Spatial regression discontinuity: Estimating effects of geographically implemented programs and policies

Christopher T. Moore
Quantitative Methods in Education (QME)
University of Minnesota
http://umn.edu/~moor0554

Introduction

- Estimating causal effects is an important aim in the field of program evaluation, and randomized field experiments are considered a gold standard (Boruch, 1991).
- However, many programs and policies are implemented in geographically defined jurisdictions, such as school districts or states, and not by randomly assigning participants to a treatment or control group.
- How might evaluators estimate causal effects in the case of treatment assignment based on geographic borders?

Spatially enabled evaluation

- Longitudinal analysis is commonly used in applied educational research, but spatial analysis is underutilized (Renger et al., 2002; Tate, 2008).
- *Spatially enabled* social science disciplines, such as public health and economics, regularly use geographic maps and spatial methods to form research questions, to sample, collect, and analyze data, and to disseminate results (Waller & Gotway, 2004).

Spatially enabled evaluation

- Patton (1997) defines evaluation as: "The systematic collection of information about the activities, characteristics, and outcomes of programs to make judgments about the program, improve program effectiveness, and/or inform decisions about future programming."
- I define spatially enabled evaluation simply as *evaluation, policy analysis, or applied research made possible with spatially referenced data*.

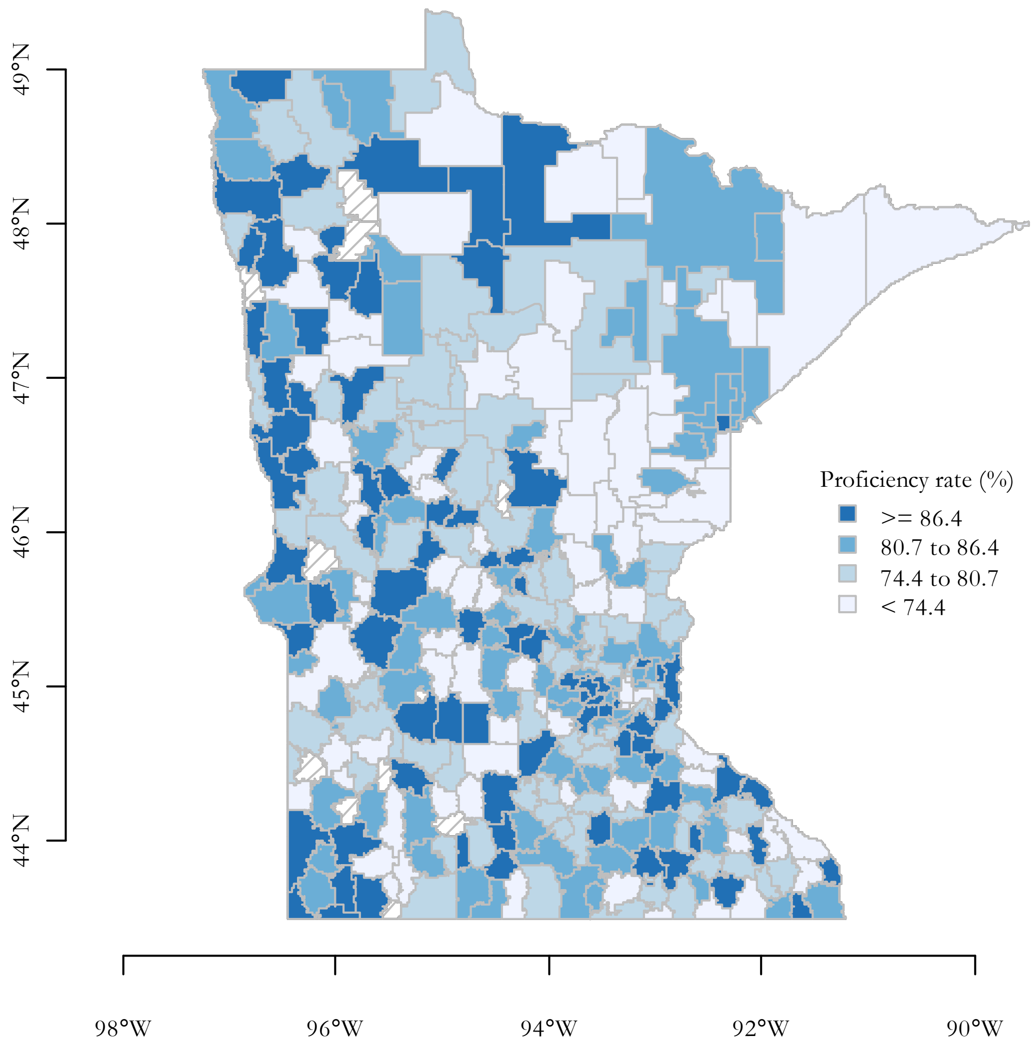
Spatially enabled evaluation: Some promising applications

- Mapping
  - Promote evaluation stakeholders' participation and understanding of evaluation findings (Craig & Elwood, 1998; Verdi & Kulhavy, 2002)
- Research design
  - Plan and implement surveys (Brown, 2005; Talen & Shah, 2007)
  - Randomly assign areas to treatment conditions (Raudenbush, 1997)
  - Propensity score matching with spatial predictors (Bondonio, 2002)
  - Spatial regression discontinuity at geographic borders (Holmes, 1998)
- Spatial statistical analysis
  - Identify geographic clusters and account for spatial dependencies (Moore, 2009)

Spatially enabled evaluation: Mitigating concerns

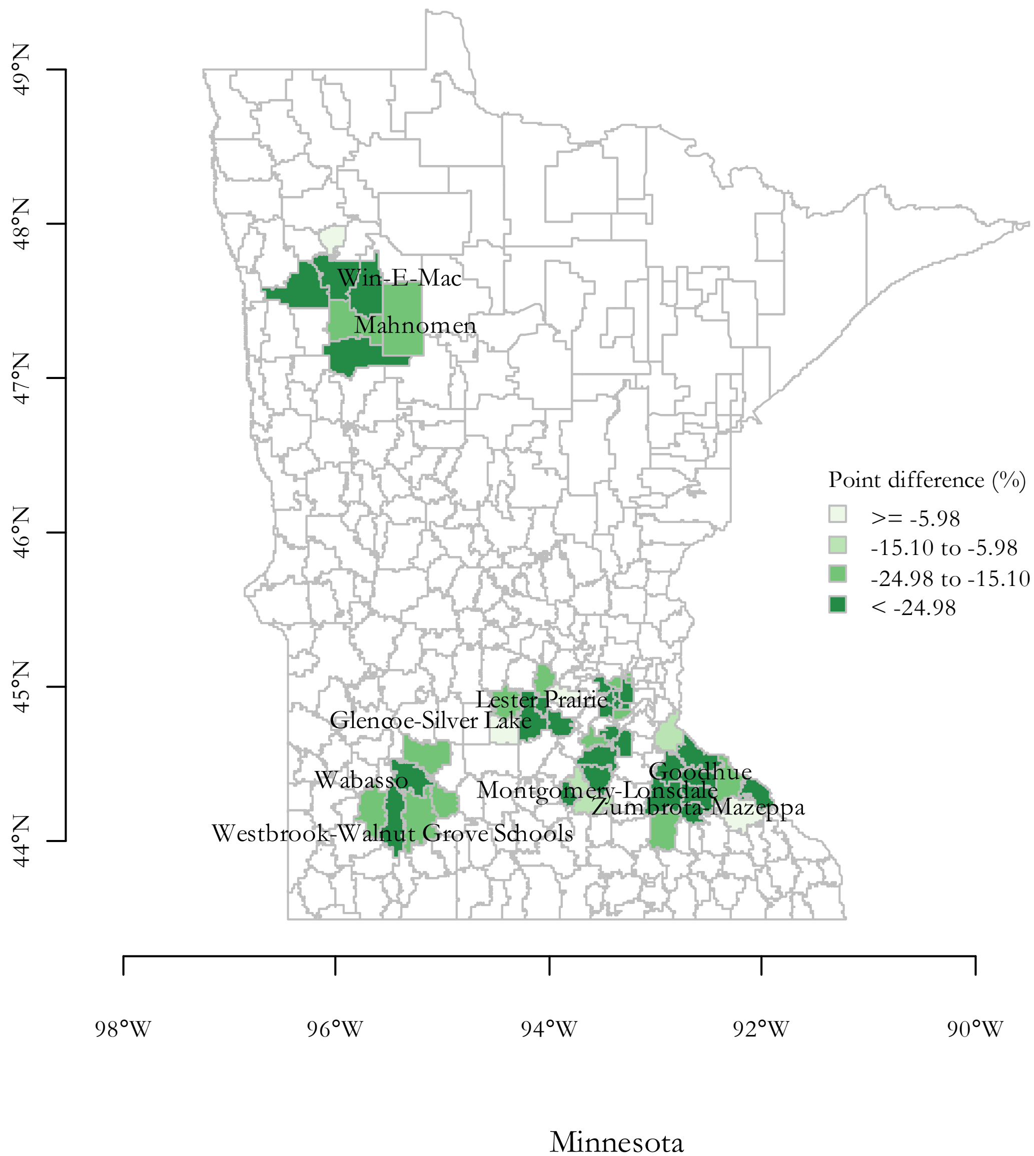
- Maps are inherently inaccurate and prone to mislead (Monmonier, 1996)
- Mere visual decoration and distraction (Carney & Levin, 2002)
- Violation of participants' privacy (where they live; Banerjee, Carlin, & Gelfand, 2004)
- Spatial autocorrelation complicates spatial statistical analysis (Anselin et al., 1996)

