"Quasi-Experimental and Experimental Approaches to Environmental Economics, with an Application to Estimating the Costs of Endangered Species Protection in North Carolina"

Michael Greenstone (MIT and NBER)
and
Ted Gayer (Georgetown University)

Resources for the Future Conference
"The Frontiers of Environmental Economics"
February, 26, 2007

Introduction

- Externalities are at the center of environmental economics.
- Without reliable estimates of the benefits and costs of pollution reduction, policies deriving from the field will be inefficient and perhaps even have negative net benefits.
  - ERT example in health policy.
- One of the frontiers of environmental economics is to improve the empirical methods to estimating benefits and costs.
- The best way forward is through increased use of quasi-experimental and experimental techniques that aim to identify exogenous variation.

Causal Hypotheses

- A causal hypothesis must contain a manipulable treatment that can be applied to a subject and an outcome that may or may not respond to the treatment.

- To isolate the treatment effect, all other determinants of the outcome must be held constant.

Fundamental Problem of Causal Inference

- The goal is to estimate $Y_{1i} – Y_{0i.}$
- But, it is impossible to observe subjects in the treated and untreated states.
- In practice, we can estimate the treatment effect as

(1) $T= E[Y_{1i}|D_i=1]-E[Y_{oi}|D_i=0]$.

Fundamental Problem of Causal Inference

This can be re-written as

(2) $E[Y_{1i}-Y_{o}|D_i=1]+\{E[Y_{0i}|D_i=1]-E[Y_{0i}|D_i=0]\}$.

→The bracketed term is called **Selection Bias**.

→ Challenge for All Empirical Work is to Identify Settings Where Selection Bias is Zero

The "Selection on Observables" Approach

- Standard Approach is to Use Observational Data to fit Regression Models like:

$$y_{ct} = X_{ct}'\beta + \theta T_{ct} + \varepsilon_{ct}, \qquad \varepsilon_{ct} = \alpha_c + u_{ct}, \quad \text{and}$$

$$T_{ct} = X_{ct}'\Pi + \eta_{ct}, \qquad \eta_{ct} = \lambda_c + v_{ct}.$$

- For consistent estimation, this approach requires $E[Y_{0i}|D_i=1, X] - E[Y_{0i}|D_i=0, X] = 0$ or Conditional Ignorability

→ Omitted factors Can Lead to Biased Estimates

→ Case of Human Health and Air Pollution, the People that are most Susceptible May Sort to "Clean Areas"

Selection on Observables (cont)

- This "selection on observables" approach is generally operationalized in three ways:
    1. Least Squares
    2. Matching (Curse of Dimensionality)
    3. Propensity Score Matching (Rubin

