MIT

CEEPR Center for Energy and Environmental Policy Research

Quasi-Experimental and Experimental Approaches to Environmental Economics

by

Michael Greenstone and Ted Gayer

07-013 December 2007

**A Joint Center of the Department of Economics,
MIT Energy Initiative, and Sloan School of Management**

Quasi-Experimental and Experimental Approaches to Environmental Economics*

Michael Greenstone [a,+] and Ted Gayer [b]

December 2007

Running Title: "Quasi-Experiments & Experiments in Enviro Econ"

[a] Department of Economics, MIT, Cambridge MA 02139; NBER, Cambridge, MA; Brookings Institution, Washington DC.
[b] Georgetown Public Policy Institute, Georgetown University, Washington DC, 20007
[+] Corresponding Author. Email address: mgreenst@mit.edu; telephone: (617) 452-4127; facsimile (617) 253-1330; street address: 50 Memorial Dr., E52-359, Cambridge, MA 02142.

Quasi-Experimental and Experimental Approaches to Environmental Economics

Abstract

This paper argues that an increased application of quasi-experimental and experimental techniques will improve understanding about core environmental economics questions. This argument is supported by a review of the limitations of associational evidence in assessing causal hypotheses. The paper also discusses the benefits of experiments and quasi-experiments, outlines some quasi-experimental methods, and highlights threats to their validity. It then illustrates the quasi-experimental method by assessing the validity of a quasi-experiment that aims to estimate the impact of the Endangered Species Act on property markets in North Carolina. The paper's larger argument is that greater application of experimental and quasi-experimental techniques can identify efficient policies that increase social welfare.

Introduction

Externalities are at the center of environmental economics.[1] A classic example is a factory's release of air pollution as a byproduct of its production of a marketable good. The air pollution may negatively impact human health, for example by raising mortality rates, but abatement raises the firm's production costs. The social problem is that firms do not internalize the health costs they impose on others through the release of air pollution. High transaction costs frequently prevent the affected parties from reaching an efficient solution that accounts for the cost and benefits [22]. In these cases, government interventions (e.g., emissions taxes or limits) can be used to achieve the level of air pollution reduction that maximizes net benefits.[2] Successful interventions, however, require estimates of the costs and benefits of environmental quality.

However, environmental policy that is not based on credible empirical research can lead to inefficient policies or even policies with negative net benefits. An example of the dangers of using unreliable empirical research concerns the recent use of estrogen replacement therapy (ERT) to mitigate the symptoms of menopause, such as hot flashes and night sweats. In the 1980s, a series of observational studies concluded that ERT did not lead to a higher rate of heart disease and that, if anything, it may decrease the incidence of heart disease. Based on these studies, ERT was widely recommended for menopausal women nationwide. Some researchers noted their concern with the ability of these studies to control for unobservable determinants of heart disease; after all, there are many reasons to believe that those receiving ERT had a healthier lifestyle than those who did not. A randomized study was finally conducted in the 1990s and was stopped three years early because the results clearly indicated that ERT substantially increased the incidence of heart disease. Millions of women had received poor medical advice

due to a policy based on associational evidence [51].

The evidence in support of the vast majority of environmental policies is associational in nature. It is therefore possible that many environmental policies fail to achieve their stated goals or fail to do so efficiently.

This paper's argument is that one of the frontiers of environmental economics is to improve the measurement of the costs and benefits of environmental quality. We contend that the best way forward is to use quasi-experimental and experimental techniques that aim to identify exogenous variation in the variable of interest. Over the last two decades, these approaches have become widely accepted in other subfields of economics (e.g., labor, public finance, and development economics). Further, we believe that environmental economics is flush with opportunities to apply these techniques.

The successful application of quasi-experimental and experimental techniques offers the promise of providing a deeper understanding of the world in which we live. Furthermore, they can lead to the identification of social welfare maximizing environmental policies, which is the aim of much of environmental economics.

This paper proceeds as follows. Section I discusses the nature of causal hypotheses, reviews the standard approaches to estimating causal relationships with observational data, and presents some evidence on the reliability of these approaches. Section II discusses the benefits of experiments and quasi-experiments, outlines some quasi-experimental methods, highlights some threats to their validity, and discusses their usefulness for answering important questions. Section III illustrates some of the challenges and issues with implementing a quasi-experiment by assessing the validity of a quasi-experiment that aims to estimate the impact of the Endangered Species Act (ESA) on property markets in North Carolina. Due to space constraints

and current data limitations, a full-scale evaluation is left for future research.

I. Causal Hypotheses and Associational Evidence

The main focus of environmental economics is to address inefficiencies caused by externalities. The usefulness of any policy prescriptions stemming from this approach rests squarely on the reliability of the empirical estimates of the benefits and costs of reducing pollution.[3] Therefore, it is important to understand the validity of the different empirical approaches to estimating the causal relationships between economic activity and externalities.

A. Causal Hypotheses and the Fundamental Problem of Causal Inference

The development of reliable estimates of the costs and benefits of environmental quality necessarily begins with the specification of a causal hypothesis or hypotheses. The key features of a causal hypothesis are that it contains a manipulable treatment that can be applied to a subject and an outcome that may or may not respond to the treatment. For a causal hypothesis to have any practical relevance, it is necessary to be able to subject it to a meaningful test. Such a test requires that all other determinants of the outcome can be held constant so that the effect of the treatment can be isolated.[4]

In the ideal it would be feasible to simultaneously observe the same "subject" in the states of the world where it received the treatment and did not receive the treatment. This would guarantee that all else is held constant. Of course, it is impossible to observe the same subject in both states simultaneously. For example, in drug trials the new drug cannot simultaneously be administered to, and withheld from, the same person. This difficulty is labeled the "Fundamental Problem of Causal Inference" and has been recognized since at least Hume [46].

More formally, it is instructive to borrow from Rubin's [70] terminology of a potential outcome. Consider the case where we are interested in measuring the impact of exposure to high levels of air pollution on human health. For ease of exposition, we assume that pollution exposure is dichotomous and there are either high or low levels of air pollution exposure. We denote the health outcome for a treated person i (i.e., someone exposed to high pollution level) as Y_{1i} and the health outcome for that same person as Y_{0i} if she is not treated (i.e., she is exposed to low pollution levels). The i subscript indexes individuals. The object of interest is $Y_{1i} - Y_{0i}$, which is the effect of exposure to high pollution levels, relative to low levels.

Since $Y_{1i} - Y_{0i}$ is not constant across individuals, it is standard to focus on average effects. For clarity, we let D represent treatment status, where D=1 designates that a person was treated and D=0 indicates that person was not treated. Thus, we would like to know $E[Y_{1i} - Y_{0i} \mid D_i=1]$, which is the average causal effect of exposure to high pollution concentrations among the treated (i.e., the treatment on the treated). It is evident that the Fundamental Problem of Causal Inference binds here, because it is impossible to observe the same individual simultaneously exposed to high and low pollution concentrations. Put another way, every individual has two potential outcomes but only one is observed.

In practice, we are faced with situations where some individuals have received the treatment and others have not. In this case, we can estimate a treatment effect

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

By adding and subtracting the unobserved quantity $E[Y_{0i} \mid D_i=1]$, which is the expected outcome for treated individuals if they had not received the treatment, we can write:

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

The first term is the average causal effect of exposure to pollution and is the quantity of

interest. The second term is the selection bias that plagues the successful estimation of average causal effects. It measures the difference in potential untreated outcomes between the individuals that did and did not receive the treatment. This term could be non-zero for a multitude of reasons. For example, Chay and Greenstone [21] have shown that residents of heavily polluted areas tend to have low incomes and that the poor may also have greater health problems. In the case where Y represents a negative health outcome (e.g., mortality), this would mean that the selection bias term is positive. In general, it is difficult to predict the direction and magnitude of selection bias.

The challenge for credible empirical research is to identify settings where it is valid to assume that the selection bias term is zero or where it can be controlled for. We now turn to a discussion of the standard approach to solving this problem.

B. The "Selection on Observables" Approach

The typical way to learn about the validity of a causal hypothesis is to use observational data to fit regression models.[5] For example, consider the following cross-sectional model for county-level infant mortality rates in year t which is based on [20]:

(3) $\quad Y_{ct} = X_{ct}'\beta + \theta T_{ct} + \varepsilon_{ct}, \qquad \varepsilon_{ct} = \alpha_c + u_{ct}$, and

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

Here Y_{ct} is the infant mortality rate in county c in year t, X_{ct} is a vector of observed determinants of Y, and T_{ct} is again a dichotomous variable that indicates exposure to high or low pollution in the county. ε_{ct} and η_{ct} are the unobservable determinants of health status and air pollution levels, respectively. Each is composed of a fixed and transitory component. The coefficient θ is the "true" effect of air pollution on measured health status. The specification of equation (4)

demonstrates that the determinants of health status, X_{ct}, may also be associated with air pollution.

For consistent estimation, the least squares estimator of θ requires that $E[\epsilon_{ct}\eta_{ct}] = 0$. If this assumption is valid, the estimated θ will provide a causal estimate of the effect of air pollution on health status. However, if there are omitted permanent (α_c and λ_c) or transitory (u_{ct} and v_{ct}) factors that covary with both air pollution and health status, then the cross-sectional estimator will be biased. Poorer individuals tend to live in more polluted areas, so it is apparent that income and perhaps income changes are potential confounders. Further, equation (3) assumes a particular functional form for the explanatory variables. In practice, the true functional form is unknown and the incorrect specification of the functional form may be an additional source of misspecification.

Another source of omitted variables bias is that air pollution may cause individuals who are susceptible to its influences to engage in unobserved (to the econometrician) compensatory behavior to mitigate its impact on their health. For example, people with respiratory diseases might migrate from polluted to clean areas to avoid the health effects of air pollution or install filters (or other devices) in their homes that clean the air they breathe. These compensatory behavioral responses to pollution will bias the estimates downward. The basic problem is that it is in people's interests to protect themselves against pollution, but these efforts undermine our ability to obtain structural estimates of the effects of pollution on health.

The data sources on pollution introduce two potential sources of bias. First, it is generally impossible to construct measures of individuals' lifetime exposure to air pollution. This is problematical if human health is a function of lifetime exposure to pollution. Since historical data on individual's exposure to pollution is generally impossible to obtain, scores of studies on adult mortality rely on the assumption that current measures of air pollution

concentrations accurately characterize past levels.

The available pollution data introduce a second bias. In particular, most cities have only a few pollution monitors. The readings from these monitors are used to develop measures of individuals' true exposure to air pollution. Since there is frequently great variation within cities in pollution concentrations and individuals spend varying amounts of time inside and outside, substantial measurement error in individuals' exposures is likely. In general, measurement error attenuates the estimated coefficient in regressions, and the degree of attenuation is increasing in the fraction of total variation in observed pollution that is due to mismeasurement.[6]

The more general problem with the approach is that it will only produce estimates of T's causal impact on Y if two assumptions hold. The first assumption is that after conditioning on a vector of observable variables, X, the treatment is "ignorable". This is often referred to as the "selection on observables" assumption. Returning to the notation from the previous subsection, this assumption implies that after conditioning on X the selection bias term is irrelevant because $E[Y_{0i} \mid D_i=1, X] - E[Y_{0i} \mid D_i=0, X] = 0$.

The second assumption is that it is necessary to assume that T causes Y, rather than Y causes T. This assumption is necessary because this approach does not rely on a manipulation of T preceding the observation of Y. So, for example, in equation (3) Y could precede T, T could precede Y, or they could occur simultaneously. In contrast, a classical experiment is structured so that T precedes the realization of Y. Even when this is the case, it is still necessary to assume that T causes Y (rather than Y causes T), because this approach does not rely on a manipulation of T preceding the observation of Y.

There are three primary ways to operationalize the "selection on observables" approach. The first method is to fit the linear regression outlined in equation (3). The second is to match

treatment and control observations with identical values of all of the components of the X vector.[7] This method is almost always infeasible when there are many variables or even a few continuous variables, because it becomes impossible to obtain matches across the full set of characteristics.

The third approach offers a solution to this "curse of dimensionality". Specifically, Rosenbaum and Rubin [69] suggest matching on the propensity score—the probability of receiving the treatment conditional on the vector X. This probability is an index of all covariates and effectively compresses the multi-dimensional vector of covariates into a simple scalar. The advantage of the propensity score approach is that it offers the promise of providing a feasible method to control for the observables in a more flexible manner than is possible with linear regression.[8]

All of these methods share a faith in the selection on observables assumption, which implies that unobservable variables do not covary with both air pollution and health status. Of course, this assumption is untestable because the unobservables are by their very nature unobservable. In the case of OLS specifications, this is sometimes referred to as "OLS and hope for the best". Indeed, there is a growing consensus in many applied microeconomic fields that the selection on observables assumption is unlikely to be valid in most settings [56, 3]. It seems reasonable to assume that environmental economics is no exception, and we provide evidence of this in the next subsection.

