RD Examples

1

$$(1) \quad y_{ija} = X_{ija}\alpha_j + f_j(a) + C_{ija}\delta_j + u_{ija} \, ,$$

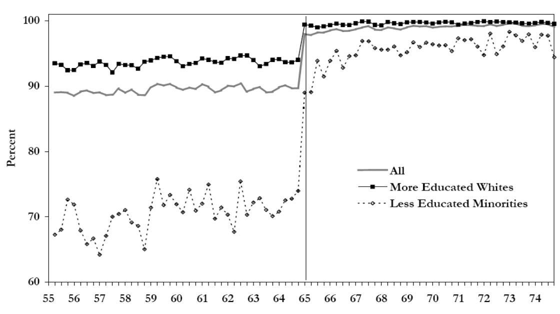
$$(2) \quad C_{ija} = X_{ija}\,\beta^C{}_j + g^C{}_j(a) + D_a\,\pi^C{}_j + v^C{}_{ija} \, ,$$

$$(3) \quad y_{ija} = X_{ija}\,\beta^y{}_j + g^y{}_j(a) + D_a\,\pi^y{}_j + v^y{}_{ija} \, ,$$

$$(4) \quad \delta^{IV}{}_j = \pi^y{}_j \, / \, \pi^C{}_j \, .$$

2

Figure 1: Health Insurance Coverage Rates by Age, 1992-2001 NHIS



3

Table 3: Regression Discontinuity Estimates of Change in Health Insurance Coverage at Age 65

	Model fit to Aggregrated Age Cells (1)	Models fit to Micro Data: Linear Prob. no controls (2)	Linear Prob. w/ controls (3)	Probit w/ controls (4)	Linear Prob. Model for Medicare (5)
All	8.41	8.41	8.60	6.05	67.32
	(0.30)	(0.26)	(0.24)	(0.20)	(1.49)
By Ethnicity/Education:					
High Education White Nonhispanic	5.28	5.28	5.28	3.95	72.59
	(0.34)	(0.26)	(0.27)	(0.32)	(1.41)
Low Education White Nonhispanic	14.21	14.21	14.43	10.86	63.26
	(1.08)	(1.42)	(1.38)	(1.21)	(1.42)
High Education Minority	9.03	9.03	9.13	8.92	57.80
	(1.47)	(1.15)	(1.13)	(1.05)	(2.53)
Low Education Minority	18.83	18.83	19.41	15.93	48.70
	(1.73)	(1.63)	(1.56)	(1.87)	(3.03)

4

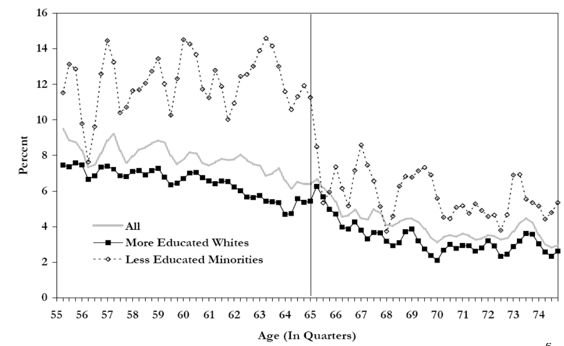
Table 3: Regression Discontinuity Estimates of Change in Health Insurance Coverage at Age 65

	Model fit to Aggregrated Age Cells (1)	Models fit to Micro Data: Linear Prob. no controls (2)	Linear Prob. w/ controls (3)	Probit w/ controls (4)	Linear Prob. Model for Medicare (5)
By Ethnicity:					
White Non-Hispanics	7.26	7.26	7.30	5.10	70.56
	(0.31)	(0.28)	(0.24)	(0.27)	(1.32)
Black Non-Hispanics	13.64	13.64	14.19	12.27	55.14
	(1.53)	(1.12)	(1.08)	(1.33)	(3.10)
Hispanics	17.30	17.30	16.88	14.25	51.87
	(2.16)	(2.02)	(1.93)	(1.77)	(2.84)
By Gender:					
Men	7.19	7.19	7.32	5.34	65.40
	(0.51)	(0.54)	(0.51)	(0.40)	(1.24)
Women	9.50	9.50	9.75	6.67	69.03
	(0.49)	(0.60)	(0.57)	(0.43)	(1.88)

Note: Table entries represent estimated coefficient of dummy for age 65 or older in models for the probability of insurance coverage (columns 1-4) or Medicare coverage (column 5). All models include quadratic in age (in quarters) fully interacted with dummy for age 65 or older. Models in columns (2)-(5) are