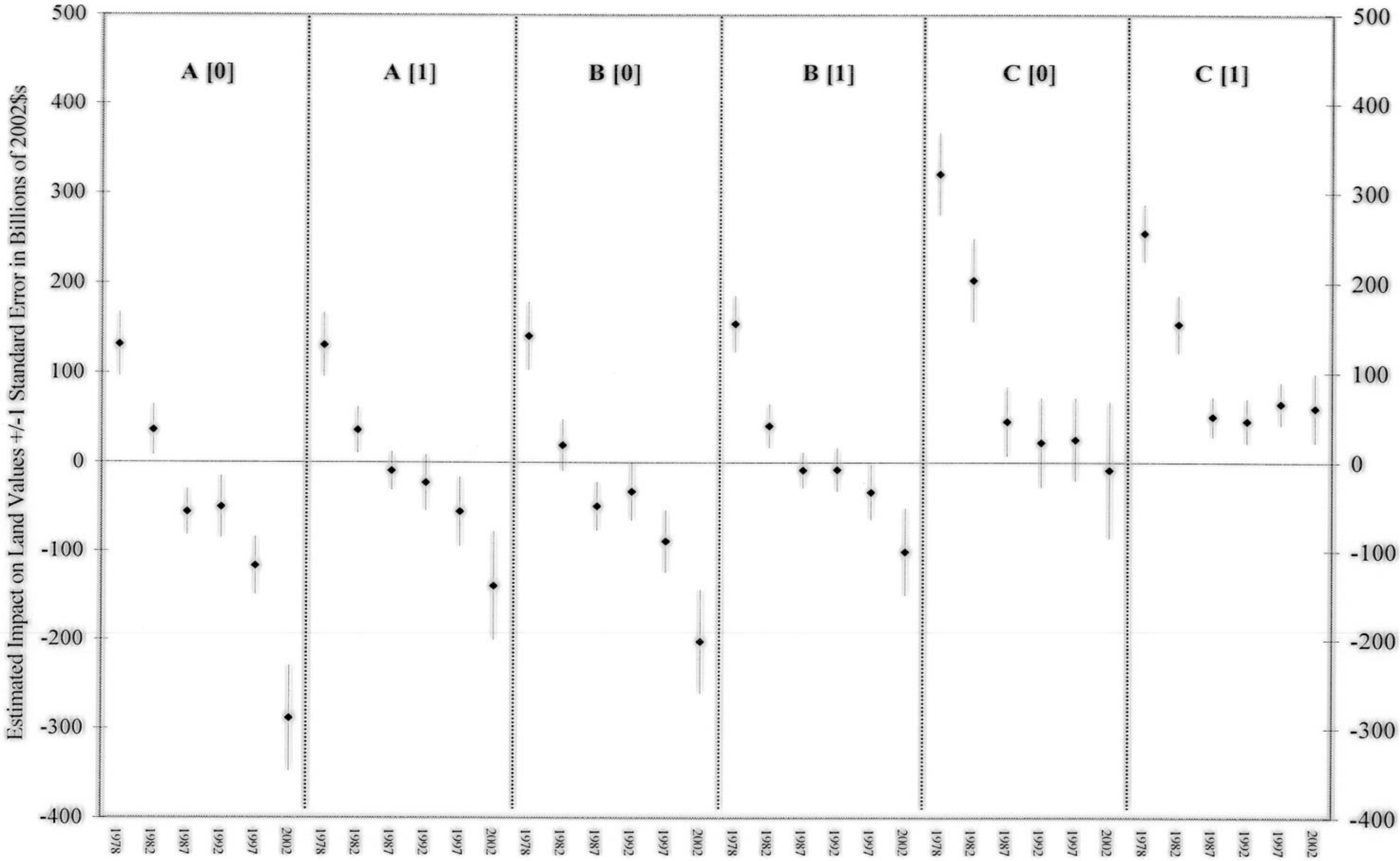
→ All Approaches Share Selection on Observables Assumption

→ Growing Consensus that this Assumption is Invalid in many Economics Settings

Table 1: Cross-Sectional Estimates of the Association between Mean TSPs and Infant Mortality Rates from Chay and Greenstone (2003)

	Infant Deaths Due to Internal Causes	
	(1)	(5)
<u>1969</u>	**2.48**	0.20
	(0.92)	(0.41)
	[412, .05]	[357,.75]
<u>1970</u>	1.30	-0.07
	(0.72)	(0.24)
	[501, .02]	[441,.67]
<u>1971</u>	1.59	0.75
	(0.98)	(0.47)
	[501, .02]	[460,.68]
<u>1972</u>	0.89	**-1.82**
	(1.20)	**(0.87)**
	[501, .00]	[455,.57]
<u>1973</u>	2.51	0.41
	(1.52)	(0.81)
	[495, .02]	[454,.66]
<u>1974</u>	**2.88**	**2.04**
	(1.34)	**(0.80)**
	[489, .03]	[455,.68]
Basic Natality	N	Y
Full Natality	N	Y
State Effects	N	Y

FIGURE 1: ± 1 STANDARD ERROR OF HEDONIC ESTIMATES OF BENCHMARK CLIMATE CHANGE SCENARIO ON VALUE OF AGRICULTURAL LAND



Notes: All dollar values are in 2002 constant dollars. Each line represents one of the 36 single year hedonic estimates of the impact of the benchmark increases of 5 degrees Fahrenheit and 8% precipitation from Table 4. The midpoint of each line is the point estimate and the top and bottom of the lines are calculated as the point estimate plus and minus one standard error of the predicted impact, respectively. See the text for further details. Taken from Deschenes and Greenstone (2007).