C. The Quality of the Evidence from the Selection on Observables Approach

Here, we present some examples based on key environmental economics questions of the quality of evidence produced by the selection on observables approach. Specifically, we assess

the sensitivity of the results to changes in the sample and in the set of conditioning covariates. We begin by borrowing from Chay and Greenstone [20], which examines the relationship between ambient concentrations of total suspended particulates (TSPs) and infant mortality. That paper presents evidence of the reliability of using a cross-sectional approach to test the causal hypothesis that TSPs exposure causes elevated infant mortality rates. While our focus for now is on the difficulties of estimating the health effects of externalities, the implications equally apply to studies of the valuation of externalities (e.g., hedonic analyses) and to studies of the costs of remediating or abating externalities [36, 21].

Table I presents regression estimates of the effect of TSPs on the number of internal infant deaths within a year of birth per 100,000 live births for each cross-section from 1969-1974. ("Internal deaths" are those due to health reasons, in contrast to death due to external, non-health reasons such as accidents and homicides.[9]) The entries report the coefficient on TSPs and its standard error (in parentheses). Column 1 presents the unadjusted TSPs coefficient; column 2 adjusts flexibly for the rich set of natality variables available from the birth certificate data files and controls for per-capita income, earnings, employment, and transfer payments by source; and column 3 adds state fixed effects.[10] The sample sizes and R^2s of the regressions are shown in brackets.

There is wide variability in the estimated effects of TSPs, both across specifications for a given cross-section and across cross-sections for a given specification. While the raw correlations in column 1 are all positive, only those from the 1969 and 1974 cross-sections are statistically significant at conventional levels. Including the controls in column 2 reduces the point estimates substantially, even as the precision of the estimates increases due to the greatly improved fit of the regressions. In fact, 4 of the 7 estimates are now negative or perversely

signed, and the only statistically significant estimate (from 1972 data) is perversely signed in that it indicates that TSPs reduce IMR. The most unrestricted specification in column 3 that also adjusts for state fixed effects produces two statistically significant estimates but one is positive and the other is negative.

The largest positive estimates from the cross-sectional analyses imply that a $1\text{-}\mu g/m^3$ reduction in mean TSPs results in roughly 3 fewer internal infant deaths per 100,000 live births. This is an elasticity of 0.14 and is broadly consistent with published estimates.

However, there is little evidence of a systematic cross-sectional association between particulates pollution and infant survival rates. While the 1974 cross-section produces estimates that are positive, significant and slightly less sensitive to specification, the 1972 cross-section provides estimates that are routinely negative. It is troubling that before looking at these results, there is no reason to believe that the 1974 data are more likely to produce a valid estimate than the 1972 data. The sensitivity of the results to the year analyzed and the set of variables used as controls suggests that omitted variables may play an important role in cross-sectional analysis. We conclude that the cross-sectional approach is unreliable in this setting.

Another example of the sensitivity of estimates in the selection on observables approach comes from a study of the impacts of climate change on agricultural land values by Deschenes and Greenstone [26]. This paper reports 36 cross-sectional estimates of the predicted changes in land values from the benchmark estimates of climate change induced increases of 5 degrees Fahrenheit in temperatures and 8% in precipitation. The different estimates are derived from six years of data and six separate specifications. These estimates along with their +/- 1 standard error range are reported in Figure 1. The figure makes clear that there is a large amount of variability across specifications, and the variability extends across the range of positive and

10

negative land value effects. In fact, the estimated changes in agricultural land values range from -$200 billion (2002$) to $320 billion or -18% to 29%. It is evident from this example too that that the cross-sectional associational approach to uncovering key relationships in environmental economics may be prone to producing unreliable estimates.

In our view, these two cases are not isolated incidents and associational evidence frequently provides mixed and/or unreliable evidence on the nature of causal relationships. In these cases, the temptation is to rely on one's prior beliefs which may not have a scientific origin. In his classic 1944 article, "The Probability Approach in Econometrics," Trygve Haavelmo described this problem of testing theories for which an experiment is unavailable, "we can make the agreement or disagreement between theory and the facts depend upon *two* things: the facts we choose to consider, as well our theory about them" [40, p. 14].

D. The Impact of Associational Evidence in the Face of Two Biases

The current reliance on associational studies may have especially pernicious consequences in the face of two biases. The first is publication bias, which exists when researchers are more likely to submit for publication – and journal editors are more likely to accept – articles that find statistically significant studies with the "expected" results. For example, the expected result in the case of air pollution and human health is that higher ambient concentrations cause increased mortality rates. The second is regulatory bias, which exists when regulators place more weight on studies that find a significant negative health impact of emissions than on other studies.

There is considerable evidence that the epidemiology literature and the economics literature both suffer from publication bias [27, 25, 18, 8]. Indeed, the leading medical journals

have recently attempted to address publication bias by requiring that all clinical trials be registered when they are begun in order to be considered for publication in the journals. This is expected to reduce the bias towards favorable results of treatments because researchers must announce their study before knowing the results [52].

There is further evidence of publication bias with respect to studies of the health effects of pollution. As noted in our discussion of Table I, cross-sectional estimates of the health/pollution relationship are not robust and indeed seem equally likely to be positive as negative. If the published papers were a random sample of the results from all estimated regressions, one would expect that they would have the same large range as the ones presented above. However, virtually all published papers report a negative relationship between pollution and health and this may reflect publication bias.[11]

Regulatory bias is also likely. For example, in their risk analyses, the EPA places greater weight on findings from epidemiological studies that find positive associations between carcinogenic risks and cancer, than studies that find no association [31]. Similarly, when evaluating the risk of methylmercury exposure to pregnant women from fish consumption, the National Research Council (NRC) explicitly downgraded a study (from the Republic of the Seychelles) that found no such risk. According to the NRC, "It would not be appropriate to base risk-assessment decisions on the Seychelles study because it did not find an association between methylmercury and adverse neurodevelopment effects" [64, p. 299].

Both the EPA and the NRC justify this bias for the sake of protecting public health. However, practices that over-estimate public health risk will reduce the ability to achieve efficient outcomes. Even if one assumes that people are risk averse, then regulatory agencies should still rely on unbiased estimates of the probability of different outcomes. Risk aversion is

reflected in the willingness-to-pay values that enter the policy benefit calculations. Perhaps more problematic than inefficient over-regulation, the inherent bias towards studies that find health effects of pollution can distort our environmental policy priorities [65]. For example, the true risk from pollutant A may be greater than the true risk of pollutant B, yet reliance on biased empirical studies could lead to more regulatory dollars devoted to the latter pollutant, all other things being equal.

II. The Experimental and Quasi-Experimental Approaches

A. Randomized Experiments

If traditional regression or associational based evidence is unlikely to allow for the identification of important causal relationships in environmental economics, what alternatives are available? The ideal solution to this problem of inference is to run a classical experiment where individuals are randomly exposed to a treatment. Due to random assignment, the treatment and control groups should be statistically identical on all dimensions, except exposure to the treatment; thus, any differences in outcomes can be ascribed to the treatment. Put another way, with a randomized experiment, it is valid to assume that the selection bias term is zero, so a comparison of outcomes among the treatment and control groups yields a credible estimate of the average causal effect of exposure to the treatment among the treated.

The use of randomized experiments in economics is growing rapidly. The fields of development and labor economics have seen a virtual explosion of experiments in the last several years [e.g., 55, 53]. In fact, there have been a few experiments in environmental economics in the last few years [e.g., 57, 54, 30]. Readers interested in learning more about a number of the subtleties in implementing and analyzing randomized experiments are directed to a recent paper

by Duflo, Glennerster, and Kremer [29].

B. Quasi-Experimental Approaches

One of this paper's primary arguments is that the quasi-experimental approach can be used to uncover causal relationships. In a quasi-experimental evaluation, the researcher exploits differences in outcomes between a treatment group and a control group, just as in a classical experiment. In the case of a quasi-experiment, however, treatment status is determined by nature, politics, an accident, or some other action beyond the researcher's control.

Despite the nonrandom assignment of treatment status, it may still be possible to draw valid inferences from the differences in outcomes between the treatment and control groups. The validity of the inference rests on the assumption that assignment to the treatment and control groups is not related to other determinants of the outcomes. In this case, it is not necessary to specify and correctly control for all the confounding variables, as is the case with the more traditional selection on observables approach. The remainder of this subsection outlines three common quasi-experimental approaches.

1. Difference in Differences (DD) and Fixed Effects. 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. For clarity, consider a canonical DD example where there are two groups or units. Neither group receives the treatment in the first period and only one group receives it in the second period. The idea is to calculate the change in the outcomes among the treated group between the two periods and then subtract the change in outcomes among the untreated group. More formally, this can be expressed as:

(5) $T^{DD} =$ {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]},

where Pd is an abbreviation for period. Importantly, all four of these means are observed in the data and can readily be estimated.

The DD estimator will produce a valid estimate of the treatment effect under the assumption that in the absence of the treatment the outcomes in the two groups would have changed identically in the treatment and control groups between periods 1 and 2. More formally, this assumption is that {E[Y_{0i} | D_i=1, Pd=2] - E[Y_{0i} | D_i=1, Pd=1]} = {E[Y_{0i} | D_i=0, Pd=2] - E[Y_{0i} | D_i=0, Pd=1]}. This assumption is not trivial and has been shown to be invalid in some settings, especially where behavioral responses are possible. For example, individuals may choose to receive the treatment in response to a shock in period 1.[12]

By using least squares regression techniques, it is possible to adjust the estimates for covariates. Further, this approach can accommodate multiple time periods and multiple treatment groups. Recent examples of DD and fixed effects estimators in environmental economics include Becker and Henderson [12], Greenstone [36], and Deschenes and Greenstone [26].[13]

2. Instrumental Variables (IV). An identification strategy based on an instrumental variable can solve the selection bias problem outlined above. The key is to locate an instrumental variable, Z, which is correlated with the treatment, but otherwise independent of potential outcomes. The availability of such a variable would allow us to rewrite equation (4) as:

(4') $T_{ct} = X_{ct}'\Pi + \delta Z_{ct} + \eta_{ct}$, $\eta_{ct} = \lambda_c + v_{ct}$.

Formally, two sufficient conditions for θ_{IV} to provide a consistent estimate of θ are that $\delta \neq 0$ and $E[Z_{ct}\varepsilon_{ct}]=0$. The first condition requires that the instrument predicts T_{ct} after conditioning on X.

The second condition requires that the Z is orthogonal to the unobserved determinants of the potential outcomes.[14]

When this strategy is implemented in a two-stage least squares (2SLS) framework, it is straightforward to understand how this strategy will produce a consistent estimate of θ. This is because the 2SLS approach (which is algebraically identical to the instrumental variable approach) estimates equation (4'), then uses the results to obtain a fitted value for T, and replaces T in equation (3) with this fitted value. The intuition is that the instrumental variable discards the variation in T that is the source of the selection bias.

An attractive feature of the IV approach is that it is straightforward to learn about the validity of its assumptions. The first assumption that the instrument is related to the endogenous variable can be directly tested. The second assumption cannot be directly tested, but it is possible to learn about the likelihood that it is valid. One method for doing this is to test for an association between the instrument and observable variables measured before the treatment was assigned. If the instrument is unrelated to observable covariates, it may be more likely that the unobservables are also orthogonal [2]. Even if this is not the case, this exercise can identify the likely sources of confounding and help to inform the choice of a statistical model. Another validity check is to test for an association between the instrument and potential outcomes in a period or place where there is no reason for such a relationship. Additionally, when multiple instruments are available one can implement a Wu-Durban-Hausman style overidentification test [41].

The IV approach has become ubiquitous in applied economics fields and is increasingly being used in environmental economics. Both Chay and Greenstone [21] and Bayer, Keohane and Timmins [11] use IV techniques to estimate the capitalization of total suspended particulates

into housing prices.

Finally, we note that there are a host of more subtle issues related to instrumental variables methods. Two especially important issues are the interpretation of IV estimates when there is heterogeneity in the treatment effect [48, 42] and the impact of "weak" instruments on the unbiasedness of IV estimators [14, 75]. Angrist and Krueger [3] are a good starting point for learning more about instrumental variable estimation.

3. Regression Discontinuity (RD). The regression discontinuity design is an increasingly popular method for solving the problem of selection bias. In the classic RD design, the assignment to the treatment is determined at least partly by the value of an observed covariate and whether that value lies on either side of a fixed threshold. For example, hazardous waste sites are eligible for federally sponsored remediation under the Superfund program if a continuous measure of site-specific risk exceeds a threshold. Similarly, counties are designated non-attainment under the Clean Air Act if ambient pollution concentrations exceed a threshold. The covariate may be associated with potential outcomes either directly or through correlation with unobservables; however, this association is assumed to be smooth. When this assumption is valid, a comparison of outcomes at the threshold after conditioning on this covariate or functions of this covariate will produce an estimate of the causal effect of the treatment.

