Applied Microeconometrics II , Lecture 7

Ciprian Domnisoru
Aalto University

Outline

- Exogeneity and overidentification tests
- Comparing OLS and IV
- Bartik instruments
- Regression discontinuity designs: functional form checks
- Application (placebo tests): Impacts of Public Transit on Traffic Congestion
- McCrary density test in Stata

Exogeneity tests

- Consider the equation of interest $y_1 = \delta X + \beta y_2 + u_1$, and the first stage $y_2 = \pi z + v_2$
- If y_2 is endogenous, $E(y_2 u_1) \neq 0$ and $E(v_2 u_1) \neq 0$
- Exogeneity test formulated as $E(v_2 u_1) = 0$. Null hypothesis is that residuals are uncorrelated.
- $u_1 = \rho v_2 + \epsilon$. Null hypothesis of exogeneity : $\rho{=}0$
- y2 is in fact exogenous, then OLS and 2SLS estimators should differ only because of sampling error - i.e. they should not give significantly different results
- $y_1 = \delta X + \beta y_2 + \rho \hat{v}_2 + \eta$

Exogeneity test

```
. ivreg2 loghoursal2 doby doby_2  (agelfted=drop16)

IV (2SLS) estimation
--------------------

Estimates efficient for homoskedasticity only
Statistics consistent for homoskedasticity only

                                                      Number of obs =    52794
                                                      F(  3, 52790) =   357.78
                                                      Prob > F      =   0.0000
Total (centered) SS     =   10802.6026                Centered R2   =  -0.1644
Total (uncentered) SS   =  300862.1476                Uncentered R2 =   0.9582
Residual SS             =  12578.16024                Root MSE      =    .4881

------------------------------------------------------------------------------
 loghoursal2 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    agelfted |  -.0281927   .0537391    -0.52   0.600    -.1335193     .077134
        doby |  -.0189836   .0119799    -1.58   0.113    -.0424638    .0044967
      doby_2 |    .000073   .0001385     0.53   0.598    -.0001984    .0003443
       _cons |   3.643409   1.075756     3.39   0.001     1.534966    5.751851
------------------------------------------------------------------------------
Underidentification test (Anderson canon. corr. LM statistic):          10.796
                                                   Chi-sq(1) P-val =    0.0010
------------------------------------------------------------------------------
Weak identification test (Cragg-Donald Wald F statistic):               10.798
Stock-Yogo weak ID test critical values: 10% maximal IV size             16.38
                                         15% maximal IV size              8.96
                                         20% maximal IV size              6.66
                                         25% maximal IV size              5.53
Source: Stock-Yogo (2005).  Reproduced by permission.
------------------------------------------------------------------------------
Sargan statistic (overidentification test of all instruments):           0.000
                                                 (equation exactly identified)
------------------------------------------------------------------------------
Instrumented:         agelfted
Included instruments: doby doby_2
Excluded instruments: drop16
------------------------------------------------------------------------------

. ivendog agelfted

Tests of endogeneity of: agelfted
H0: Regressor is exogenous
   Wu-Hausman F test:                  5.63426  F(1,52789)    P-value = 0.01762
   Durbin-Wu-Hausman chi-sq test:      5.63419  Chi-sq(1)     P-value = 0.01761
```

Multiple instruments and overidentification tests

- More (relevant) instruments can increase the first stage F-stat, reducing the variance of 2SLS estimates.
- If there are more instruments than endogenous regressors (the model if "overidentified"), it is possible to test – partially – for instrument exogeneity.
- Consider two relevant instruments: if 2SLS estimates using instruments separately are very different, then one or the other (or both) instruments must be failing the exogeneity restriction.
- Overidentification (or "J", or "Sargan") test: regress **residuals** from 2SLS equation (using X, not $\hat{X}$) on instruments; compute F-stat that the coefficients on all instruments are zero; J-statistic is mF, where m is the number of instruments. Under null of exogeneity, $\sim \chi^2$, reject if larger than some critical value. If m=k, J=0.

IV vs. OLS

- IV and OLS estimates sometimes vary widely.
- Think about the sign of the expected bias in OLS and the potential magnitude.
- Add multiple controls to OLS and see how OLS coefficient changes.
- If you think the OLS estimate is biased upwards (coefficient is overestimated), you would expect to find the IV is smaller than OLS. However measurement error (attenuation bias) can depress OLS, so IV may be larger.
- Do you have any reason to believe compliers are special? Might particularly benefit from the policy?

Bartik instruments

Bartik Instruments

- In economics, instrumental variables originated as an attempt to to isolate exogenous supply or demand shifters in some particular market.
- Last time, you saw the example using shifts in supply of fish caused by storms to recover demand parameters.

TABLE 5

Two-stage-least-squares estimates of demand function with stormy and mixed as instruments

Variable	est.	(s.e.)	est.	(s.e.)
Av. price effect	–1·01	(0·42)	–0·947	(0·46)
Monday			–0·013	(0·18)
Tuesday			–0·51	(0·18)
Wednesday			–0·56	(0·17)
Thursday			0·10	(0·17)
Weather on shore			0·02	(0·16)
Rain on shore			0·07	(0·16)

Note: Standard errors are reported in parentheses.

ıs of instrumental variables with the demand function equal to

$$\ln q_t^d(p) = \beta_0 + \beta_1 \cdot \ln p + \beta_2 \cdot x + \varepsilon_t^d,$$

Bartik Instruments

- $\Delta WAGEit = \alpha + \beta_{it}\Delta Employment_{it} + \epsilon_{it}$
  - β_{it} Inverse elasticity of labor supply. Regression motivated by local economic development policies. Bartik: "In response to an employment shock, employers are more likely to promote less-skilled individuals to avoid raising the occupation's real wage. An increase in occupational real wages would be needed to attract individuals of "normal" skill levels from outside the labor force or from other metropolitan areas."
- Goal: isolate local labor demand that is unrelated to changes in local labor supply.
- $\Delta Employment_{it} = \sum_{k=1}^{K} z_{ik} g_{ik}$
- z_{ik} industry k share in local employment i
- g_{ik} growth rate of industry k in location i.
- g_{ik} has a nationwide component g_k and a local component
- Use $\Delta B_{it} = \sum_{k=1}^{K} z_{ik} g_k$ as an instrument.
- "Local employment growth rate predicted by interacting local industry employment shares with national industry employment growth rates."
- labor, public, development, macroeconomics, international trade, and finance.

