EPIC | ENERGY POLICY INSTITUTE AT THE UNIVERSITY OF CHICAGO

Does Hazardous Waste Matter? Evidence from the Housing Market and the Superfund Program

Michael Greenstone, University of Chicago

Justin Gallagher, Case Western Reserve University

Rationale for Superfund Program

"The Reagan Administration believes that a policy of straightforward regulation and careful resource management, combined with an unshakable environmental commitment, is our mandate from the American people. And I believe that this policy will result in a swift cleanup of existing environmental hazards. Waste site cleanup will be our environmental legacy to future generation."

Anne M. Gorsuch, Administrator of US EPA, Mar. 12, 1982

"With approximately 73 million people living fewer than 4 miles from one or more of the nation's active Superfund sites, they present some of the most complex and diverse of all health and environmental pollution problems facing us today."

Carol Browner, Administrator of US EPA, March 1995

Great Uncertainty About Costs and Benefits of the Clean-Ups

"CERCLA, or Superfund as it is popularly known, is widely regarded as a wasteful and ineffective program in dire need of substantive reform."

John Shanahan (Policy Analyst, Heritage Foundation)

"Although the program has been in existence for over 14 years, we still know very little about the benefits of site cleanup or about the associated costs."

Katherine Probst (Senior Fellow, Resources for the Future)

"At the median site expenditure in our sample, the cost per case of cancer prevented is in excess of $6 billion."

Viscusi and Hamilton, AER (1999)

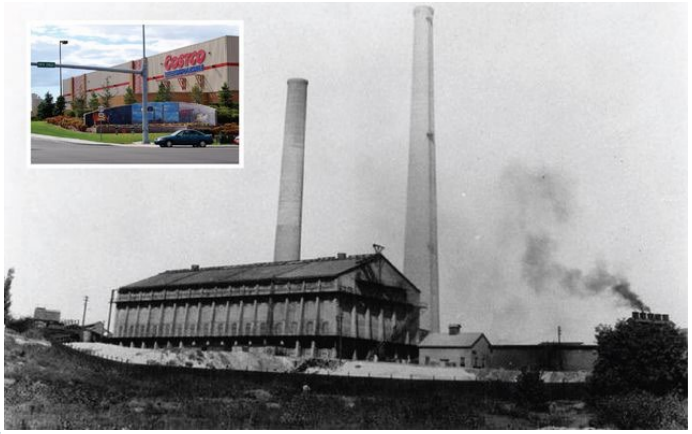
Superfund "Success Stories"

Source: U.S. EPA

The East Ditch, located at the south end of Bowers Landfill, contained discarded tires and debris.

Source: U.S. EPA

1993: After excacating soil, the west field was graded and seeded to support wetlands.



Outline of Approach

1) Use housing market to estimate individuals' valuation of clean-ups
2) Econometric identification problem. Is there a counterfactual for Love Canal (or the "Valley of the Drums")?
3) Use two approaches to solve identification problem:
   - » Linear regression to adjust for heterogeneity
   - » Exploit research design implicit in first years of Superfund, including availability of regression discontinuity design
4) Assembled comprehensive data set on progression of clean-ups, house prices, and clean-up costs
5) Advantages
   - » Not a case study
   - » Quasi-experiment
   - » Use markets to infer valuations, rather than EPA estimates of health risks

Plan of Talk

I. Introduction

II. The Superfund Program

III. Hedonic Method and Research Design

IV. Data Sources and Summary Statistics

V. Econometric Methods

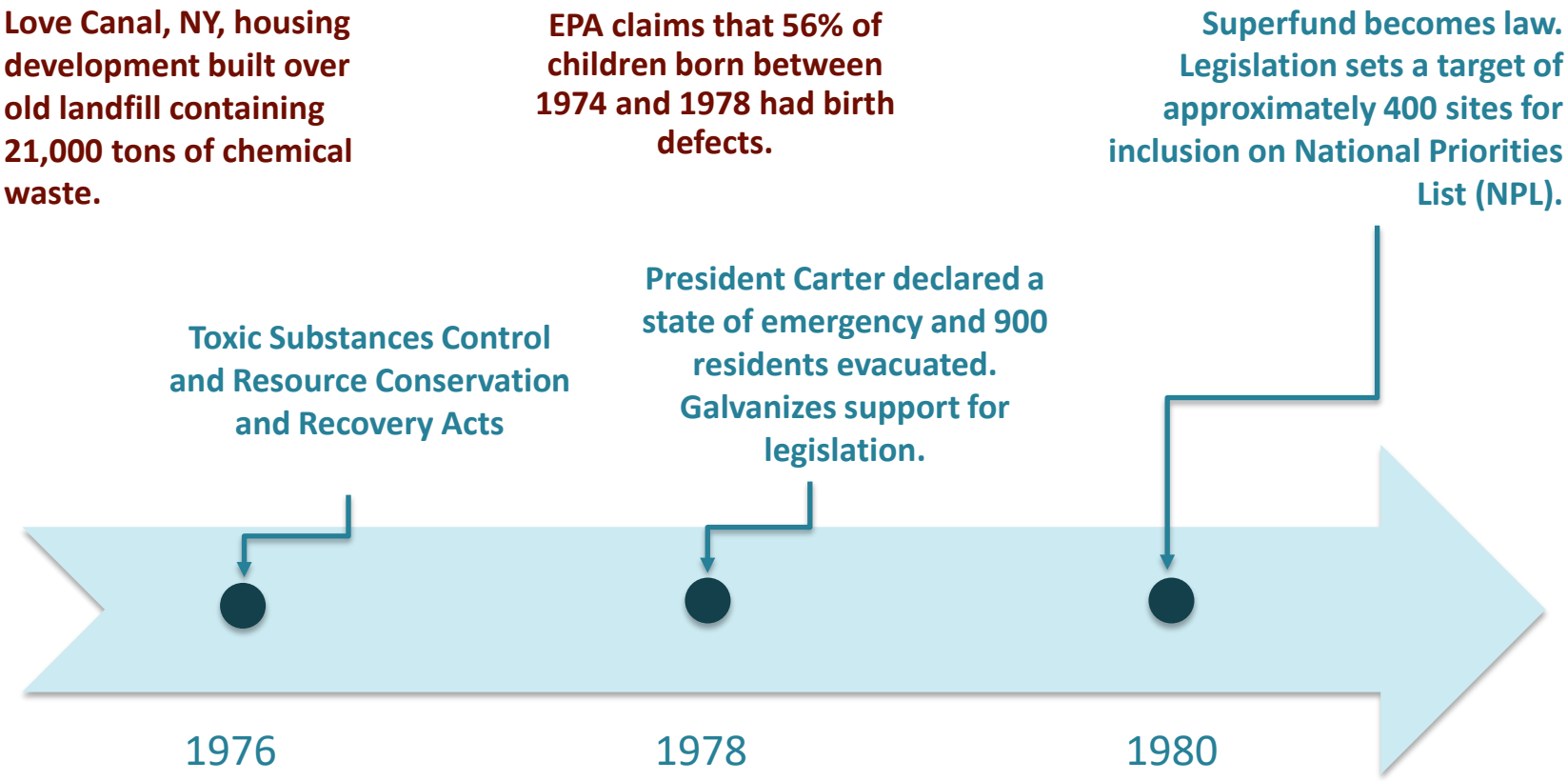
VI. Empirical Results

VII. Conclusions/Future Directions

Findings

1) **Median time until Superfund clean-up completion is 12-13 years**
2) **Expected costs of Superfund clean-ups are about $28 million (2000$) versus actual costs of ~$43 Million**
3) **Superfund sites are in neighborhoods that differ substantially from the rest of the country**
   - » E.g. higher % of mobile homes; lower population density etc.
   - » Confounding mitigated with quasi-experiment based on 1982 HRS scores
4) **Conventional estimates suggest gains in property values of 7% to 9% or $60 million in neighborhoods around sites nearly 20 years later**
5) **Quasi-experimental approach finds:**
   - » No meaningful effect on housing prices
   - » No meaningful effect on rental rates
   - » No meaningful effect on population's demand shifters
   - » No meaningful supply response

Timeline of Superfund Program (Part 1)



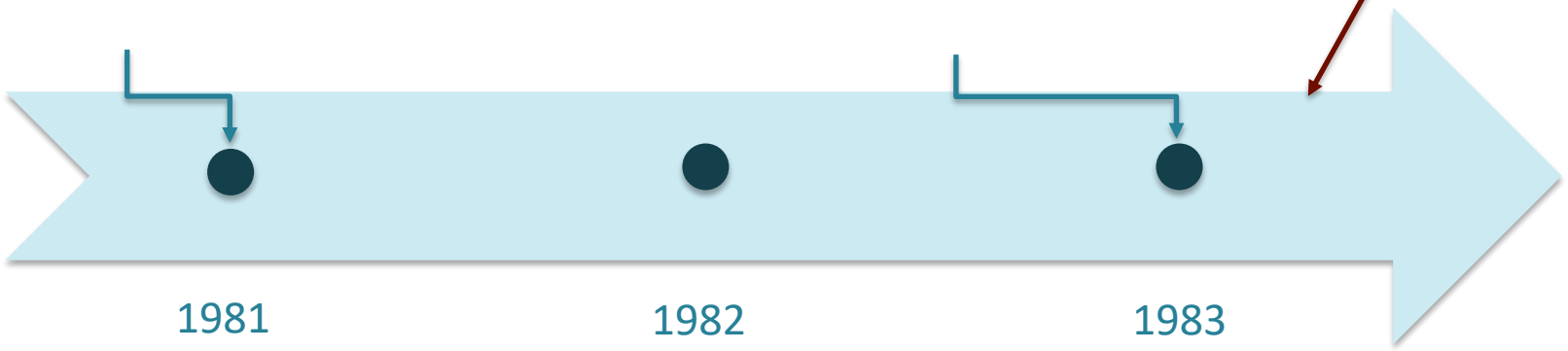
Timeline of Superfund Program (Part 2)

EPA develops hazardous ranking system (HRS) to rank the 690 worst sites. Values range from 0 to 100. Cut-off for top 400 is HRS score of 28.5.

Roughly 15,000 hazardous waste sites referred to EPA for possible Superfund remediation.

EPA announces 1st NPL in 1983. Only NPL sites eligible for federally financed remedial clean-ups that aim to return sites to "natural state"

Post-1983 cleanup process is slow.