To make this assumption clear, consider a sharp RD design where assignment of the treatment is a deterministic function of one of the covariates. Define R as this covariate and c as the threshold, so the unit is assigned to the treatment when $R \geq c$. The identifying assumption is that the distribution functions of Y_0 and Y_1 conditional on R are continuous at $R = c$. Under this assumption,

$$E[Y_{0i} \mid R = c] = \lim_{R \uparrow c} E[Y_{0i} \mid R = r] = \lim_{R \uparrow c} E[Y_{0i} \mid D_i = 0, R = r] = \lim_{R \uparrow c} E[Y_i \mid R = r]$$

17

and similarly

$$E[Y_{1i} \mid R = c] = \lim_{R \downarrow c} E[Y_i \mid R = r].$$

Thus, the selection bias term is zero at $R = c$ (and possibly in some range around c) because control units with values of R just below c can be used to form a valid counterfactual for the treated units with values of R just above c. In this case, the average treatment effect at $R = c$ is

$$\lim_{R \downarrow c} E[Y_i \mid R = r] - \lim_{R \uparrow c} E[Y_i \mid R = r].$$

This estimand is the difference of two regression functions at $R = c$. It is generally implemented by making parametric assumptions about the relationship between Y and R.

Much of the appeal of this approach is that, in principle, it offers a rare opportunity to know precisely the rule that determines the assignment of the treatment. In addition to the Greenstone and Gallagher [38] paper that uses a RD design to estimate the benefits of Superfund remediations, this approach has been implemented in several other settings. For example, Angrist and Lavy [4] examine the impact of class size on student achievement, DiNardo and Lee [28] explore the effects of unionization on plant outcomes, Card, Dobkin, and Maestas [17] examine the impacts of eligibility for medical services under the Medicare program, and Chay and Greenstone [20, 21] use a RD design to assess the benefits of clean air regulations. As these examples illustrate, RD designs are pervasive and can be used to answer a wide range of questions. Their pervasiveness is because administrators frequently use discrete cutoffs for program eligibility to ensure that benefits are distributed in a fair and transparent manner.

There are a number of subtleties involved with the implementation of RD designs. One immediate issue is whether one is evaluating a "sharp" discontinuity or a "fuzzy" one. In a sharp RD design, the probability of receiving the treatment goes from zero to one at the discontinuity, while in a fuzzy one the probability of receiving the treatment changes discontinuously at the

threshold but increases by less than one. Both sharp and fuzzy designs can be used to obtain causal estimates of the treatment effect, but there are some differences in implementation and interpretation.

An especially appealing feature of RD designs is that they lend themselves to graphical analyses that can display the results in a powerful and easy to understand manner. In making these graphs, researchers face an important set of choices on how to best implement the nonparametric regressions that underlie the graphs, including bandwidth choice. Other issues that bear more attention than is feasible to discuss in this paper are how best to implement RD estimators, the asymptotic properties of RD estimators, and specification tests. Imbens and Lemieux [49] provide an accessible summary of many of the practical issues associated with the implementation of RD designs.

Finally, we emphasize that valid RD designs only provide estimates of the average of the treatment for the subpopulation where $R = c$. To extend the external validity of estimates from RD designs, it is necessary to make assumptions (e.g., homogeneity of the treatment effect) that may not be justified. In many situations, however, the estimated treatment effect may be of interest in its own right; for example, policymakers may be considering expanding a program so that the treatment effect for participants with a value of R slightly less than c is the policy relevant treatment effect (rather than the population average treatment effect). In summary, the RD design is frequently an effective method for addressing issues of internal validity, but its estimates may have limited external validity.

C. Threats to the Validity of Experiments and Quasi-Experiments and Assessing Their Importance

There are a series of pitfalls that can undermine the value of experiments and quasi-experiments. In their classic work, Cook and Campbell [23] call these pitfalls "threats to validity," where validity is the truth of a proposition or conclusion.[15] They divide these into threats to internal, external, and construct validity.[16] The first applies only to quasi-experiments, while the latter two are relevant for randomized experiments and quasi-experiments.

Internal validity refers to whether it is possible to validly draw the inference that the difference in the dependent variables is due to the explanatory variable of interest. Cook and Campbell [23] and Meyer [61] provide exhaustive lists of these threats, but they can largely be summarized as instances where treatment status may be related to the post-treatment outcome for reasons other than the treatment.[17] This could be due to omitted variables, inadequate controls for pre-period trends, and/or the selection rule that determines treatment status. The key feature of a threat to internal validity is that it causes the selection bias term in equation (2) to be nonzero. This issue is not a concern in the case of randomized experiments but it is a central concern with quasi-experiments.

External validity is applicable to cases where the treatment effect is heterogeneous and refers to whether an experiment's or quasi-experiment's results can be generalized to other contexts. People, places, and time are the three major threats to external validity. For example, the individuals in the treatment group may differ from the overall population (perhaps they are more sensitive to air pollution) so that the estimated treatment effect is not informative about the effect of the treatment in the overall population. Other examples of cases where external validity is compromised are when the estimated treatment effect may differ across geographic or institutional settings or if it differs across years (e.g., if in the future a pill is invented that protects individuals from air pollution, then the installation of scrubbers would have a different

effect on health).[18]

An issue that is closely related to external validity is that a treatment's effect may depend on whether it is implemented on a small or large scale. As an example, consider estimating the impact of air pollution on human health when people have sorted themselves into locations of the country based on the sensitivity of their health to air pollution. If the government implements a program that improves air quality in some regions, but not others, this may induce individuals to change their locations. The estimated treatment effect that is based on the distribution of the population before the re-sorting is likely to differ from the longer run effect of the policy that depends on the degree and type of general equilibrium sorting.[19] This example underscores that the interpretation of a reliable quasi-experimental analysis requires careful economic modeling.

Finally, construct validity refers to whether the researcher correctly understands the nature of the treatment. Returning to the above example, suppose that scrubbers reduce emissions of TSPs and other air pollutants, but the researcher is only aware of the reduction in TSPs. If these other pollutants are important predictors of human health, then a post-adoption comparison of human health in the two areas would be unable to separate the effect of TSPs from the effect of the other air pollutants. In this case, the researcher's inadequate understanding of the treatment would cause her to conclude that TSPs affects human health when the effect on human health may be due to the reduction of the other pollutants. Notably, it is still possible to obtain unbiased estimates of the overall effect of the new abatement technology—that is, the properly understood treatment.

The prior discussion naturally raises the question of how to find an experiment or quasi-experiment that sheds light on the question of interest. The first-order issue in relation to quasi-experiments is to assess their internal validity. Unfortunately, there is not a handy recipe or

statistical formula that can be taken off the shelf. Since the treatment in a quasi-experiment is not assigned randomly, they are by their very nature messy and lack the sharpness and reliability of classical experiments. Consequently, the keys to a good quasi-experimental design are to be watchful (bordering on paranoid) about the threats to its internal validity and to leave no stone unturned in testing whether its assumptions are valid. Although these assumptions cannot be tested directly, careful research or "shoe leather" – as the statistician David Freedman calls it – can help to understand the source of the variation that determines the explanatory variable of interest and assess a quasi-experiment's validity [33].

As we discussed in the context of instrumental variables strategies, one informal method for assessing the internal validity of a quasi-experiment is to test whether the distributions of the observable covariates are balanced across the treatment and control groups. If the observable covariates are balanced, then it may be reasonable to presume that the unobservables are also balanced (and that selection bias therefore is not a concern). A related method of assessing the internal validity of a quasi-experiment is to test whether the estimated effect is sensitive to changes in specification. This is particularly relevant for the variables that are not well balanced. In the case when all the observables are balanced, it is unnecessary to adjust the treatment effect for observables. Since this is rarely the case, an examination of the sensitivity of the estimated treatment effect is an important part of any analysis.

Of course, the balancing of the observables and/or regression adjustment does not guarantee the internal validity of a quasi-experiment. This is because the unobservables may still differ across the treatment and control groups. Economic reasoning or models can often help to identify cases where this is especially likely to be the case. For example, a thorough understanding of the assignment rule might reveal that individuals with a particular characteristic

(e.g., high incomes) are more likely to receive the treatment. Such a finding would undermine the credibility of any results.

A thorough understanding of the experiment or quasi-experimental assignment rule and the nature of the treatment can also help to assess the external and construct validity. For example, one might be interested in determining the effect of a nationwide 10% reduction in TSPs. However, the available quasi-experimental evidence on the health effects of TSPs may be derived from areas with high TSPs concentrations. If individuals sort themselves across the country based on their susceptibility to TSPs, then these quasi-experimental results may not be informative about the health effects in relatively clean areas. Similarly, if TSPs do not affect human health below some concentration, then estimates of the gradient at high concentrations will not answer the broader question. Regarding construct validity, if the treatment (e.g., the installation of a scrubber) affects emissions of multiple pollutants then the available quasi-experimental evidence may be uninformative about the health effects of TSPs.

These threats to external and construct validity highlight a concern with experiments and quasi-experiments. They are unable to determine the welfare effects of policies that have not previously been implemented or to estimate treatment effects for populations that have not previously been treated. Indeed, all empirical studies face these limitations, which can only be overcome by relying on theoretical modeling of behavioral relationships in order to make predictions under these circumstances. This type of modeling is frequently necessary to answer relevant questions, but the appeal of quasi-experiments (and experiments) is that they make transparent which predictions are based on data and which require further assumptions.[20]

In summary, the appeal of the experimental and quasi-experimental approaches is that they rely on transparent variation in the explanatory variable of interest. The transparency of the

variation (or the mere labeling of the variation as quasi-experimental) does not remove concerns about threats to validity. Extensive expenditures of "shoe leather" are necessary to allay these concerns. As Meyer [61] wrote, "If one cannot experimentally control the variation one is using, one should understand its source." We would add that one must also work to consider the appropriate economic model of behavior that is consistent with the quasi-experimental or experimental findings.

D. Can the Experimental and Quasi-Experimental Approaches Answer Important Questions?

In many instances randomized experiments are unavailable to answer questions where answers are important. Further, quasi-experiments are determined by nature, politics, an accident, or some other action beyond the researcher's control. A key limitation of these approaches is that there is no guarantee that these actions will help to inform the most interesting or important questions. Put another way, an exclusive reliance on experimental and quasi-experimental methods places researchers in the uncomfortable position of relying on nature to set their research agendas. This is frustrating because it raises the possibility that many of the most important questions cannot be answered. One economist expressed this concern, "if applied to other areas of empirical work [quasi-experiments] would effectively stop estimation" [47].

We find this criticism largely unmerited and potentially detrimental to progress in understanding the world for at least two reasons. First, experimental and quasi-experimental approaches have been successful in furthering understanding about a number of important topics that ex ante might not have seemed amenable to analysis with these methods. An exhaustive listing is beyond the scope of this paper, but one would include quasi-experiments that estimate: how cholera is transmitted [74]; the effects of anti-discrimination laws on African-American's

earnings and health outcomes [44, 19, 1]; the labor supply consequences of unemployment insurance benefits [60]; the effect of minimum wage laws on employment [15]; the returns to an additional year of schooling [9, 16]; the effect of class size on scholastic achievement [4]; the effect of pre-kindergarten on test scores [35]; the impact of mandatory disclosure laws on equity markets [39]; and individuals' willingness to pay for school quality [13] and clean air [21].

The point is that we believe that a much greater emphasis should be placed on implementing experiments and quasi-experiments to answer important environmental economics questions. The successes in implementing these techniques in other areas of economics (notably labor, development and public economics) underscore that with sufficient expenditures of "shoe leather" it is possible to use them to answer important questions. In fact, researchers should encourage and work with governments to evaluate new policies by implementing randomized or quasi-randomized assignments of pilot programs.[21]

Our second disagreement with this criticism of experimental and quasi-experimental approaches is that it can distract from the main empirical concern of estimating causal relationships. Our view is that empirical research's value should be based on its credibility in recovering the causal relationship of interest. In contrast to associational studies, the methods we have described above place primary emphasis on recovering causal relationships. If in some instances current research is unable to shed much light on the causal relationships of interest, then researchers should be clear and transparent about the current state of knowledge (rather than treating associational evidence as if it were causal). This can motivate future researchers to uncover causal tests of the relevant hypothesis.[22]

III. The Costs of the Endangered Species Act: Evidence from North Carolina

We now demonstrate the quasi-experimental approach by assessing the validity of a quasi-experiment that aims to evaluate the welfare costs of the Endangered Species Act in North Carolina.

A. Introduction

The Endangered Species Act of 1973 (hereafter, the Act) is perhaps the most far-reaching environmental statute in the United States, as it makes it unlawful for any private landowner to "take" a fish or wildlife species that is designated as endangered by the Department of Interior's Fish and Wildlife Services (for terrestrial and freshwater species) or by the Department of Commerce's National Oceanic and Atmospheric Administration's Fisheries Service (for marine species). In practice, the government defines "take" in a way that places much of the burden of species protection on private landowners by restricting opportunities to develop their land.

Not surprisingly, this onus placed on private landowners makes the Act highly controversial, with one side claiming that the burden is too restrictive and the other side claiming that the restrictions offer the only effective way to protect species from extinction. The controversy is apparent with respect to many of the Fish and Wildlife Service's individual decisions on which species to list for protection (see, for example, [68]), and manifests itself in more general calls for reforming the Act entirely (see, for example, [73] and H.R. 3824 in 2005 sponsored by Representative Richard Pombo of California).