School districts: Math proficiency among third graders



Minnesota

Early reading intervention clusters: Economic proficiency gap



Modern causal theory

- Conditions necessary for causal inference (Mill, 1846; Shadish, Cook, & Campbell, 2002)
  - Theorized cause must *precede* observed effect.
  - Cause and effect must be *related*.
  - *No competing explanations* can falsify the inference.
- Rubin's (1974) *potential outcomes model*
  - *Counterfactuals* impossible to observe
  - *Average treatment effect*
    - An accurate estimate of the true causal effect if the treatment mechanism is *strongly ignorable*
    - Randomized experiments: the *simplest* but not the only design that supports valid causal inference (Holland, 1986)

Regression discontinuity

- A quasi-experimental design and statistical modeling approach.
- Uses a continuous assignment variable exhibited by participants (e.g., test scores).
- A cutoff point along the assignment variable determines their treatment assignment.
- Yields a *local average treatment effect* (Imbens & Lemieux, 2008).
  - High internal validity
  - Low external validity

Regression discontinuity

- Federal agencies and evaluation theorists favor regression discontinuity (U.S. Department of Education, 2005; Cook, 1991).
- Well-designed regression discontinuity studies yield causal estimates that are comparable to those derived from randomized controlled trials (Galindo and Shadish, 2008; Cook, Shadish, & Wong, 2008).
- The treatment assignment process is completely known and perfectly measured—a feature that prospective regression discontinuity shares with randomized controlled trials (Shadish, Cook, and Campbell, 2002).

Regression discontinuity: Functional form

- Specifying a good fitting functional form is key to unbiased estimates.
- If the data do not adhere to a functional form, then a non-parametric "smoother" approach may be more appropriate.

Regression discontinuity: Functional form

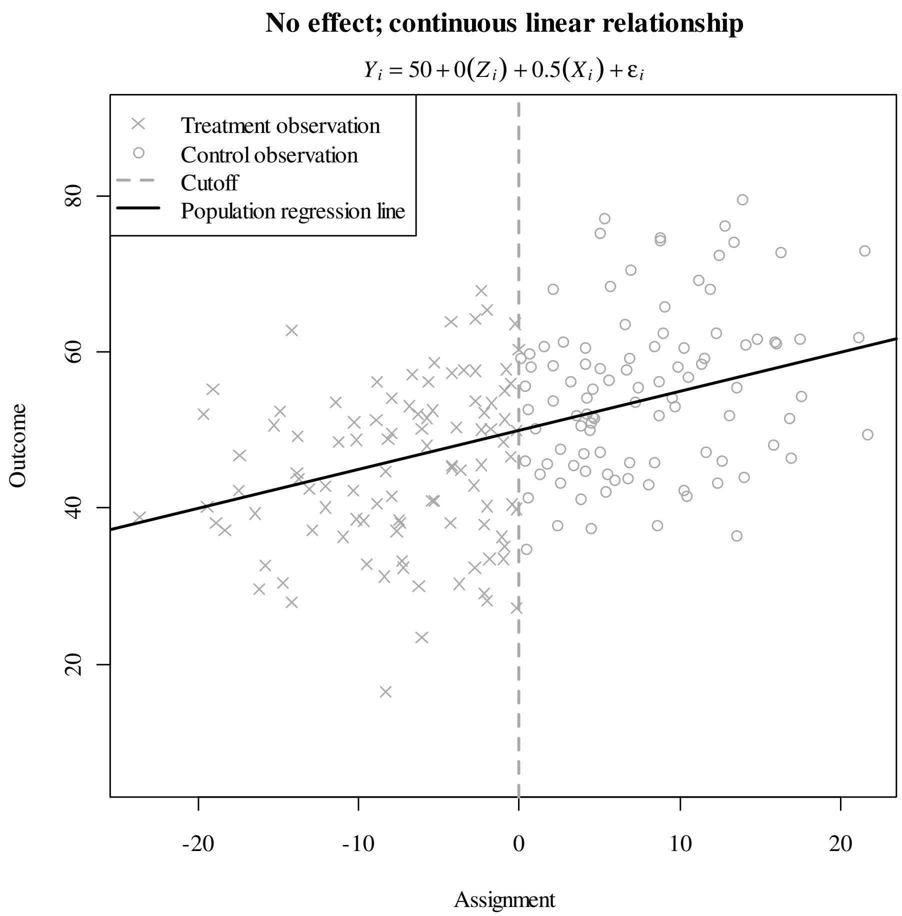
Common initial specification: cubic functional form with interactions, followed by backward elimination of statistically insignificant parameters:

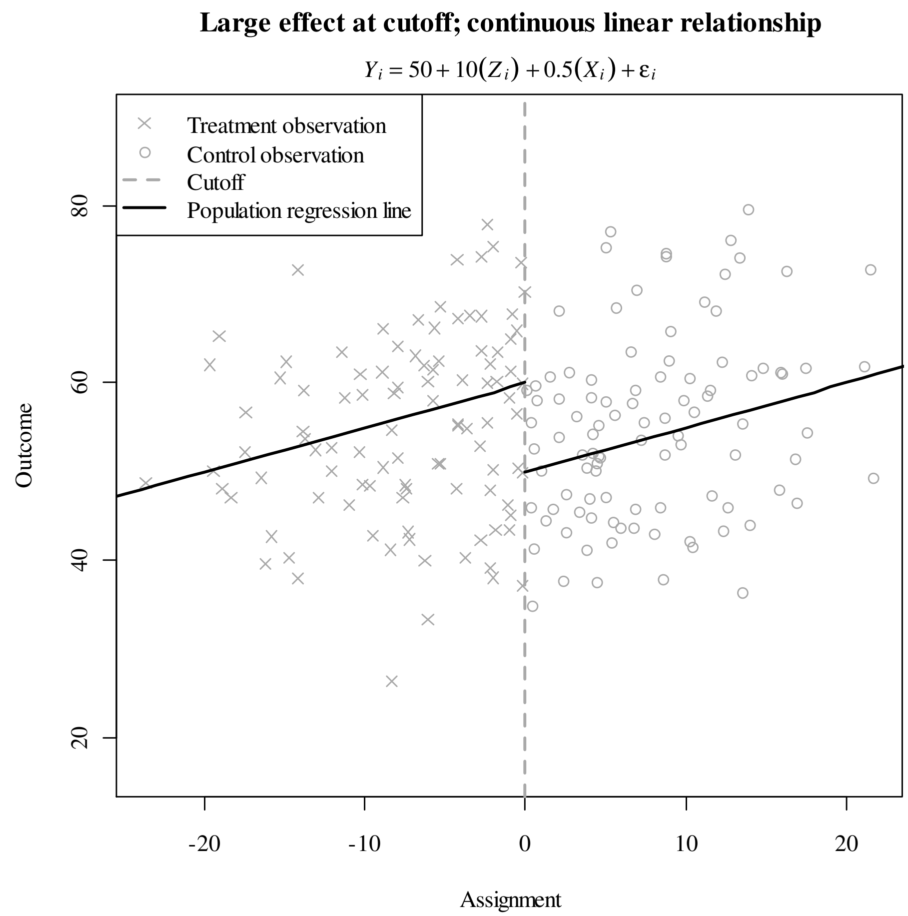
$$Y_i = \beta_0 + \beta_1 Z_i + \beta_2 X_i + \beta_3 Z_i X_i + \beta_4 X_i^2 + \beta_5 Z_i X_i^2 + \beta_6 X_i^3 + \beta_7 Z_i X_i^3 + \varepsilon_i$$

where Y is the dependent variable, Z is the treatment dummy variable, and X is distance centered at the cutoff (i.e., X_i - X_{cutoff}).