Bartik/Shift-share instruments

- Trade: impact of Chinese imports on manufacturing employment in U.S. cities (denoted by i)
  $\Delta MANUF_{it} = \alpha + \beta ImportExposure_{it} + \epsilon_{it}$
- $ImportExposure_{it} = \sum_{k=1}^{K} z_{ikt} g_{kt}^{US}$
- Import Exposure correlated with unobservables that also impact manufacturing employment.
- Autor et al. (2013) instrument:
  $B_{it} = \sum_{k=1}^{K} z_{ik(t-1)} g_{kt}^{OTHER}$
- Lagged ("initial") shares of employment in city i, g_{kt}^{OTHER} growth of Chinese imports in other high-income countries.
- weighted average of a "shift": how much China is exporting in different *k* product categories, with "shares" coming from initial industry composition.

Bartik/Shift-share instruments

- $\Delta NATIVEWAGES_{it} = \alpha + \beta Immigration_{it} + \epsilon_{it}$
- Concern: local demand shocks
- Use instrument $B_{it}{=}\sum_{k=1}^{K} z_{ik(t-1)} g_{kt}^{OVERALL}$
- z are the lagged shares of immigrants from source country k in city i, and g is the normalized change in overall immigration from country k into the U.S. Weighted average of the national inflow rates from each country ("the shift"), with weights depending on the initial distribution of immigrants ("the shares").

Criticisms

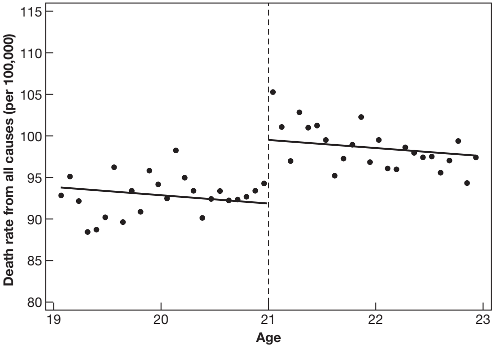
- Goldsmith-Pinkham, Sorkin and Swift (2020, AER). Why are the initial shares exogeneous? Check how much the initial shares are correlated with other potential confounders in the initial year. E.g. (computer manufacturing and education)
- Borusyak, Hull, and Jaravel (2018). Exogenous shares sufficient, but not necessary; Can identify effects if shocks are "as good as random"
- Jaeger et al. (2018). In immigration literature, Bartik instrument supposed to be exogenous to local demand shocks. However, if adjustment to shocks long-term, Bartik instrument biased. More reliable for initial immigration shares that rely on older, idiosyncratic policies.

Regression discontinuity designs: functional form specification

Sharp Regression Discontinuity (RD) Design

Things to consider: 1) Is the outcome continuous around the threshold? 2) Is there any manipulation? 3) Are there competing factors/policies? 4) How does the outcome evolve in the absence of the policy: linearly? non-linearly?

FIGURE 4.2
A sharp RD estimate of MLDA mortality effects



Notes: This figure plots death rates from all causes against age in months. The lines in the figure show fitted values from a regression of death rates on an over-21 dummy and age in months (the vertical dashed line indicates the minimum legal drinking age (MLDA) cutoff).

Sharp Regression Discontinuity (RD) Design

- A simple RD design analysis of the MLDA estimates causal effects using a regression like

$$\bar{M}_a = \alpha + \rho D_a + \gamma a + e_a$$

  - M_a is the death rate in month a
    - month is a 30-day interval counting from the 21st birthday
  - D_a is the treatment dummy

$$D_a = \begin{cases} 1 & \text{if } a \geq 21 \\ 0 & \text{if } a < 21 \end{cases}$$

  - a is a linear control for age measured in months

Sharp Regression Discontinuity (RD) Design

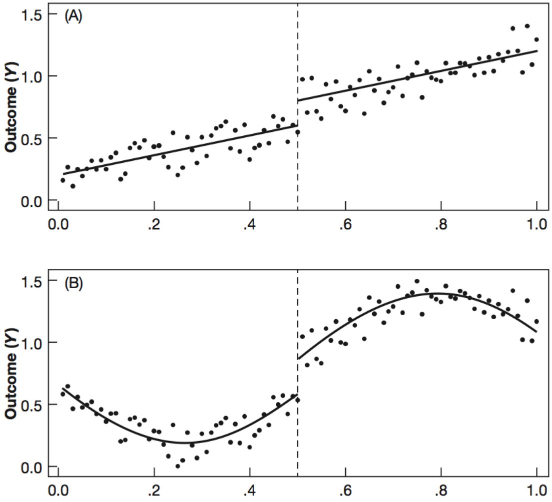
- Mortality clearly changes with the running variable, *a*, for reasons unrelated to the MLDA
  - deaths rates from disease-related causes like cancer (known as internal causes) are low but increasing for those in their late teens and early 20s
  - deaths from external causes, primarily car accidents, homicides, and suicides, fall
- To separate this trend variation from any possible MLDA effects, an RD analysis controls for smooth variation in death rates generated by *a*
- the negative slope captured by γ reflects smoothly declining death rates among young people as they mature