Assignment of HRS Scores, 1981-2

1) **Upon passage of Superfund, there was no existing measure to judge risk of proximity to a hazardous waste site**

2) **EPA developed HRS in about a year:**
   - » Guidelines for scoring a site were 30 pages; today it exceeds 500 pages
   - » Nonlinear function of ground water, surface water, and air migration

$$\textbf{(1)} \quad \textbf{HRS Score} = [(S^2_{gw} + S^2_{sw} + S^2_{a}) / 3]^{½}$$

3) **EPA Aware of Arbitrariness in Test:**
   - » "EPA cannot develop new standards for the hundreds of substances [present]…such a task would also be enormous, costly and time-consuming, and would unduly hamper the clean-up of releases" (EPA 1982).
   - » "EPA has not made a determination that sites scoring less than 28.5 do not present a significant risk to human health, welfare, or the environment. …HRS does not measure absolute risks associated with a site…[this] would require significantly greater time and funds" (EPA 1982).

Bottom line: HRS score is noisy measure of risk

Do Housing Prices Rise With Clean-ups of Hazardous Waste Sites?

Key Questions

- What is the economic value of clean-ups to individuals?
- What are the monetary “benefits” of the Superfund program?

Approach: Hedonic Method

- Housing market as implicit market for inferring value of clean-ups.

Do Housing Prices Rise With Clean-ups of Hazardous Waste Sites?

Two Stages

First Stage:

- Relate differences in housing prices to differences in housing characteristics:

$$P = P(c_1, c_2, \ldots, c_n)$$ Hedonic Price Schedule (HPS)

- $\frac{\partial P}{\partial c_n}$ is the marginal implicit price of the nth characteristic
- Rosen (1974) provided a model that showed that the HPS is determined by the equilibrium interactions of consumers and suppliers.
- At each point along the function:

$\frac{\partial P}{\partial c_n}$ = MC of a supplier = MWTP for a consumer

MWTP is an interesting parameter in some contexts

Do Housing Prices Rise With Clean-ups of Hazardous Waste Sites?

HPS is generated by the equilibrium interactions of consumers and producers

1) **Consumer (Renters) Side**:
   - » Maximize: $u = u(X, C) \quad s.t. \quad I - P - X = 0$
   - » Individuals satisfy: $\frac{\partial U}{\partial c_j} / \frac{\partial U}{\partial x} = \frac{\partial P}{\partial c_j}$
   - » Resulting bid curve: $B_j = B_j(I - P, c_j, C^*_{-j}, u^*)$
2) **Supplier (Landowners) Side**:
   - » Maximize: $\pi = P(C) - \tau_i(C)$
   - » Invert π o derive offer curve: $O_j = O_j(c_j, C^*_{-j}, \pi^*)$

In equilibrium: $$\frac{\partial U}{\partial c_j} / \frac{\partial U}{\partial x} = \frac{\partial P}{\partial c_j} = \frac{\partial \tau}{\partial c_j}$$

Welfare Effects of Superfund Clean-Ups

Consider a Superfund clean-up at a site that raises local environmental quality from c_j^1 to c_j^3.

1) **Partial** Equilibrium Case (HPS Doesn't Shift):
   - » Price at site will rise
   - » If zero moving costs, then zero welfare effects for renters
     - Renters at original site leave
     - New renters are richer and have greater tastes for environmental quality
   - » All benefits accrue to landowners at the improved site; possible supply response

Welfare Effects of Superfund Clean-Ups

2) **General Equilibrium Case (HPS Shifts):**
   - » Additional benefits scattered throughout the economy

3) **Expresses for Full Welfare Benefits:**

$$\Delta \text{Consumer WTP} = \sum_i \left[B_i(c_{ij}^{*\text{post}}, C_{-ij}^*, u_j^*) - B_i(c_{ij}^{*\text{pre}}, C_{-ij}^*, u_j^*)\right] - \sum_i \left[P_i^{\text{post}}(c_{ij}^{*\text{post}}, C_{-ij}^*) - P_i^{\text{pre}}(c_{ij}^{*\text{pre}}, C_{-ij}^*)\right]$$

$$\Delta \text{Supplier WTP} = \sum_i \left[P_i^{\text{post}}(c_{ij}^{*\text{post}}, C_{-ij}^*) - P_i^{\text{pre}}(c_{ij}^{*\text{pre}}, C_{-ij}^*)\right] - \sum_i \left[\tau_i(c_{ij}^{*\text{post}}, C_{-ij}^*) - \tau_i(c_{ij}^{*\text{pre}}, C_{-ij}^*)\right]$$

The Hedonic Method in Practice

1) **In numerous settings, 1st stage of hedonic approach produces weak results (Omitted Variable Bias likely explanation):**
   - » Air pollution and house prices (Chay & Greenstone 2004)
   - » Job risk and wages (Black & Kneisner 2003)
   - » School quality and house prices (Black 1999)
   - » Climate and land values (Deschenes & Greenstone 2004)
2) **Small (1975):**
   - » "I have entirely avoided…the important question of whether the empirical difficulties, especially correlation between pollution and unmeasured neighborhood characteristics, are so overwhelming as to render the entire method useless. I hope that…future work can proceed to solving these practical problems….The degree of attention devoted to this [problem]…is what will really determine whether the method stands or falls…" [p. 107]."
3) **Consequences are biased estimates of MWTP and WTP**

The Hedonic Method in Practice (cont.)

4) **Quasi-experiments as a solution:**
   - » Air pollution and house prices (Chay & Greenstone 2004)
   - » VSL (Ashenfelter & Greenstone 2002)
   - » School quality and house prices (Black 1999)

5) **Near consensus that the 2nd step of estimating bid and offer functions has never been successfully accomplished**

Ideal Analysis

1) Randomly assign clean-ups
2) Observe owners' valuations

Proposed "Solution"

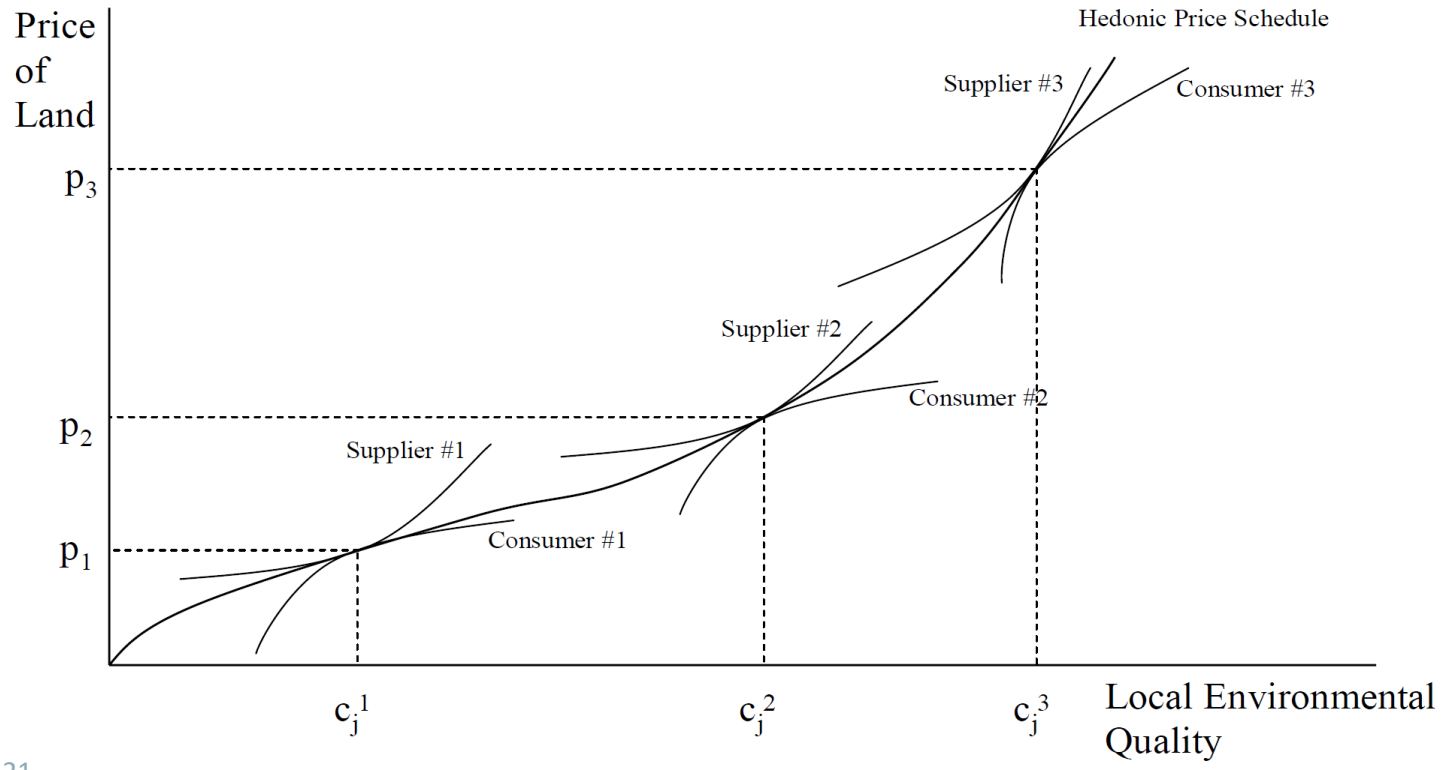
1) Exploit **initial evaluation of potential Superfund sites in 1981-2**
2) Mandatory NPL listing if HRS score > 28.5
3) Use 1(HRS Score > 28.5) as "Instrument" for NPL status in equations for 1990 and 2000 housing prices
   - » Attempt to identify causal effect of NPL Status on housing prices