Despite all the controversy, relatively little is known about the extent to which species protection restricts development and housing supply, resulting in welfare costs. There is certainly the potential for high costs stemming from the ESA given that it places the onus for protection primarily on private landowners through restrictions on development. The potential

for high costs also exists because the decision to list a species for protection (at the federal level) is to be made "without reference to possible economic or other impacts" (50 CFR 424.11(b)). In 1978, the Supreme Court ruled that this language means that the "value of endangered species is incalculable," and a qualified species must be protected "whatever the cost," and that this language "admits no exception" (Tennessee Valley Authority v. Hill, 437 U.S. 153).[23] Even if current statutes rule out consideration of the costs, future ones might include them.

Nonetheless, there are other considerations that might mitigate the welfare costs of species protection. First, some studies have found that landowners preemptively destroy habitat in order to avoid ESA repercussions [58, 72], which in addition to diminishing the benefits of protection might also serve to mitigate the costs to foregone development.[24] Second, development restrictions stemming from species protection might serve as a type of zoning restriction. Localities frequently impose land use regulations that can provide public good benefits to landowners. Such landowners might jointly prefer to restrict local development, but are individually provided with an incentive to develop. In effect, the ESA might substitute for other land use regulations, thus lessening (or perhaps even eliminating) their regulatory costs relative to the counterfactual land use restrictions.

It is therefore an open empirical question whether – and to what extent – species protection restricts development, resulting in costs to landowners, developers, and homebuyers. There is, of course, a strong incentive for landowners to overstate the cost of the Act as a means to receive statutory relief. However, there are relatively few empirical studies of its costs. Some studies [e.g., 58] have examined whether the ESA leads to preemptive, cost-avoiding, behavior by private landowners. Other studies [e.g., 76, 59, 78] examine the impacts of the government's designation of some areas as "critical habitats" for protected species.[25] To date, the critical

habitat designations have only been applied to a subset of protected species; and among the species that receive these designations, they are only applied to a subset of their habitats. Further, they do not provide any further restrictions beyond those embodied in the ESA. Consequently, these studies are only able to provide a partial picture of the ESA's impacts. To summarize, there are no studies that attempt to provide comprehensive estimates of the welfare costs associated with the ESA's restrictions on development.

The main empirical complication in addressing this research question is that areas containing one or more protected species differ from areas that do not contain such a species, and these differences are associated with differences in housing market outcomes. If these differences are unobservable, then a simple regression analysis will lead to biased results. This is the classic selection bias discussed in the first half of the paper.

In the remainder of this paper we describe a quasi-experiment that has the potential to solve the selection problem. The goal is to highlight the steps involved in identifying a quasi-experiment and to demonstrate some ways to assess its validity. Due to space constraints and the absence of comprehensive outcomes data, we leave the full-blown estimation of the welfare impacts for future research.

The proposed quasi-experiment is based on a non-profit conservation organization's (NatureServe) assessment of the relative rarity or imperilment of every species present in North Carolina. NatureServe's assessments are scientifically based and are <u>not</u> based on whether a species is protected under the ESA. Specifically, NatureServe assigns each species to one of the following Global Conservation Status Rank (GCSR) categories: Possibly Extinct, Critically Imperiled, Imperiled, Vulnerable, Apparently Secure, and Demonstrably Secure. The key feature of this quasi-experiment is that within these ranks there are species that are protected and

unprotected by the ESA. The basis of the analysis is a comparison of housing market outcomes in census tracts that include the habitats of protected species to outcomes in census tracts that include the habitats of unprotected species of the same Global Conservation Status Rank. If within a GCSR rank, the tracts with the unprotected species are a valid counterfactual for the tracts with the protected species, then this quasi-experiment will produce unbiased estimates of the impacts of ESA protections.

B. Background on the Endangered Species Act and the Listing Process

1. Legislative Background. Quasi-experiments that are based on a law must understand the nature of the law. Here, we describe the workings of the ESA. The Endangered Species Preservation Act of 1966 was the first U.S. statute that aimed systematically to protect endangered species. This legislation was inspired by the plight of the whooping crane. This law was rather limited in scope; its main focus was to authorize the Secretary of Interior to identify native fish and wildlife that were threatened with extinction and to allow the Fish and Wildlife Service to spend up to $15 million per year to purchase habitat for listed species. In addition, the statute directed federal land agencies to protect these endangered species and their habitats "insofar as is practicable and consistent with [the agencies'] primary purpose" (Public Law 89-669, 80 Stat. 926).

In 1969, pressured by the growing movement to save the whales, Congress supplemented the statute with the Endangered Species Conservation Act, which expanded the protected species list to include some invertebrates, authorized the listing of foreign species threatened with extinction, and banned the importation of these species except for specified scientific purposes.

The Endangered Species Act of 1973, which passed in the Senate by a voice vote and in

the House of Representatives by a 355-to-4 vote, substantially changed the structure of the law on protection of endangered species. Among other things, the new law distinguished between threatened and endangered species, allowed listing of a species that is in danger in just a significant portion of its range, extended protection eligibility to all wildlife (including invertebrates) and plants, and defined species to include any subspecies or distinct population segment of a species.

Sections 7 and 9 are the two major components at the heart of the 1973 statute. Section 7 requires federal agencies to "consult" with the Secretary of Interior (or the Secretary of Commerce for marine species) in order to "insure that any action authorized, funded, or carried out by such agency … is not likely to jeopardize the continued existence of any endangered species or threatened species or result in the destruction or adverse modification of habitat of such species …" (ESA 7(2)). If the agency action is found to place an endangered or threatened species in jeopardy or to result in adverse modification of habitat for the species, the Secretary must suggest "reasonable and prudent alternatives" (ESA 7(b)(3)(A)). While the primary impact of this section of the Act is on the actions undertaken on federal lands, it could restrict activities on private lands if such activity requires a federal permit. Importantly, this section of the Act applies to both endangered and threatened species and it applies to both animals (fish and wildlife) and plants.

Section 9 most directly impacts private landowners. It makes it illegal to "take" a listed fish or wildlife species, where "take means to harass, harm, pursue, hunt, shoot, wound, kill, trap, capture, or collect or attempt to engage in any such conduct" (ESA 3(19)). While this prohibition might imply restrictions only on direct physical harm to a species, a landmark Supreme Court ruling in 1995 (Babbitt v. Sweet Home, 515 U.S. 687) deferred to the

Department of Interior's more expansive definition of "take" 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" (50 CFR 17.3).[26]

Section 9 is the source of most of the controversy surrounding the Act. Property rights advocates complain that it is highly burdensome and that it violates their Fifth Amendment rights on the taking of private property without just compensation. In contrast, conservationists argue that it is the key component for species protection, without which many valued species will be lost to extinction.

The "taking" prohibition listed under section 9 applies only to fish and wildlife species, not to plants. Section 9 does offer some protections for endangered plants, such as restrictions on importing and interstate sales, but it does not provide the prohibition on taking plants. This suggests that, all things equal, we should expect smaller (if any) economic impacts for areas that contain endangered plants compared to areas that contain endangered animals.

The "taking" prohibition listed under section 9 applies to endangered species, and not necessarily to threatened species. An endangered species is "any species which is in danger of extinction throughout all or a significant portion of its range …." (ESA 3(6)), and a threatened species is "any species which is likely to become an endangered species within the foreseeable future throughout all or a significant portion of its range" (ESA 3(20)).

For threatened species, section 4(d) of the Act gives the Secretary of Interior discretion on which section 9 prohibitions to apply. So, for example, the Secretary may exempt a limited range of activities from take prohibitions for certain threatened species. Section 4(d) of the Act also allows the Secretary to tailor "protective regulations" that it "deems necessary and advisable

to provide for the conservation of [the threatened] species" (ESA 4(d)). Thus, with respect to threatened species, the Fish and Wildlife Service (acting with authority from the Secretary of Interior), has greater flexibility about which prohibitions described in section 9 of the Act to apply to the species, and which other protective regulations to apply to the species. This flexibility could conceivably result in greater (and thus costlier) protections for threatened species, although this is neither the intent nor the implementation of the statute.

The 1973 Act has been re-authorized eight times and significantly amended three times, most recently in 1988. Among the more significant changes, the 1978 Amendments to the Act established a Cabinet-level Endangered Species Committee that can exempt federal actions from the prohibitions of section 7. The 1982 Amendments instituted a permit system that allows the "incidental taking" of a listed species provided that the permit holder implements a habitat conservation plan for the species. Nonetheless, the overall framework of the Act has remained essentially unchanged since 1973. The Act was due for reauthorization in 1993, but such legislation has not yet been enacted.[27]

2. The Listing Process. Section 4 of the Act requires that the decision to list a species as endangered or threatened be based on the following factors: "the present or threatened destruction, modification, or curtailment of its habitat or range; the overutilization for commercial, recreational, scientific, or educational purposes; disease or predation; the inadequacy of existing regulatory mechanisms; or other natural or manmade factors affecting its continued existence" (ESA 4(a)(1)).

While the listing determination must be made "solely on the basis of the best scientific and commercial data available" (ESA 4(b)(1)(A)), the Act does allow some prioritizing in the determination process. That is, the Fish and Wildlife Service can decide that a species warrants

inclusion on the list, but that it is "precluded by pending proposals" for other species (ESA 4(b)(3)(B)). Species that are deemed as "warranted but precluded" are designated as candidate species. They do not receive the statutory protections of sections 7 and 9, even though the Fish and Wildlife Agency believes that they are possibly endangered or threatened.

C. Research Design

This legislative background and the availability of new data files on species conservation statuses and habitats provide an unique opportunity for a quasi-experimental evaluation of the economic costs of the ESA. Specifically, we learned about NatureServe, which is a nonprofit conservation organization dedicated to collecting and managing information about thousands of rare and endangered species of plants and animals. As part of their mission, NatureServe collects information on the relative imperilment of these species and maps each species' habitats throughout the world.

The imperilment data is based on scientific criteria and is completely unrelated to ESA provisions. NatureServe summarizes the status of each species with their measure of imperilment called Global Conservation Status Rank (GCSR). The GCSR takes on the following values (with the meaning in parentheses): G1 (Critically Imperiled), G2 (Imperiled), G3 (Vulnerable), G4 (Apparently Secure), G5 (Demonstrably Secure), G6 (Unranked), and G7 (Possibly Extinct).[28] Species that are imperiled or vulnerable everywhere they occur have GCSRs or G-ranks of G1, G2, or G3.

The second feature of these data is that they contain detailed GIS maps of the habitats of each of the species. To date, NatureServe has only provided us with habitat data from North Carolina, although they have promised to work to provide data from additional states in the

future. We overlaid the North Carolina habitat maps on 2000 census tract boundaries for North Carolina to determine the tracts that overlap with each of these species' habitats. Notably, the NatureServe data on habitats are widely considered to be the most reliable information available. For example, the Fish and Wildlife Service often relies on NatureServe habitat maps to determine where ESA regulations should apply.

The key feature of this quasi-experiment is that within GCSRs, there are species that are protected and unprotected by the ESA. The key assumption is that the observed and unobserved characteristics of a census tract that make it hospitable for a species may also determine housing market outcomes. Thus, by holding constant GCSRs, it may be possible to avoid confounding the impacts of ESA protections with other factors.

D. Theoretical Implications

The key feature of the ESA is that it removes land from possible development. Quigley and Swoboda [67] present a theoretical model of the welfare costs of species protection, in which they assume that residents cannot move outside of the market but can move costlessly within it. With respect to the newly regulated land, it is clear that the value of this land should decline, especially for undeveloped parcels. Thus, the decline in the value of this land is a welfare loss.

Now consider land parcels that are in the same market as the newly regulated land but are not subject to ESA restrictions themselves. The quantity demanded for this land increases, leading to an increase in price. This increase benefits owners of these land parcels, but harms consumers/renters. In the absence of consumer disutility associated with congestion, this increase in unregulated land values simply amounts to a transfer from consumers/renters of land to owners of land.

Finally, the model suggests that, at least in a closed market, there will be overall price impacts. In particular, the reduction in the amount of land available for conversion to housing will lead to an increase in land and house prices. These price increases would reflect a loss of welfare to producers and consumers of houses, with the distribution of the losses depending on elasticities.

Due to space constraints and the absence of comprehensive price and quantity data, we leave the full-blown estimation of the welfare impacts for future research. However, we briefly note a few empirical predictions. If the ESA is strictly enforced and does not substitute for other land use regulations, we expect that the ESA's impacts would be most directly felt on land parcels that have not yet been converted from some other use (predominantly agriculture) to land for housing. Among these parcels that overlap with newly protected species, we would expect a reduction in their value as the option value of converting to housing will be substantially reduced. Further, we expect that fewer of them will be converted to housing, reducing the supply of housing. To investigate these possibilities, we are in negotiations to obtain data on the value and conversion of land parcels using US Department of Agricultural data files on land use and values. We are also exploring methods to define local housing markets so that we can test for increases in the market-wide price level of houses.