RD Specifics

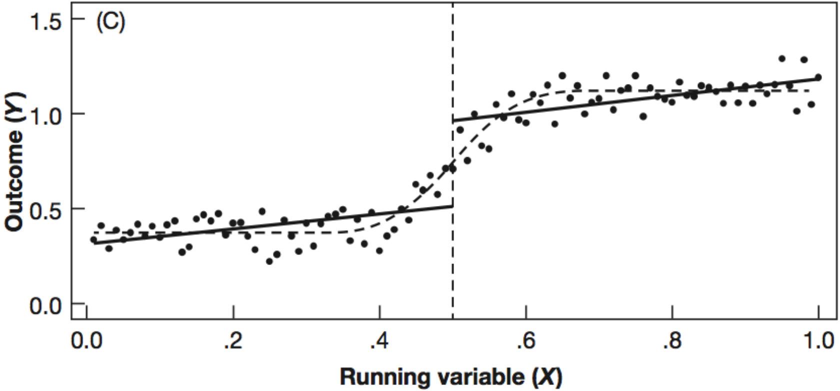
- RD tools aren't guaranteed to produce reliable causal estimates
  - challenge: mistaking a nonlinearity for a discontinuity
  - figure may exhibit a nonlinear trend with sharp turns to the left and right of the cutoff – like a discontinuity – but there is none
- Figure 4.3 shows three cases
  - *Panel A*: relationship between *running variable* (X) and the *outcome* (Y) is *linear*, with a clear jump in $E[Y|X]$ at cutoff
  - *Panel B*: relationship between X and Y is *nonlinear*, but jump in $E[Y|X]$ at the cutoff is still plain to see
  - *Panel C*: relationship between X and Y is *nonlinear*, with no jump in $E[Y|X]$ at the cutoff (RD challenge)

RD Specifics

FIGURE 4.3
RD in action, three ways



RD Specifics



Notes: Panel A shows RD with a linear model for $E[Y_i|X_i]$; panel B adds some curvature. Panel C shows nonlinearity mistaken for a discontinuity. The vertical dashed line indicates a hypothetical RD cutoff.

RD Specifics

- Two strategies to reduce likelihood of RD mistakes
  - 1st: modeling nonlinearity directly
  - 2nd: focusing solely on observations near the cutoff
- Nonlinear modeling strategy
  - typically, polynomial functions of the running variable
  - example: model with quadratic running variable control

$$\bar{M}_a = \alpha + \rho D_a + \gamma_1 a + \gamma_2 a^2 + e_a$$

  - ideally, results are insensitive to degree of nonlinearity
  - Gelman and Imbens (2019) warning on using higher degree polynomials.
  - Polynomial selection procedures developed by Pei et al. (2018) - asymptotic mean squared error.
  - you must report how RD estimates change with model

RD Specifics

- Alternative strategy: different slopes left and right of cutoff
  - in practice, allow interactions of running variable a with D_a

$$\bar{M}_a = \alpha + \rho D_a + \gamma(a - a_0) + \delta[(a - a_0)D_a] + e_a$$

  - running variable centered around the cutoff $(a_0 = 21)$
  - subtle implication: away from cutoff a_0, MLDA treatment effect is given by $\rho + \delta(a - a_0)$
  - estimates away from cutoff constitute bold extrapolation
  - no data on counterfactual death rates in a world where drinking at ages substantially older than 21 is forbidden

RD Specifics

- Mixing both strategies

$$\bar{M}_a = \alpha + \rho D_a + \gamma_1(a - a_0) + \delta_1[(a - a_0)D_a] \\ + \gamma_2(a - a_0)^2 + \delta_2[(a - a_0)^2 D_a] + e_a$$

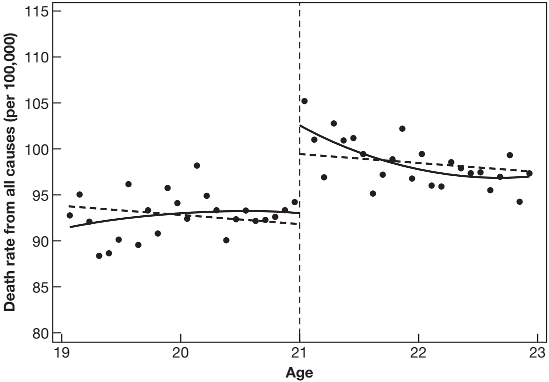
- Treatment effect away from cutoff a_0 is now given by

$$\rho + \delta_1(a - a_0) + \delta_2(a - a_0)^2$$

- Figure 4.4 shows trend function estimated by equation above
  - which model is better, fancy or simple?
  - no rules here, just thoughtful look at the data
  - ideally, results not highly sensitive to modeling choices

RD Specifics

Figure 4.4
Quadratic control in an RD design



Notes: This figure plots death rates from all causes against age in months. Dashed lines in the figure show fitted values from a regression of death rates on an over-21 dummy and age in months. The solid lines plot fitted values from a regression of mortality on an over-21 dummy and a quadratic in age, interacted with the over-21 dummy (the vertical dashed line indicates the minimum legal drinking age [MLDA] cutoff).