Advantages of Approach

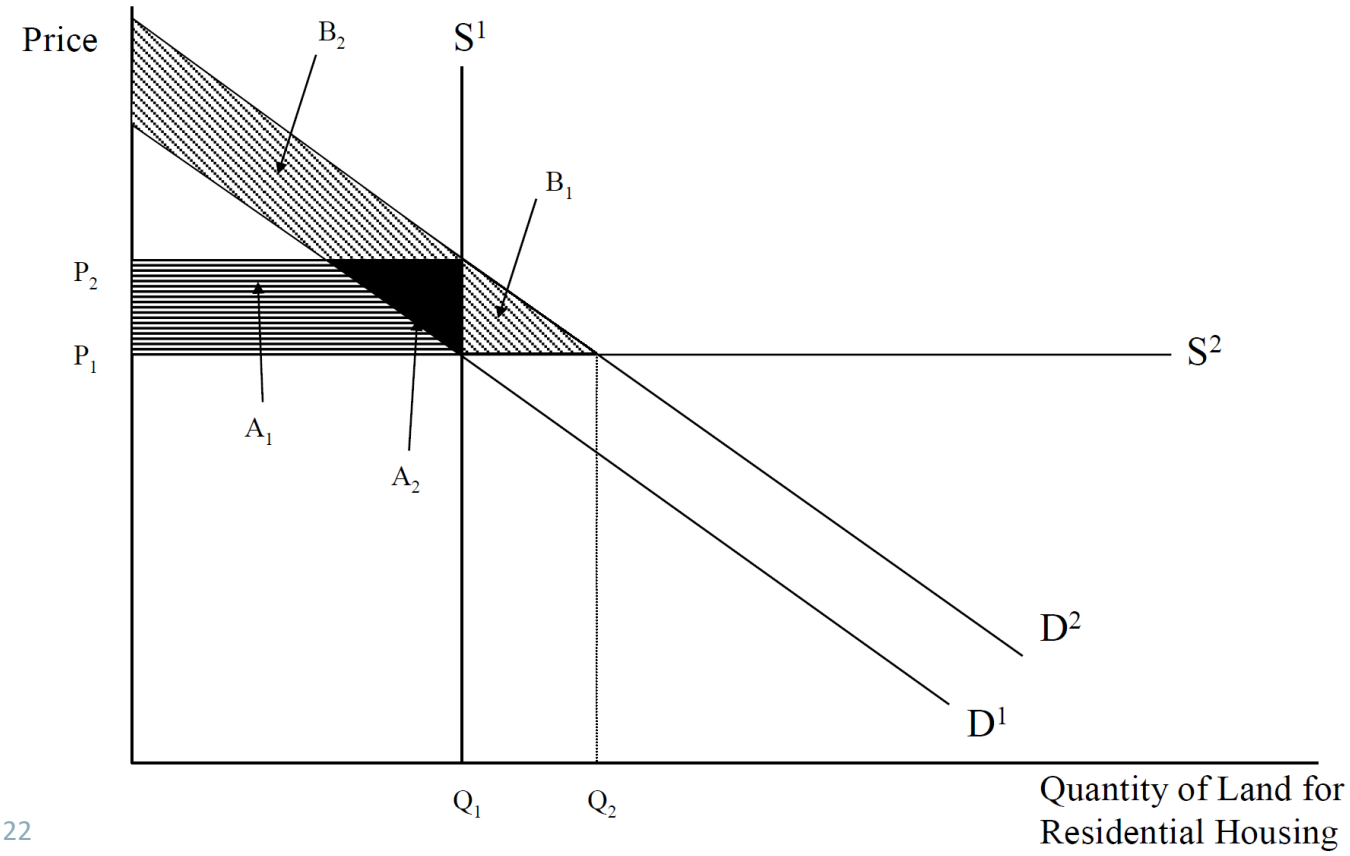
1) 28.5 cut-off established after testing of 690 sites. Consequently, unlikely that scores were based on expected costs and benefits of clean-up.
2) Tremendous scientific uncertainty about the health effects of the tens of thousands of chemicals present at sites, implying that HRS score is noisy
3) Availability of regression discontinuity design

Hedonic Market for Local Environmental Quality

Bid Curves, Offer Curves, and the Equilibrium Hedonic Price Schedule



Welfare Gains Due to Amenity Improvements with Two Supply Curves



Data Sources

1) 690 **hazardous waste sites to be considered for NPL:**
   - » HRS and pathway scores
   - » For all NPL sites listed by 2000, we collected:
     a) Dates of NPL listing, clean-up initiated, construction complete, and deletion
     b) Size of site
     c) Expected and actual costs of clean-up
2) Housing price and characteristics data, 1980 – 2000:
   - » Unit of observation is census tract, which contains roughly 4,000 people and 1,350 housing units (65,443 total tracts in 2000)
   - » 2000 census tracts matched across censuses with Geolytics' Neighborhood Change database
3) Place hazardous waste sites in census tracts:
   - » NPL sites (latitude and longitude were given)
   - » Non-NPL sites (used latitude and longitude and street addresses)

Econometric Methods

1) **"Naïve Estimator": Least squares estimation with data from entire U.S.**

$$y_{c2000} = \theta\, 1(\mathrm{NPL}_{c2000}) + X_{c80}\beta + \epsilon_{c2000},$$
$$1(\mathrm{NPL}_{c2000}) = X'_{c80}\Pi + \eta_{c2000}.$$

y_{c2000} = log of the median property value in census tract c in 2000,

$1(\mathrm{NPL}_{c2000})$ = 1 for observations from census tracts that contain a hazardous waste site placed on the NPL by 2000,

X_{c80} = Vector of determinants of housing prices, measured in 1980 (include 1980 Housing Prices) agnostic about what variables belong in X,

ϵ_{c2000} = unobservable determinants of housing prices

η_{c2000} = the unobservable determinants of NPL status

θ = 'true' effect of NPL status on property values

Econometric Methods

2) Quasi-Experimental: IV Estimate with Sample of 487 Initial HRS Sites

$$y_{c2000} = \theta\, 1(\mathrm{NPL}_{c2000}) + X_{c80}\beta + \epsilon_{c2000},$$
$$1(\mathrm{NPL}_{c2000}) = X'_{c80}\Pi + \delta\, 1(\mathrm{HRS}_{c82} > 28.5) + \eta_{c2000}.$$

"Sufficient conditions":

- $\delta \neq 0$
- $E[1(\mathrm{HRS}_{C82} > 28.5)\ \epsilon_{c2000}] = 0$

Econometric Methods

3) Regression Discontinuity Model
   - » Even if $E[1(\mathrm{HRS}_{C82} > 28.5)\ \epsilon_{c90}] \neq 0$ consistent inference may be possible
   - » Control for quadratic of HRS score
   - » Restrict sample to sites with 16.5 < 1982 HRS < 40.5
   - » Directly control for 3 pathway scores

TABLE I
SUMMARY STATISTICS ON THE SUPERFUND PROGRAM

	All NPL sites w/ nonmissing house price data (1)	1982 HRS sites w/ nonmissing house price data (2)	1982 HRS sites w/ missing house price data (3)
Number of sites	985	487	189
1982 HRS score above 28.5	—	306	95
A. Timing of placement on NPL			
Total	985	332	111
# 1981–1985	406	312	97
# 1986–1989	340	14	9
# 1990–1994	166	4	3
# 1995–1999	73	2	2
B. HRS information			
Mean scores | HRS $\geq$ 28.5	41.89	44.47	43.23
Mean scores | HRS $<$ 28.5	—	15.54	16.50
C. Size of site (in acres)			
Number of sites with size data	920	310	97
Mean (median)	1,187 (29)	334 (25)	10,507 (35)
Maximum	195,200	42,560	405,760

TABLE I
SUMMARY STATISTICS ON THE SUPERFUND PROGRAM

	All NPL sites w/ nonmissing house price data (1)	1982 HRS sites w/ nonmissing house price data (2)	1982 HRS sites w/ missing house price data (3)
D. Stages of cleanup for NPL sites			
Median years from NPL listing until:			
ROD issued	—	4.3	4.3
Cleanup initiated	—	5.8	6.8
Construction complete	—	12.1	11.5
Deleted from NPL	—	12.8	12.5
1990 status among sites NPL by 1990			
NPL only	394	100	31
ROD issued or cleanup initiated	335	210	68
Construction complete or deleted	22	16	7
2000 status among sites NPL by 2000			
NPL only	137	15	3
ROD issued or cleanup initiated	370	119	33
Construction complete or deleted	478	198	75

TABLE I
SUMMARY STATISTICS ON THE SUPERFUND PROGRAM

	All NPL sites w/ nonmissing house price data (1)	1982 HRS sites w/ nonmissing house price data (2)	1982 HRS sites w/ missing house price data (3)
E. Expected costs of remediation (millions of year-2000 $)			
No. of sites with nonmissing costs	753	293	95
Mean (median)	$28.3 ($11.0)	$27.5 ($15.0)	$29.6 ($11.5)
95th percentile	$89.6	$95.3	$146.0
F. Actual and expected costs conditional on construction complete (millions of year-2000 $)			
Sites w/both costs nonmissing	477	203	69
Mean (median) expected costs	$15.5 ($7.8)	$20.6 ($9.7)	$17.3 ($7.3)
Mean (median) actual costs	$21.6 ($11.6)	$32.0 ($16.2)	$23.3 ($8.9)

Notes: Column (1) includes information for sites placed on the NPL before December 31, 1999. The estimated cost information is calculated as the sum across the first RODs for each operating unit associated with a site. See the Data Appendix for further details.

TABLE II

MEAN CENSUS TRACT CHARACTERISTICS BY CATEGORIES OF THE 1982 HRS SCORE

	NPL site by 2000 (1)	No NPL site by 2000 (2)	HRS < 28.5 (3)	HRS > 28.5 (4)	HRS > 16.5 & < 28.5 (5)	HRS > 28.5 & < 40.5 (6)	*p*-Value (1) vs. (2) (7)	*p*-Value (3) vs. (4) (8)	*p*-Value (5) vs. (6) (9)
No. of Census tracts	985	41,989	181	306	90	137	—	—	—
Superfund cleanup activities									
Ever NPL by 1990	0.7574	—	0.1271	0.9902	0.2222	0.9854	—	**.000**	**.000**
Ever NPL by 2000	1.0000	—	0.1602	0.9902	0.2667	0.9854	—	**.000**	**.000**
1980 mean housing prices									
Site's Census tract	58,045	69,904	45,027	52,137	46,135	50,648	**.000**	**.000**	.084
2-mile-radius circle around site	56,020	—	48,243	53,081	48,595	52,497	—	.016	.179
3-mile-radius circle around site	56,839	—	51,543	54,458	49,434	53,868	—	.257	.126
1980 housing characteristics									
Total housing units	1,392	1,350	1,357	1,353	1,367	1,319	.039	.951	.575
Mobile homes (%)	0.0862	0.0473	0.0813	0.0785	0.0944	0.0787	**.000**	.792	.285
Occupied (%)	0.9408	0.9330	0.9408	0.9411	0.9412	0.9411	**.000**	.940	.989
Owner occupied (%)	0.6818	0.6125	0.6792	0.6800	0.6942	0.6730	**.000**	.959	.344
0–2 bedrooms (%)	0.4484	0.4722	0.4691	0.4443	0.4671	0.4496	**.000**	.107	.417
3–4 bedrooms (%)	0.5245	0.5016	0.5099	0.5288	0.5089	0.5199	**.000**	.202	.586
Built last 5 years (%)	0.1434	0.1543	0.1185	0.1404	0.1366	0.1397	**.006**	.050	.844
Built last 10 years (%)	0.2834	0.2874	0.2370	0.2814	0.2673	0.2758	.506	.012	.723
No air-conditioning (%)	0.4903	0.4220	0.5058	0.4801	0.5157	0.5103	**.000**	.253	.870
Units attached (%)	0.0374	0.0754	0.0603	0.0307	0.0511	0.0317	**.000**	.040	.297