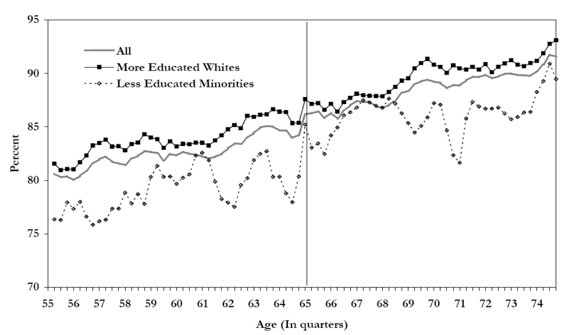
5

Figure 4: Percent Who Delayed Medical Care Last Year for Cost Reasons



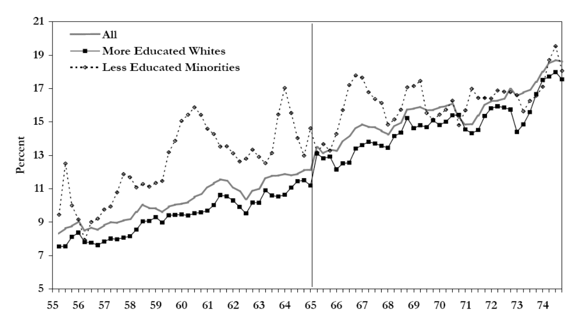
6

Figure 6: Percent with At Least One Doctor Visit in Past Year by Age, 1992-2001 NHIS



7

Figure 7: Percent with One or More Hopital Stays in Past Year by Age, 1992-2001 NHIS



8

Table 4: Regression Discontinuity Estimates of Effect of Reaching Age 65 on Health Care Access and Utilization

	Delayed Care Last Year for Cost Reasons (1)	Did Not Get Care Last Year for Cost Reasons (2)	Needed to See Dr. Last Year, Didn't Because of Cost (3)	Has Usual Place for Preventative Care (4)	Saw Doctor at Least Once Last Year (5)	Spent 1 or More Nights in Hospital Last Year (6)
All	-0.56	-0.23	-2.70	1.40	0.97	1.18
	(0.64)	(0.52)	(0.56)	(0.86)	(0.68)	(0.59)
By Ethnicity/Education:						
High Education White Nonhispanic	0.31	0.21	-1.76	0.36	0.16	1.07
	(0.83)	(0.46)	(0.47)	(1.18)	(0.93)	(0.72)
Low Education White Nonhispanic	-1.12	0.20	-4.60	3.88	1.79	0.95
	(1.40)	(1.65)	(1.95)	(2.46)	(1.34)	(1.58)
High Education Minority	-1.77	-0.60	-3.54	3.74	-0.22	2.11
	(1.61)	(1.19)	(1.36)	(1.53)	(1.91)	(1.15)
Low Education Minority	-5.06	-4.39	-8.90	0.99	5.63	1.09
	(2.02)	(1.41)	(2.42)	(2.04)	(2.06)	(1.93)

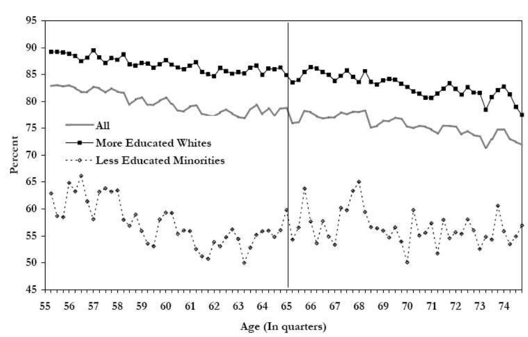
9

Table 5: Regression Discontinuity Estimates of Effect of Reaching Age 65 on Preventative Care

