Empirical Methods for Political Economy Analyses of Environmental Policy

Daniel L. Millimet
Southern Methodist University & IZA

Synopsis

Theoretical models highlighting the salient role of features of the political system in the determination of environmental policy have grown in sophistication and popularity. In turn, this has lead to a burgeoning empirical literature testing the predictions emanating from these models. However, proper testing requires use of observational data, compelling one to confront many statistical issues – nonrandom selection, endogeneity, and measurement error – and choose among many different estimation methods. This article provides an overview of some of these estimation methods, emphasizing the differences among them. In particular, focus is given to the estimation of the causal effect of political institutions on environmental policy, and the role of interactions and spillovers between political agents in the determination of environmental policy.

Keywords: Environmental Federalism, Environmental Regulation, Instrumental Variables, Measurement Error, Nonrandom Selection, Political Economy, Potential Outcomes, Selection on Observables, Selection on Unobservables, Spatial Econometrics, Spillovers, Treatment Effects

*Corresponding author: Daniel Millimet, Department of Economics, Box 0496, Southern Methodist University, Dallas, TX 75275-0496. Tel: (214) 768-3269. Fax: (214) 768-1821. E-mail: millimet@smu.edu.
Glossary

**Instrumental Variable.** Refers to a variable satisfying three criteria: (i) it is correlated with the endogenous regressor in the empirical model, (ii) it is uncorrelated with the error term in the outcome equation (i.e., it is exogenous), and (iii) it only affects the outcome variable via its effect on the endogenous regressor.

**Panel Data.** Refers to a particular class of data structures whereby the researcher has repeated observations of the primary sampling unit. For example, if the primary sampling unit is a country, a panel data set may contain multiple years of data for each country.

**Potential Outcomes.** Refers to the outcome that would be realized for a particular observation under a certain state of nature. For states of nature that are not realized, the potential outcome is not observed. For states of nature that are realized, the potential outcome is the outcome that is observed by the researcher.

**Propensity Score.** Refers to the probability of a particular observation receiving the treatment in question. It is typically unknown and estimated by the researcher by specifying a model that relates treatment assignment to observable attributes.

**Selection on Observables.** Refers to situations where observations under study self-select into the treatment or control group only on the basis of attributes observable to the empirical researcher.

**Selection on Unobservables.** Refers to situations where observations under study self-select into the treatment or control group on the basis of at least some attributes that are unobservable to the empirical researcher.
Cross References

14. Strategic Environmental Policy: Theory
15. Environmental Federalism: Empirics
16. Lobbying, Voting and Environmental Policy: Theory
20. Political Economy of Trade and the Environment
21. Political Economy of International Environmental Agreements
22. Democracy, Political Institutions and Environmental Policy
1 Introduction

Empirical researchers are often interested in testing theories concerning the political economy of environmental policy. To do so requires one to confront many statistical issues and choose among many different estimation methods. This article provides an overview of some of these estimation methods and highlights the differences among them. In the first section, the literature concerned with the causal effect of political institutions on environmental policy is examined. Since many of these institutions are binary in nature (e.g., democratic versus non-democratic regimes) and researchers are (rightfully) concerned with nonrandom selection into different institutional arrangements, the empirical methods developed in the program evaluation literature are extremely useful. This section contains a brief overview of many empirical methods from this area.

In the second section, the literature concerning the role of interactions and spillovers between political agents in the determination of environmental policy is discussed. The practical importance of such interactions is central to the debate over environmental federalism (i.e., the appropriate bureaucratic level to determine environmental policy). Since these interactions are modeled as the dependence of one agent’s choices on the choices of others, where all agents act simultaneously, empirical tests of the importance of such interactions are not straightforward. However, the literature has successfully relied on advancements in spatial econometrics to empirically test for such interactions. This section provides a short introduction to some of the empirical methods from this area.

2 Political Institutions and Environmental Policy

2.1 Background

The theoretical and empirical literature examining the causal effect of political institutions on environmental policy is quite vast. Rules governing the election of policymakers, bureaucratic organization, and international dealings may all play crucial roles in shaping policy outcomes in general, and environmental policy in particular.

Electoral rules that have been linked to environmental policy include majoritarian versus proportional electoral regimes and political term limits. Proportional regimes are hypothesized to generate more stringent environmental policy since such electoral rules compel policymakers to internalize the benefits from the provision of public goods that are national in scope. Majoritarian electoral rules may also induce majority bias, thereby distorting environmental policy from its optimum. Existant empirical research utilizes cross-country data to assess the causal effect of electoral regime type on several measures of environmental
policy.