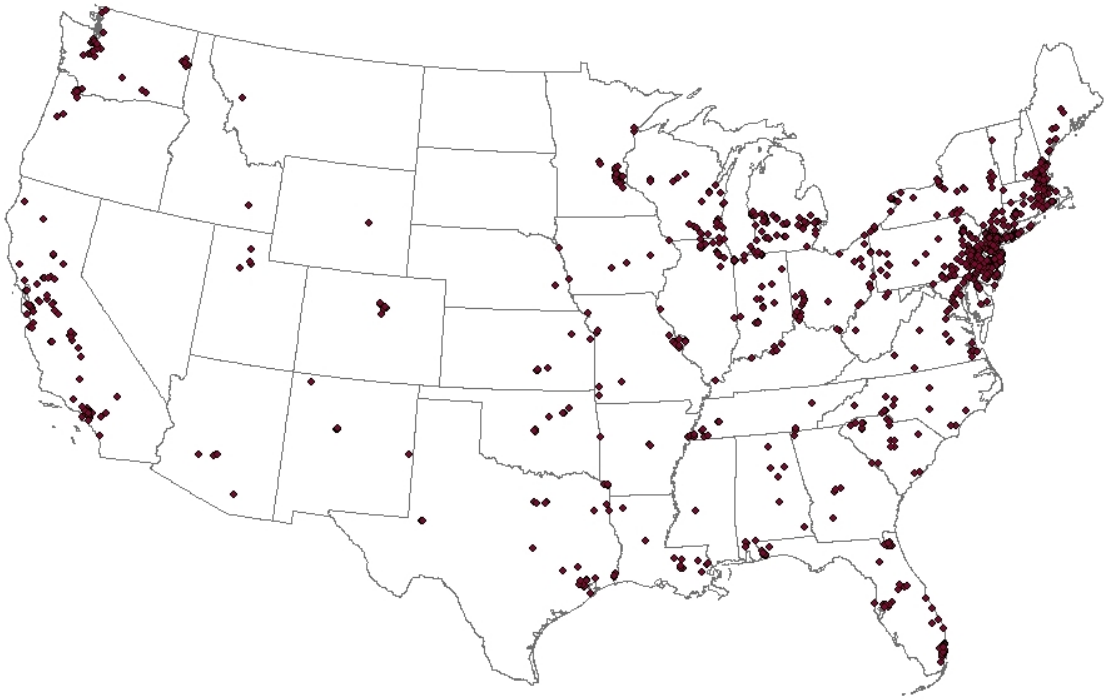
TABLE II
(CONTINUED)

	NPL site by 2000 (1)	No NPL site by 2000 (2)	HRS < 28.5 (3)	HRS > 28.5 (4)	HRS > 16.5 & < 28.5 (5)	HRS > 28.5 & < 40.5 (6)	*p*-Value (1) vs. (2) (7)	*p*-Value (3) vs. (4) (8)	*p*-Value (5) vs. (6) (9)
1980 demographics & economic characteristics									
Population density	1,407	5,786	1,670	1,157	1,361	1,151	**.000**	.067	.570
Black (%)	0.0914	0.1207	0.1126	0.0713	0.0819	0.0844	**.000**	.037	.926
Hispanic (%)	0.0515	0.0739	0.0443	0.0424	0.0309	0.0300	**.000**	.841	.928
Under 18 (%)	0.2939	0.2780	0.2932	0.2936	0.2885	0.2934	**.000**	.958	.568
Female head HH (%)	0.1616	0.1934	0.1879	0.1576	0.1639	0.1664	**.000**	.017	.862
Same house 5 yrs ago (%)	0.5442	0.5127	0.6025	0.5623	0.5854	0.5655	**.000**	**.001**	.244
> 25 no HS diploma (%)	0.3427	0.3144	0.4053	0.3429	0.3881	0.3533	**.000**	**.000**	.060
> 25 B.A. or better (%)	0.1389	0.1767	0.1003	0.1377	0.1092	0.1343	**.000**	**.000**	.036
< Poverty line (%)	0.1056	0.1141	0.1139	0.1005	0.1072	0.1115	**.003**	.109	.716
Public assistance (%)	0.0736	0.0773	0.0885	0.0745	0.0805	0.0755	.084	.041	.578
Household income	20,340	21,526	19,635	20,869	19,812	20,301	**.000**	.013	.486
1980 geographic distribution across Census regions									
Northeast (%)	0.3797	0.2116	0.3315	0.4771	0.3889	0.4234	**.000**	**.001**	.6063
Midwest (%)	0.2183	0.2320	0.3481	0.2255	0.3222	0.2847	.302	**.004**	.5507
South (%)	0.2355	0.3227	0.2155	0.1928	0.1889	0.2044	**.000**	.552	.7744
West (%)	0.1665	0.2337	0.1050	0.1046	0.1000	0.0876	**.000**	.989	.7565

Notes: Columns (1)–(6) report the means of the variables listed in the row headings across the groups of Census tracts listed at the top of the columns. In all of these columns, the sample restriction that the Census tract must have nonmissing house price data in 1980, 1990, and 2000 is added. Columns (7)–(9) report the *p*-values from tests that the means in different sets of the subsamples are equal. The 1980 mean housing price in the two-mile- and three-mile-radius circles are calculated as the weighted mean across Census tracts that fall within the circle, where the weight is the fraction of the tract's land area inside the circle multiplied by the tract's 1980 population. All other entries in the table refer to characteristics of the tracts where the sites are located (except the column (2) entries, which report the means in tracts without a site). The *p*-values less than .01 are denoted in bold. For the air-conditioning and bath questions, the numerator is year-round housing units and the denominator is all housing units. For all other variables in the "housing characteristics" category, the denominator is all housing units. In contrast to the remainder of the paper, the dollar figures are not adjusted for inflation.

Geographic Distribution of Hazardous Waste Sites

NPL Hazardous Waste Sites in the All NPL Sample



Geographic Distribution of Hazardous Waste Sites

HRS Sample with 1982 HRS Scores Exceeding 28.5

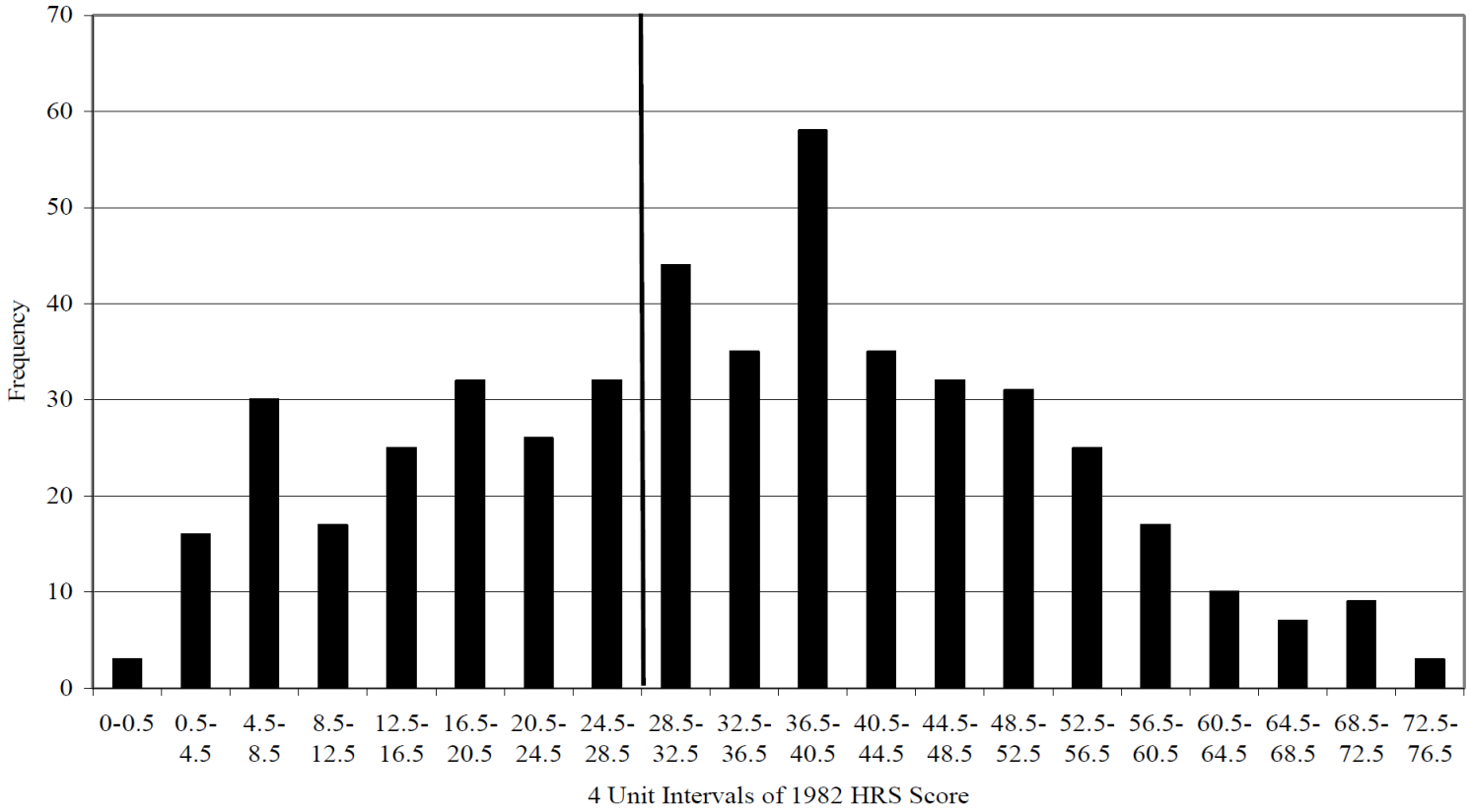
- Sites are spread throughout the United States