Finally, we note that data from the decennial population censuses are unlikely to be very useful for assessing the impacts of the ESA restrictions. This is because they are designed to be informative about people, not about land. Consequently, they do not provide information on the undeveloped land parcels that are most directly impacted by the legislation. In fact, this is why the Department of Agriculture data are well suited to infer the welfare impacts of ESA restrictions.

E. Data Sources

This subsection details the data that we have collected to date. They allow for an assessment of the validity of our quasi-experiment based on species' GCSRs.

NatureServe and ESA Status Data. The data on species' habitats comes from NatureServe's Natural Heritage Program. NatureServe works in partnership with a network of 74 independent Natural Heritage member programs that gather scientific information on rare species and ecosystems in the United States, Latin America, and Canada. The NatureServe data file is so valuable because it has been developed centrally by NatureServe, which ensures that a common methodology is used in determining species' habitats. It was first initiated in 1974 and has been updated regularly since.

We purchased data on the habitats of all species tracked by NatureServe in North Carolina. The specific data reveal which 2000 census tracts overlap with the current habitat for each species. In the data file that NatureServe provided, there are 1,227 species in North Carolina.[29] Further, the data file contains the Global Conservation Status Ranks, which as discussed above, are an essential component of the quasi-experiment.[30]

By supplementing the NatureServe data with information from various issues of the Federal Register, we obtained the ESA regulatory status of each of the 1,227 species in North Carolina. Specifically, we determined which of the species fell into the endangered, threatened, candidate, and unregulated categories. Recall, the strictest restrictions apply to areas that include the habitats of endangered species and no restrictions apply to areas that encompass the habitats of candidate and unregulated species.

Census Data. The housing, demographic and economic data come from Geolytics's

Neighborhood Change Database, which includes information from the 1970, 1980, 1990, and 2000 Censuses. Here, we focus on the 1990 data to assess whether observable determinants of housing market outcomes are balanced in tracts with and without protected species prior to the ESA's designations in the 1990s. The decennial population census data is critical for assessing the validity of the GCSR-based quasi-experiment (even though it does not provide good outcomes data for a welfare analysis).

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. Census tracts are the smallest geographic unit that can be matched across the 1970-2000 Censuses. The Census Bureau placed the entire country in tracts in 2000. Geolytics fit 1970, 1980, and 1990 census tract data to the year 2000 census tract boundaries to form a panel. The primary limitation of this approach is that in 1970 and 1980, the US Census Bureau only tracted areas that were considered "urban" or belonged to a metropolitan area. The result is that the remaining areas of the country cannot be matched to a 2000 census tract, so the 1970 and 1980 values of the Census variables are missing for 2000 tracts that include these areas.

To assess the validity of our quasi-experiment, we use the rich set of covariates available in the Geolytics data file. The baseline covariates are measured at the census tract level and include information on housing, economic, and demographic characteristics from the US Census. The Data Appendix provides a complete description of these data.

Finally, we note that the Geolytics data file is unlikely to be very useful for assessing the impacts of the ESA restrictions because it is based on the population census. Consequently, it does not provide information on the undeveloped land parcels directly impacted by the legislation.

F. Summary Statistics.

Table II provides some summary statistics on the species that are present in North Carolina. Panel A reports that there are 1,227 different species tracked by NatureServe in the state. Of these, 408 are animals, 803 are plants, and 16 are fungi. Each species is assigned a Global Conservation Status Rank designated its level of imperilment. Of the 1,227 species present, there are 95 that are critically imperiled, 153 that are imperiled, 250 that are vulnerable, 328 that are apparently secure, 365 that are secure, 34 that are unranked, and 2 that are possibly extinct.

Panel B in Table II contains information on each species by their ESA regulatory designation (endangered, threatened, candidate, or unregulated). There are 62 species located in North Carolina that are designated as either endangered, threatened, or candidate species. Most of these are endangered, and most of the endangered species are animals. Of the unregulated species, more than half are fungi and over thirty percent are animals. Panel B also shows the decade of listing for the endangered, threatened, and candidate species. Of the endangered species, 4 were listed in the 1960s, 9 in the 1970s, 16 in the 1980s, 12 in the 1990s, and 1 in the current decade.

The focus of this quasi-experiment is an evaluation of the assignment of the endangered classifications in the 1990s. Panel C in Table II contains information on the number of census tracts that contain species that were designated as endangered, threatened, or candidate in the 1990s.[31] For example, there were 85 census tracts that had an animal designated as endangered in the 1990s, and 219 census tracts that had an animal designated as threatened in the 1990s. Similarly, there were 97 census tracts that had a plant designated as endangered in the 1990s, and

61 census tracts that had a plant designated as threatened in the 1990s. For each of the designation types in the 1990s, Panel C also shows the number of census tracts that contain species of different GCSRs. For example, there are 85 census tracts that had an endangered species designated in the 1990s with a GCSR of G1.

Although North Carolina is just a single state, it is evident that there are hundreds of tracts with species covered by the ESA. Further, there are even more tracts that contain the habitats of similarly imperiled species that are not protected by the ESA. Thus, there is reason for some optimism that there will be enough power to detect changes in the outcomes.

G. Assessing the Validity of a Quasi-Experiment Based on GCSRs

The difficulty in estimating the impact of species protection on property market outcomes is that census tracts that did not have a species listed during the decade might not form a suitable counterfactual by which to compare the "treated" census tracts that did have a species listed during this decade. Table III provides an opportunity to assess this possibility for the 1990s for animal species. Specifically, column (1) reports on the means of a series of determinants of market outcomes measured in 1990 among the 85 census tracts with at least one animal species listed as endangered by the ESA during the 1990s. Column (2) presents the means of these same variables among the 1,170 tracts without a single animal species listed as endangered or threatened during this period. Column (3) reports the difference between the columns, as well as the heteroskedastic-consistent standard error associated with the difference. The differences that are statistically significant at the five percent level are in bold typeface.

The entries reveal that the tracts with endangered species differ substantially from tracts without them. For example, 1980 and 1990 mean housing prices are $6,300 and $16,800 dollars

higher in the tracts without listed species, respectively. Further, the population density is much higher in these tracts as are the proportion of households headed by a female and the proportion of adults with a college degree or higher. In fact, 19 of the 29 covariates are statistically different between the two sets of tracts. In summary, the areas without an endangered species appear to be much more urbanized. This finding is not terribly surprising since urban environments are not hospitable places for animals.

Column (4) provides an opportunity to assess whether the quasi-experiment helps to reduce this confounding. Specifically, it reports the mean differences (and the heteroskedastic-consistent standard error) between these two sets of tracts after conditioning on separate indicators for each of the GCSRs. These indicators take on a value of 1 if the tract contains a species of the relevant GCSR, regardless of whether the species is protected by the ESA.

The results are striking. After conditioning on these indicators, only 3 of the 29 variables statistically differ across the two sets of tracts. For example, the statistically meaningful differences in 1980 and 1990 mean housing values, as well as the difference in population density, are no longer evident. Notably, the raw differences have declined substantially so this finding is not simply due to the higher standard errors.

Table IV repeats this analysis for *plant* species. The entries in columns (1) and (2) are statistically different in 22 of the 29 cases. In this case, the conditioning on GCSRs again helps to mitigate the confounding but it is not as effective as in Table III. Thus, we have less confidence in the validity of the quasi-experiment for plant species. For this reason and because the development restrictions associated with ESA protection of plants are less severe, it seems appropriate to analyze the impacts of animal and plant species protections in North Carolina on outcomes separately.[32]

A threat to internal validity that is not directly examined in Tables III and IV is that for reasons unrelated to the ESA there may be differences in local economic shocks between the "treatment" and "control" tracts, even within a G-score. This could occur if, for example, the treatment group is disproportionately comprised of tracts in the Western region of North Carolina where furniture plants and textile mills dominate, while the control group contains more tracts in Eastern North Carolina where there are many pork producers. We explored this possibility by assessing the degree of balance among the covariates in Tables III and IV after conditioning on G-Rankings and county fixed effects. The intuition is that the adjustment for the county fixed effects will absorb differences in local economic shocks between the treatment and control tracts. The results reveal that this effort to control for local shocks improves the balance between the treatment and control tracts for plants (Table IV) and somewhat worsens it for animals (Table III). In both cases, however, this empirical strategy greatly reduces the confounding associated with the unadjusted comparisons (i.e., a comparison of columns 1 and 2 in Tables III and IV).[33]

Overall, this section's results suggest that the least squares estimation of equation (6) without the GCSR indicators may not produce reliable estimates of the impact of the ESA's protection of species on housing prices. However, the tables provide support for the GCSR quasi-experiment for animals but less persuasive evidence for plants. A full-scale evaluation involving data from multiple states is left for future research.

In the context of this paper's broader message, this section has demonstrated that potentially valid quasi-experiments to solve the vexing problem of selection bias are more readily available than may be widely believed. Further, it has highlighted some ways to judge a quasi-experiment's validity without peeking at data on potential outcomes.

IV. Conclusions

Environmental economics can help to understand the welfare implications of pollution and design optimal policy in response to pollution. However, the practical importance of this contribution rests squarely on the ability of researchers to estimate causal relationships of the benefits and costs of emission reductions. Unfortunately, the traditional associational evidence of the benefits and costs of emission reductions can be highly misleading and can therefore lead to poor policies.

In order to advance the field of environmental economics, it is important that researchers and policymakers place greater emphasis on credible empirical approaches. The ideal way to achieve this is through a classical experiment in which individuals are randomly selected into treatment or control groups. Recent research suggests that it may be possible to implement randomized experiments in more settings than is commonly assumed, but in many instances these experiments are not feasible.

This paper has demonstrated that the quasi-experimental approach can be an appealing alternative. Specifically, it can successfully eliminate selection bias. The greater application of quasi-experimental techniques has the potential to improve our understanding of core environmental economics questions. Ultimately, this may lead to more efficient policies that increase social welfare.

Data Appendix

A. NatureServe Data

The following is a subset of the fields contained within the NatureServe data.

Census Tract/Census Block – The U.S. Census Tracts that overlap with each species' habitat.

Element Code – Unique record identifier for the species that is assigned by the NatureServe central database staff. It consists of a ten-character code that can be used to create relationships between all data provided.

Element Global ID - Unique identifier for the species in the Biotics database system; used as the primary key.

Global Common Name - The global (i.e., range-wide) common name of an element adopted for use in the NatureServe Central Databases (e.g., the common name for *Haliaeetus leucocephalus* is bald eagle). Use of this field is subject to several caveats: common names are not available for all plants; names for other groups may be incomplete; many elements have several common names (often in different languages); spellings of common names follow no standard conventions and are not systematically edited.

Global Conservation Status Rank - The conservation status of a species from a global (i.e., range-wide) perspective, characterizing the relative rarity or imperilment of the species. The basic global ranks are: GX - Presumed Extinct, GH - Possibly Extinct, G1 - Critically Imperiled, G2 – Imperiled, G3 – Vulnerable, G4 - Apparently Secure, and G5 – Secure. For more detailed definitions and additional information, please see: http://www.natureserve.org/explorer/granks.htm.

Global Rank Date - The date on which the Global Conservation Status Rank (GRANK) of an element was last reviewed and updated by NatureServe scientists. If an Element Rank is reaffirmed but not changed, then the date does not change.

Global Rank Review Date - Date on which the Global Conservation Status Rank (GRANK) was last reviewed (i.e., assigned, reaffirmed, or changed) by NatureServe scientists. Note that the Rank Review Date is updated each time that a global rank is reviewed, regardless of whether the rank is changed.

Global Scientific Name - The standard global (i.e., range-wide) scientific name (genus and species) adopted for use in the Natural Heritage Central Databases based on standard taxonomic references.

Subnation – Abbreviation for the subnational jurisdiction (state or province) where the Element Occurrence is located.

U.S. Endangered Species Act Status - Official federal status assigned under the U.S. Endangered Species Act of 1973. Basic USESA status values include: LE – Listed endangered, LT - Listed threatened, PE - Proposed endangered, PT – Proposed threatened, C – Candidate, PDL - Proposed for delisting, LE(S/A) – Listed endangered because of similarity of appearance, LT(S/A) - Listed threatened because of similarity of appearance, XE - Essential experimental population, XN - Nonessential experimental population. For additional information about how NatureServe manages US ESA status information, please see: http://www.natureserve.org/explorer/statusus.htm

U.S. Endangered Species Act Status Date - The date of publication in the Federal Register of notification of an official status for a taxon or population. Dates appear only for taxa and populations that

are specifically named under the U.S. Endangered Species Act.

B. Geolytics Census Data

The following are the covariates used to assess the validity of the quasi-experiment based on GCSRs and data on species habitats. All of the variables are measured in 1990 and are measured at the census tract level.