In terms of political term limits, theoretical arguments have been put forth suggesting that incumbents seeking re-election may manipulate secondary policy issues—such as the environment—to persuade voters. Consequently, binding term limits, by precluding incumbents from seeking re-election, break this link between policy outcomes and the preferences of voters. This model predicts substantial variation in policy outcomes across periods depending on whether key policymakers are eligible or not eligible for re-election. Empirical researchers have employed state-level panel data from the U.S. to assess variation in environmental policy across regimes defined on the basis of whether the governor faces a binding term limit.

Examples of institutional arrangements analyzed include democratic versus autocratic political regimes, presidential-congressional versus parliamentary constitutional regimes, and unicameral versus bicameral legislative regimes. Theoretical models typically imply a beneficial effect of democracy on environmental policy due to the greater participation of the electorate in the political process. Early empirical assessments of this conjecture use cross-country data, where regime type is measured using a binary indicator for democracy and environmental policy is measured by participation in international environmental agreements.

In terms of constitutional regimes, theoretical studies have contrasted tax setting behavior in presidential-congressional, as opposed to parliamentary, systems. According to some models, the greater separation of powers under a presidential-congressional regime enables voters to better discipline policymakers and thereby limit taxes (and hence opportunities for policymakers to extract rents and/or engage in redistribution). Other models offer a different mechanism, focusing on the effects of lobbying under the two constitutional regimes. These models predict that presidential-congressional regimes lead to greater inequality in the provision of public goods and that lobbying reinforces such inequality; lobbying mitigates inequities under parliamentary regimes. As above, empirical researchers have exploited cross-country data on regime type to study the causal effects on environmental policy.

A similar literature has emerged analyzing the effect of legislative organization—specifically, bicameral versus unicameral regimes—on policy outcomes. Theoretical research predicts that a bicameral legislature reduces the influence of lobbying due to the greater number of veto players involved in the determination of policy. As a result, environmental policy is closer to the social optimum in bicameral regimes. Again, empirical researchers have relied on cross-country data to assess the causal effect of legislative regime type on the level and dispersion of several measures of environmental policy.

Finally, the role of international agreements in the determination of environmental policy has proven to be quite contentious. The primary example of this is the effect of membership in the World Trade Organi-
zation (WTO) and its predecessor, the General Agreement on Tariffs and Trade (GATT), on environmental policy. Membership may alter a country’s environmental policy due to WTO restrictions on domestic policy that has (intended or unintended) consequences on trade. Moreover, membership may limit a country’s ability to use trade sanctions to enforce other international agreements such as environmental treaties. On the other hand, the WTO may offer a framework to prevent a ‘race-to-the-bottom’ in environmental regulatory stringency. Exploitation of country-level panel data is typically used to empirically assess the impact of WTO membership on environmental outcomes.

2.2 Empirical Methods

As the preceding discussion highlights, researchers seeking to empirically assess the relationship between political institutions and environmental policy are often interested in the causal effect of a binary treatment (e.g., majoritarian or proportional elections, democratic or autocratic, bicameral or unicameral, WTO member or nonmember) on policy outcomes. Fortunately, the methodological literature in statistics and economics has witnessed profound growth in this area. This growth has not only led to new estimation techniques, but also the revitalization of older techniques as well as improved understanding of how different estimators relate to one another. Moreover, many of the empirical methods discussed here apply or have been extended to the case of non-binary treatments. This is crucial since many other political attributes not discussed here are typically measured in a non-binary fashion (e.g., corruption or political instability); even those considered here are not always measured in a binary fashion (e.g., democracy may be measured using a categorical or continuous index). For ease of exposition, however, I focus on binary treatments.

Empirical methods useful for estimating causal effects of a binary treatment can be grouped into three categories: (i) randomization, (ii) selection on observables, and (iii) selection on unobservables. Random experiments have been in use for more than a century and underwent great advancements after the early part of the 20th century. Often randomization is used to ensure comparability (at least in expectation) across the treatment and control groups. However, when randomization is not feasible – and, clearly, the analysis of political institutions is such a case – researchers must rely on non-experimental or observational data. In such situations, nonrandom selection of subjects (in this case, political units) into the treatment or control group becomes a paramount concern and the demands placed on the data are heightened.