Geographic Distribution of Hazardous Waste Sites

HRS Sample with 1982 HRS Scores Below 28.5

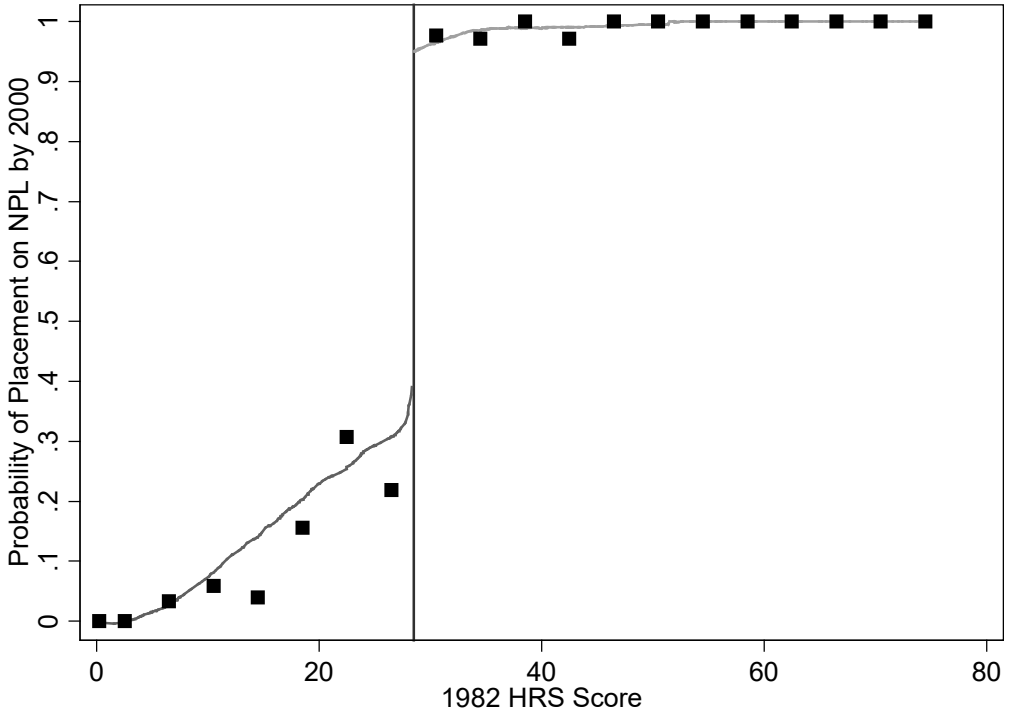
- Spread in fewer states
- Poses a problem for identification in the localized housing market shocks
- Use models that include state fixed effects

Distribution Of 1982 HRS Scores In the 1982 HRS Sample



HRS score above 28.5 Predicts NPL Status

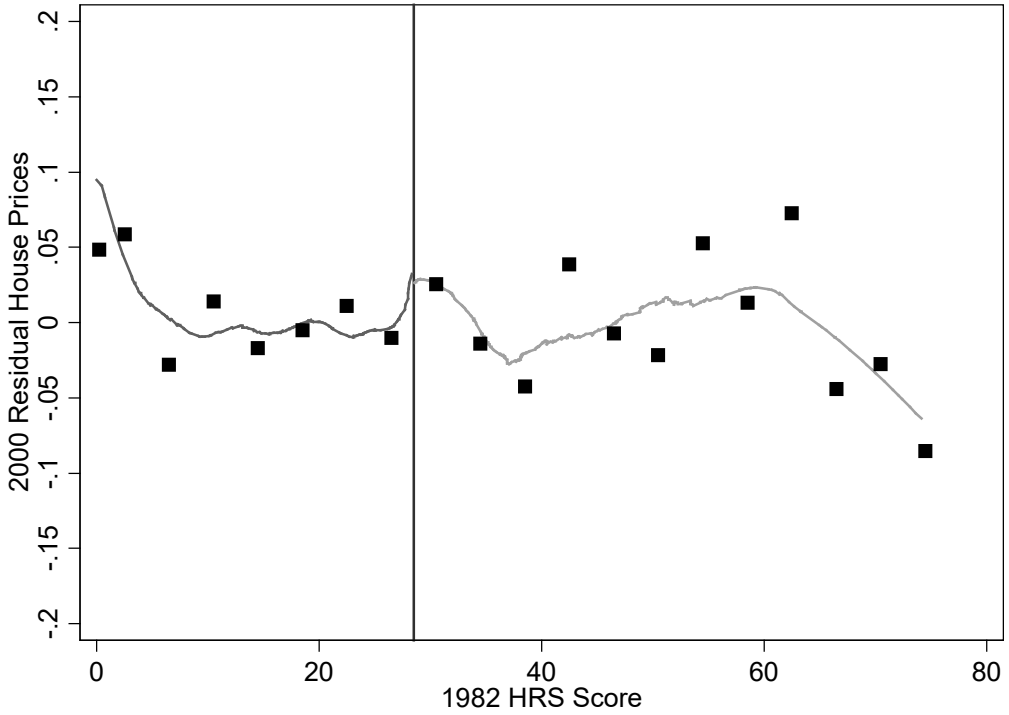
Probability of Placement on the NPL by 1982 HRS Score



- Virtually all sites with initial scores greater than 28.5 were placed on the NPL by 2000
- A statistical model reveals that a HRS score above 28.5 is associated with an 83% rise in the probability of NPL placement

Weak link between year 2000 residual housing prices and 1982 HRS score

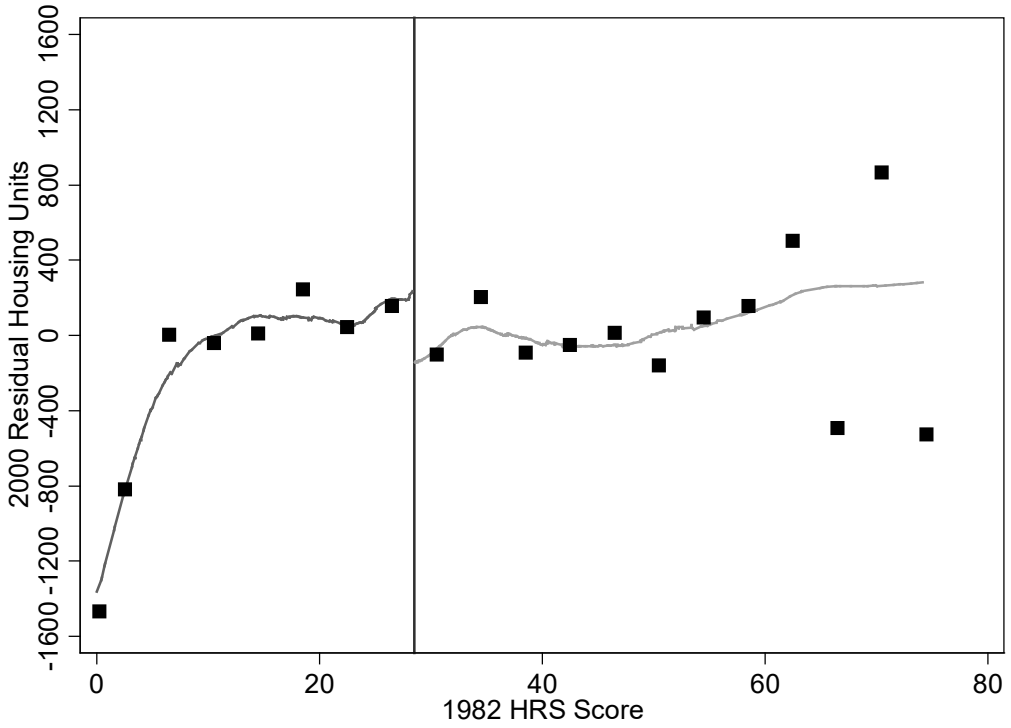
Year-2000 Residual House Prices by 1982 HRS Score



- Based on nonparametric regressions adjusting for the column (4) covariates of Table 4.
- Estimated separately below (dark line) and above (light line) the 28.5 threshold.

Similar Results for Housing Supply

Year-2000 Residual Housing Units by 1982 HRS Score



- Based on nonparametric regressions adjusting for the column (4) covariates of Table 7.
- Estimated separately below (dark line) and above (light line) the 28.5 threshold.

TABLE III

Conventional Estimates of the Association between NPL Status and House Prices with Data from the Entire United States

	(1)	(2)	(3)
A. All NPL sample, own Census tract observation			
1(NPL status by 2000)	0.040	0.046	0.067
	(0.012)	(0.011)	(0.009)
R^2	0.579	0.654	0.779
B. All NPL sample, 3-mile-radius circle sample observation			
1(NPL status by 2000)	0.030	0.060	0.106
	(0.011)	(0.013)	(0.011)
Ho: > 0.046, p-value	.061	.862	.999
R^2	0.580	0.652	0.776
1980 ln house price	Yes	Yes	Yes
1980 housing characteristics	No	Yes	Yes
1980 economic and demographic variables	No	No	Yes
State fixed effects	No	No	Yes

TABLE III

Conventional Estimates of the Association between NPL Status and House Prices with Data from the Entire United States

	(1)	(2)	(3)
C. Restrict NPL sites to those in 1982 HRS sample, own Census tract observation			
1(NPL status by 2000)	0.071	0.076	0.057
	(0.016)	(0.015)	(0.013)
R^2	0.581	0.655	0.780
D. Restrict NPL sites to those in 1982 HRS sample, 3-mile-radius circle sample observation			
1(NPL status by 2000)	0.046	0.143	0.191
	(0.015)	(0.021)	(0.021)
Ho: > 0.058, p-value	.215	.999	.999
R^2	0.580	0.653	0.777
1980 ln house price	Yes	Yes	Yes
1980 housing characteristics	No	Yes	Yes
1980 economic and demographic variables	No	No	Yes
State fixed effects	No	No	Yes

Notes: The entries report the coefficient and heteroscedastic-consistent standard error (in parentheses) on the NPL indicator, as well as the R^2 statistic, from twelve separate regressions. The controls are listed in the row headings at the bottom of the table. Panels B and D also report p-values from tests of whether the NPL parameters multiplied by the 1980 aggregate value of the housing stock exceed the average cost of a cleanup, which is $39 million and $43 million in Panels B and D, respectively. The aggregate values of the housing stock in the three-mile-radius circles around NPL sites in Panels B and D are $855 million and $736 million, respectively. The sample size is 42,974 in Panels A and B and 42,321 in Panels C and D. In Panel A/B (C/D) 985 (332) observations are from an area containing a hazardous waste site that had been on the NPL at any time prior to the year-2000 observation on housing prices. The difference between A/B and C/D is that in C/D observations from areas with the 653 NPL sites that were not tested for inclusion in the initial NPL are dropped. The remainder of the sample is composed of the 41,989 observations on Census tracts with complete housing price data that neither have an NPL site nor are adjacent to a tract with an NPL site. In Panels A and C, the unit of observation is the tract that contains the site and in B and D it is based on the Census tracts that fall within circles centered at the site with a radius of three miles. A few Census tracts have multiple sites. Both here and in the subsequent tables, observations from these tracts are weighted by the square root of the number of sites in the regressions. See the text and Data Appendix for further details.