Sharp RD Estimates

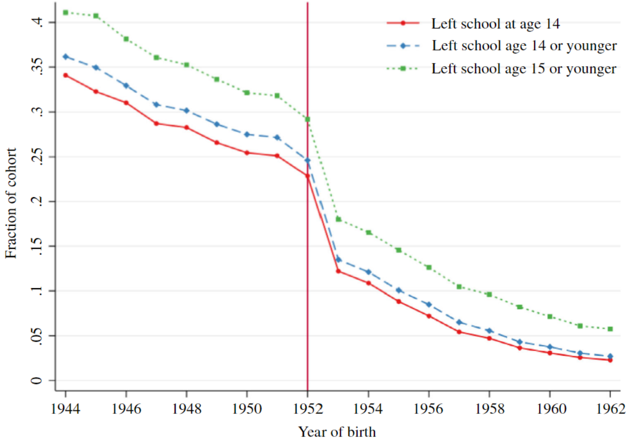
Table 4.1
Sharp RD estimates of MLDA effects on mortality

Dependent variable	Ages 19–22		Ages 20–21	
	(1)	(2)	(3)	(4)
All deaths	7.66	9.55	9.75	9.61
	(1.51)	(1.83)	(2.06)	(2.29)
Motor vehicle accidents	4.53	4.66	4.76	5.89
	(.72)	(1.09)	(1.08)	(1.33)
Suicide	1.79	1.81	1.72	1.30
	(.50)	(.78)	(.73)	(1.14)
Homicide	.10	.20	.16	−.45
	(.45)	(.50)	(.59)	(.93)
Other external causes	.84	1.80	1.41	1.63
	(.42)	(.56)	(.59)	(.75)
All internal causes	.39	1.07	1.69	1.25
	(.54)	(.80)	(.74)	(1.01)
Alcohol-related causes	.44	.80	.74	1.03
	(.21)	(.32)	(.33)	(.41)
Controls	age	age, age^2, interacted with over-21	age	age, age^2, interacted with over-21
Sample size	48	48	24	24

Notes: This table reports coefficients on an over-21 dummy from regressions of month-of-age-specific death rates by cause on an over-21 dummy and linear or interacted quadratic age controls. Standard errors are reported in parentheses.

Fuzzy RD example

(a) Effects of the reform on the fraction of the cohort leaving school before age 16



Fuzzy RD example

The first stage specification regresses the age at which individuals report leaving full time education (S_i) on the policy instrument variable, an indicator for whether their cohort was affected by the school leaving age increase (Z_i), controlling for f^1 and f^2, functions of the year of birth cohort before and respectively after the reform, as well as for survey year fixed effects λ_t.

$$S_i = \alpha_0 + \alpha_1 Z_i + f^1(B_i - C) + f^2(B_i - C) + \lambda_t + \epsilon_i \quad (1)$$

The 2SLS estimates are obtained by regressing the log of wages on years of completed schooling S_i, which are instrumented using the post-reform cohort indicator variable Z_i.

$$lnW_i = \gamma_0 + \gamma_1 \hat{S}_i + f^1(B_i - C) + f^2(B_i - C) + \lambda_t + \epsilon_i \quad (2)$$

Fuzzy RD example: Grenet(2013)

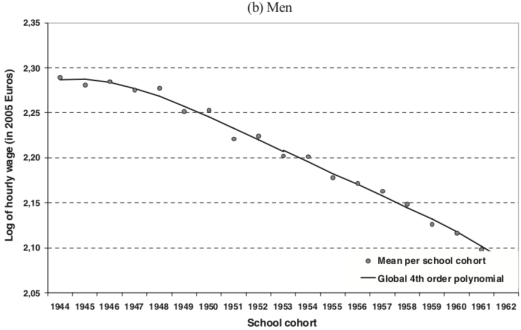
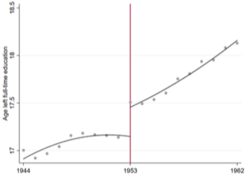
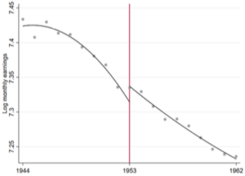


Fig. 5. Impact of the 1967 Berthoin reform in France on the log hourly wages (in 2005 euros), calculated for female and male wage earners separately (school cohorts 1944–1962)
Notes: The dots show the average log of hourly wage in France, grouped at the school cohort cell for the subsample of female and male wage earners who were born between 1944 and 1962. The solid line represents the fitted values from a global fourth-order polynomial regression, allowing for an intercept shift at the 1953 school cohort.

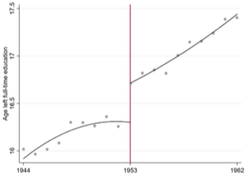
Fuzzy RD example: Domnisoru (2021)



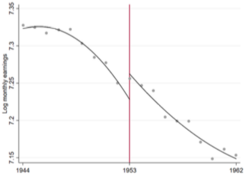
(a) Schooling, full sample



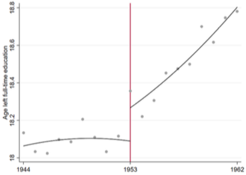
(b) Earnings, full sample



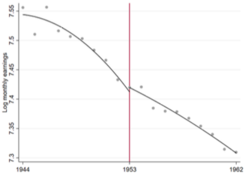
(c) Schooling, lower parental education



(d) Earnings, lower parental education

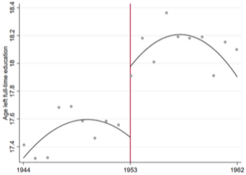


(e) Schooling, higher parental education

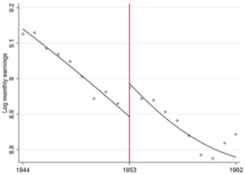


(f) Earnings, higher parental education

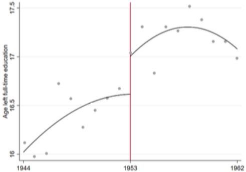
Fuzzy RD example: Domnisoru(2021)



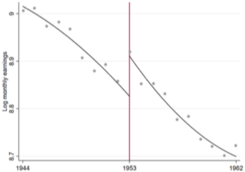
(a) Schooling, full sample



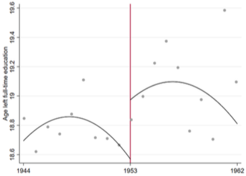
(b) Earnings, full sample



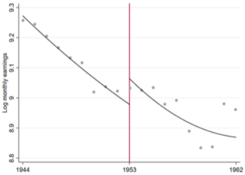
(c) Schooling, lower parental education



(d) Earnings, lower parental education



(e) Schooling, higher parental education



(f) Earnings, higher parental education

Nonparametric RD

- *Parametric* RD: straightforward regression estimation
- *Nonparametric* RD: estimation focusing on points close to the cutoff
  - compares averages in a narrow window just to the left and just to the right of the cutoff
  - problem of distinguishing jumps from nonlinear trends becomes less important as we zero in on points close to the cutoff
  - drawback: if window is too narrow, estimates are likely to be too imprecise to be useful
  - trade-off: reduction in bias near the boundary increases variance from throwing data away