Empirical methods used to address nonrandom selection in observational data are divided into the two remaining strands: selection on observables and selection on unobservables techniques. If subjects self-select into the treatment or control group only on the basis of attributes observable to the researcher, this is referred to as the case of selection on observables. On the other hand, if subjects self-select into the treatment or control group on the basis of at least some attributes that are unobservable to the researcher,
but correlated with the outcome of interest, this is known as the case of selection on unobservables. To understand the concept underlying this distinction, as well as provide a framework for understanding precisely what different estimation techniques actually estimate, it is useful to review the potential outcomes framework attributed to the seminal work of Neyman, Fisher, Roy, Rubin, and Heckman. After that, estimators falling within each of the two categories are briefly presented. Although clearly crucial for meaningful empirical work, issues associated with inference are not discussed here.

2.2.1 Potential Outcomes Framework

Consider a random sample of $N$ observations from a large population indexed by $i = 1, ..., N$. Let $Y_i(T)$ denote the potential outcome of observation $i$ under treatment $T$, $T \in \{0, 1\}$. Implicit in this specification is the Stable Unit Treatment Value Assumption (SUTVA), whereby the potential outcomes of observation $i$ do not depend on the treatment assignment of other observations in the population. In the context of environmental policy, this assumption should not be overlooked. For example, the policy choices of country $i$ under, say, a democracy may depend on whether or not its neighbors are democratic or autocratic. In this case, the potential outcomes do not just depend on $T_i$, but also $T_j (i \neq j)$.

Under SUTVA, the observation-specific causal effect of the treatment ($T = 1$) relative to the control ($T = 0$) is defined as the difference between the corresponding potential outcomes. Formally,

$$\tau_i = Y_i(1) - Y_i(0).$$

This simple, yet powerful, framework already illuminates one salient point: without further assumptions, there is no reason to expect the causal effect of the treatment to be identical across all observations. Thus, attention typically focuses on estimation of specific aspects of the distribution of $\tau$ in the population. Several population parameters are of potential interest. The three most common are the average treatment effect (ATE), $\tau_{ATE}$, the average treatment on the treated (ATT), $\tau_{ATT}$, and the average treatment on the untreated (ATU), $\tau_{ATU}$. Distributional treatment effects, such as quantile treatment effects, are often useful in cases where economic theory predicts differential effects of a treatment across the distribution or policymakers are particularly concerned with outcomes associated with a certain segment of the population. While the conditions for identification often do not differ from those considered here, I do not discuss estimation of such parameters.
The three average treatment effect parameters are defined as follows:

\[ \tau_{ATE} = \mathbb{E}[\tau_i] = \mathbb{E}[Y_i(1) - Y_i(0)] \]
\[ \tau_{ATT} = \mathbb{E}[\tau_i|T_i = 1] = \mathbb{E}[Y_i(1) - Y_i(0)|T_i = 1] \]
\[ \tau_{ATU} = \mathbb{E}[\tau_i|T_i = 0] = \mathbb{E}[Y_i(1) - Y_i(0)|T_i = 0]. \]

The ATE is the expected treatment effect of an observation chosen at random from the population; the ATT (ATU) is the expected treatment effect of an observation chosen at random from the subpopulation of treated (controls). The parameter that is most interesting depends on the question one wishes to answer.

The data contain \( \{Y_i, T_i, X_i\}_{i=1}^N \), where \( Y_i \) is the observed outcome, \( T_i \) is a binary indicator of the treatment received, and \( X_i \) is a vector of covariates. The relationship between the potential and observed outcomes is given by

\[ Y_i = T_iY_i(1) + (1 - T_i)Y_i(0) \]

which makes clear that only one potential outcome is observed for any individual; the missing potential outcome is referred to as the counterfactual. As such, estimating \( \tau_i \) (or any aspect of the distribution of \( \tau \)) is not trivial as there is an inherent missing data problem.

Random experiments circumvent this missing data problem since the randomization process ensures that

\[ Y(0), Y(1) \perp T, \]

where \( \perp \) denotes independence. In other words, randomization implies that the treatment is assigned independently of potential outcomes. While this unconditional independence is unlikely to hold in observational data, a weaker version of conditional independence,

\[ Y(0), Y(1) \perp T \mid W, \]

may hold, where \( W \) is a vector of predetermined, observation-specific attributes. If \( W \subseteq X \), then all of the covariates required for (5) to hold are observed in the data and the researcher is in the selection on observables world. If \( W \not\subseteq X \), then at least one covariate required for (5) to hold is unobserved and the researcher is in the selection on unobservables world.

