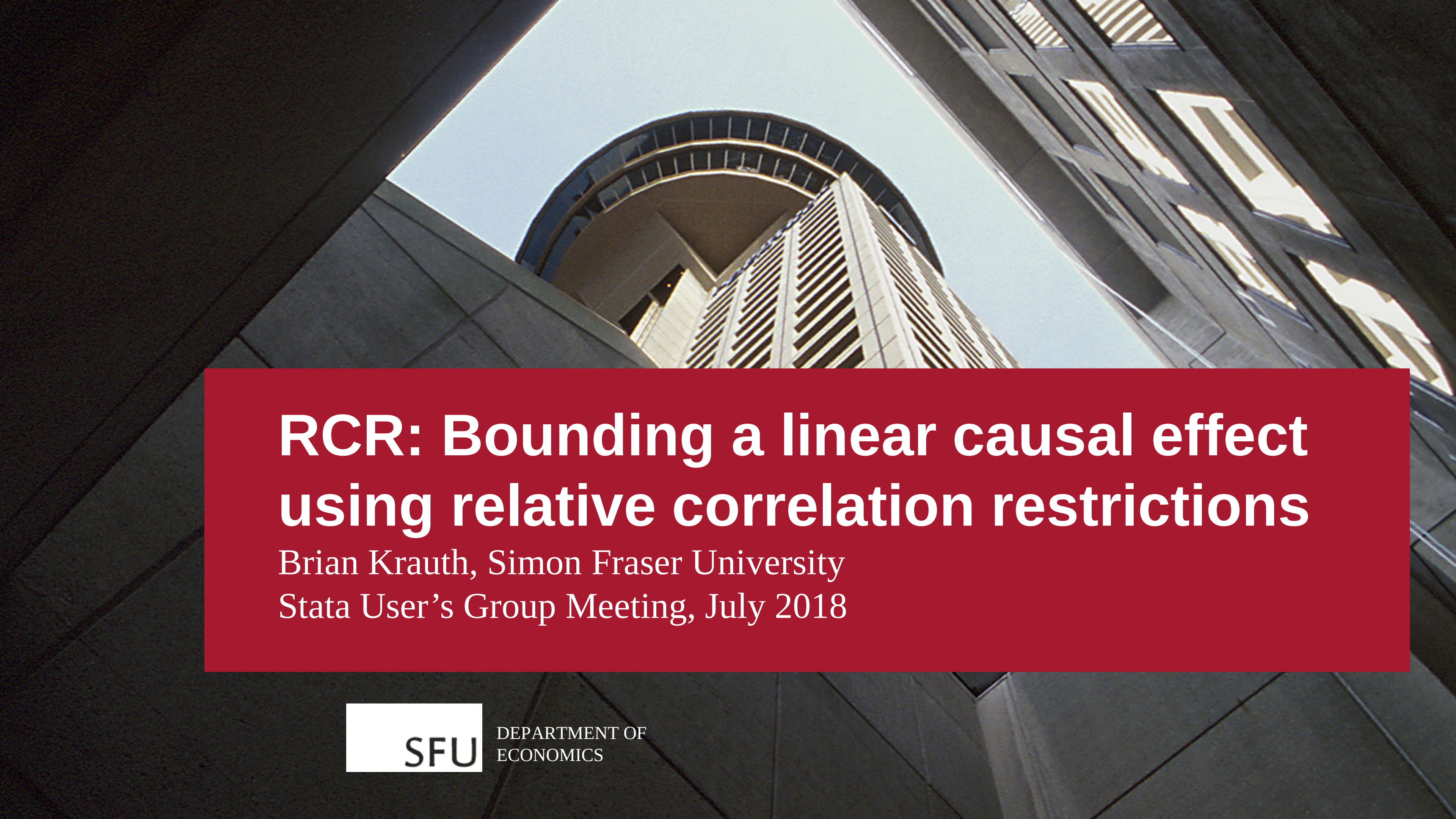

RCR: Bounding a linear causal effect using relative correlation restrictions

Brian Krauth, Simon Fraser University

Stata User’s Group Meeting, July 2018

SFU DEPARTMENT OF ECONOMICS

RCR

- `rcr` is a package that implements the estimator from "Bounding a linear causal effect using relative correlation restrictions," *Journal of Econometric Methods* 2016.