Nonparametric RD

- Nonparametric RD amounts to estimating equation below in a narrow window around the cutoff, that is, $a_0 - b \leq a \leq a_0 + b$

$$\bar{M}_a = \alpha + \rho D_a + \gamma a + e_a$$

  - b describes the width of the window and is called *bandwidth*
  - local-linear kernel regression: you fit linear regressions to each observation in the data and their neighbouring observations, weighted by a smooth kernel distribution. The further away from the observation in question, the less weight the data contribute to that regression. When all the little linear components are added together, the resulting function is smooth.
  - Note we are not making any assumptions about the functional form
  - High bandwidth: high bias, low variance (more data points, farther from the cutoff)
  - Low bandwidth: low bias, high variance (fewer data points, closer to the cutoff)

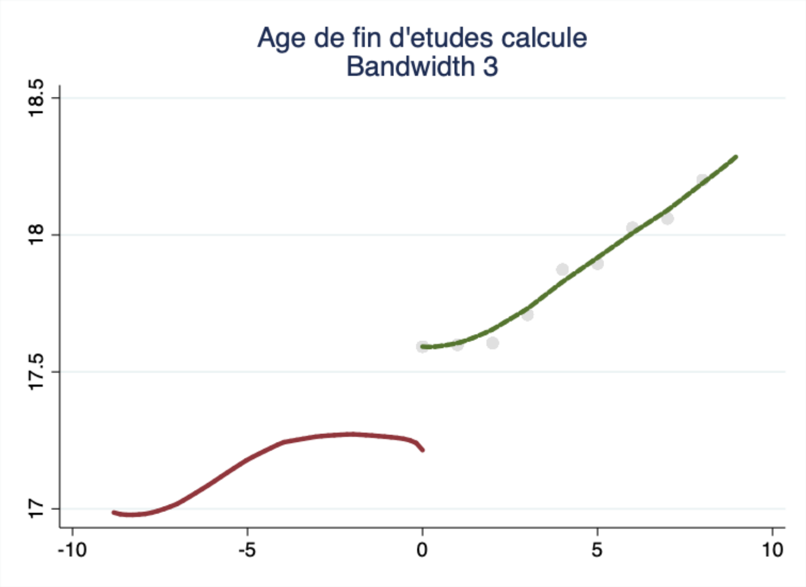

Age de fin d'etudes calcule
Bandwidth 3
18.5
18
17.5
17
-10
-5
0
5
10

Table 2: Effects of the Berthoin Reform on Educational Attainment and Monthly Earnings, LFS data, Local Linear Regression Estimates, Men

Bandwidth	First stage	(s.e.)	Reduced form	(s.e.)	2SLS	(s.e.)
A. Full sample						
4	.342***	(.032)	.009**	(.004)	.028**	(.011)
5	.320***	(.036)	.011***	(.003)	.034***	(.011)
6	.297***	(.043)	.011***	(.003)	.038***	(.010)
7	.270***	(.050)	.010**	(.004)	.038***	(.012)
8	.247***	(.053)	.008*	(.004)	.034**	(.014)
9	.230***	(.055)	.007	(.004)	.030*	(.016)
10	.217***	(.055)	.006	(.004)	.029*	(.017)
B. Parents in lower education occupations						
4	.370***	(.054)	.016***	(.004)	.042***	(.008)
5	.361***	(.047)	.019***	(.005)	.053***	(.009)
6	.345***	(.042)	.024***	(.005)	.069***	(.014)
7	.318***	(.049)	.025***	(.005)	.079***	(.018)
8	.298***	(.052)	.024***	(.006)	.080***	(.019)
9	.282***	(.053)	.021***	(.005)	.076***	(.019)
10	.282***	(.053)	.019***	(.005)	.076***	(.019)
C. Parents in higher education occupations						
4	.232***	(.065)	-.006	(.004)	-.027**	(.013)
5	.219**	(.095)	-.008	(.008)	-.038	(.046)
6	.190**	(.094)	-.009	(.007)	-.050	(.053)
7	.167*	(.092)	-.010	(.006)	-.064	(.061)
8	.140	(.088)	-.012*	(.006)	-.088	(.079)
9	.122	(.085)	-.012**	(.005)	-.100	(.093)
10	.109	(.082)	-.011	(.008)	-.105	(.103)

```
. rd learn agelfted dobyc [fweight=weight], kernel(tri) bw(6) cluster(clust) covar(duhab1 duhab2 duhab3 duhab4 duhab5 duhab6
Three variables specified; jump in treatment
at Z=0 will be estimated. Local Wald Estimate
is the ratio of jump in outcome to jump in treatment.

 Assignment variable Z is dobyc
 Treatment variable X_T is agelfted
 Outcome variable y is learn

(17,706 missing values generated)
(17,706 missing values generated)
(17,706 missing values generated)
Estimating for bandwidth 6
Estimating for bandwidth 3
Estimating for bandwidth 12
```

learn	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
numer	.0114453	.0036412	3.14	0.002	.0043087	.0185819
denom	.2975762	.0434297	6.85	0.000	.2124556	.3826968
lwald	.0384618	.0107652	3.57	0.000	.0173623	.0595612
numer50	.0186326	.0018255	10.21	0.000	.0150547	.0222105
denom50	.41717	.0015626	266.97	0.000	.4141074	.4202327
lwald50	.0446643	.0044652	10.00	0.000	.0359126	.0534159
numer200	.0058644	.004771	1.23	0.219	-.0034865	.0152153
denom200	.2053791	.0562355	3.65	0.000	.0951595	.3155987
lwald200	.0285541	.0183195	1.56	0.119	-.0073515	.0644597