	Had Flu Shot in Past Year (1)	Had Blood Cholesterol Checked in Past 2 Yrs (2)	Mammogram in Past 2 Yrs (3)	Ever Had Mammogram (4)	Had PSA or Rectal Exam in Past 2 Yrs (5)	Ever Had PSA or Rectal Exam (6)	Ever Diagnosed with Hypertension (7)
All	1.03	1.45	1.02	-0.18	1.08	0.21	0.66
	(1.03)	(0.64)	(0.82)	(0.58)	(1.64)	(1.07)	(1.85)
By Ethnicity/Education:							
High Education White Nonhispanic	0.74	0.40	1.86	0.58	1.00	-0.01	-1.66
	(1.23)	(0.93)	(0.97)	(0.69)	(1.42)	(1.10)	(2.42)
Low Education White Nonhispanic	4.91	9.99	0.02	-2.72	-3.56	-3.43	2.78
	(1.15)	(1.74)	(2.75)	(2.15)	(5.12)	(3.43)	(3.59)
High Education Minority	-0.32	0.57	-4.42	-2.33	4.85	3.58	9.15
	(2.70)	(2.23)	(2.30)	(1.25)	(2.26)	(1.54)	(3.96)
Low Education Minority	2.32	3.73	5.83	0.56	4.49	0.90	3.34
	(3.11)	(2.31)	(2.55)	(1.43)	(5.20)	(4.37)	(5.19)

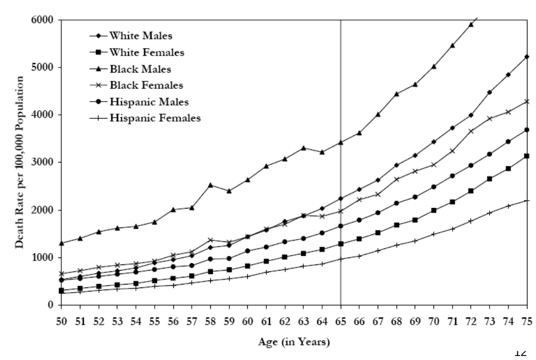
10

Figure 13: Percent in Good, Very Good, or Excellent Health by Age, 1992-2001 NHIS



11

Figure 14. Age-Specific Death Rates by Sex and Race



12

F-test

- $y_i = \beta_0 + x_{1i}\beta_1 + x_{2i}\beta_2 + x_{3i}\beta_3 + x_{4i}\beta_4 + x_{5i}\beta_5 + \varepsilon_i$
- Want to test the restriction that a set of parameters equals zero
- H_o: $\beta_3 = \beta_4 = \beta_5 = 0$
- Model above is unrestricted model

13

- N observations in the unrestricted model
- K parameters in the unrestricted model
- DOF = N-K
- J restrictions (count the equal signs)
- If null is true, the F test statistic should be distributed as an F with (J,N-K) degree of freedom

14

- Restricted model
- $y_i = \beta_0 + x_{1i}\beta_1 + x_{2i}\beta_2 + \varepsilon_i$
- $F = [(SSE_R - SSE_u)/J]/[SSE_u/(N-k)]$
- SSE_j sum of squared error in the restricted (R) and unrestricted (u) model

15

- . * run basic log wage regression controlling for;
- . * quadratic in age, education, nonwhite and union;
- . * dummy variables. There are 3 smsa levels and;
- . * four census regions. to add these effects to the model;
- . * would require 2 smsa and 3 regional effects;
- . * we want to test that the coefficients in these 5;
- . * coefficients are all zero;

16

- . * construct smsa and region effects;
- . xi i.smsa i.region;
- i.smsa _Ismsa_1-3 (naturally coded; _Ismsa_1 omitted)
- i.region _Iregion_1-4 (naturally coded; _Iregion_1 omitted)

- **. * the unrestricted model;**
- **. * run regression with smsa and region and race fixed-effects;**
- . * run regression with smsa and region and race fixed-effects;
- . reg earnwkl age age2 educ nonwhite union _I*;

SSE_u N-K N

```
      Source |       SS       df       MS              Number of obs =   19906
-------------+------------------------------           F( 10, 19895) = 1011.40
       Model |  1765.82128    10  176.582128           Prob > F      =  0.0000
    Residual |  3473.51741 19895  .174592481           R-squared     =  0.3370
-------------+------------------------------           Adj R-squared =  0.3367
       Total |  5239.33869 19905  .263217216           Root MSE      =  .41784

------------------------------------------------------------------------------
     earnwkl |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0702362   .0019649    35.75   0.000     .0663848    .0740876
        age2 |   -.000706    .000024   -29.40   0.000    -.0007531    -.000659
        educ |   .0638027   .0011181    57.06   0.000     .0616111    .0659943
    nonwhite |  -.2118622   .0090022   -23.53   0.000    -.2295073    -.194217
       union |   .1120935   .0072515    15.46   0.000     .0978798    .1263071
    _Ismsa_2 |  -.1235448   .0072032   -17.15   0.000    -.1376637    -.109426
    _Ismsa_3 |  -.2067638   .0078561   -26.32   0.000    -.2221623   -.1913653
  _Iregion_2 |  -.0096336   .0085916    -1.12   0.262    -.0264739    .0072068
  _Iregion_3 |  -.0298032   .0084302    -3.54   0.000    -.0463272   -.0132793
  _Iregion_4 |    .034886   .0089323     3.91   0.000     .0173781     .052394
       _cons |   3.775609   .0390658    96.65   0.000     3.699037    3.852181
------------------------------------------------------------------------------
```