TABLE IV

QUASI-EXPERIMENTAL ESTIMATES OF THE EFFECT OF NPL STATUS ON HOUSE PRICES, SAMPLES BASED ON THE 1982 HRS SAMPLE SITES

					RD-style estimators		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	A. Own Census tract						
1(NPL status by 2000)	0.035	0.037	0.043	0.047	0.007	0.022	0.027
	(0.031)	(0.035)	(0.031)	(0.027)	(0.063)	(0.042)	(0.038)
1980 ln house price	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Instrument for 1 (NPL 2000)	No	Yes	Yes	Yes	Yes	Yes	Yes
1980 housing charasteristics	No	No	Yes	Yes	Yes	Yes	Yes
1980 economic and demographic variables	No	No	No	Yes	Yes	Yes	Yes
State fixed effects	No	No	No	Yes	Yes	Yes	Yes
Quadratic in 1982 HRS score	No	No	No	No	Yes	No	No
Control for pathway scores	No	No	No	No	No	Yes	No
RD sample	No	No	No	No	No	No	Yes

TABLE IV

Quasi-Experimental Estimates of the Effect of NPL Status on House Prices, Samples Based on the 1982 HRS Sample Sites

					RD-style estimators		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	B. Adjacent Census tracts						
1(NPL status by 2000)	0.071	0.066	0.012	0.015	−0.006	−0.002	0.001
	(0.031)	(0.035)	(0.029)	(0.022)	(0.056)	(0.035)	(0.035)
1980 ln house price	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Instrument for 1 (NPL 2000)	No	Yes	Yes	Yes	Yes	Yes	Yes
1980 housing charasteristics	No	No	Yes	Yes	Yes	Yes	Yes
1980 economic and demographic variables	No	No	No	Yes	Yes	Yes	Yes
State fixed effects	No	No	No	Yes	Yes	Yes	Yes
Quadratic in 1982 HRS score	No	No	No	No	Yes	No	No
Control for pathway scores	No	No	No	No	No	Yes	No
RD sample	No	No	No	No	No	No	Yes

TABLE IV

QUASI-EXPERIMENTAL ESTIMATES OF THE EFFECT OF NPL STATUS ON HOUSE PRICES, SAMPLES BASED ON THE 1982 HRS SAMPLE SITES

					RD-style estimators		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
C. Two-mile radius from hazardous waste sites							
1(NPL status by 2000)	0.021	0.019	0.011	0.001	0.023	–0.018	–0.007
	(0.028)	(0.032)	(0.029)	(0.023)	(0.054)	(0.035)	(0.034)
Ho: > 0.138, *p*-value	.000	.000	.000	.000	.018	.000	.000
1980 ln house price	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Instrument for 1 (NPL 2000)	No	Yes	Yes	Yes	Yes	Yes	Yes
1980 housing charasteristics	No	No	Yes	Yes	Yes	Yes	Yes
1980 economic and demographic variables	No	No	No	Yes	Yes	Yes	Yes
State fixed effects	No	No	No	Yes	Yes	Yes	Yes
Quadratic in 1982 HRS score	No	No	No	No	Yes	No	No
Control for pathway scores	No	No	No	No	No	Yes	No
RD sample	No	No	No	No	No	No	Yes

TABLE IV

QUASI-EXPERIMENTAL ESTIMATES OF THE EFFECT OF NPL STATUS ON HOUSE PRICES, SAMPLES BASED ON THE 1982 HRS SAMPLE SITES

					RD-style estimators		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
D. Three-mile radius from hazardous waste sites							
1(NPL status by 2000)	0.059	0.055	0.035	–0.004	–0.027	–0.024	–0.006
	(0.033)	(0.038)	(0.031)	(0.022)	(0.051)	(0.034)	(0.034)
Ho: > 0.058, *p*-value	.483	.467	.236	.003	.048	.007	.031
1980 ln house price	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Instrument for 1 (NPL 2000)	No	Yes	Yes	Yes	Yes	Yes	Yes
1980 housing charasteristics	No	No	Yes	Yes	Yes	Yes	Yes
1980 economic and demographic variables	No	No	No	Yes	Yes	Yes	Yes
State fixed effects	No	No	No	Yes	Yes	Yes	Yes
Quadratic in 1982 HRS score	No	No	No	No	Yes	No	No
Control for pathway scores	No	No	No	No	No	Yes	No
RD sample	No	No	No	No	No	No	Yes

Notes: The entries report the results from 28 separate regressions. The ln (2000 median house price) is the dependent variable throughout the table. In Panels A and B (C and D) the samples are based on the 1982 HRS sites and *N* is 487 (483) in columns (1) through (6). Column (7) utilizes the RD sample of the 227 (226) sites in Panels A and B (C and D) with 1982 HRS scores between 16.5 and 40.5. The entries are the regression coefficients and heteroscedastic consistent standard errors (in parentheses) associated with the NPL indicator. The NPL indicator is instrumented with an indicator for whether the tract had a hazardous waste site with a 1982 HRS score exceeding 28.5 in columns (2) through (7); in column (1) the results come from an OLS approach. Panels C and D also report *p*-values from tests of whether the NPL parameters multiplied by the value of the housing stock in 1980 exceeds $43 million, which is our best estimate of the cost of the average cleanup. The values of the housing stocks in 1980 in the four panels are roughly $75, $552, $311, and $736 (millions of year-2000 $), respectively. The units of observation are the Census tract that contains the site (Panel A), tracts that share a border with the site (Panel B), the areas within a circle of two-mile radius from the site (Panel C), and the areas within a circle of three-mile radius from the site (Panel D). See the notes to Table III, the text, and the Data Appendix for further details.

TABLE V

QUASI-EXPERIMENTAL ESTIMATES OF STAGES OF SUPERFUND CLEANUPS ON HOUSING RENTAL RATES, SAMPLE OF TWO-MILE-RADIUS CIRCLES AROUND 1982 HRS SAMPLE SITES

			RD-style estimators		
	(1)	(2)	(3)	(4)	(5)
1(NPL only)	0.126	−0.018	−0.040	−0.054	−0.043
[115 sites, mean HRS = 40.2]	(0.046)	(0.033)	(0.049)	(0.037)	(0.051)
1(ROD & incomplete remediation)	0.106	−0.017	−0.045	−0.059	−0.075
[329 sites, mean HRS = 44.3]	(0.030)	(0.022)	(0.041)	(0.028)	(0.032)
1(construction complete or NPL deletion)	0.062	0.002	−0.023	−0.036	−0.034
[214 sites, mean HRS = 41.6]	(0.032)	(0.021)	(0.041)	(0.028)	(0.031)
p-value from *F*-test of equality	.22	.59	.51	.47	.37
1980 rental rate	Yes	Yes	Yes	Yes	Yes
1980 housing characteristics of rental units	No	Yes	Yes	Yes	Yes
1980 economic and demographic variables	No	Yes	Yes	Yes	Yes
State fixed effects	No	Yes	Yes	Yes	Yes
Quadratic in 1982 HRS score	No	No	Yes	No	No
Control for pathway scores	No	No	No	Yes	No
RD sample	No	No	No	No	Yes

Notes: The entries report the results from five separate instrumental variable regressions. The ln (median rental rate) is the dependent variable throughout the table. The sample is composed of two-mile-radius circles around the 1982 HRS sample sites. There are two observations per circle, one for 2000 and one for 1990. The resulting sample sizes in columns (1) through (4) are 966, 960, 960, and 452, respectively. Here, the indicator variable for NPL status has been replaced by three independent indicator variables. They are equal to 1 for sites that by 1990 or 2000 were placed on the NPL but no ROD had been issued, had been issued an ROD but remediation was incomplete, and had been diagnosed "construction complete" or deleted from the NPL, respectively. The instruments are the interactions of the indicator for a 1982 HRS score above 28.5 and these three independent indicators. The table reports the instrumental-variable parameter estimates and standard errors clustered at the site level for the three indicators of cleanup status. The table also reports the p-value associated with an F-test that the three parameters are equal. The effect of all of the controls listed in the row headings are allowed to differ in 1990 and 2000. See the notes to the previous tables and the text for further details.

TABLE VI

QUASI-EXPERIMENTAL ESTIMATES OF 2000 NPL STATUS ON 2000 DEMAND SHIFTERS, SAMPLE OF TWO-MILE-RADIUS CIRCLES AROUND 1982 HRS SAMPLE SITES

			RD-style estimators		
	(1)	(2)	(3)	(4)	(5)
	A. Income and wealth				
Household income	2,698	1,431	−1,232	123	−593
[1980 mean: 42,506;	(1237)	(1302)	(3130)	(1900)	(2227)
2000–1980 mean: 14,301]					
Public assistance (%)	−0.007	−0.005	0.008	0.003	0.004
[1980 mean: 0.078;	(0.003)	(0.003)	(0.007)	(0.004)	(0.005)
2000–1980 mean: 0.000]					
College graduates (%)	0.001	−0.001	−0.009	−0.005	−0.010
[1980 mean: 0.134;	(0.007)	(0.007)	(0.019)	(0.011)	(0.013)
2000–1980 mean: 0.082]					
1980 dependent variable	Yes	Yes	Yes	Yes	Yes
State fixed effects	No	Yes	Yes	Yes	Yes
Quadratic in 1982 HRS score	No	No	Yes	No	No
Control for pathway scores	No	No	No	Yes	No
RD sample	No	No	No	No	Yes

TABLE VI

QUASI-EXPERIMENTAL ESTIMATES OF 2000 NPL STATUS ON 2000 DEMAND SHIFTERS, SAMPLE OF TWO-MILE-RADIUS CIRCLES AROUND 1982 HRS SAMPLE SITES

			RD-style estimators		
	(1)	(2)	(3)	(4)	(5)
B. Demographics demand shifters					
Population under age 6 (%)	0.000	−0.000	0.002	0.000	0.001
[1980 mean: 0.086;	(0.001)	(0.001)	(0.003)	(0.002)	(0.002)
2000–1980 mean: –0.019]					
Population over age 65 (%)	−0.000	−0.003	−0.014	−0.007	−0.005
[1980 mean: 0.106;	(0.004)	(0.004)	(0.009)	(0.005)	(0.005)
2000–1980 mean: 0.019]					
Black (%)	−0.015	−0.016	−0.007	−0.012	−0.008
[1980 mean: 0.088;	(0.008)	(0.007)	(0.018)	(0.010)	(0.009)
2000–1980 mean: 0.026]					
1980 dependent variable	Yes	Yes	Yes	Yes	Yes
State fixed effects	No	Yes	Yes	Yes	Yes
Quadratic in 1982 HRS score	No	No	Yes	No	No
Control for pathway scores	No	No	No	Yes	No
RD sample	No	No	No	No	Yes

TABLE VI

QUASI-EXPERIMENTAL ESTIMATES OF 2000 NPL STATUS ON 2000 DEMAND SHIFTERS, SAMPLE OF TWO-MILE-RADIUS CIRCLES AROUND 1982 HRS SAMPLE SITES

			RD-style estimators		
	(1)	(2)	(3)	(4)	(5)
	C. Total population				
Total population	1,864	514	−2,342	−23	−289
[1980 mean: 18,038;	(526)	(522)	(1,556)	(809)	(811)
2000–1980 mean: 1,226]					
1980 dependent variable	Yes	Yes	Yes	Yes	Yes
State fixed effects	No	Yes	Yes	Yes	Yes
Quadratic in 1982 HRS score	No	No	Yes	No	No
Control for pathway scores	No	No	No	Yes	No
RD sample	No	No	No	No	Yes

Notes: The entries report the results from 35 separate instrumental variable regressions. The year-2000 values of the italicized variables are the dependent variables. The unit of observation is the area within a circle of a two-mile radius around the 1982 HRS sample sites. The sample sizes are 483 in columns (1) through (4) and 226 in column (5). The variable of interest is an indicator that equals 1 for observations from tracts with a hazardous waste site that was placed on the NPL by 2000, and this variable is instrumented with an indicator for whether the tract had a hazardous waste site with a 1982 HRS score exceeding 28.5. The entries are the regression coefficients and heteroscedastic consistent standard errors (in parentheses) associated with the NPL indicator. The mean of the dependent variable in 1980 and the mean change between 2000 and 1980 are reported in square brackets (household income is reported in year-2000 dollars). See the notes to the previous tables and the text for further details.