Associational Evidence and Two Biases

- Publication Bias: Researchers are more likely to submit for publication (and journal editors are more likely to accept) articles that find statistically significant results with expected signs.
- Regulatory Bias: Regulators place more weight on studies that find a significant negate health impact of pollution than on other studies.

→ These biases can lead to inefficient regulation and and misallocation of regulatory dollars across environmental risks.

Randomized Experiments

- In these settings, it is generally valid to assume

$E[Y_{0i}|D_i=1] - E[Y_{0i}|D_i=0] = 0$

We now turn to cases where experiments are unavailable.

Quasi-Experimental Approaches

- In a quasi-experimental evaluation, the researcher exploits differences in outcomes between a treatment group and a control group, just as in a classical randomized experiment.
- However, treatment status is determined by nature, politics, an accident, or some other action beyond the researcher's control.
- The validity of the quasi-experiment rests on the assumption that assignment to treatment and control groups is not related to other determinants of the outcome.

Quasi-Experimental Approaches: DD

- This approach exploits the availability of panel data that covers at least one period before the assignment of the treatment and one period after its assignment. Treatment effect estimated as: $\{E[Y_{1i}|D_i=1, Pd=2]-E[Y_{1i}|D_i=1, Pd=1]\}$ –
  $\{E[Y_{0i}|D_i=0, Pd=2]-E[Y_{0i}|D_i=0, Pd=1]\}$.
- DD produces valid estimate under the assumption that in the absence of the treatment the outcomes in the two groups would have changed identically in both groups between periods 1 and 2.

Quasi-Experimental Approaches: Instrumental Variables

- IV uses a variable that is correlated with the treatment but otherwise independent of the outcome.
- By isolating the exogenous variation in the treatment variable, IV solves the selection bias.
- Instrument relevance can be tested directly. Instrument exogeneity cannot be directly tested, but it can be examined through such things as examining the association between the instrument and the observable variables measured before treatment was assignment.

Quasi-Experimental Approaches: Regression Discontinuity

- The key feature of this approach is that the selection for the treatment is a nonlinear function of an observable variable. Intuition is to control for "running" variable and exploit discontinuity in selection rule.
- Chay and Greenstone (2005); Greenstone and Gallagher (2007)

Threats to Validity

- Internal Validity
- External Validity
- Construct Validity

Conquering the Threats to Validity

- Understand the Source of Variation in the Explanatory Variable of Interest
- Massive Expenditures of Shoe Leather

Can Quasi-Experimental Approaches Answer Important Questions?

- One limitation of experiments and quasi-experiments is that they restrict which research questions can be addressed.
- Yet they have been used for a wide variety of economic settings in recent years. Other approaches are likely to lack internal validity.
- A greater emphasis should be placed on these research designs in environmental economics
  - More randomized pilot evaluations of policies.
  - Greater emphasis on experiments and quasi-experiments in regulatory process.
  - More information presented on reliability of estimates in regulatory process.

Background on ESA

- The goal of the Endangered Species Act (ESA) is to protect animal and plant species from extinction.
- Highly controversial:
  - Opponents claim that is costly to landowners and inhibits development. Some claim that it presents perverse incentives for species protection.
  - Proponents claim it is needed in order to protect species from extinction. But no good time-series data on species.
  - No comprehensive study of the costs and benefits of species protection.
- **Research question:** What are the welfare costs of protecting a species under ESA?
  - Comprehensive, nationwide estimate of costs of ESA. Building data set state by state, starting with NC.
  - Use property market outcomes to infer welfare costs.

Background Literature on ESA

- Some empirical studies of preemptive habitat destruction:
  - Lueck and Michael (2003) – Foresting decisions in NC and the red-cockaded woodpecker.
- Some empirical studies on "critical habitat" effect on local and state markets:
  - Sunding (2003) – Coastal southern California.
  - Zabel and Paterson (2005) – Housing permits in California.
  - Margolis, Osgood, and List (2004) – Construction permits and the pygmy owl.