- . * do the f-test that the 5 smsa/region effects are;
- . * all zero;
- . test _Ismsa_2 _Ismsa_3 _Iregion_2 _Iregion_3 _Iregion_4;
- (1) _Ismsa_2 = 0
- (2) _Ismsa_3 = 0
- (3) _Iregion_2 = 0
- (4) _Iregion_3 = 0
- (5) _Iregion_4 = 0
- F(5, 19895) = 171.17
- Prob > F = 0.0000

SSE_r

```
. * do the test brute force;
. *run the restricted model without smsa/region effects;
. reg earnwkl age age2 educ nonwhite union;

      Source |       SS       df       MS              Number of obs =   19906
-------------+------------------------------           F(  5, 19900) = 1775.70
       Model |  1616.39963     5  323.279927           Prob > F      =  0.0000
    Residual |  3622.93905 19900  .182057239           R-squared     =  0.3085
-------------+------------------------------           Adj R-squared =  0.3083
       Total |  5239.33869 19905  .263217216           Root MSE      =  .42668

------------------------------------------------------------------------------
     earnwkl |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0679808   .0020033    33.93   0.000     .0640542    .0719075
        age2 |  -.0006778   .0000245   -27.69   0.000    -.0007258   -.0006299
        educ |    .069219   .0011256    61.50   0.000     .0670127    .0714252
    nonwhite |  -.1716133   .0089118   -19.26   0.000    -.1890812   -.1541453
       union |   .1301547   .0072923    17.85   0.000     .1158612    .1444481
       _cons |   3.630805   .0394126    92.12   0.000     3.553553    3.708057
------------------------------------------------------------------------------
```

- SSE_u = 3473.52
- N = 19906
- K = 11
- N-K = 19895
- J =5
- SSE_R = 3622.94
- F = [(3473.52-3622.94)/5]/(3473.52/19895)
- F = 171.16

21

What order polynomial?

- Model from Card/Lee paper
- (2) $Y_{ij} = D_j\beta_o + h(x_j) + \varepsilon_{ij}$
- x takes on discrete values $[x_1, x_2 \ldots x_j]$
- x_k=0 at discontinuity
- $D_j=1[x_j>0]$
- $h(x_j)$ polynomial in x

22

- (2) is a restricted model – parametric form of the regression with k parameters (k-1 in h and one for β_o)
- Unrestricted model, run regression with complete set of dummies for x_j
- (2a) $Y_{ij} = u_j + \varepsilon_{ij}$
- J dummies in total, N-J DOF

23

- Goodness of fit test

$$G = \frac{(ESS_r - ESS_u)/(J-K)}{ESS_u/(N-J)}$$

- Under null that h(x) captures the time series characteristics, G is distributed as an F(J-K,N-J)

24

- One alteration – what if there are m other right hand side variables?
- (2) $Y_{ij} = D_j\beta_o + Z_{ij}\gamma + h(x_j) + \varepsilon_{ij}$

- where Z is 1 x M vector.

$$G = \frac{(ESS_r - ESS_u)/(J-K)}{ESS_u/(N-J-M)}$$

25

Restricted model

```
. reg insured male white black hispanic _Ie* _Iyear* index index2 index_age65
> index2_age65 age65;

      Source |       SS       df       MS              Number of obs =   46950
-------------+------------------------------           F( 19, 46930) =  260.17
       Model |  334.010957    19  17.5795241           Prob > F      =  0.0000
    Residual |   3170.9949 46930  .067568611           R-squared     =  0.0953
-------------+------------------------------           Adj R-squared =  0.0949
       Total |  3505.00586 46949  .074655602           Root MSE      =  .25994
```

N-1 (arrow to 46949)

SSE_R (arrow to 3170.9949)

N-K-M (arrow to 46930)