The problem

- You want to measure the effect (β_x) of some variable of interest (x_i) on some outcome (y_i).
- You estimate a linear regression model of the form:

$$y_i = x_i\beta_x + c_i\beta_c + v_i$$

  where c_i is a vector of control variables.
- For this to work, you need:

$$corr(x_i, v_i) = 0$$

  but you aren't quite sure that's true

RCR analysis: Model

Ordinary regression analysis

- Model is:

$$y_i = x_i \beta_x + c_i \beta_c + v_i$$

- We assume:

$$corr(x_i, v_i) = 0$$

RCR analysis

- Model is:

$$y_i = x_i \beta_x + c_i \beta_c + v_i$$

- We assume:

$$corr(x_i, v_i) = \lambda\, corr(x_i, c_i \beta_c)$$

where $\lambda \in [\lambda_L, \lambda_H]$

- Ordinary regression is a special case in which $\lambda_L = \lambda_H = 0$.

RCR analysis: Identification

Ordinary regression analysis

- For β_x we can obtain
  - A point estimate $\hat{\beta}_x$
  - A standard error for $\hat{\beta}_x$
  - A confidence interval $[\hat{\beta}_x^L, \hat{\beta}_x^H]$
  - Hypothesis tests

RCR analysis

- For β_x we can obtain
  - Bounds $[\hat{\beta}_x^L, \hat{\beta}_x^H]$
  - Standard errors for $\hat{\beta}_x^L$ and $\hat{\beta}_x^H$
  - A confidence interval $[\hat{\beta}_x^L, \hat{\beta}_x^H]$
  - Hypothesis tests

RCR analysis: Implementation

Ordinary regression analysis

- Implementation in Stata:

```
reg y x c
```

RCR analysis

- Implementation in Stata:

```
rcr y x c, lambda(0 1)
```

if you assume $\lambda \in [0,1]$

What should λ be?

- Strong assumptions (small $\lambda_H - \lambda_L$) mean strong identification (small $\hat{\beta}_x^H - \hat{\beta}_x^L$).
  - If $[\lambda_L, \lambda_H]$ is wide enough, then $\left[\hat{\beta}_x^L, \hat{\beta}_x^H\right] = (-\infty, \infty)$
  - If $[\lambda_L, \lambda_H]$ is narrow, then $\left[\hat{\beta}_x^L, \hat{\beta}_x^H\right]$ is narrow too.
- Natural points of interest:
  - If $\lambda = 0$, then $corr(x_i, v_i) = 0$. Usual assumption.
  - If $\lambda = 1$, then $corr(x_i, v_i) = corr(x_i, c_i\beta_c)$. (Unmeasureable) correlation with unobservables same as (measurable) correlation with control variables.
  - So $[\lambda_L, \lambda_H] = [0,1]$ is a reasonable starting point benchmark.
  - Can also ask how small λ needs to be to preserve OLS results

RCR options

- **`lambda(numlist)`**: Allows you to specify the range for the λ parameter (default = **`lambda(0 1)`**)
- **`cluster(varname)`**: cluster-robust standard errors (default = no clustering)
- **`vceadj(#)`**: the covariance matrix is multiplied by this number, allows you to implement degrees-of-freedom corrections if you are using transformed data (default = **`vceadj(1)`**)
- **`citype(string)`**: Confidence interval method (default = **`citype("conservative")`**)
- **`level(#)`**: confidence level (default = **`level(95)`**)
- Can handle all **`if`**, **`in`** and **`weight`** options.

Applications

- Application #1: Project STAR
  - Krueger (1999) study of effect of class size on test scores in Tennessee elementary schools
  - Imperfectly implemented experimental design
- Application #2: Malaria
  - Bleakley (2010) study of effect of malaria on labor productivity in US states
  - Observational design with lots of control variables

Project STAR: Implementation

Ordinary regression analysis

- Implementation in Stata:

```
reg SAT Small_Class White_Asian
Girl Free_Lunch White_Teacher
Teacher_Experience Masters_Degree
if Grade == 0, cluster(TCHID)
absorb(SCHID)
```

RCR analysis

- Implementation in Stata:

```
rcr SAT Small_Class White_Asian
Girl Free_Lunch White_Teacher
Teacher_Experience Masters_Degree
if Grade == 0, cluster(TCHID)
vceadj(1.013558143577264)
lambda(0 1) citype("imbens-
manski")
```

Project STAR: Results

Ordinary regression analysis

Linear regression

Number of obs =	5839
F(7, 322) =	62.24
Prob > F =	0.0000
R-squared =	0.0927
Root MSE =	22.202

(Std. Err. adjusted for 323 clusters in TCHID)

SAT	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
Small_Class	5.201503	1.031215	5.04	0.000	3.172733	7.230272
White_Asian	8.388708	1.35401	6.20	0.000	5.724884	11.05253
Girl	4.382026	.625326	7.01	0.000	3.151785	5.612266
Free_Lunch	-13.07747	.7688515	-17.01	0.000	-14.59008	-11.56486
White_Teac~r	-1.08975	2.170195	-0.50	0.616	-5.359302	3.179802
Teacher_Ex~e	.2650214	.103897	2.55	0.011	.0606187	.469424
Masters_De~e	-.5996357	1.053635	-0.57	0.570	-2.672514	1.473243
_cons	47.09615	2.405294	19.58	0.000	42.36407	51.82822