1990 Ln House Price
ln median value of owner occupied housing units in 1990

1990 Housing Characteristics
total housing units (rental and owner occupied); % of total housing units (rental and owner occupied) that are occupied; total housing units owner occupied; % of owner occupied housing units with 0 bedrooms; % of owner occupied housing units with 1 bedroom; % of owner occupied housing units with 2 bedrooms; % of owner occupied housing units with 3 bedrooms; % of owner occupied housing units with 4 bedrooms; % of owner occupied housing units with 5 or more bedrooms; % of owner occupied housing units that are detached; % of owner occupied housing units that are attached; % of owner occupied housing units that are mobile homes; % of owner occupied housing units built within last year; % of owner occupied housing units built 2 to 5 years ago; % of owner occupied housing units built 6 to 10 years ago; % of owner occupied housing units built 10 to 20 years ago; % of owner occupied housing units built 20 to 30 years ago; % of owner occupied housing units built 30 to 40 years ago; % of owner occupied housing units built more than 40 years ago; % of all housing units without a full kitchen; % of all housing units that have no heating or rely on a fire, stove, or portable heater; % of all housing units without air conditioning; % of all housing units without a full bathroom

1990 Economic Conditions
mean household income; % of households with income below poverty line; unemployment rate; % of households that receive some form of public assistance

1990 Demographics
population density; % of population Black; % of population Hispanic; % of population under age 18; % of population 65 or older; % of population foreign born; % of households headed by females; % of households residing in same house as 5 years ago; % of individuals aged 16-19 that are high school dropouts; % of population over 25 that failed to complete high school; % of population over 25 that have a BA or better (i.e., at least 16 years of education)

References
[1]   D. Almond, K. Chay and M. Greenstone, Civil Rights, the War on Poverty, and Black-White Convergence in Infant Mortality in the Rural South and Mississippi, MIT Department of Economics Working Paper No. 07-04, 2007.
[2]   J. G. Altonji, T. E. Elder and C. R. Taber, Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools, Journal of Political Economy 113 (1) (2005) 151-184.
[3]   J. D. Angrist, and A. B. Krueger, Empirical Strategies in Labor Economics, in O. Ashenfelter and D. Card (Eds.) Handbook of Labor Economics, Volume 3A, North-Holland, Oxford, 1999 pp. 1277-1366.
[4]   J. D. Angrist, and V. Lavy, Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement, Quarterly Journal of Economics, 114 (2) (1999) 533-75.
[5]   J. D. Angrist, and V. Lavy, Does Teacher Training Affect Pupil Learning? Evidence from Matched Comparisons in Jerusalem Public Schools, Journal of Labor Economics 19 (2) (2001) 343-369.

[6] O. Ashenfelter, Estimating the Effect of Training Programs on Earnings, Review of Economics and Statistics 60 (1) (1978) 47-57.

[7] O. Ashenfelter, and D. Card, Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs, Review of Economics and Statistics 67 (4) (1985) 648-60.

[8] O. Ashenfelter, C. Harmon and H. Oosterbeek, A Review of Estimates of the Schooling/Earnings Relationship, with Tests for Publication Bias, Labour Economics 6 (4) (1999) 453-470.

[9] O. Ashenfelter, and A. B. Krueger, Estimates of the Economic Return to Schooling from a New Sample of Twins, American Economic Review 84 (5) (1994) 1157-1173.

[10] W. J. Baumol, and W. E. Oates, The Theory of Environmental Policy: 2nd Edition, Cambridge University Press, Cambridge, 1988.

[11] P. Bayer, N. Keohane and C. Timmins, Migration and Hedonic Valuation: The Case of Air Quality, NBER Working Paper No. 12106, 2006.

[12] R. Becker, and V. Henderson, Effects of Air Quality Regulations on Polluting Industries, Journal of Political Economy 108 (2000) 379-421.

[13] S. E. Black, Do Better Schools Matter? Parental Valuation Of Elementary Education, Quarterly Journal of Economics 114 (2) (1999) 577-599.

[14] J. Bound, D. A. Jaeger, and R. M. Baker, Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak, Journal of the American Statistical Association 90 (430) (1995) 443-450.

[15] D. Card, Using Regional Variation in Wages to Measure the Effects of the Federal Minimum Wage, Industrial and Labor Relations Review 46 (1) (1992) 22-37.

[16] D. Card, Using Geographic Variation in College Proximity to Estimate the Return to Schooling, In L.N. Christofides, E.K. Grant, and R. Swidinsky (Eds.) Aspects of Labor Market Behaviour: Essays in Honour of John Vanderkamp, University of Toronto Press, Toronto, 1995, pp. 201-222.

[17] D. Card, C. Dobkin, and N. Maestas, The Impact of Nearly Universal Insurance Coverage on Health Care Utilization and Health: Evidence from Medicare, NBER Working Paper No. 10365, 2004.

[18] D. Card, and A. Krueger, Time-Series Minimum-Wage Studies: A Meta-Analysis, American Economic Review 85 (2) (1995) 238-243.

[19] K. Chay, The Impact of Federal Civil Rights Policy on Black Economic Progress: Evidence from the Equal Employment Opportunity Act of 1972, Industrial and Labor Relations Review 51(4) (1998) 608-632.

[20] K. Chay, and M. Greenstone, The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession, Quarterly Journal of Economics 118 (3) (2003) 1121-1167.

[21] K. Chay, and M. Greenstone, Does Air Quality Matter? Evidence from the Housing Market, Journal of Political Economy 113 (2) (2005) 376-424.

[22] R. Coase, The Problem of Social Cost, Journal of Law and Economics 3 (1960) 1-44.

[23] T. D. Cook, and D. T. Campbell, Quasi-Experimentation: Design and Analysis Issues for Field Settings, Houghton Mifflin Co., Boston, 1979.

[24] M. L. Cropper, and W. E. Oates, Environmental Economics: A Survey, Journal of Economic Literature 30 (2) (1992) 675-740.

[25] J. B. DeLong, and K. Lang, Are All Economic Hypotheses False?" Journal of Political Economy 100 (6) (1992) 1257-1272.

[26] O. Deschenes, and M. Greenstone, The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather, American Economic Review 97 (1) (2007) 354-385.

[27] K. Dickersin, The Existence of Publication Bias and Risk Factors for Its Occurrence, Journal of the American Medical Association 263 (10) (1990) 1385-1389.

[28] J. DiNardo, and D. S. Lee, Economic Impacts of New Unionization on Private Sector Employers: 1984-2001, Quarterly Journal of Economics 119 (2004) 1383-1441.

[29] E. Duflo, R. Glennerster and M. Kremer, Using Randomization in Development Economics: A Toolkit, in Handbook of Development Economics, Vol. 4 forthcoming (2007).

[30] E. Duflo, M. Greenstone and R. Hanna, The Health and Economic Impacts of Indoor Air Pollution: Evidence from a Randomized Study in Orissa, India, Working Paper, 2007.

[31] Environmental Protection Agency, Guidelines for Carcinogen Risk Assessment, Risk Assessment Forum, US EPA, Washington, DC, 2005.

[32] P. J. Ferraro, C. McIntosh, and M. Ospina, The Effectiveness of the US Endangered Species Act: An Econometric Analysis Using Matching Methods, Journal of Environmental Economics and Management Forthcoming (2007).

[33] D. A. Freedman, Statistical Models and Shoe Leather, Sociological Methodology 21 (1991) 291-313.

[34] T. Gayer, and J. K. Horowitz, Market-based Approaches to Environmental Regulation, Foundations and Trends in Microeconomics 1 (4) (2006) 1-129.

[35] W. T. Gormley, and T. Gayer, Promoting School Readiness in Oklahoma: An Evaluation of Tulsa's Pre-K Program, Journal of Human Resources 40 (3) (2005) 533-558.

[36] M. Greenstone, The Impacts of Environmental Regulations on Industrial Activity: Evidence from the 1970 and 1977 Clean Air Act Amendments and the Census of Manufactures, Journal of Political Economy 110 (2002) 1175-1219.

[37] M. Greenstone, Did the Clean Air Act Cause the Remarkable Decline in Sulfur Dioxide Concentrations? Journal of Environmental Economics and Management 47 (2004) 585-611.

[38] M. Greenstone, and J. Gallagher, Does Hazardous Waste Matter? Evidence from the Housing Market and the Superfund Program, MIT Economics Working Paper #05-27, 2006.

[39] M. Greenstone, P. Oyer and A. Vissing-Jorgensen, Mandated Disclosure, Stock Returns, and the 1964 Securities Act, Quarterly Journal of Economics 121 (2) (2006) 399-460.

[40] T. Haavelmo, The Probability Approach in Econometrics, Econometrica 12 (1944) 1-115.

[41] J. Hausman, Specification Tests in Econometrics, Econometrica 46 (1978) 1251-1271.

[42] J. J. Heckman, Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture Journal of Political Economy 109 (4) (2001) 673-748.

[43] J. J. Heckman, L. Lochner and C. Taber, Human Capital Formation and General Equilibrium Treatment Effects: A Study of Tax and Tuition Policy, American Economic Review 88 (2) (1998) 381-386.

[44] J. J. Heckman, and B. S. Payner, Determining the Impact of Federal Antidiscrimination Policy on the Economic Status of Blacks: A Study of South Carolina, American Economics Review 79 (1) (1989) 138-177.

[45] J. J. Heckman, and E. Vytlacil, Instrumental Variables Methods for the Correlated Random Coefficient Model: Estimating the Average Rate of Return Schooling When the Return is Correlated with Schooling, Journal of Human Resources 33 (4) (1999) 974-987.

[46] P. W. Holland, Statistics and Causal Inference, Journal of the American Statistical Association 81 (1986) 945-960.

[47] M. D. Hurd, Research on the Elderly: Economic Status, Retirement, and Consumption and Savings, Journal of Economic Literature 28 (2) (1990) 565-637.

[48] G. Imbens, and J. D. Angrist, Identification and Estimation of Local Average Treatment Effects, Econometrica 62 (2) (1994) 467-475.

[49] G. Imbens, and T. Lemieux, Regression Discontinuity Designs: A Guide to Practice, NBER Working Paper No. 13039, 2007.

[50] M. P. Keane, Structural vs. Atheoretic Approaches to Econometrics, Mimeograph, 2007.

[51] G. Kolata, In Public Health, Definitive Data Can Be Elusive, New York Times (April 23, 2002) F1.

[52] M. Krakovsky, Register or Perish, Scientific American 291 (2004) 18-20.

[53] M. Kremer, and E. Miguel, Worms: Identifying Impacts on Education and Health in the presence of Treatment Externalities, Econometrica 72 (1) (2004) 159-217.

[54] M. Kremer, J. Leino, E. Miguel and A. Zwane, Spring Cleaning: A Randomized Evaluation of

Source Water Quality Improvement, Mimeograph, 2007.

[55] A. B. Krueger, Experimental Estimates of Education Production Functions, Quarterly Journal of Economics 114 (2) (1999) 497-532.

[56] R. J. LaLonde, Evaluating the Econometric Evaluations of Training Programs Using Experimental Data, American Economic Review 76 (4) (1986) 602-620.

[57] J. List, Neoclassical Theory Versus Prospect Theory: Evidence from the Marketplace, Econometrica 72 (2) (2004) 615-625.

[58] D. Lueck, and J. A. Michael, Preemptive Habitat Destruction under the Endangered Species Act, Journal of Law and Economics 46 (1) (2003) 27-60.

[59] M. Margolis, D. E. Osgood and J. A. List, Measuring the Preemption of Regulatory Takings in the U.S. Endangered Species Act: Evidence from a Natural Experiment, Working Paper, 2004.

[60] B. D. Meyer, Unemployment Insurance and Unemployment Spells, Econometrica 58 (1990) 757-782.

[61] B. D. Meyer, Natural and Quasi-Experiments in Economics, Journal of Business & Economic Statistics 13 (2) (1995) 151-161.

[62] J. S. Mill, System of Logic, John W. Parker, West Strand, London, 1843.

[63] National Association of Home Builders, Developer's Guide to Endangered Species Regulation, Home Builder Press, Washington, DC, 1996.

[64] National Research Council (NRC), Toxicological Effects of Methylercury, National Academy Press, Washington, DC, 2000.

[65] A. L. Nichols, and R. J. Zeckhauser, The Perils of Prudence: How Conservative Risk Assessments Distort Regulation, Regulation 10 (2) (1986) 13-24.

[66] Office of Management and Budget (OMB), Circular A-4: Regulatory Analysis. September 17, 2003.

[67] J. M. Quigley, and A. M. Swoboda, The Urban Impacts of the Endangered Species Act: A General Equilibrium Analysis, Journal of Urban Economics 61 (2007) 299-318.

[68] R. Reinhold, Tiny Songbird Poses Big Test of U.S. Environmental Policy, New York Times, March 16, 1993.

[69] P. R. Rosenbaum, and D. B. Rubin, The Central Role of the Propensity Score in Observational Studies for Causal Effects, Biometrika 70 (1983) 41-55.

[70] D. B. Rubin, Estimating Causal Effects of Treatments in Randomized and Non-Randomized Studies, Journal of Educational Psychology 66 (1974) 688-701.

[71] D. B. Rubin, Assignment to a Treatment Group on the Basis of a Covariate, Journal of Educational Statistics 2 (1977) 1-26.