Statutory Mechanisms of ESA

- **Section 7** of ESA: Federal agencies must "consult" with the Secretary of the Interior in order to "insure that any action authorized, funded of carried out by such agency … is not likely to jeopardize the continued existence of any endangered species or threatened species or results in the destruction or adverse modification of habitat of such species…"
- **Section 9** of ESA: It is illegal to "take" a protected species, where "take means to harass, harm, pursue, hunt, shoot, wound, kill, trap, capture, or collect or attempt to engage in any such conduct." DOI defined "harm" (and was upheld by SCOTUS in 1995) as "an act [that] may include significant habitat modification or degradation where it actually kills or injures wildlife significantly impairing essential behavioral patterns, including breeding, feeding, or sheltering."

Theoretical Implications

- ESA reduces demand for newly regulated land, which reduces the price of this land.
- Now consider land in the same market but unregulated:
  - Quigley and Swoboda's (2007) model assumes costless moving within market and no moving outside of market. Regulation increases demand for the land and housing of unregulated part of market, leading to price increases (and transfer from renters to owners).
- We cannot observe the fraction of the census tract that is covered by ESA restrictions. There is also no concrete way to define the market.
- By focusing on particular census tracts, we will have an incomplete picture of the full welfare implications b/c we do not capture impacts on housing and land values on unregulated land in other census tracts within the same market.
- Model predicts housing price increase in regulated tracts. Need land measures to obtain full welfare impacts.

Research Design

- We implement a quasi-experimental evaluation that relies on the Global Conservation Status Rank (GCSR) of each species in the state.
  - GCSR (collected by NatureServe) measures the relative rarity or imperilment of each species. Ranks: Possibly extinct, critically imperiled, imperiled, vulnerable, apparently secure, demonstrably secure, and unmeasured or unmeasurable.
  - NatureServe also provides us with location data for each species in the state. We overlaid the habitat maps onto 2000 census tract boundaries for NC to determine the tracts that overlap with each of the species' habitats.
- Within each rank there are protected and unprotected species.
- Within each rank, the tracts that contain unprotected species may form a valid counterfactural for tracts that contain protected species.
- Thus, within the set of tracts that cover habitats of species with the same GCSR, we compare the evolution of property market outcomes in tracts with protected species to outcomes in tracts with species that are not protected.

Data Sources

NatureServe Data

- Data on species' habitats in NC comes from NatureServe's Natural Heritage Program.
- There are 1,227 species in the NC data set. The data file also contains the GCSR for each species.

Census Data

- The housing, demographic, and economic data come from Geolytics *Neighborhood Change Database*, which includes information from the 1970, 1980, 1990, and 2000 Censuses. We focus on the changes in census-tract level outcomes between 1990 and 2000.
- We use the Geolytics data to form a panel of census tracts based on 2000 census tract boundaries, which are drawn so that they include approximately 4,000 people in 2000.
- Outcome measures are median and mean housing value in a census tract in 2000. (Currently pursing land value data.)

Table 2: Summary Statistics of NatureServe's North Carolina Species Data

A. Full NatureServe Species Information	
# of NatureServe Species	1,227
Kingdom	
Animalia	408
Plantae	803
Fungi	16
Global Conservation Status Rank	
G1 (Critically Impreiled)	95
G2 (Imperiled)	153
G3 (Vulnerable)	250
G4 (Apparently Secure)	328
G5 (Secure)	365
G6 (Unranked)	34
G7 (Possibly Extinct)	2