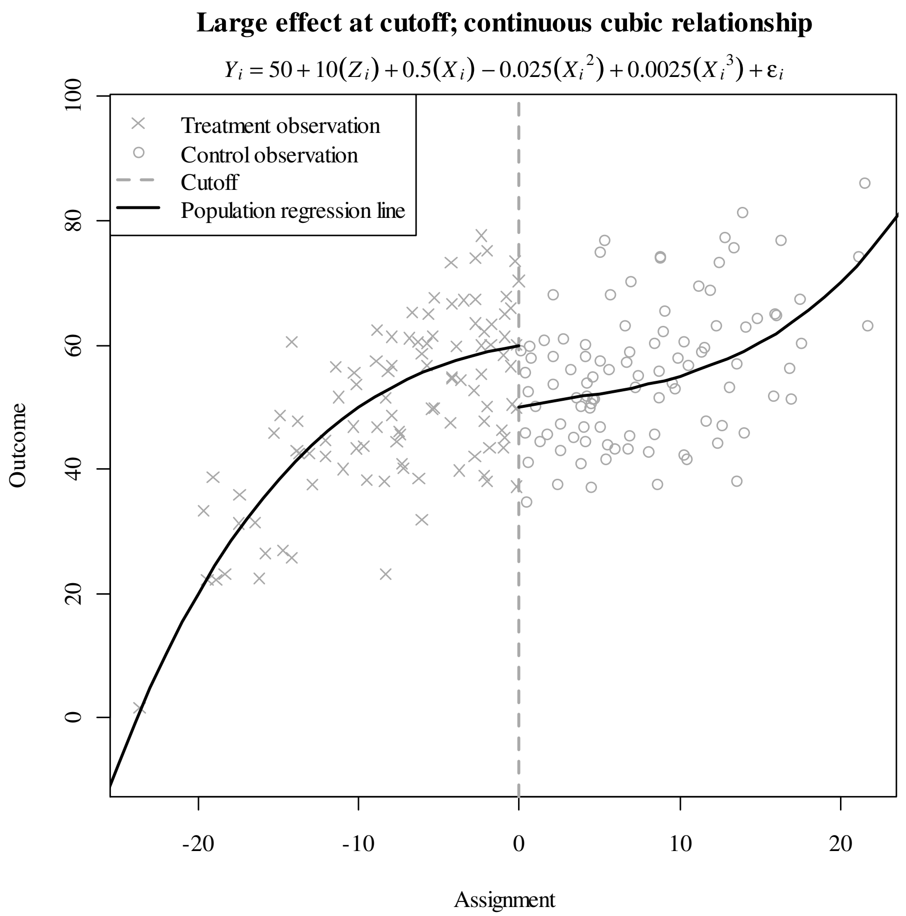
$\hat{\beta}_1$ is the estimated causal effect.

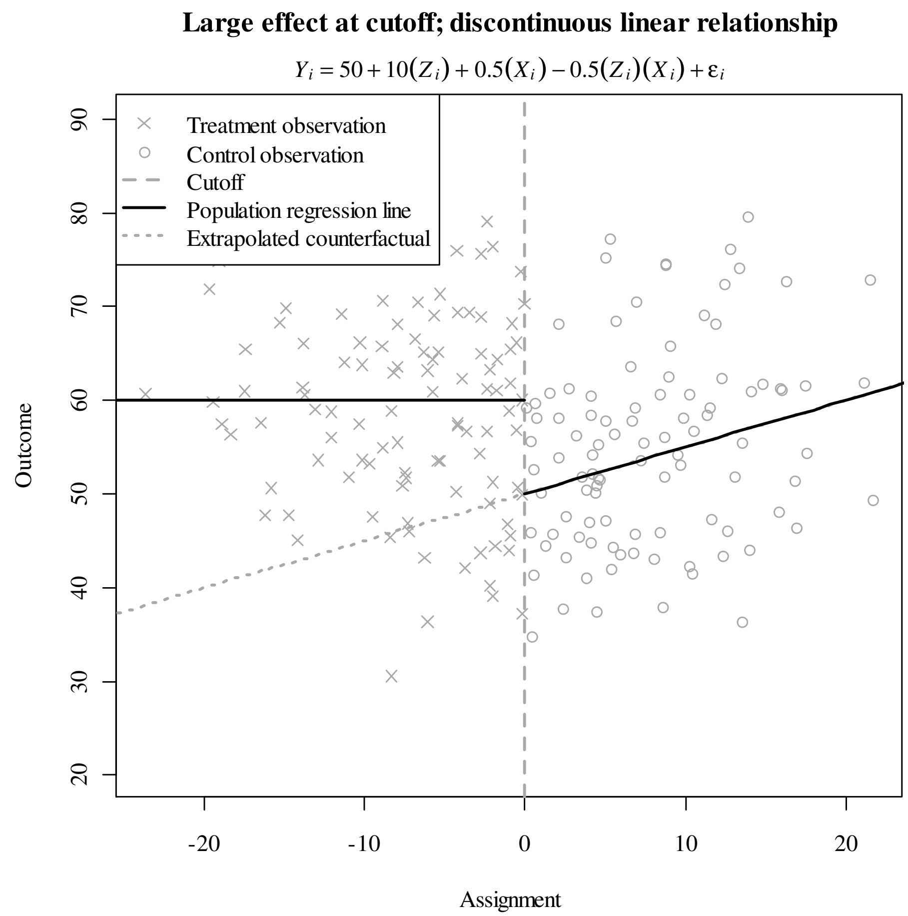
Regression discontinuity simulations





Regression discontinuity simulations





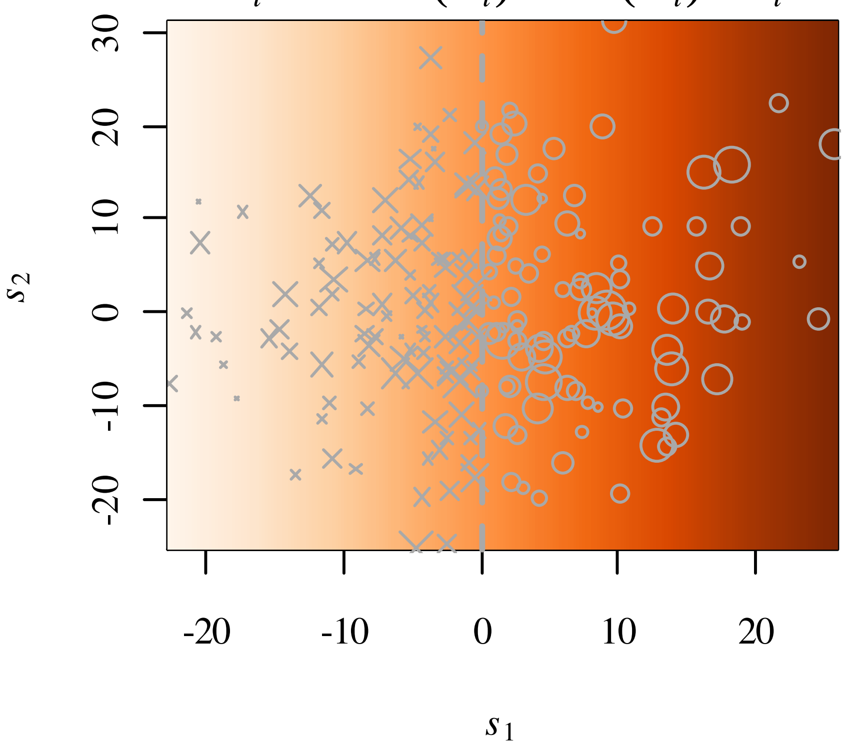
Spatial regression discontinuity

- Spatial regression discontinuity is a special case that recognizes geographic borders as sharp cutoff points.
- Participants in program evaluation studies commonly receive new program services or experience policy changes because they reside in a particular city, school district, or state and not because they were randomly assigned.
- Geographic distance represents the assignment variable; treatment-defining border represents the cutoff.
- Two-dimensional space (e.g., latitude and longitude) must be reduced to one one-dimensional distance.