TABLE VII

Quasi-Experimental Estimates of the Effect of Year-2000 NPL Status on Housing Supply, Samples of Two-Mile- and Three-Mile-Radius Circles around 1982 HRS Sample Sites

			RD-style estimators		
	(1)	(2)	(3)	(4)	(5)
	Total housing units				
Two-mile radius from hazardous waste sites	332	94	−829	−208	−255
[1980 mean: 6,835; 2000–1980 mean: 853]	(139)	(147)	(349)	(210)	(187)
1980 dependent variable and ln house price	Yes	Yes	Yes	Yes	Yes
1980 housing characteristics	No	Yes	Yes	Yes	Yes
1980 economic and demographic variables	No	Yes	Yes	Yes	Yes
State fixed effects	No	Yes	Yes	Yes	Yes
Quadratic in 1982 HRS score	No	No	Yes	No	No
Control for pathway scores	No	No	No	Yes	No
RD sample	No	No	No	No	Yes

TABLE VII

QUASI-EXPERIMENTAL ESTIMATES OF THE EFFECT OF YEAR-2000 NPL STATUS ON HOUSING SUPPLY, SAMPLES OF TWO-MILE- AND THREE-MILE-RADIUS CIRCLES AROUND 1982 HRS SAMPLE SITES

			RD-style estimators		
	(1)	(2)	(3)	(4)	(5)
	Total housing units				
Three-mile radius from hazardous waste sites	1,046	292	−903	61	−77
[1980 mean: 15,657; 2000–1980 mean: 1,960]	(317)	(278)	(669)	(408)	(356)
1980 dependent variable and ln house price	Yes	Yes	Yes	Yes	Yes
1980 housing characteristics	No	Yes	Yes	Yes	Yes
1980 economic and demographic variables	No	Yes	Yes	Yes	Yes
State fixed effects	No	Yes	Yes	Yes	Yes
Quadratic in 1982 HRS score	No	No	Yes	No	No
Control for pathway scores	No	No	No	Yes	No
RD sample	No	No	No	No	Yes

Notes: The entries report the results from ten separate instrumental variable regressions. The dependent variables are the number of housing units. The results are reported for the cases in which the units of observation are the areas within a circle of two- and three-mile radius around the 1982 HRS sample sites. The samples sizes are 483 in columns (1) through (4) and 226 in column (5). The dependent and independent variables are calculated as weighted means across the relevant Census tracts, where the weight is the fraction of the tract that falls within the circle multiplied by the tract's 1980 population. The variable of interest is an indicator for NPL status and this variable is instrumented with an indicator for whether the tract had a hazardous waste site with a 1982 HRS score exceeding 28.5. The entries are the regression coefficients and heteroscedastic consistent standard errors (in parentheses) associated with the NPL indicator. The means of the dependent variable in 1980 and the mean change between 2000 and 1980 are reported in square brackets. See the notes to the previous tables, the text, and the Data Appendix for further details.

Interpretation: Three Possible Explanations

1) Heterogeneity in individuals' willingness to pay; individuals with low valuations live near hazardous waste sites
2) Imperfect information or failure to believe that clean-ups reduce health risks
3) The clean-ups don't substantially change the health risks

Summary of Findings

1) **Median time until Superfund clean-up completed is 12-13 years**
2) **Expected costs of Superfund clean-ups are about $28 million (2000$); actual costs about 1.5 X larger**
3) **Superfund sites are in neighborhoods that differ substantially from the rest of the country (e.g.** higher % of mobile homes; lower population density etc.)
4) **Confounding appears to be less important with quasi-experiment based on 1982 HRS scores**
5) **Conventional estimates suggest gains in property values of 7-9% in census tracts containing the site and neighboring tracts nearly 20 years later**
6) **Quasi-experimental estimate suggest that Superfund clean-ups have little effect on nearby property values.**
7) **No immediate decline in rental rates and little difference in effects on rental rates across stages of clean-up**
8) **No evidence of increases in population of tract containing site. Demographic composition of site is largely unchanged.**

Future Directions

1) **Directly estimate the health benefits of clean-ups**

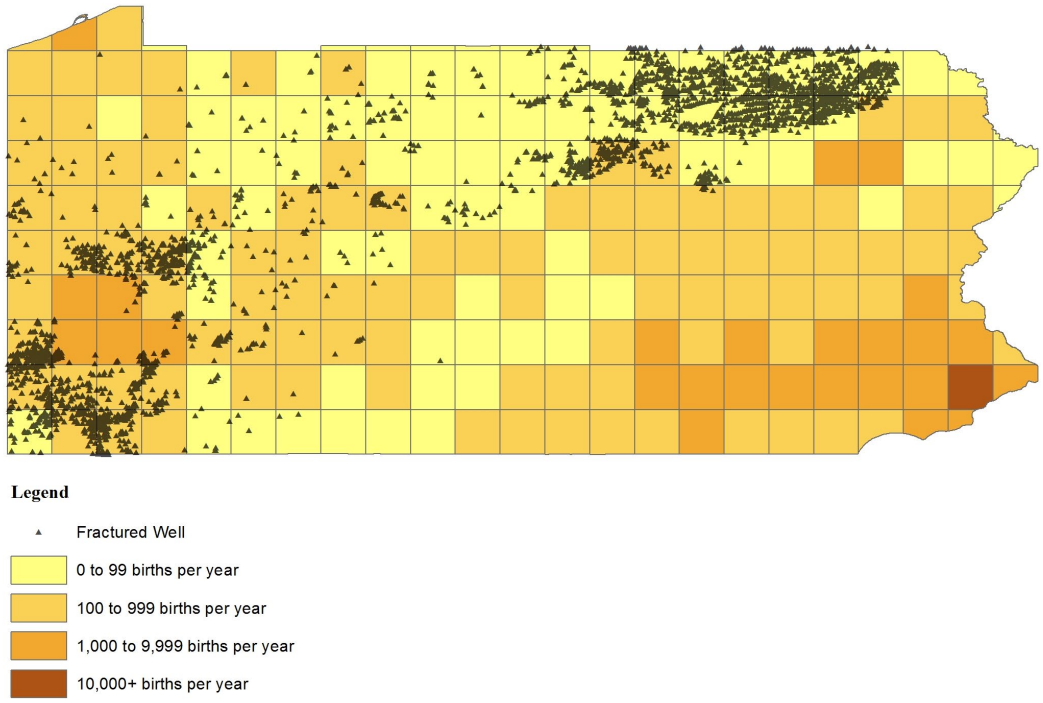
Hydraulic Fracturing and Infant Health: New Evidence from Pennsylvania