RCR analysis

RCR model

Number of obs =	5839
Lower bound on lambda =	0
Upper bound on lambda =	1

(Std. Err. adjusted for 323 clusters in TCHID)

SAT	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
lambdaInf	12.3106	8.372929	1.47	0.141	-4.10004	28.72124
betaxInf	8.16971	43.9407	0.19	0.853	-77.95248	94.2919
lambda0	28.93549	146.1582	0.20	0.843	-257.5293	315.4003
betaxL	5.135044	1.367873	3.75	0.000	2.454062	7.816026
betaxH	5.201503	1.03756	5.01	0.000	3.167923	7.235082
betax	(Imbens-Manski confidence interval)				2.4865	7.210477

Treatment Variable: Small_Class
Control Variables : White_Asian Girl Free_Lunch White_Teacher
Teacher_Experience Masters_Degree

Project STAR: Interpretation

Ordinary regression analysis

- If class size is uncorrelated with unobservables, then
  - Being in a small class increases test scores by 5.2 percentile points
    - 95% CI = (3.2,7.2)

RCR analysis

- If class size is no more correlated with unobservables than it is with the control variables,
  - Being in a small class increases test scores by 5.1 to 5.2 percentile points
    - 95% CI = (2.5,7.2)
- If class size is less than 12.3 times as correlated with unobservables
  - The class size effect is positive.

Malaria: Implementation

Ordinary regression analysis

- Implementation in Stata:

```
reg docc malmort1890 south
lebergott99 [aw=wtbpl]
```

RCR analysis

- Implementation in Stata:

```
rcr docc malmort1890 south
lebergott99 [aw=wtbpl], lambda(0
1)
```

Malaria: Results

Ordinary regression analysis

```
      Source |       SS       df       MS              Number of obs =      48
-------------+------------------------------           F(  3,    44) =   15.52
       Model |  .196184286     3  .065394762           Prob > F      =  0.0000
    Residual |  .185372534    44  .004213012           R-squared     =  0.5142
-------------+------------------------------           Adj R-squared =  0.4810
       Total |  .381556819    47   .00811823           Root MSE      =  .06491

------------------------------------------------------------------------------
        docc |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 malmort1890 |   .1124381   .0476992     2.36   0.023     .0163066    .2085695
       south |     .04293   .0370928     1.16   0.253    -.0318257    .1176857
 lebergott99 |  -.0288005   .0227818    -1.26   0.213    -.0747141    .0171131
       _cons |   .2702744   .0131139    20.61   0.000     .2438451    .2967038
------------------------------------------------------------------------------
```

RCR analysis

```
RCR model                                                 Number of obs =        48
                                                   Lower bound on lambda  =         0
                                                   Upper bound on lambda  =         1


------------------------------------------------------------------------------
        docc |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   lambdaInf |   .9345488   .1745427     5.35   0.000     .5924514    1.276646
    betaxInf |   .2822161   .0554518     5.09   0.000     .1735326    .3908997
     lambda0 |   .3165399   .1179677     2.68   0.007     .0853274    .5477523
      betaxL |  -9.0e+306          .        .       .            .           .
      betaxH |   9.0e+306          .        .       .            .           .
-------------+----------------------------------------------------------------
       betax |   (Conservative confidence interval)     -9.0e+306    9.0e+306
------------------------------------------------------------------------------
Treatment Variable:    malmort1890
Control Variables :    south lebergott99
```

Malaria: Interpretation

Ordinary regression analysis

- If malaria is uncorrelated with unobservables, then:
  - Malaria eradication increases the occupational index score by 0.11 points per unit of pre-eradication malaria mortality
    - 95% CI = (0.03,0.19)

RCR analysis

- If malaria is no more correlated with unobservables than it is with the control variables, then:
  - We cannot put nontrivial bounds on the malaria eradication effect
- If class size is less than 0.32 times as correlated with unobservables, then:
  - The malaria eradication effect is positive.

Limitations

- Estimator is partly implemented in Fortran
  - Ready-to-run for Windows
  - Tyler Ransom has code and libraries for Linux (https://github.com/tyleransom/rcr)

Availability

- Full package available with
  - Ado-files for
    - `rcr` command
    - Postestimation commands for hypothesis testing, graphical tools, etc.
  - Help for all commands
  - Example code and data
- Go to http://www.sfu.ca/~bkrauth/code.htm
  - or just Google "rcr Krauth"
- I'm happy to provide technical support, just email me.