Spatial regression discontinuity simulations

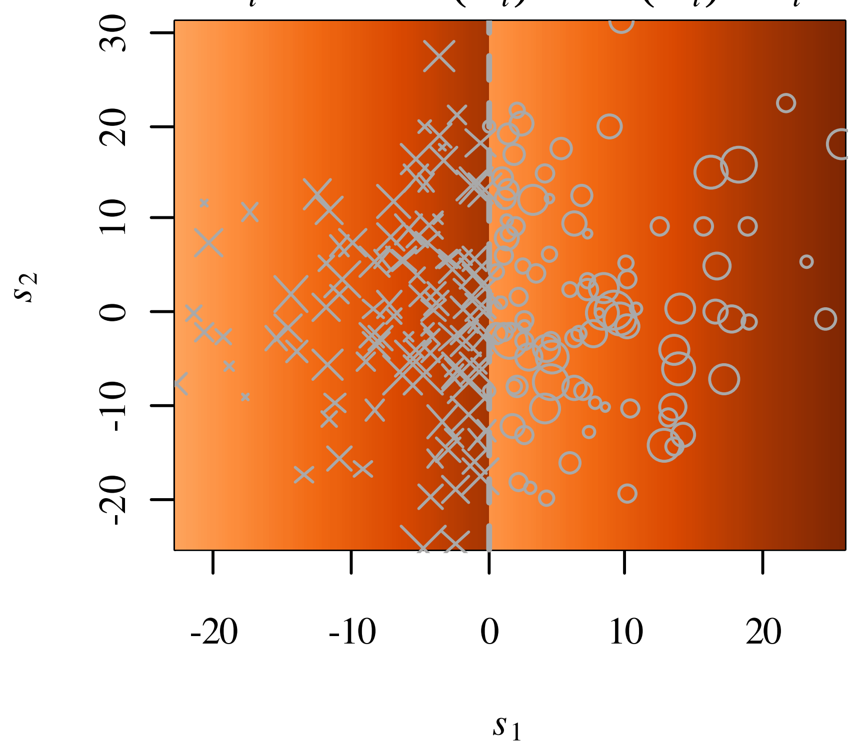
No effect; continuous linear relationship

$$Y_i = 50 + 0(Z_i) + 0.5(X_i) + \varepsilon_i$$



Large effect at cutoff; continuous linear relationship

$$Y_i = 50 + 10(Z_i) + 0.5(X_i) + \varepsilon_i$$



Spatial regression discontinuity

- Only two refereed studies:
  - Holmes (1998) examined the influence of labor union policies on manufacturing activity at borders separating "probusiness" and "antibusiness" states.
  - Black (1999) examined the localized effect of test scores on home values at school districts borders in order to estimate the monetary value of school improvement while controlling for the influence of neighborhood characteristics that confound and typically overestimate the value of better schools.
- Both performed sensitivity analyses, including modeling nonequivalent dependent variables.
- Neither considered curvilinear transformations of distance from the border.

Spatial regression discontinuity: Strengths and difficulties

- Strengths include:
  - Individuals tend to be similar both sides of a border.
  - High cost of moving to a new location helps prevent fuzzy regression discontinuity (i.e., helps enforce fidelity to treatment assignment).
  - Same strengths as standard approach if applied prospectively.
    - Both treatment and border
    - No known examples
  - Helps avoid ecological fallacy by localizing estimates.
  - Widely applicable

Spatial regression discontinuity: Strengths and difficulties

- Difficulties include:
  - Convenient borders
    - Re-use of previously established border known to participants precludes full knowledge of the assignment process.
    - Perform pretest comparison; adjust for dissimilarities at border at pretest.
    - Consider non-equivalent dependent variables.
      - Program/policy should only affect intended outcomes.
      - Change in unintended outcomes may reflect a confounding variable.
  - Geographic assignment of service providers
    - Individuals are free to cross borders to experience a different treatment condition, despite their residence.
  - What measure of distance?
    - Nearest distance? Weighted average? Mahalanobis?
  - Requires geographic competencies [e.g., geographic information systems (GIS) skills]

Spatial regression discontinuity: Minimum wage example

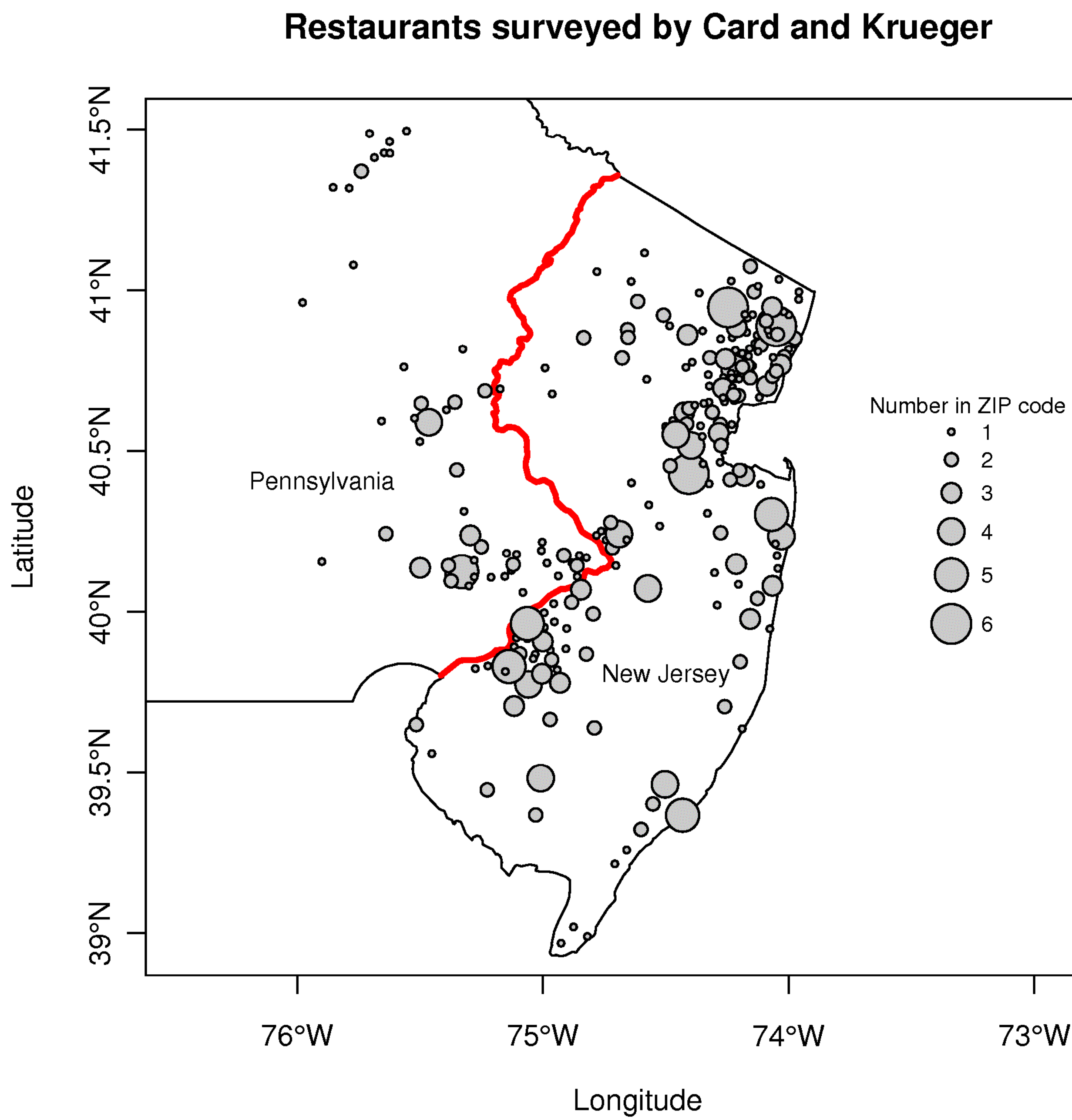
- Card and Krueger (1994) conducted a *natural experiment*. They surveyed fast food restaurants in Pennsylvania (PA) and New Jersey (NJ) before and after New Jersey raised its minimum wage from \$4.25 to \$5.05 in 1992.
- Economic theory asserts that increasing minimum wage should cause employment to decrease.
- Lack of *average treatment effect* on employment in New Jersey relative to Pennsylvania, a finding that contributed to the passage of the Fair Minimum Wage Act of 2007.