Janet Currie, Princeton University

Michael Greenstone, University of Chicago

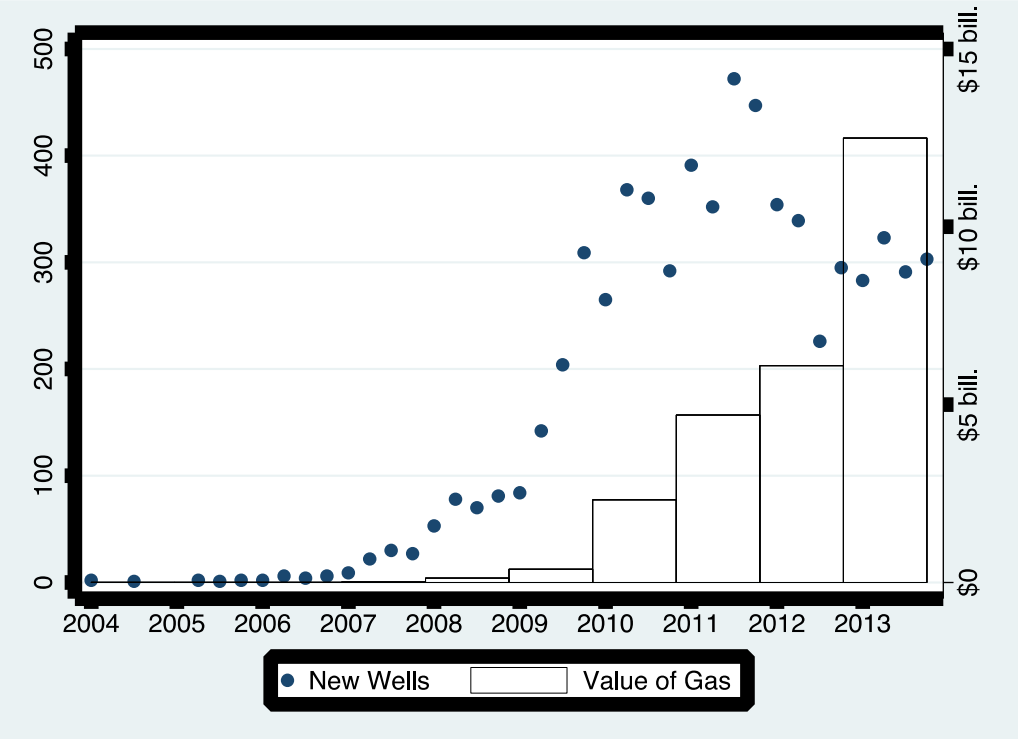
Katherine Meckel, Texas A&M University

Locations of Births and Fractured Wells, Pennsylvania



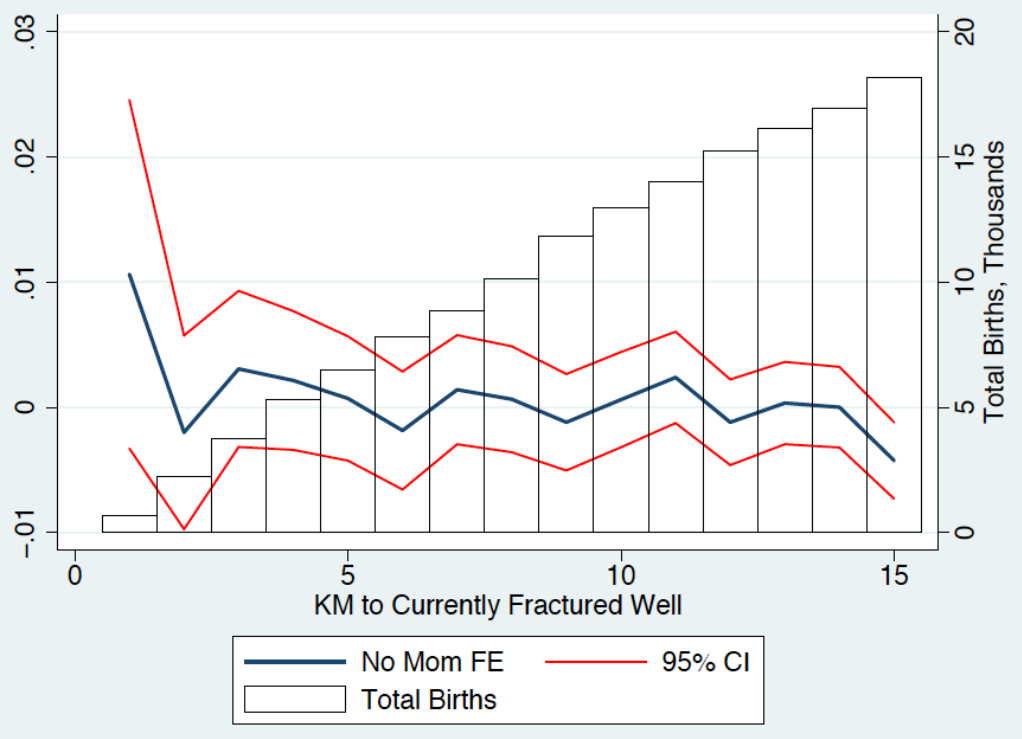
- Each square displayed above is 0.25 degrees latitude by 0.25 degrees longitude.
- Use the universe of birth certificates in Pennsylvania, 2004–2013, which includes maternal address, to calculate average yearly births per square.

Number of Fractured Wells and Value of all Drilling in PA, 2004–2013



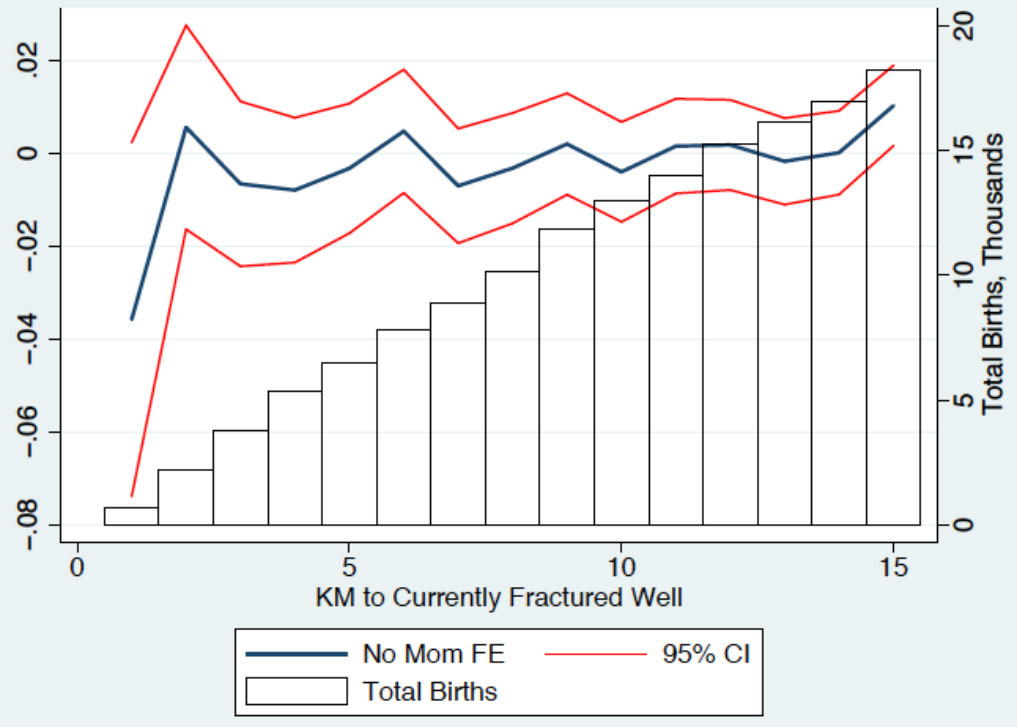
- Blue dots indicate total fractured wells in PA by spud year and quarter
- Bar graphs indicate annual value of gas from fractured wells in PA

Effect of Fracturing on Low Birth Weight, County Fixed Effects



- Lines indicate in utero effect on infant health of hydraulic fracturing at 1 km intervals from the well site
- Bars indicate average yearly births at each distance from a well site

Effect of Fracturing on Infant Health Index, County Fixed Effects



- Lines indicate in utero effect on infant health of hydraulic fracturing at 1 km intervals from the well site.
- Bars indicate average yearly births at each distance from a well site.

Table 1: Difference in Means

	Near: 0-1 km		Far: 3-15 km		P-Values		
	Before (1)	After (2)	Before (3)	After (4)	(5) (1) - (2)	(6) (3) - (4)	(7) D-in-D
Mother Characteristics							
Married	0.68	0.61	0.62	0.63	0.00	0.00	0.00
Black	0.01	0.02	0.12	0.06	0.01	0.00	0.00
Hispanic	0.01	0.01	0.02	0.02	0.66	0.00	0.57
< High School	0.11	0.12	0.11	0.10	0.27	0.00	0.01
High School	0.28	0.31	0.26	0.25	0.02	0.00	0.00
Some College	0.32	0.32	0.28	0.29	0.89	0.00	0.43
College	0.22	0.17	0.22	0.23	0.00	0.00	0.00
Advanced	0.08	0.08	0.13	0.13	0.74	0.69	0.74
Less than 20	0.06	0.06	0.06	0.05	0.89	0.00	0.11
20-24	0.21	0.26	0.22	0.21	0.00	0.00	0.00
25-29	0.30	0.31	0.29	0.29	0.42	0.00	0.86
30-34	0.26	0.23	0.27	0.28	0.01	0.00	0.00
Over 35	0.17	0.14	0.17	0.16	0.01	0.00	0.07

Table 1: Difference in Means

	Near: 0-1 km		Far: 3-15 km		P-Values		
	Before (1)	After (2)	Before (3)	After (4)	(5) (1) - (2)	(6) (3) - (4)	(7) D-in-D
Infant Characteristics							
Male	0.50	0.52	0.51	0.52	0.24	0.27	0.32
1st born	0.42	0.41	0.43	0.43	0.66	0.15	0.84
2nd born	0.33	0.34	0.33	0.33	0.46	0.05	0.68
3rd born	0.16	0.16	0.15	0.15	0.79	0.64	0.73
4th and up	0.09	0.09	0.09	0.09	0.90	0.51	0.98
Health Outcomes							
Low Birth Weight	0.05	0.07	0.06	0.06	0.04	0.00	0.01
Birth Weight	3354.35	3312.81	3316.94	3331.08	0.01	0.00	0.00
Health Index	0.05	0.01	0.01	0.02	0.01	0.00	0.00
N	4,871	1,798	133,107	78,366			

Notes: The data source is the universe of birth certificates in PA, 2004-2013, matched to the PA DEP Internal Operator Well Inventory, which is described in the notes to Figure 1. Maternal and infant demographic indicators and health outcomes are recorded at the time of birth. "Near: 0-1 km" indicates that the mother lives within 0-1 km of at least one well site. "Far: 3-15 km" indicates that the mother lives 3-15 km from the nearest well site. Columns 5-8 report p-values from t-tests of equality of means across the different samples indicated. Column 8 tests whether (2)-(1)=(4)-(3).

Table 2: Effect of Fracturing on Infant Health

	(Near = 0-1 km) x After			(Near = 1-2 km) x After			(Near = 2-3 km) x After		
Dependent Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Low Birth Weight	0.016**	0.015**	0.012	0.006+	0.005	0.004	0.009***	0.008***	0.007
(mean = 0.065)	(0.007)	(0.007)	(0.014)	(0.004)	(0.004)	(0.007)	(0.003)	(0.003)	(0.005)
Birth Weight	-38.654**	-36.707**	-13.034	-3.534	-2.023	-10.439	-7.092	-5.294	0.803
(mean = 3319.6)	(15.558)	(15.595)	(31.137)	(8.487)	(8.530)	(14.349)	(6.515)	(6.575)	(10.608)
Health Index	-0.054***	-0.052***	-0.004	-0.020**	-0.018+	-0.018	-0.028***	-0.025***	-0.015
(mean = 0.000)	(0.019)	(0.019)	(0.040)	(0.010)	(0.011)	(0.020)	(0.008)	(0.008)	(0.015)
N	1,086,917	231,578	231,578	1,102,424	247,085	247,085	1,117,919	262,580	262,580
Mother FE	No	No	Yes	No	No	Yes	No	No	Yes
Under 15 km	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes

Each coefficient and standard error are from a different regression and represent the effect on the given infant health outcome of in utero exposure to fracturing (conception occurs after well spud date) within the indicated distance. The data sources for the regression are the universe of birth certificates issued in PA from 2004 to 2013 and the PA DEP Internal Operator Well Inventory, which is described in the notes to Figure 1. We calculate distance between maternal residence and well sites using Vincenty's formula. The infant health index ranges from 0-1; an increase indicates poorer health. Each regression specification includes region of maternal residence*year FE, year*month of birth FE, and county of maternal residence FE. Also included are the following demographic controls: mother is married, marital status missing, maternal race and ethnicity (black, Hispanic, missing), maternal education (no HS, HS diploma, some college, college, advanced degree, missing), maternal age (<20, 20-24, 25-29, 30-34, 35+, missing), child is male, child sex missing, and child parity (first, second, third, fourth and up, parity missing). Where indicated, we include a vector of maternal ID fixed effects ("mother FE"). "Under 15 km" indicates the subset of mothers living less than 15 km from the nearest well site. Standard errors are clustered on maternal ID.