26

Unrestricted Model

```
. * run unrestricted model that has dummies for all quarters instead of polyno
> mial;
. * use for test in card/lee, equation (3);
. reg insured male white black hispanic _Ie* _Iyear* _Iage*;

      Source |       SS       df       MS              Number of obs =   46950
-------------+------------------------------           F( 93, 46856) =   54.01
       Model |   339.36356    93  3.64907053           Prob > F      =  0.0000
    Residual |   3165.6423 46856  .067561087           R-squared     =  0.0968
-------------+------------------------------           Adj R-squared =  0.0950
       Total |  3505.00586 46949  .074655602           Root MSE      =  .25993

------------------------------------------------------------------------------
     insured |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        male |   .0079442   .0024272     3.27   0.001     .0031869    .0127014
```

SSE_u (arrow to 3165.6423)

N-J-M (arrow to 46856)

J+M-1 (no constant) (arrow to 93)

27

- SSE_u = 3165.64
- N-J-M = 46856
- J+M = 93
- SSE_r =3170.99
- K+M = 19
- J – K = 93-19 = 74
- F = 1.07
- = [(3170.99-3165.64)/74)/[3165.64/(46856)]

28

- P-value F=1.07 w/ 74 and 46,856 Degrees of freedom is 0.31
- Cannot reject null that h(x) and full set of effects for x explain the same time pattern in y

29

```
. gen age65=age_qtr>259;

. * scale the age in quarters index so that it equals 0;
. * in the month you become eligible for Medicare;
. gen index=age_qtr-260;

. gen index2=index*index;

. gen index3=index*index*index;

. gen index4=index2*index2;

. gen index_age65=index*age65;

. gen index2_age65=index2*age65;

. gen index3_age65=index3*age65;

. gen index4_age65=index4*age65;
```

30

```
. reg insured male white black hispanic _Ie* _Iyear* index index2 index_age65
> index2_age65 age65, cluster(index);

                                                       R-squared     =  0.0953
Number of clusters (index) = 80                        Root MSE      =  .25994
------------------------------------------------------------------------------
             |               Robust
     insured |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        male |   .0077525   .0026721     2.90   0.005     .0024337    .0130712
       white |    .039912   .0074113     5.39   0.000     .0251602    .0546637
                              Delete parameters

_Iyear_2001 |  -.0002411   .0031174    -0.08   0.939    -.0064461    .0059638
       index |   .0008314   .0006618     1.26   0.213    -.0004858    .0021486
      index2 |   .0000103   .0000163     0.64   0.526     -.000022    .0000427
 index_age65 |   .0009125   .0009527     0.96   0.341    -.0009838    .0028088
index2_age65 |  -.0000391   .0000223    -1.75   0.083    -.0000834    5.27e-06
       age65 |   .0910622   .0085334    10.67   0.000     .0740769    .1080474
       _cons |   .6820244   .0164487    41.46   0.000      .649284    .7147647
------------------------------------------------------------------------------
```

Components of h(x), quadratic in the index
Quadratic index after quarter 259

Coefficient on D
In Card/Lee
Notation

31

```
. reg hosp_12m male white black hispanic _Ie* _Iyear* index index2 index_age65
>  index2_age65 age65, cluster(index);

Regression with robust standard errors                 Number of obs =   46950
                                                       F( 19,    79) =   34.44
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.0128
Number of clusters (index) = 80                        Root MSE      =  .33722

------------------------------------------------------------------------------
             |               Robust
    hosp_12m |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        male |   .0146667   .0034806     4.21   0.000     .0077388    .0215947

                                 delete some results

index2_age65 |   .0000438   .0000274     1.60   0.114    -.0000108    .0000984
       age65 |   .0253064   .0085476     2.96   0.004     .0082929    .0423199
       _cons |   .1134856   .0065795    17.25   0.000     .1003894    .1265818
------------------------------------------------------------------------------
```

32

Results for different outcomes

Outcome	Coef (std error) on AGE 65
Have Insurance	0.091 (0.0085)
In good health	0.0070 (0.0102)
Delayed medical care	-0.0048 (0.0064)
Did not get medical care	-0.0021 (0.0050)
Hosp visits in 12 months	0.0253 (0.0085)

33

Sensitivity of health insurance results

Order of index	Order of post-65 index	Coef (std error) on age 65 Insured	Hosp in 12
2	2	0.091 (0.0085)	0.0250 (0.0086)
1	0	0.101 (0.0058)	0.0147 (0.0061)
3	3	0.084 (0.0106)	0.0098 (0.0094)
4	4	0.073 (0.013)	0.020 (0.0109)

34