Spatial regression discontinuity: Minimum wage example

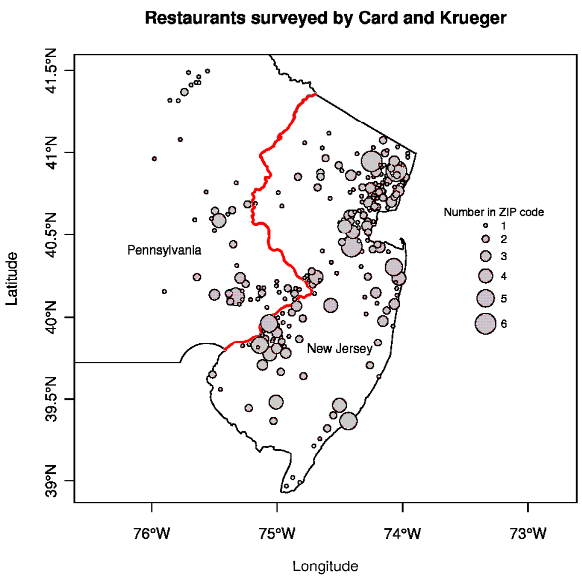
- Restaurants were assigned to the treatment condition based on their location relative to the PA-NJ border.
- A slight shift in the state border would have changed treatment assignment.
- Restaurants located in close proximity on each side of the border (e.g., in the Philadelphia metropolitan area) should have more in common than restaurants separated by long distances.

Geoprocessing

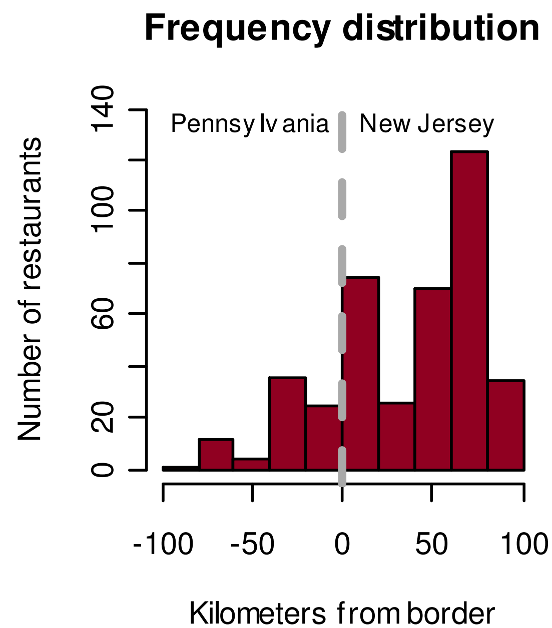
- R statistical software (R Development Core Team, 2009)
  - Joined ZIP code centroid coordinates to survey data
  - Created a line shapefile where the PA and NJ borders intersect
  - Mapped locations
  - Performed statistical analyses
- ArcGIS Near Tool (Environmental Systems Research Institute, 2008)
  - Measured each restaurant's distance from the PA-NJ border


Restaurants surveyed by Card and Krueger
Latitude
41.5°N
41°N
40.5°N
40°N
39.5°N
39°N
Longitude
76°W
75°W
74°W
73°W
Pennsylvania
New Jersey
Number in ZIP code
1
2
3
4
5
6