Application (placebo tests): Impacts of Public Transit on Traffic Congestion

Application: Impacts of Public Transit on Traffic Congestion

- Public transit in the U.S.
  - 1 percent of passenger miles traveled
  - *but* it attracts strong public support
- Anderson's (2014) simple choice model
  - prediction: transit riders likely to be individuals commuting along routes with severe roadway delays
  - thus, riders have high marginal impacts on traffic congestion
- Testing model prediction
  - data from a strike in 2003 by Los Angeles transit workers
    - on October 14, 2003, Metropolitan Transportation Authority (MTA) workers began a strike that lasted 35 days and shut down MTA bus and rail lines
  - leveraging hourly data on traffic speeds for all major Los Angeles freeways, Anderson (2014) estimates a *sharp* RD design using time as the running variable

Application: Impacts of Public Transit on Congestion

- Findings
  - abrupt increase in average delays of 47 percent – 0.19 minutes per mile – during peak periods
    - impact many times larger than estimates in the literature
  - largest effects on freeways which parallel transit lines with heavy ridership
  - no effects during the same period in neighboring counties unaffected by the transit strike
- Implications
  - annualized congestion relief benefit of operating the LA transit system between $1.2 to $4.1 billion
    - $1.20 to $4.10 per peak-hour transit passenger mile
  - net benefits of transit systems much larger than expected

Application: Impacts of Public Transit on Congestion

- *Sharp* RD design estimating equation

$$y_{it} = \alpha + \beta strike_{it} + f(date_{it}) + \delta X_{it} + \epsilon_{it}$$

  - y_{it} is the average delay (in minutes per mile) for detector i during hour t
  - $strike_{it}$ is a binary variable equals to one when the strike is in effect and zero otherwise
  - $date_{it}$ is the date measured in days from the beginning of the strike
  - function $f(date_{it})$ is specified as $\gamma_1 date_{it} + \gamma_2(date_{it} \times strike_{it})$
  - X_{it} represents several control variables to increase the precision of the estimates (β will be unbiased even without the controls)

Application: Impacts of Public Transit on Congestion

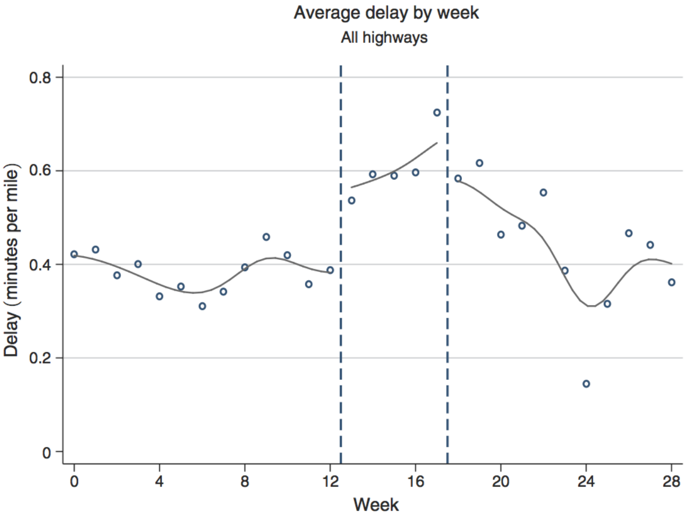


FIGURE 2. WEEKLY PEAK HOUR DELAY ON MAJOR LOS ANGELES FREEWAYS, 7/14/2003 TO 1/30/2004

Application: Impacts of Public Transit on Congestion

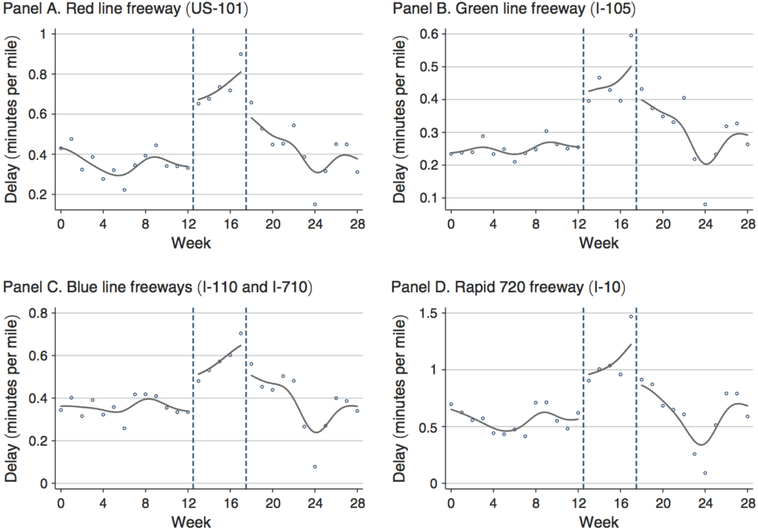


FIGURE 3. AVERAGE WEEKLY PEAK HOUR DELAY ON SPECIFIC LOS ANGELES FREEWAYS, 7/14/2003 TO 1/30/2004

Application: Impacts of Public Transit on Congestion

TABLE 4—EFFECT OF STRIKE ON DELAYS DURING ALL PEAK HOURS

Average delay (in minutes per mile)	(1)	(2)	(3)	(4)	(5)	(6)
Strike	0.194	0.332	0.218	0.190	0.357	0.125
	(0.041)	(0.076)	(0.052)	(0.051)	(0.128)	(0.042)
Date	−0.004	−0.003	−0.002	−0.003	−0.005	−0.005
	(0.002)	(0.003)	(0.002)	(0.002)	(0.004)	(0.002)
Date × strike	0.007	0.006	−0.001	0.007	0.012	0.007
	(0.002)	(0.003)	(0.002)	(0.003)	(0.007)	(0.002)
Average delay prestrike	0.409	0.369	0.264	0.357	0.600	0.434
Freeways	All	101	105	110 and 710	10	Other
Parallel transit line		Red line	Green line	Blue line	Rapid 720	
Sample size	178,549	15,854	31,058	19,152	15,357	97,128