### 2.2.2 Selection on Observables Estimators

Assuming that \( W \subseteq X \) is referred to as the conditional independence or unconfoundedness assumption.
Assumption 1 \( Y(0), Y(1) \perp T \mid X, \) where \( X \) is a set of pre-determined covariates.

However, Assumption 1 is not sufficient to identify any of the average treatment effect parameters defined above.

Traditional regression-based estimators proceed by placing additional structure on the data-generating process.

Assumption 2 Potential outcomes and latent treatment assignment are additively separable in observables and unobservables

\[
Y_i(0) = \alpha_0 + X_i \beta_0 + \varepsilon_{0i} \\
Y_i(1) = \alpha_1 + X_i \beta_1 + \varepsilon_{1i} \\
T_i = \mathbb{I}[X_i \gamma + \eta_i > 0]
\]

where \( \mathbb{I}[\cdot] \) is the indicator function. \( \varepsilon_0, \varepsilon_1, \eta \sim \mathcal{N}_3(0, \Sigma) \), where

\[
\Sigma = \begin{bmatrix}
\sigma_0^2 & \rho & 0 \\
\rho & \sigma_1^2 & 0 \\
0 & 0 & 1
\end{bmatrix}.
\]

Under Assumption 2,

\[
\tau_{ATE} = \mathbb{E}[Y_i(1) - Y_i(0)] = \mathbb{E}[\{(\alpha_1 - \alpha_0) + X(\beta_1 - \beta_0) + \varepsilon_{1i} - \varepsilon_{0i}\}] = \alpha_1 - \alpha_0 + \mathbb{E}[X](\beta_1 - \beta_0)
\]

\[
\tau_{ATT} = \alpha_1 - \alpha_0 + \mathbb{E}[X|T = 1](\beta_1 - \beta_0)
\]

\[
\tau_{ATU} = \alpha_1 - \alpha_0 + \mathbb{E}[X|T = 0](\beta_1 - \beta_0).
\]

Moreover, if one restricts \( \beta_1 = \beta_0 = \beta \), then all three average treatment effects parameters are identical (although treatment effects still vary across observations unless one additionally assumes \( \varepsilon_{1i} = \varepsilon_{0i} \)). Estimation proceeds by substituting these functional form assumptions into (3). The resulting estimating equation is

\[
Y_i = \alpha_0 + (\alpha_1 - \alpha_0)T_i + X_i \beta_0 + T_i X_i (\beta_1 - \beta_0) + [\varepsilon_{0i} + (\varepsilon_{1i} - \varepsilon_{0i})T_i].
\]

Under Assumptions (1) and (2), Ordinary Least Squares (OLS) estimation of the parameters in (7), along with the replacement of \( \mathbb{E}[X], \mathbb{E}[X|T = 1] \), and \( \mathbb{E}[X|T = 0] \) with their sample counterparts, provides an unbiased estimate of the average treatment effect parameters in (6).
Many alternative estimators that do not rely on the functional form restrictions in Assumption (2) instead make use of the propensity score in combination with a common support or overlap assumption.

**Assumption 3** $0 < \Pr(T = 1|X) < 1$ for all $X$, where $P(X_i) \equiv \Pr(T_i = 1|X_i)$ is the propensity score.

The propensity score represents the conditional probability of receiving the treatment given the covariates. Assumption (3) is needed to ensure at least some treated and control observations for every realization of $X$; failure of this assumption implies that the average treatment effect parameters are not well defined. To be precise, failure of the common support assumption does not ‘invalidate’ the methods relying on it. It does, however, alter the interpretation of the parameters being estimated. For example, estimators of $\tau_{ATE}$ that require Assumption (3) would be biased when this assumption fails, but would be unbiased for the alternative parameter $\bar{\tau}_{ATE} = \mathbb{E}[\tau|X \subset \mathcal{S}]$, where $\mathcal{S}$ denotes the region of common support. $\bar{\tau}_{ATE}$ represents the expected treatment effect for an observation chosen at random from the subpopulation with values of $X$ contained in the common support.