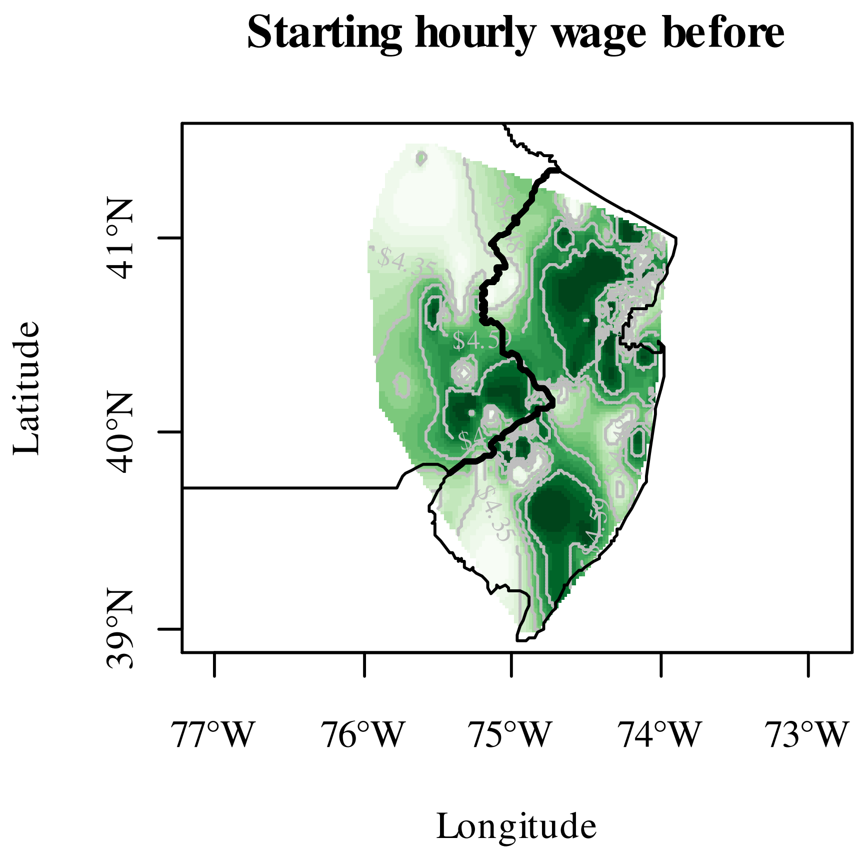
Geoprocessing

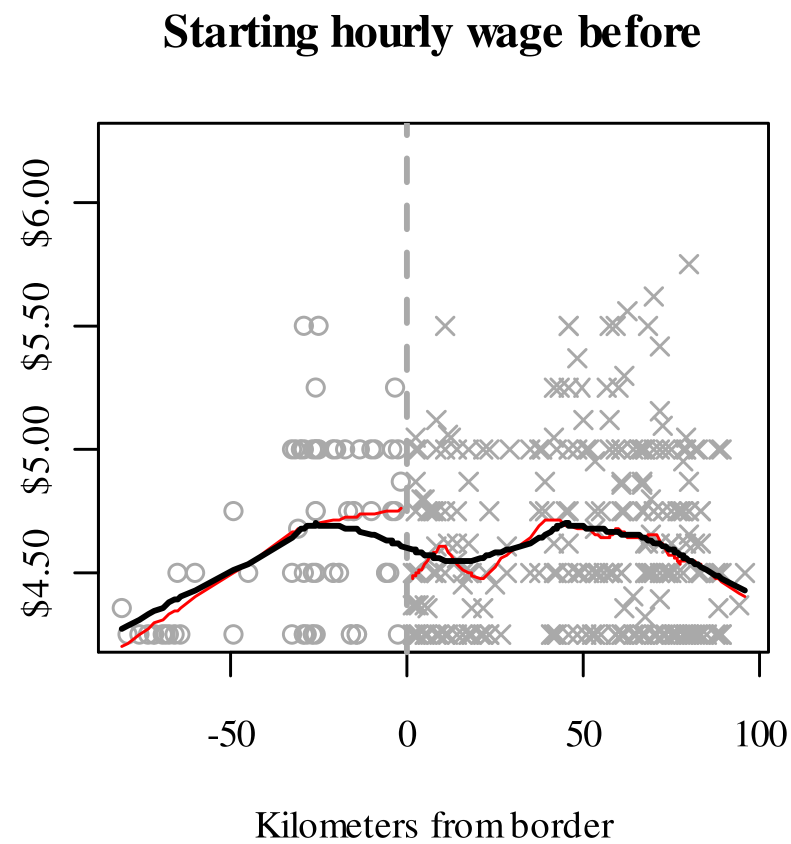


GIS-enabled measurement of distance from border, re-centered at cutoff



Exploratory plots:
3- and 2-d interpolations

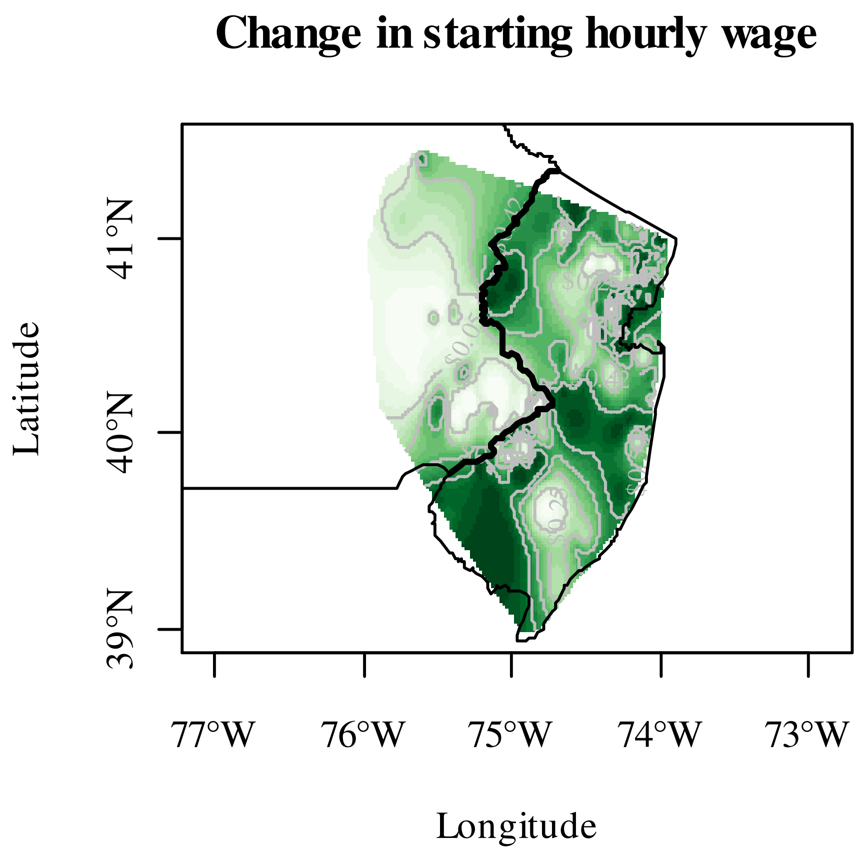


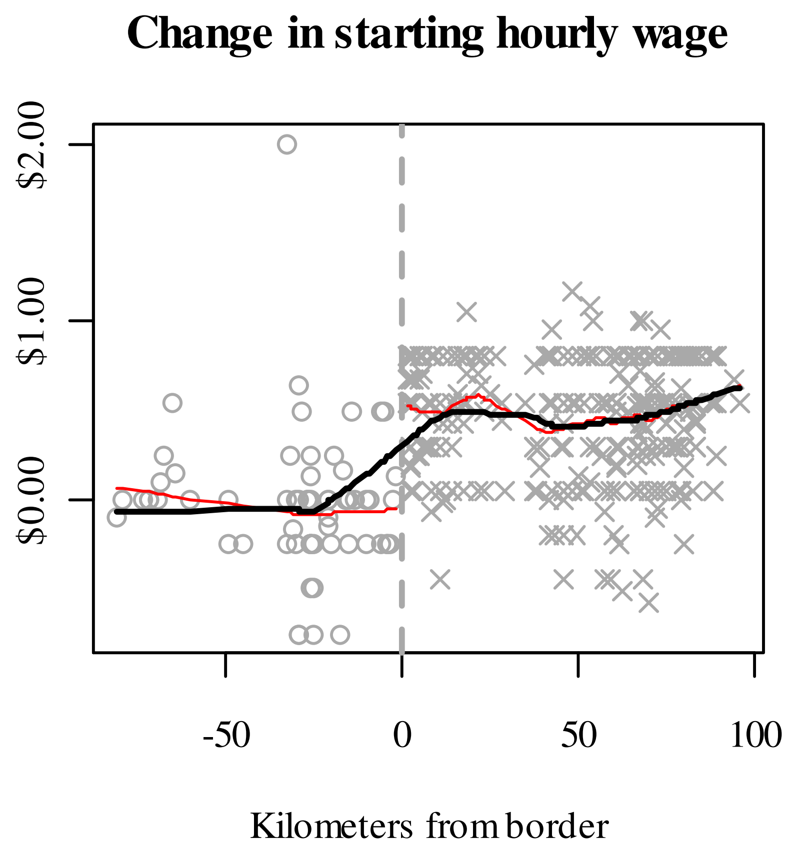


Exploratory plots: 3- and 2-d interpolations

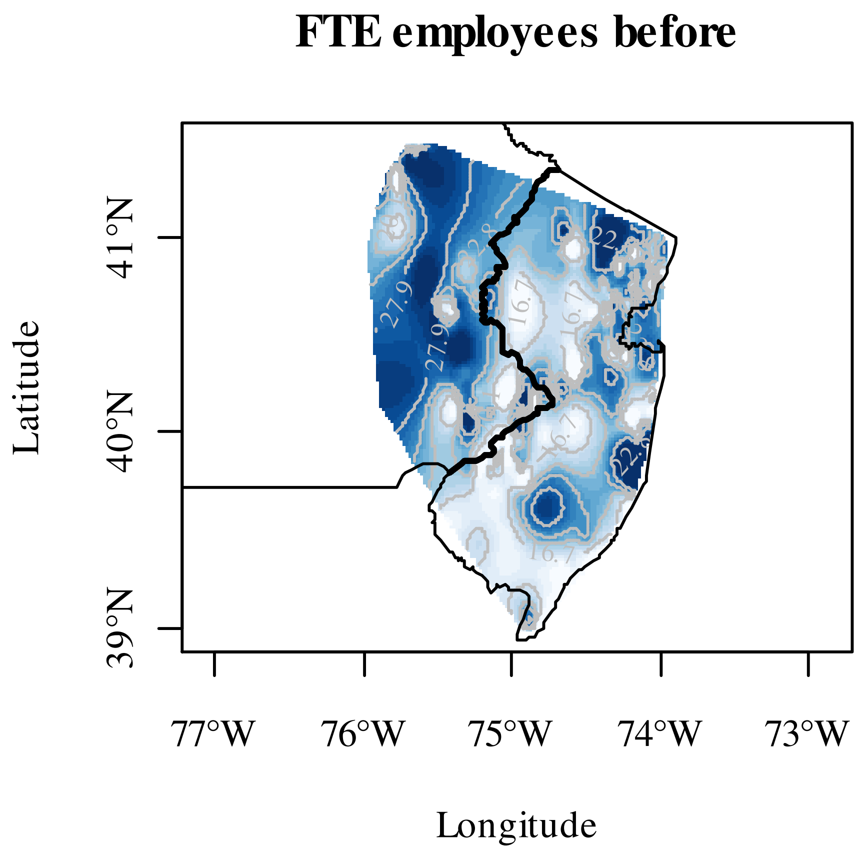
- Starting hourly wages appear to differ at the border at pretest.
  - "Regression to the mean" validity threat
  - Posttest wage would be inappropriate dependent variable; adjustment required
  - Dependent variable: change in starting hourly wage

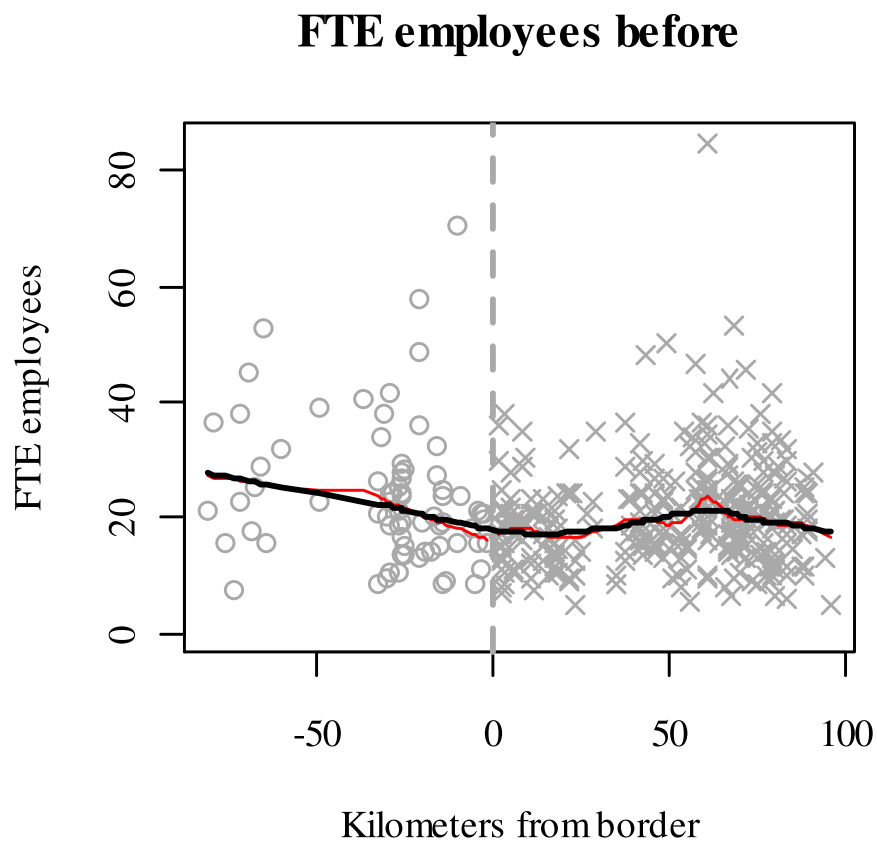
Exploratory plots:
3- and 2-d interpolations