Notes: Each column represents a separate VMT-weighted regression, with weights equal to (length of highway covered by detector i) × (average prestrike traffic flow over detector i). The observation is the detector-hour, and the sample is limited to weekdays from 7–10 AM and 2–8 PM within 28 days of the strike's beginning. Parentheses contain clustered standard errors that are robust to within-day and within-detector serial correlation. All regressions include day-of-week and detector fixed effects.

Application: Impacts of Public Transit on Congestion

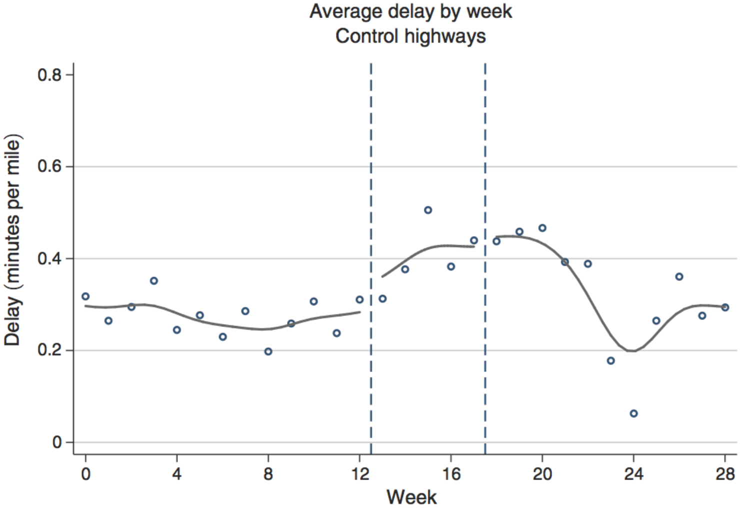


FIGURE 5. WEEKLY PEAK HOUR DELAY ON ORANGE/VENTURA COUNTY FREEWAYS, 7/14/2003 TO 1/30/2004

Application: Impacts of Public Transit on Congestion

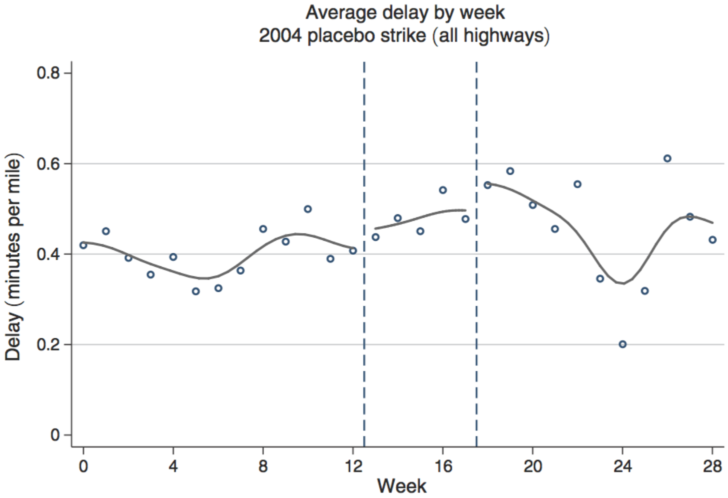


FIGURE 6. WEEKLY PEAK HOUR DELAY ON MAJOR LOS ANGELES FREEWAYS ONE YEAR LATER, 7/14/2004 TO 1/30/2005

Application: Impacts of Public Transit on Congestion

Table 9—Effect of Placebo Strikes on Delays

Falsification sample:	Orange and Ventura counties			October/November 2004		
Dependent variable: Average delay (in minutes per mile)	(1)	(2)	(3)	(4)	(5)	(6)
Strike	0.024	0.025	0.026	0.060	0.082	0.045
	(0.027)	(0.037)	(0.037)	(0.042)	(0.054)	(0.064)
Date	0.000	−0.001	0.000	−0.002	−0.002	−0.003
	(0.001)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)
Date × strike	0.004	0.000	0.006	0.005	0.002	0.007
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
Average delay prestrike	0.205	0.170	0.219	0.433	0.539	0.386
Hours	All peak	AM peak	PM peak	All peak	AM peak	PM peak
Sample size	13,149	4,296	8,853	177,572	59,532	118,034

Notes: Each column represents a separate VMT-weighted regression, with weights equal to (length of highway covered by detector i) × (average prestrike traffic flow over detector i). The observation is the detector-hour, and the sample is limited to weekdays from 7–10 AM and 2–8 PM within 28 days of the strike's beginning. Parentheses contain clustered standard errors that are robust to within-day and within-detector serial correlation. All regressions include day-of-week and detector fixed effects. In columns 1–3 the strike variable is defined normally but the sample contains detectors in neighboring counties not subject to the strike. In columns 4–6 the strike variable equals zero prior to October 12, 2004 and unity after October 12, 2004.