[72] D. R. Simmons, and R. T. Simmons, The Endangered Species Act Turns 30, Regulation 26 (4) (2003) 6-8.

[73] R. J. Smith, The Endangered Species Act: Saving Species or Stopping Growth? Regulation 15 (1) (1992) 83-87.

[74] J. Snow, On the Mode of Communication of Cholera, John Churchill, London, 1855.

[75] D. Staiger, and J. H. Stock, Instrumental Variables Regression with Weak Instruments," Econometrica, 65 (3) (1997) 557-586.

[76] D. L. Sunding, Economic Impacts of Critical Habitat Designation for the Coastal California Gnatcatcher, California Resource Management Institute Working Paper, 2003.

[77] W. K. Viscusi, and J. T. Hamilton, Are Risk Regulators Rational? Evidence from Hazardous Waste Cleanup Decisions, American Economic Review 89 (4) (1999) 1010-1027.

[78] J. E. Zabel, and R. W. Paterson, The Effects of Critical Habitat Designation on Housing Supply: An Analysis of California Housing Construction Activity, Working Paper, 2005.

Table I: Cross-Sectional Estimates of the Association between Mean TSPs and Infant Deaths Due to Internal Causes per 100,000 Live Births (estimated standard errors in parentheses)

	(1)	(2)	(3)
1969 Cross-Section	**2.48**	-0.14	0.20
	(0.92)	(0.38)	(0.41)
	[412,.05]	[357,.69]	[357,.75]
1970 Cross-Section	1.30	0.26	-0.07
	(0.72)	(0.28)	(0.24)
	[501,.02]	[441,.60]	[441,.67]
1971 Cross-Section	1.59	-0.05	0.75
	(0.98)	(0.44)	(0.47)
	[501,.02]	[460,.62]	[460,.68]
1972 Cross-Section	0.89	**-1.32**	**-1.82**
	(1.20)	**(0.65)**	**(0.87)**
	[501,.00]	[455,.48]	[455,.57]
1973 Cross-Section	2.51	-1.06	0.41
	(1.52)	(0.79)	(0.81)
	[495,.02]	[454,.59]	[454,.66]
1974 Cross-Section	**2.88**	1.01	**2.04**
	(1.34)	(0.67)	**(0.80)**
	[489,.03]	[455,.61]	[455,.68]
1969-1974 Pooled	**2.54**	0.16	0.22
	(0.84)	(0.22)	(0.20)
	[2899,.04]	[2622,.58]	[2622,.61]
Basic Natality Vars.	N	Y	Y
Unrestricted Natality	N	Y	Y
Income, Employment	N	Y	Y
Income Assist. Sources	N	Y	Y
State Effects	N	N	Y

Notes: Numbers in brackets are the number of counties and R-squareds of the regressions, respectively. The potential sample is limited to the 501 counties with TSPs data in 1970, 1971 and 1972. In a given year, the sample is further restricted to counties with nonmissing covariates. Sampling errors are estimated using the Eicker-White formula to correct for heteroskedasticity. The sampling errors in the "1969-1974 Pooled" row are also corrected for county-level clustering in the residuals over time. Regressions are weighted by numbers of births in each county. Internal causes of death arise from common health problems, such as respiratory and cardiopulmonary deaths. The control variables are listed in the Data Appendix and in Table I. State Effects are separate indicator variables for each state. Bold text indicates that the null hypothesis that the estimate is equal to zero can be rejected at the 5% level. See the text and Chay and Greenstone [20] for further details.

Table II: 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 Imperiled)	95
G2 (Imperiled)	153
G3 (Vulnerable)	250
G4 (Apparently Secure)	328
G5 (Secure)	365
G6 (Unranked)	34
G7 (Possibly Extinct)	2

B. Listed Species Information

	Endangered Species	Threatened Species	Candidate Species	Unregulated Species
# of Species for Each Designation	42	16	4	1,165

Kingdom

	Endangered Species	Threatened Species	Candidate Species	Unregulated Species
Animalia	24	7	1	376
Plantae	17	9	3	15
Fungi	1	0	0	774

Informal Taxonomic Group

Animalia:	Endangered Species	Threatened Species	Candidate Species	Unregulated Species
Amphibians	0	0	0	27
Amphipods	0	0	0	1
Birds	3	1	0	47
Butterflies and Skippers	1	0	0	35
Caddisflies	0	0	0	13
Crayfishes	0	0	0	13
Crocodilians	0	1	0	0
Dragonflies and Damselflies	0	0	0	11
Fishes	2	2	1	56

Freshwater Mussels	7	0	0	34
Freshwater Snails	0	0	0	8
Giant Silkworms	0	0	0	1
Grasshoppers	0	0	0	4
Isopods	0	0	0	1
Mammals	6	0	0	21
Mayflies	0	0	0	12
Other Beetles	0	0	0	1
Other Crustaceans	0	0	0	2
Other Insects	0	0	0	2
Other Molluscs	0	0	0	1
Other Moths	0	0	0	24
Papaipema Moths	0	0	0	1
Reptiles	0	0	0	13
Spiders	1	0	0	4
Stoneflies	0	0	0	6
Terrestrial Snails	0	1	0	26
Tiger Moths	0	0	0	2
Turtles	4	2	0	5
Underwing Moths	0	0	0	5
Plantae:				
Conifers and Relatives	0	0	0	2
Ferns and Relatives	0	0	0	33
Flowering Plants	17	9	3	559
Hornworts	0	0	0	2
Liverworts	0	0	0	61
Mosses	0	0	0	117
Fungi:				
Lichens	1	0	0	15

Global Conservation Status Rank

G1 (Critically Imperiled)	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

C. Census Tract Species Information

# of Tracts without a NatureServe Species	41
# of Tracts with a NatureServe Species	1,522

	# of Tracts with 1990s Endangered Species that are …	# of Tracts with 1990s Threatened Species that are …	# of Tracts with 1990s Candidate Species that are …	# of Tracts with No 1990s Listed Species, but with Species that are …
Kingdom				
Animalia	85	219	0	1,449
Plantae	97	61	28	1,170
Fungi	31	0	0	57
Global Conservation Status Rank				
G1 (Critically Imperiled)	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

Table III: Census Tract Summary Statistics for Treatment and Control Groups (**Animals**)

	(T) 1 or More Endangered Species Listed in 1990s (1)	(C) No Species Listed as Endangered or Threatened Species in 1990s (2)	Difference (robust SEs) (3)	Difference Conditional on G-Rankings (robust SEs) (4)
Number of Census Tracts	85	1,170		
1980 Housing Characteristics				
Mean Housing Value	24,790	31,048	**-6,258** **(3,035)**	-511.7 (3,336)
1990 Housing Characteristics				
Median Housing Value	57,835	61,797	-3,962 (2,955)	-3,821 (4,528)
Mean Housing Value	42,948	59,750	**-16,802** **(3,137)**	-2,027 (3,706)
Proportion Mobile Homes	0.217	0.122	**0.094** **(0.011)**	0.016 (0.015)
Proportion Occupied	0.857	0.912	**-0.054** **(0.013)**	0.018 (0.020)
Proportion Owner-Occupied	0.648	0.602	**0.046** **(0.016)**	0.035 (0.023)
Proportion with 0 to 2 Bedrooms	0.357	0.423	**-0.065** **(0.010)**	**-0.049** **(0.015)**
Proportion with 3 to 4 Bedrooms	0.620	0.556	**0.064** **(0.010)**	**0.047** **(0.014)**
Proportion with 5 Plus Bedrooms	0.023	0.021	0.002 (0.003)	0.002 (0.464)
Proportion with No Full Kitchen	0.021	0.011	**0.010** **(0.003)**	0.002 (0.003)
Proportion Attached	0.013	0.027	**-0.014** **(0.005)**	-0.001 (0.005)
Proportion Detached	0.701	0.664	**0.037** **(0.013)**	0.029 (0.018)
1990 Demographic Characteristics				
Population Density (100k per sq. m.)	9.760	44.270	**-34.510**	-7.700

|			(2.740)	(4.740)
Proportion Black	0.186	0.231	-0.045	-0.003
			(0.024)	(0.029)
Proportion Hispanic	0.008	0.010	-0.002	0.001
			(0.002)	(0.002)
Proportion under 18	0.246	0.239	0.006	0.011
			(0.005)	(0.006)
Proportion 65 or Older	0.136	0.125	0.010	-0.008
			(0.006)	(0.008)
Proportion Foreign Born	0.009	0.019	**-0.010**	-0.003
			(0.001)	(0.002)
Proportion of HHs Headed by a Female	0.202	0.247	**-0.045**	-0.018
			(0.013)	(0.018)
Proportion in Same House in Last 5 Years	0.601	0.534	**0.067**	0.013
			(0.013)	(0.016)
Proportion HS Dropouts	0.142	0.134	0.008	0.016
			(0.009)	(0.012)
Proportion No HS	0.181	0.173	0.007	-0.012
			(0.006)	(0.008)
Proportion with BA or Better	0.124	0.185	**-0.061**	-0.017
			(0.010)	(0.015)
Species Listed in 1980s	0.271	0.050	**0.220**	**-0.187**
			(0.049)	**(0.066)**
Species Listed in 1970s	0.318	0.150	**0.167**	0.010
			(0.052)	(0.065)
1990 Economic Characteristics				
Mean HH Income	29,921	33,580	**-3,659**	36
			(1,274)	(1369)
Proportion below Povery Level	0.152	0.134	0.018	0.004
			(0.010)	(0.012)
Proportion on Public Assistance	0.087	0.071	**0.016**	0.015
			(0.006)	(0.008)
Unemployment Rate	0.057	0.051	0.006	0.007
			(0.004)	(0.004)

Notes: See the text for details.

Table IV: Census Tract Summary Statistics for Treatment and Control Groups (**Plants**)

	(T) 1 or More Endangered Species Listed in 1990s (1)	(C) No Species Listed as Endangered or Threatened Species in 1990s (2)	Difference (robust SEs) (3)	Difference Conditional on G-Rankings (robust SEs) (4)
Number of Census Tracts	97	1,026		
1980 Housing Characteristics				
Mean Housing Value	25,493	30,054	**-4,560** **(1,798)**	**-3,698** **(1,864)**
1990 Housing Characteristics				
Median Housing Value	54,109	62,521	**-8,414** **(2,079)**	**-7,379** **(2,372)**
Mean Housing Value	43,125	58,188	**-15,063** **(2,532)**	**-27,763** **(2,079)**
Proportion Mobile Homes	0.202	0.143	**0.059** **(0.011)**	-0.006 (0.013)
Proportion Occupied	0.869	0.908	**-0.039** **(0.012)**	-0.005 (0.011)
Proportion Owner-Occupied	0.683	0.621	**0.062** **(0.015)**	**0.050** **(0.016)**
Proportion with 0 to 2 Bedrooms	0.365	0.406	**-0.040** **(0.010)**	-0.014 (0.011)
Proportion with 3 to 4 Bedrooms	0.615	0.573	**0.042** **(0.010)**	0.016 (0.011)
Proportion with 5 Plus Bedrooms	0.019	0.022	-0.002 (0.002)	-0.002 (0.002)
Proportion with No Full Kitchen	0.016	0.013	0.003 (0.002)	0.003 (0.002)
Proportion Attached	0.012	0.025	**-0.013** **(0.004)**	-0.007 (0.004)
Proportion Detached	0.722	0.673	**0.049** **(0.012)**	**0.048** **(0.013)**
1990 Demographic Characteristics				
Population Density (100k per sq. m.)	6.350	36.150	**-29.800** **(1.760)**	**-14.670** **(2.230)**
Proportion Black	0.155	0.217	**-0.063**	-0.022

54

			(0.022)	(0.022)
Proportion Hispanic	0.007	0.009	**-0.003**	**-0.004**
			(0.001)	**(0.001)**
Proportion under 18	0.246	0.241	0.005	0.004
			(0.004)	(0.004)
Proportion 65 or Older	0.132	0.124	0.008	0.005
			(0.005)	(0.006)
Proportion Foreign Born	0.011	0.017	**-0.006**	**-0.006**
			(0.001)	**(0.001)**
Proportion of HHs Headed by a Female	0.175	0.235	**-0.060**	**-0.034**
			(0.011)	**(0.012)**
Proportion in Same House in Last 5 Years	0.612	0.543	**0.069**	**0.051**
			(0.012)	**(0.013)**
Proportion HS Dropouts	0.148	0.134	0.014	**0.022**
			(0.009)	**(0.009)**
Proportion No HS	0.197	0.173	**0.024**	**0.024**
			(0.006)	**(0.006)**
Proportion with BA or Better	0.118	0.185	**-0.067**	**-0.052**
			(0.010)	**(0.012)**
Species Listed in 1980s	0.155	0.048	**0.107**	0.009
			(0.037)	(0.039)
Species Listed in 1970s	0.000	0.006	**-0.006**	**-0.016**
			(0.002)	**(0.007)**
1990 Economic Characteristics				
Mean HH Income	30,957	33,386	**-2,429**	-619
			(2,429)	(970)
Proportion below Poverty Level	0.124	0.137	-0.013	**-0.023**
			(0.008)	**(0.008)**
Proportion on Public Assistance	0.065	0.072	-0.006	-0.009
			(0.004)	(0.005)
Unemployment Rate	0.044	0.052	**-0.008**	**-0.012**
			(0.003)	**(0.003)**

Notes: See the text for details.