Exploratory plots:
3- and 2-d interpolations

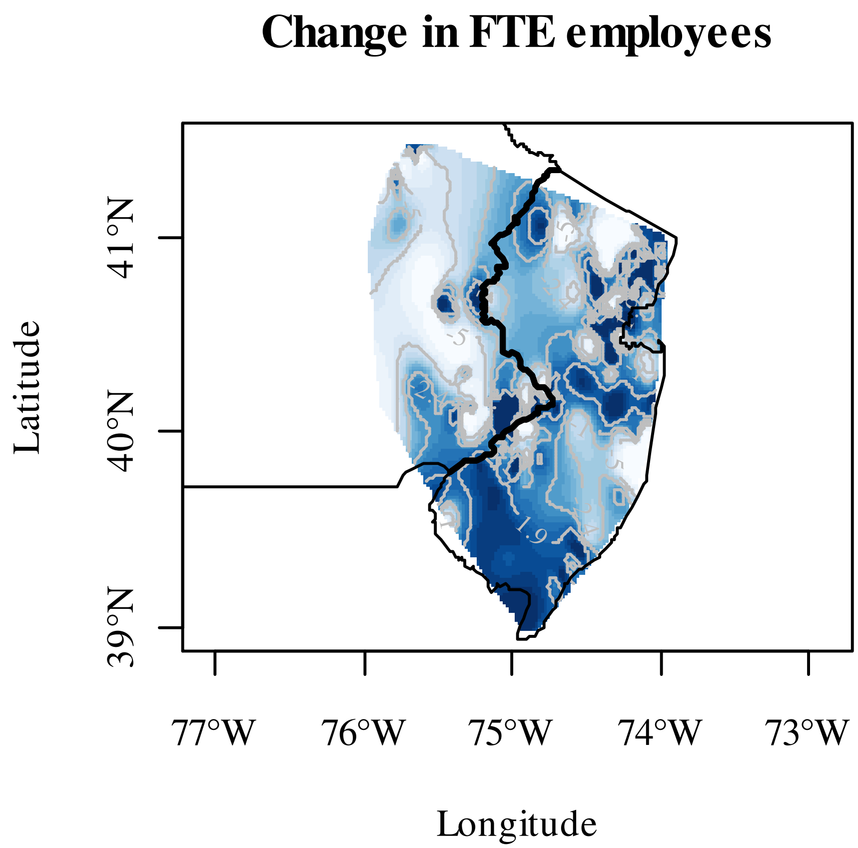


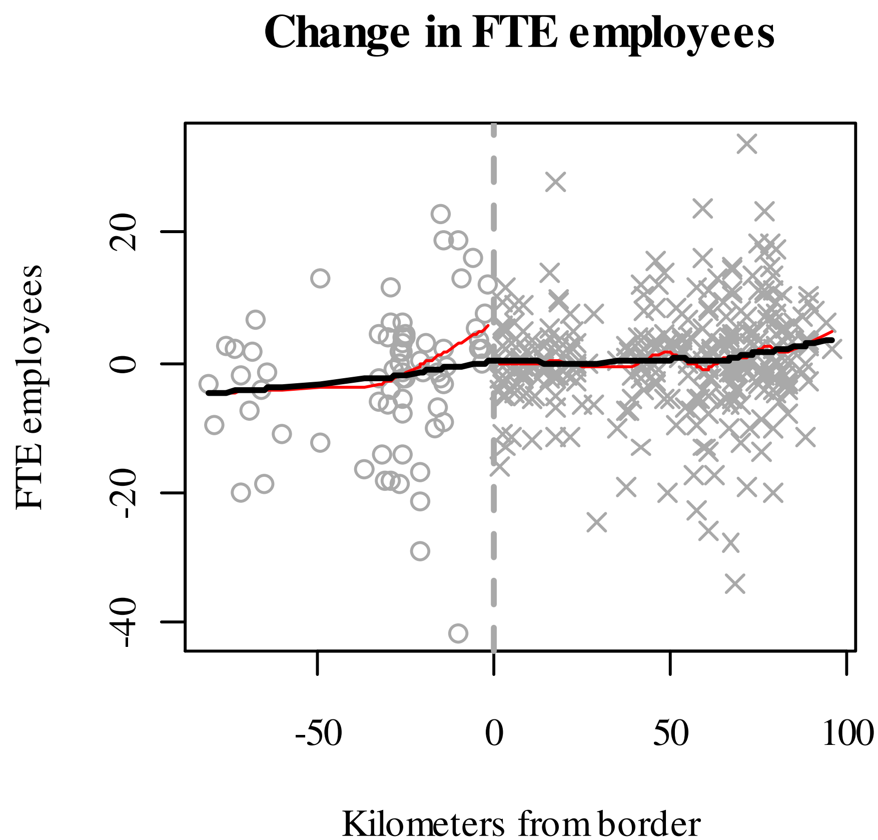


Exploratory plots: 3- and 2-d interpolations

- FTE employees appear similar at the border at pretest.
  - Posttest FTE employees would be sufficient dependent variable; no adjustment required.
  - Dependent variable: change in FTE employees, for comparability with wage outcome

Exploratory plots:
3- and 2-d interpolations





Intra-cluster correlation

Cluster	Change in starting hourly wage			Change in FTE employees		
	N	ICC	Design effect	*N*	ICC	Design effect
ZIP	224	0.16	1.19	227	0.05	1.06
Metropolitan Statistical Area (MSA)	5	0.05	7.93	5	0.00	1.13
Restaurant chain	4	0.03	4.06	4	0.00	1.38

- These results suggest the need for a mixed effects model (hierarchical linear model) of the wage outcome to avoid type I error.

Initial specification

Level 1: Restaurants

$$\Delta Y_{ijkl} = \beta_{0jkl} + \beta_{1jkl}(NJ_{ijkl}) + e_{ijkl}$$

Level 2: ZIP codes

$$\beta_{0jkl} = \beta_{00kl} + \beta_{01kl}(MINDIST_{jkl}) + \beta_{02kl}(MINDIST^{2}_{jkl}) + \beta_{03kl}(MINDIST^{3}_{jkl})$$

$$\beta_{1jkl} = \beta_{10kl} + \beta_{11kl}(MINDIST_{jkl}) + \beta_{12kl}(MINDIST^{2}_{jkl}) + \beta_{13kl}(MINDIST^{3}_{jkl})$$

Level 3: MSAs and chains

$$\beta_{00kl} = \beta_{000} + \beta_{001}(MHHINCOME) + \beta_{002}(EMPRATE) + r_{00k} + r_{00l}$$

$$\beta_{10kl} = \beta_{100} + r_{10k}$$

- Control variables: median household income and employment rate in metropolitan statistical area
- No random effects for change in FTE employees

Change in starting hourly wage

Fixed effect	Coefficient	*SE*	*t*-value
Constant	-0.073	0.065	-1.123
NJ	0.530	0.052	10.134

Random effect	Variance component
MSA	0.002
Chain	0.007
Residual	0.115

Pseudo $R^2_{y\hat{y}}$	0.277

Change in FTE employees

	Coefficient	Robust *SE*	*t*-value	*p*-value
Constant	2.428	2.019	1.203	0.230
NJ	-1.495	2.104	-0.711	0.478
Distance	0.229	0.080	2.846	0.005
Distance^2	0.002	0.001	2.047	0.042
NJ*Distance	-0.341	0.142	-2.404	0.017
Adjusted R^2		0.032		
***F* (*df*)**		4.093 (4, 374)		0.003