Table 2: Summary Statistics of NatureServe's North Carolina Species Data

B. Listed Species Information

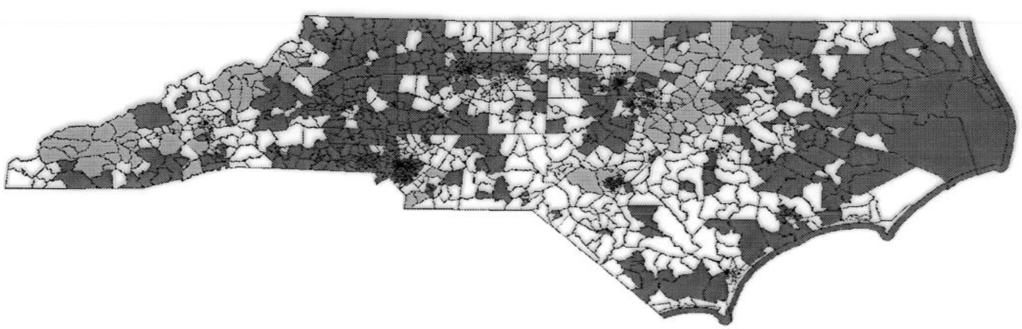
	Endangered	Threatened	Candidate	Unregulated
# of Species	42	16	4	1,165
Kingdom				
Animalia	24	7	1	376
Plantae	17	9	3	15
Fungi	1	0	0	774
Global Conservation Status Rank				
G1 (Critically Impreiled)	14	4	2	75
G2 (Imperiled)	19	6	2	126
G3 (Vulnerable)	9	4	0	238
G4 (Apparently Secure)	1	0	0	327
G5 (Secure)	0	2	0	363
G6 (Unranked)	0	0	0	34
G7 (Possibly Extinct)	0	0	0	2
Decade of Listing				
1960s	4	0	0	
1970s	9	3	0	
1980s	16	7	1	
1990s	12	6	2	
2000s	1	0	1	

Table 2: Summary Statistics of NatureServe's North Carolina Species Data

C. Census Tract Species Information

# without NatureS Species	41
# with NatureS Species	1,522

	Tracts with 1990s			
	Endangered Species	Threatened Species	Candidate Species	Non-Candidate Species
Kingdom				
Animalia	85	219	0	1,449
Plantae	97	61	28	1,170
Fungi	31	0	0	57
G1 (Critically Impreiled)	85	0	2	253
G2 (Imperiled)	65	61	28	633
G3 (Vulnerable)	50	81	0	1,012
G4 (Apparently Secure)	0	0	0	1,420
G5 (Secure)	0	141	0	944
G6 (Unranked)	0	0	0	287
G7 (Possibly Extinct)	0	0	0	3



Econometric Methods

- Least squares model:

$$y_{c2000} = X_{c1990}'\beta + \theta T_c + \varepsilon_{c2000}$$

$$T_c = X_{c1990}'\Pi + \eta_{c2000}$$

- Biased and inconsistent estimate of treatment effect if

$$E[\varepsilon_{c2000}\eta_{c2000}] \neq 0.$$

- We add a full set of GCSR dummy variables to equation at top. θ is identified from comparisons of tracts with protected species to tracts with unprotected species, for species with identical likelihoods of imperilment.

The Case for a Quasi-Experiment

- The main empirical complication in estimating costs of species protection is that areas (e.g., census tracts) that contain one or more protected species are likely different from areas that do not contain such a species.
- If these differences are unobserved and contribute to property market outcomes, a “selection on observables” approach, such as linear regression, will yield biased cost estimates.
- Column 3 of Table 3 suggests that this problem is likely to arise. Column 4 suggests that our quasi-experiment helps to reduce this problem.

Table 3a: Comparison of Ex-Ante Means between Tracts with and withouth Endangered Species (Animals)

	Difference (3)	Difference| G-Rankings (4)
1980 Mean House Value	**-6,258**	-511.7
	(3,053)	(3,336)
1990 Mean House Value	**-16,802**	-2,027
	(3,137)	(3,706)
% Mobile Homes	**0.094**	0.016
	(0.011)	(0.015)
% Occupied	**-0.054**	0.018
	(0.013)	(0.020)
% Attached	**-0.014**	-0.001
	(0.005)	(0.005)
% Detached	**0.037**	0.029
	(0.013)	(0.018)
1990 Demographics		
Population Density	**-34.510**	-7.700
	(2.740)	(4.740)
% BA or Better	**-0.061**	-0.017
	(0.010)	(0.015)
Species Listed in 1980s	**0.220**	**-0.187**
	(0.049)	**(0.066)**
Species Listed in 1970s	**0.167**	0.010
	(0.052)	(0.065)
1990 Economic		
Mean HH Income	**-3,659**	36
	(1,274)	(1369)
% Public Assistance	**0.016**	0.015
	(0.006)	(0.008)