Estimators relying on Assumptions (1) and (3) are intuitively very appealing. Assumptions (1) ensures that $T$ is assigned randomly across observations in the population conditional on $X$. Thus, conditional on $X$, the data mimic a random experiment. Moreover, Assumption (3) ensures that once $X$ is conditioned upon, neither the treatment nor the control group will be empty (at least in the population). Together, then, the average treatment effect is (nonparametrically) identified at every value of $X$ – denoted by $\tau(X)$ – and the average treatment effect parameters in (2) can be estimated by averaging over the correct distribution of $X$. Specifically, for the ATE, one averages $\tau(X)$ over the distribution of $X$ in the population (as estimated by the sample counterpart). For the ATT (ATU), one averages over the distribution of $X$ in the subpopulation of treated (controls), again as estimated by the sample counterparts.

In theory, $\tau(X)$ can be estimated by matching or grouping observations with identical values of $X$, and then computing the mean difference in outcomes across treated and control observations within groups. Formally, this amounts to replacing the missing counterfactual for observation $i$ in (1) by the mean outcome of observations with $T_i = 1 - T_i$ and $X_i = X_i$. In practice, however, estimators of this type suffer from the curse of dimensionality; as the number of covariates in $X$ increase, the likelihood of observing treated and control observations with identical covariates tends to zero. Fortunately, it has been shown that if Assumption (1) holds, then

$$Y(0), Y(1) \perp T \mid P(X).$$

As a result, instead of estimating average treatment effects at every value of $X$, one needs only an estimate at each value of $P(X)$. The average treatment effect parameters in (2) can then be estimated by averaging over the correct distribution of $P(X)$.
Prior to continuing, it is important to discuss three practical issues confronted when applying any estimator that utilizes the propensity score. First, despite the gains from reducing the conditioning set from the (typically) high dimensional $X$ to the single index, $P(X)$, it remains likely that either the treatment or control group will not contain observations for all values of $P(X)$ in finite samples. Many commonly employed estimators differ simply in the means by which they address this problem. Consequently, all such estimators are asymptotically equivalent, but may differ markedly in finite samples (as they typically represent different trade-offs along the bias-variance spectrum). Second, as the propensity score is typically unknown, it must be estimated. Moreover, even if the propensity score were known, research has shown that using the true propensity score is inefficient. In practice, a probit or logit model is used to estimate the unknown propensity score; others consider nonparametric estimation of the propensity score.

Finally, the covariates included in the propensity score model (including higher order terms and interactions in the absence of nonparametric estimation) must be decided upon. In the interest of brevity, I simply note two important results. First, while there is some tension between accumulating a set of covariates such that Assumptions (1) and (3) hold simultaneously – adding additional covariates makes the unconfoundedness assumption more plausible, it may invalidate the common support assumption – this should play no role in deciding which covariates belong in $X$. The only factor for deciding whether a covariate belongs in $X$ is whether it is necessary to condition on that variable for unconfoundedness to hold. If there is no set of covariates that simultaneously satisfy Assumptions (1) and (3), then one must turn to estimators that do not require the common support assumption. Second, there is little cost to over-specifying the propensity score model (in terms of including irrelevant higher order and/or interaction terms or even irrelevant variables), but a substantial cost to under-specifying the propensity score model. Thus, while so-called balancing tests are often utilized, researchers are advised to err on the side of a less parsimonious specification.

The first set of propensity score estimators considered are matching estimators. All matching estimators entail estimating the missing counterfactual for each observation and then estimating the average treatment effect parameters in (2) using the estimates of $\hat{\tau}_i$. Moreover, all estimators of the missing counterfactual in this set can be written in the following form:

$$\hat{Y}_i(0) = \frac{1}{\sum_{l \in \{T_l = 0\}} \omega_{il}} \sum_{l \in \{T_l = 0\}} \omega_{il} Y_l(0) \quad \text{if } i \in \{T_i = 1\} \quad (9)$$

$$\hat{Y}_i(1) = \frac{1}{\sum_{l \in \{T_l = 1\}} \omega_{il}} \sum_{l \in \{T_l = 1\}} \omega_{il} Y_l(1) \quad \text{if } i \in \{T_i = 0\} \quad (10)$$
where $\omega_{il}$ is the weight given by observation $i$ to observation $l$ when estimating $i$’s missing counterfactual. In other words, the missing counterfactual for treatment (control) observations is given by a weighted average of the outcomes of control (treatment) observations. Upon estimation of the missing counterfactuals,

$$
\tilde{\tau}_i = \begin{cases} 
Y_i(1) - \hat{Y}_i(0) & \text{if } T_i = 1 \\
\hat{Y}_i(1) - Y_i(0) & \text{if } T_i = 0 
\end{cases}
$$

and the ATE, ATT, and ATU are estimated by the sample average of $\tilde{\tau}$ in the full sample, treatment group, and control group respectively.