Manipulation of the running variable

Manipulation of the running variable

- Manipulation of the running variable generates selection bias
- Hypothetical example (McCrary, 2008): a doctor plans to randomly assign heart patients to a statin and a placebo to study the effect of the statin on heart attack within 10 years
  - the doctor randomly assigns patients to two different waiting rooms, A and B, and plans to give those in A the statin and those in B the placebo
  - if some of the patients learn of the planned treatment assignment mechanism, we would expect them to proceed to waiting room A
  - if the doctor does not learn about the patients' actions and follows the original protocol, random assignment of patients to separate waiting rooms may be undone by patient sorting after random assignment

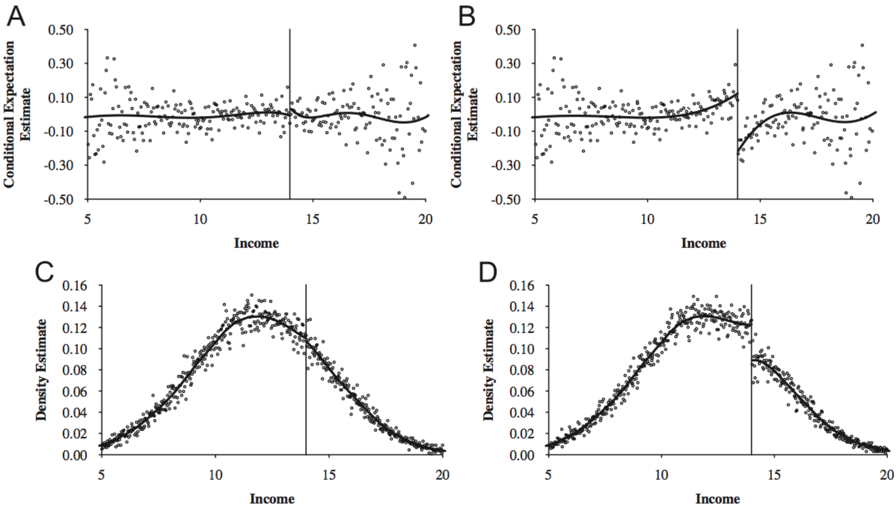
Manipulation of the running variable: McCrary density test

- McCrary (2008) proposes a formal test for selection bias
  - test is based on the intuition that, in the previous example, we would expect for waiting room A to become crowded
  - in the RD context, this is analogous to expecting the running variable to be discontinuous at the cutoff, with surprisingly many individuals just barely qualifying for a desirable treatment assignment and surprisingly few failing to qualify

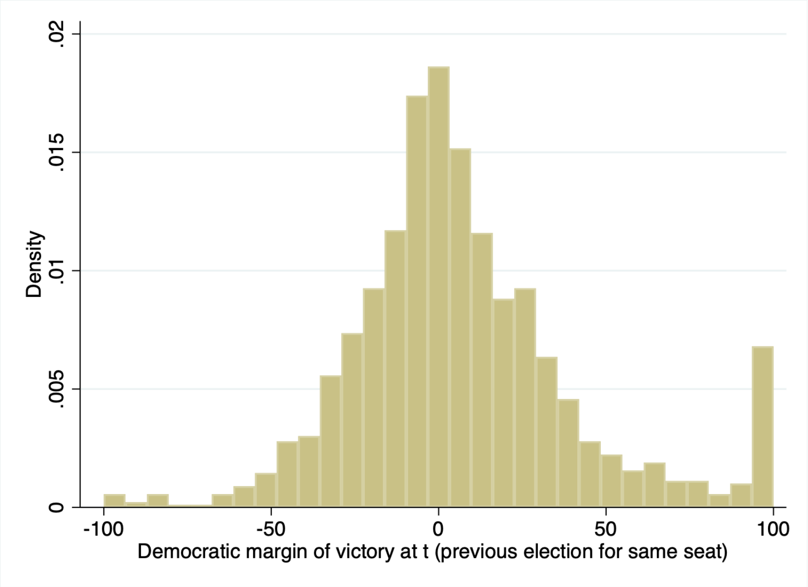
Manipulation of the running variable: McCrary density test

- Illustration with another hypothetical example: gaming the system with an income-tested job training program (figure depicted in the next slide)
  (A) labor supply response to treatment with no pre-announcement and no manipulation
  (B) labor supply response to treatment with pre-announcement and manipulation
  (C) density of income with no pre-announcement and no manipulation
  (D) density of income with pre-announcement and manipulation

Manipulation of the running variable: McCrary density test



Labor Supply Reduction or Selection Bias?



Density
0
.005
.01
.015
.02
-100
-50
0
50
100
Democratic margin of victory at t (previous election for same seat)

```
.  use rddensity_senate.dta

. rddensity margin
Computing data-driven bandwidth selectors.

Point estimates and standard errors have been adjusted for repeated observations.
(Use option nomasspoints to suppress this adjustment.)

RD Manipulation test using local polynomial density estimation.

     c =       0.000 | Left of c  Right of c          Number of obs =         1390
---------------------+----------------------          Model         = unrestricted
     Number of obs   |       640         750          BW method     =         comb
Eff. Number of obs   |       408         460          Kernel        =   triangular
    Order est. (p)   |         2           2          VCE method    =    jackknife
    Order bias (q)   |         3           3
       BW est. (h)   |    19.841      27.119

Running variable: margin.
--------------------------------------------
              Method |       T          P>|T|
---------------------+----------------------
              Robust |   -0.8753      0.3814
--------------------------------------------

P-values of binomial tests. (H0: prob = .5)
-------------------------------------------------------
 Window Length / 2   |      <c          >=c  |    P>|T|
---------------------+-----------------------+---------
               0.430 |       8           12  |  0.5034
               0.861 |      17           25  |  0.2800
               1.291 |      25           34  |  0.2976
               1.722 |      45           47  |  0.9170
               2.152 |      51           55  |  0.7709
               2.583 |      66           65  |  1.0000
               3.013 |      79           71  |  0.5678
               3.444 |      94           86  |  0.6020
               3.874 |     105           94  |  0.4785
               4.305 |     115          107  |  0.6386
-------------------------------------------------------
```

Placebo test on pretreatment X- Lee, Moretti, and Butler (2004)