Final models

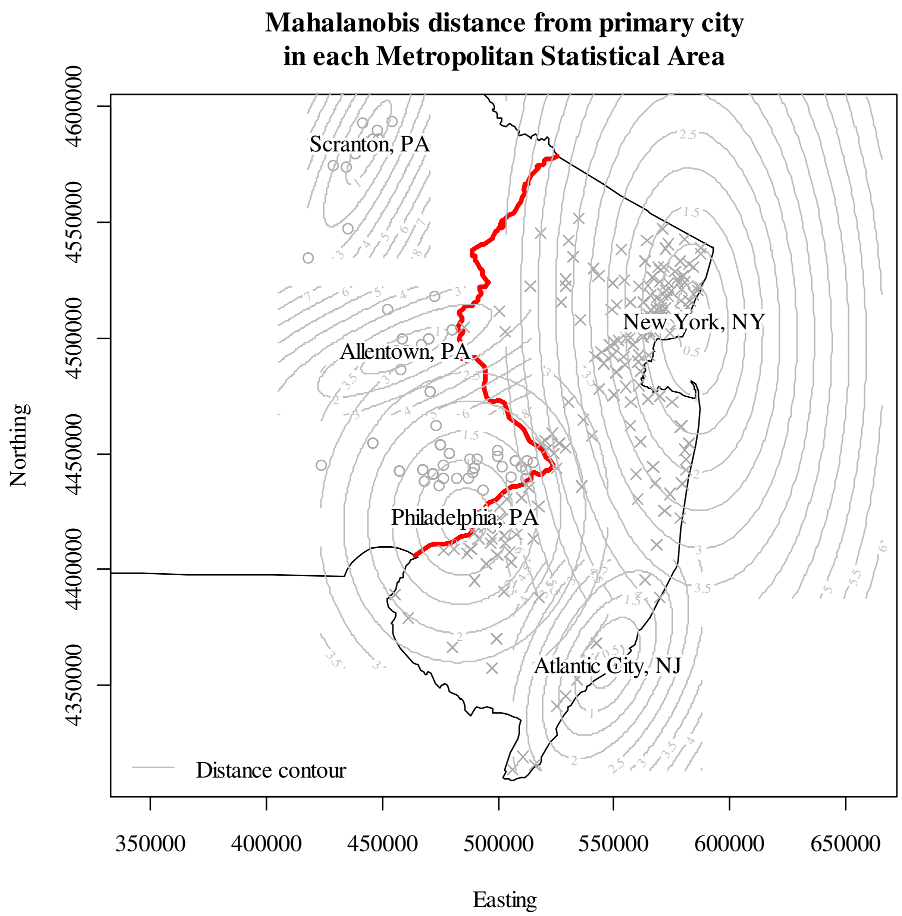
Conclusions

- Raising the minimum wage had a significant, positive *average treatment effect* on starting hourly wages in the short run.
- Not enough evidence that raising the minimum wage had a *local average treatment effect* on the number of FTE employees at the PA-NJ border in the short run.

Conclusions

- Retrospective spatial regression discontinuity is weaker than prospective regression discontinuity and random assignment, but validity checks and adjustments can help rule out competing explanations for observed effects.
- Prediction: evaluators will build geographic competencies and apply spatial regression discontinuity broadly.

Next step: Sensitivity analyses and articulate prospective approach



How to characterize prospective spatial regression discontinuity?

- Convenient border + new treatment ≠ *prospective*?
- Or should *precedence* and *ignorability* be emphasized?

	Precedence conditions	
Ignorability conditions	**No manipulation of treatment variable in treatment area before outcome measurement**	**Manipulation of treatment variable occurred in treatment area before outcome measurement**
Familiar/established border has previously determined assignment to program or policy changes or permitted selection/moving into area	• Implausible precedence • Nonignorable assignment mechanism • Black's (1990) study of the effect of school improvement on home values	• Precedence established • Nonignorable assignment mechanism • Holmes' (1989) study of the effect of labor policies on manufacturing
Newly established treatment assignment border	• Implausible precedence • Strongly ignorability • No known studies	• Precedence established • Strongly ignorability • No known studies

Borders

- Can serve useful purposes (e.g., efficient public administration).
- More likely to persist if they reflect a static, latent border between two dissimilar areas (e.g., geophysical or cultural).
- Can persist despite tension brought on by underlying similarity between areas if
  - spillovers and movement allowed (democratic consent)
  - powerful entities enforce the border, until...

Berlin Wall

References

Anselin, L., Bera, A. K., Florax, R., & Yoon, M. J. (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics, 26*(1), 77-104.

Environmental Systems Research Institute (2008). *ArcGIS 9.2*. Redlands, CA: Author.

Banerjee, S., Carlin, B. P., & Gelfand, A. E. (2004). *Hierarchical modeling and analysis for spatial data*. Boca Raton, FL: Chapman & Hall/CRC.

Black, S. E. (1999). Do better schools matter? parental valuation of elementary education. *Quarterly Journal of Economics, 114*(2), 577-599.

Bondonio, D. (2002). Evaluating decentralized policies: A method to compare the performance of economic development programmes across different regions or states. *Evaluation, 8*(1), 101-124.

Boruch, R. F. (1991). The president's mandate: Discovering what works and what works better. In M. W. McLaughlin & D. C. Phillips (Eds.), *Ninetieth yearbook of the National Society for the Study of Education, Part II. Evaluation and education: At quarter century* (pp. 147-167). Chicago, IL: The University of Chicago Press.

Brown, T. L. (2005). Evaluating geographic program performance analysis. *Public Performance & Management Review, 29*(2), 164-190.

Card, D., & Krueger, A. B. (1994). Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania. *The American Economic Review, 84*(4), 772-793.

Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review, 14*(1), 5-26.

Cook, T. D., Shadish, W. R., & Wong, V. C. (2008). Three conditions under which experiments and observational studies produce comparable causal estimates: New findings from within-study comparisons. *Journal of Policy Analysis and Management, 27*(4), 724-750.

Craig, W. J. & Elwood, S. A. (1998). How and why community groups use maps and geographic information. *Cartography and Geographic Information Systems, 25*(2), 95-104.

Galindo, R., & Shadish, W. R., (2008). *A randomized experiment comparing random to cutoff-based assignment*. Paper presented at the annual conference of the American Evaluation Association, Denver, CO.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association, 81*(396), 945-960.

Holmes, T. J. (1998). The effect of state policies on the location of manufacturing: Evidence from state borders. *Journal of Political Economy, 106*(4), 667-705.

References

Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics, 142*(2), 615-635.

Mill, J. S. (1846). *A system of logic ratiocinative and inductive: Being a connected view of the principles of evidence and the methods of scientific investigation*. New York: Harper. Retrieved from http://books.google.com/books?id=LVAYAAAAIAAJ

Monmonier, M. (1996). *How to lie with maps*. Chicago: University Of Chicago Press.

Moore, C. T. (2009a). *Spatially enabled evaluation: Toward a grounded theory of spatial methods in applied educational research*. Paper presented at the annual conference of the American Educational Research Association, San Diego, CA.

Patton, M. Q. (1997). *Utilization-focused evaluation: The new century text*. Thousand Oaks, CA: Sage.

Raudenbush, S. W. (1997). Statistical analysis and optimal design for cluster randomized trials. *Psychological Methods, 2*(2), 173-185.

Renger, R., Cimetta, A., Pettygrove, S., & Rogan, S. (2002). Geographic information systems (GIS) as an evaluation tool. *American Journal of Evaluation, 23*(4), 469.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology, 66*(5), 688-701.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.

Talen, E., & Shah, S. (2007). Neighborhood evaluation using GIS: An exploratory study. Environment and Behavior, 39(5), 583-615.

Tate, William F. (2008). Geography of opportunity: Poverty, place, and educational outcomes. *Educational Researcher, 37*(7), 397-411.

R Development Core Team (2009). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

U.S. Department of Education. (2005). Scientifically based evaluation methods: Notice of final priority. *Federal Register, 70*(15), 3585-3589. Retrieved from http://www.ed.gov/legislation/FedRegister/finrule/2005-1/012505a.html

Verdi, M. P., & Kulhavy, R. W. (2002). Learning with maps and texts: An overview. *Educational Psychology Review, 14*(1), 27-46.

Waller, L. A., & Gotway, C. A. (2004). *Applied spatial statistics for public health data*. Hoboken, NJ: Wiley.