Table 4: Least Squares Estimates of Endangered Species Listing in Census Tract in 1990s

				Animals			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Outcome = Ln Mean House Value (2000)							
	No G-Rank FEs			Include G-Rank Fixed Effects			
Endangered Species Listing in 1990s	**-0.195**	0.044	0.041	-0.041	-0.002	-0.005	-0.027
	(0.048)	(0.023)	(0.023)	(0.067)	(0.029)	(0.030)	(0.050)
Observations	1,247	1,245	1,245	1,247	1,245	1,245	881
G-Ranking Indicators	N	N	N	**Y**	**Y**	**Y**	**Y**
Include Covariates	N	Y	Y	N	Y	Y	Y
Control for Listing in Previous Decades	N	N	Y	N	N	Y	Y
Include 1980 Outcome Variable as Covariate	N	N	N	N	N	N	Y

Table 4: Least Squares Estimates of Endangered Species Listing in Census Tract in 1980s

				Animals			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Outcome = Ln Mean House Value (2000) 1990s							
	No G-Rank FEs			Include G-Rank Fixed Effects			
Endangered Species Listing in 1980s	**-0.333**	-0.028	**-0.095**	-0.018	-0.008	-0.087	-0.100
	(0.049)	(0.046)	**(0.034)**	(0.061)	(0.058)	(0.045)	(0.054)
Observations	1,334	917	917	1,334	917	917	828
G-Ranking Indicators	N	N	N	**Y**	**Y**	**Y**	**Y**
Include Covariates	N	Y	Y	N	Y	Y	Y
Control for Listing in Previous Decades	N	N	Y	N	N	Y	Y
Include 1980 Outcome Variable as Covariate	N	N	N	N	N	N	Y

Table 4: Least Squares Estimates of Endangered Species Listing in Census Tract in 1990s

				Plants			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Outcome = Ln Mean House Value (2000)							
	No G-Rank FEs			Include G-Rank Fixed Effects			
Endangered Species Listing in 1990s	**-0.197**	0.006	0.005	**-0.137**	-0.015	-0.016	-0.015
	(0.049)	(0.023)	(0.023)	**(0.053)**	(0.025)	(0.025)	(0.034)
Observations	1,114	1,113	1,113	1,114	1,113	1,113	720
G-Ranking Indicators	N	N	N	**Y**	**Y**	**Y**	**Y**
Include Covariates	N	Y	Y	N	Y	Y	Y
Control for Listing in Previous Decades	N	N	Y	N	N	Y	Y
Include 1980 Outcome Variable as Covariate	N	N	N	N	N	N	Y

Table 4: Least Squares Estimates of Endangered Species Listing in Census Tract in 1980s

	Plants						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Outcome = Ln Mean House Value (2000)							
	No G-Rank FEs			Include G-Rank Fixed Effects			
Endangered Species Listing in 1980s	**-0.245**	-0.038	-0.038	-0.122	-0.070	-0.070	-0.068
	(0.061)	(0.049)	(0.049)	(0.066)	(0.051)	(0.051)	(0.055)
Observations	1,111	717	717	1,111	717	717	632
G-Ranking Indicators	N	N	N	**Y**	**Y**	**Y**	**Y**
Include Covariates	N	Y	Y	N	Y	Y	Y
Control for Listing in Previous Decades	N	N	Y	N	N	Y	Y
Include 1980 Outcome Variable as Covariate	N	N	N	N	N	N	Y

Conclusions

- Associational evidence of benefits and costs of pollution reduction can be highly misleading and can therefore lead to poor policies.
- It is important that researchers and policymakers place greater emphasis on credible empirical approaches, such as quasi-experiments and experiments.
- We provided some preliminary quasi-experimental estimates of the property market impacts of ESA protections. These preliminary results indicate that species protection has little impact on housing values.