FIGURE 1: ±1 STANDARD ERROR OF HEDONIC ESTIMATES OF BENCHMARK CLIMATE CHANGE SCENARIO ON VALUE OF AGRICULTURAL LAND
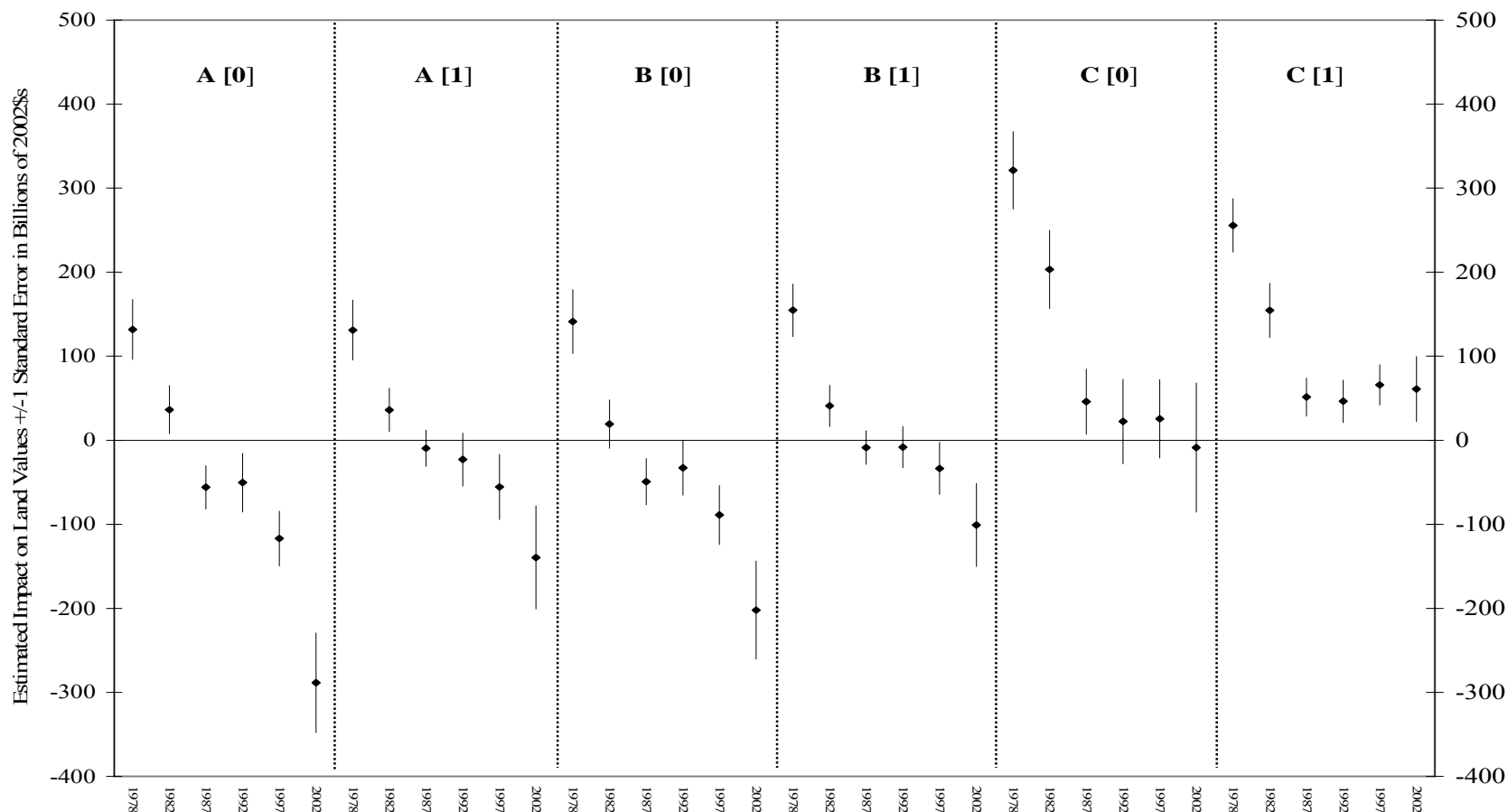
Notes: All dollar values are in 2002 constant dollars. Each of the 36 lines represents an estimate of the impact of increases of 5 degrees Fahrenheit and 8% precipitation due to climate change on agricultural land values in the US. The midpoint of each line is the point estimate and the top and bottom of the lines are calculated as the point estimate plus and minus one standard error of the predicted impact, respectively. The A, B, and C panels correspond to three sets of control variables. In the A columns, the climate variables (i.e., temperature and precipitation) are the only regressors. The entries in the B panels are adjusted for soil characteristics, as well as per capita income and population density and their squares. The specification associated with the C panels adds state fixed effects to the B specification. Among the A, B, and C panels, the panel "[0]" regression equations are unweighted and the panel "[1]" entries are the result of weighting by the square root of acres of farmland. Within each of the six panels, the lines report the estimated impact from estimating the relevant model on data from one of the 6 years indicated on the x-axis. See the text and Deschenes and Greenstone [26] for further details (including a list of the exact covariates).

[1] For a formal analysis of the inefficiency caused by externalities, see Baumol and Oates [10] or Cropper and Oates [24].

[2] See Gayer and Horowitz [34] for a discussion of design issues for policy instruments.

[3] We use the terms "pollution" or "emissions" broadly to mean reductions in environmental quality. They include externalities due to such things as over-fishing, deforestation, and species extinction.

[4] This definition of causality is certainly not original. Philosophers in the 19th century used similar definitions. For example, John Stuart Mill [62] wrote: "If an instance in which the phenomenon under investigation occurs, and an instance in which it does not occur, have every circumstance in common save one, that one occurring in the former: the circumstances in which alone the two instances differ, is the effect, or the cause, or a necessary part of the cause of the phenomenon."

[5] See Freedman [33] for a criticism of the regression approach to testing causal propositions.

[6] One potential solution for dealing with omitted variables problems is to use a fixed effects model that removes all permanent determinants of mortality as potential sources of bias. However, the cost of this approach can be quite steep because fixed effects are likely to exacerbate the attenuation bias due to mismeasurement. This is because in the fixed effects case, the magnitude of the attenuation bias also depends on the correlation across years in the "true" measure of air pollution. Specifically, a high correlation in the "true" year to year values of air pollution greatly exacerbates the attenuation bias. It is reasonable to assume a county's true air pollution concentrations are highly correlated across years. So, although fixed effects remove permanent unobserved factors as a source of bias, it is likely to exacerbate the attenuation bias due to measurement error.

[7] See Angrist and Lavy [4] and Rubin [71] for applications.

[8] See Greenstone [37] for an application in environmental economics.

[9] "Internal" and "external" deaths span all possible causes of death. Deaths with 9th International Classification of Diseases (ICD) codes from 001 to 799 are classified as internal, while those with ICD codes from 800 to 999 are in the external category.

[10] The control variables included in each specification are listed in the Data Appendix of Chay and Greenstone [20].

[11] The presence of publication bias would also imply that the findings of meta-analyses of published papers are unlikely to produce reliable results.

[12] The tendency for individuals that have received a negative labor market shock to sign up for job training programs is a particularly well known example that is often referred to as the "Ashenfelter [pre-program earnings] dip" and makes it extremely difficult to estimate the impact of training programs in the absence of experimental or quasi-experimental variation in program participation [6, 7].

[13] Angrist and Krueger [3] provide a more extensive treatment of the DD and fixed effects approaches.

[14] More precisely, the requirement is that $(Z-E(Z|X))$ is orthogonal to the unobserved components of potential outcomes.

[15] These "threats to validity" apply to all empirical studies, but this subsection discusses them in the context of quasi-experiments.

[16] This subsection draws on Cook and Campbell [23] and Meyer [61].

[17] Other threats to internal validity include misspecified variances that lead to biased standard errors, sample attrition, and changes in data collection that cause changes in the measured variables.

[18] See Imbens and Angrist [48] and Heckman and Vytlacil [45] on the interpretation of instrumental variables estimates in the presence of heterogeneous responses.

[19] See Heckman, Lochner, and Taber [43] on estimating general equilibrium treatment effects.

[20] See Keane [50] for the vigorous expression of an alternative viewpoint.

[21] One frequent criticism of randomized evaluations of policies is that it is unethical to deny some members of the population the treatment. We do not find this argument compelling when the treatment effect is unknown, which is often the case. Moreover, the limited set of resources available for many environmental programs means that treatments in environmental programs may be assigned on the basis of political concerns that are not welfare enhancing. For example, Viscusi and Hamilton [77] claim that political criteria impact Superfund decisions such that "Superfund expenditures do not fare well when evaluated in terms of cancer protection" [77, p. 1012].

[22] Policymakers can also improve the dissemination of whether a causal relationship is reliably estimated. For example, within the Office of Management and Budget (OMB), the Office of Information and Regulatory Affairs (OIRA) is charged with overseeing the regulatory process. Under Executive Order 12866, OIRA verifies that each major regulation has benefits that "justify" its costs. OIRA recently published guidelines to the agencies on how to conduct cost-benefit analyses [66]. It would be a relatively simple task for the guidelines to place more emphasis on

experimental and quasi-experimental studies and to include criteria for assessing the validity of associational studies. Efforts like these would spur researchers to present meaningful tests of validity and in an ideal world the candor would strengthen scientists' credibility with policymakers in future interactions.

[23] See Ferraro, McIntosh and Ospina [32] for an analysis of the ESA's effectiveness at protecting species. They find mixed evidence on the legislation's effectiveness although like all other work on this question they do not have access to a time series of population counts and must rely on the qualitative outcome "change in endangerment status" between 1993 and 2004.

[24] Preemptive habitat destruction constitutes a change in land management decisions in order to avoid potential endangered species problems. As such, it is not a violation of the law. Indeed the National Association of Home Builders [63] stated in one of its guidance documents, "Unfortunately, the highest level of assurance that a property owner will not face an ESA issue is to maintain the property in a condition such that protected species cannot occupy the property. Agricultural farming, denuding of property, and managing the vegetation in ways that prevent the presence of such species are often employed in areas where ESA conflicts are known to occur." Of course, there is also the concern that the ESA provides incentives for illegally destroying protected species or habitats, a practice referred to as "shoot, shovel, and shut up".

[25] A critical habitat is a formally-designated geographic area, whether occupied by a protected species or not, that is considered to be essential for a given protected species' conservation. These designations are required by law; however, in practice, due to a lack of information, of resources, and of any additional statutory protection they provide to species, the government has frequently failed to make such designations.

[26] The Supreme Court decision upholding this definition was a six-to-three vote. In his dissenting opinion, Justice Antonin Scalia argued that "the Court's holding that the hunting and killing prohibition incidentally preserves habitat on private lands imposes unfairness to the point of financial ruin – not just upon the rich, but upon the simplest farmer who finds his land conscripted to national zoological use."

[27] Former Representative Richard Pombo of California proposed H.R. 3824 in 2005, which passed the House of Representatives by a 229-to-193 vote. The bill attempts to overhaul the Act. However, the Senate did not consider the bill, and its near-term prospects are doubtful.

[28] For more information on the measure, see http://www.natureserve.org/explorer/granks.htm.

[29] In maintaining its database, NatureServe devotes the most resources to accurately tracking the habitats of the species that are imperiled or vulnerable everywhere they occur, which are the species with ranks of G1, G2, and G3. These resource allocation decisions are unaffected by whether a species is protected under the ESA and are instead based on NatureServe's assessment of the level of imperilment.

[30] Currently, we only have access to each species' most recent G-rank. Although G-ranks change infrequently, it is possible that the G-ranks in our data file differ from those that applied for some species in 1990. For this reason, it is especially important that we chose to base the quasi-experiment on species' global status or G-rank, rather than the national or sub-national ranks which are also available from NatureServe. Specifically, the global ranks are an "assessment of the condition of the species or ecological community across its entire range." (see http://www.natureserve.org/explorer/ranking.htm) In contrast, the national or sub-national ranks document a species' condition nationally or at the sub-national (primarily states) level. The most recent national or sub-national ranks may reflect the impact of the ESA during the 1990s. It is unlikely that the G-ranks would reflect the ESA's impacts. Nevertheless, we are working to obtain the full history of global, national, and sub-national ranks from NatureServe and will use them in our full evaluation of the ESA.

[31] A tract is coded as containing a species if there is any overlap between the species' habitat and the tract's "footprint". Consequently, an entire tract is considered part of a species' habitat even if the species' habitat only partially overlaps the tract's 'footprint'.

[32] In a separate unreported analysis, we defined the treatment as those census tracts that contained *only one* species listed as endangered in the 1990s. This further restriction on the treatment definition did not materially alter any results.

[33] These results are available from the authors upon request.