Different matching estimators utilize different weighting schemes. Single nearest neighbor (SNN) matching sets $\omega_{il} = 1$ for the observation with the closest propensity score to $P(X_i)$ but with $T_i = 1 - T_i$ and zero otherwise. SNN is typically performed with replacement (i.e., an observation $l$ may be matched with observations $i$ and $i'$ if it is the closest match to both). SNN minimizes the bias due to inexact matches (since it only uses the closest possible match), but does so at the expense of greater variance (since it uses only one match). Alternative weighting schemes utilize more observations to reduce variance despite the introduction of additional bias.

$K$ nearest neighbor (KNN) matching sets $\omega_{il} = 1/K$ for the $K$ observations with the closest propensity score to $P(X_i)$ but with $T_i = 1 - T_i$ and zero otherwise, where $K$ is some finite integer. As opposed to fixing the number of matches regardless of difference in the propensity scores, $|\bar{P}(X_i) - \bar{P}(X_l)|$, radius matching fixes the maximum allowable difference in propensity scores and utilizes all matches with the specified radius (or caliper). Formally, $\omega_{il} = 1/K_i$ if $|\bar{P}(X_i) - \bar{P}(X_l)| \leq \kappa$ but with $T_i = 1 - T_i$ and zero otherwise, where $K_i$ is the number of observations with $T \neq T_i$ and $\kappa$, the caliper, is a real number inside the unit interval.

The stratification estimator is typically not thought of as a matching estimator, although it can be derived using a particular weight scheme. The stratification estimator partitions the sample into groups (or strata) based on different ranges of the propensity score (e.g., ten strata based on deciles of the propensity score distribution). Within each strata, the mean difference in outcomes across the treatment and control groups is estimated. Finally, the average treatment effect parameters in (2) are obtained as a weighted average of the within-strata average treatment effects. Specifically, for the ATE, one averages the within-strata average treatment effects using the population proportion located within each strata (estimated by the sample counterparts). For the ATT (ATU), one averages over the proportion of treated (control) observations located within each strata (estimated by the sample counterparts).

Equivalently, one may implement the stratification estimator as follows. Partition the sample in $K$ strata based on the propensity score. Define $D_{ki} = 1$ if $P(X_i)$ lies in strata $k$, $k = 1, \ldots, K$, and zero
otherwise. Estimate the following model via OLS:

$$Y_i = \sum_{k=1}^{K} \beta_k D_{ki} + \sum_{k=1}^{K} \alpha_k T_i D_{ki} + \nu_i.$$  \hfill(12)

The stratification estimate of the ATE is given by \( \hat{\tau}_{ATE} = \sum_{k=1}^{K} (N_k/N) \hat{\alpha}_k \), where \( N_k \) is the sample size of strata \( k \) and \( N \) is the total sample size. The estimate of the ATT is given by \( \hat{\tau}_{ATT} = \sum_{k=1}^{K} (N_{1k}/N_1) \hat{\alpha}_k \), where \( N_{1k} \) is the number of treated observations in strata \( k \) and \( N_1 \) is the total sample size of the treatment group. The ATU is estimated analogously using a similarly defined \( N_{0k} \) and \( N_0 \). Finally, the stratification estimator can be obtained as a matching estimator setting \( \omega_{il} = 1 \) for observations in the same strata as \( i \) but with \( T_l = 1 - T_i \), where \( N_i \) is the number of observations in same strata as \( i \) but with \( T \neq T_i \).

Kernel matching (KM) potentially differs from all the previous estimators in that it may give – depending on the kernel – positive weight to all observations with \( T_l = T_i \). Specifically, under KM, the weights have the form

$$\omega_{il} = \max \left\{ 0, \frac{G\left(\frac{P(X_i) - P(X_l)}{a_N}\right)}{\sum_{l' \in \{T_l = 1 - T_i\}} G\left(\frac{P(X_{l'}) - P(X_l)}{a_N}\right)} \right\}$$ \hfill(13)

where \( G(\bullet) \) is the kernel function and \( a_N \) is the bandwidth. In practice, choice of the bandwidth is more important than choice of kernel.

The next set of estimators also rely solely on the propensity score and are known as inverse propensity score weighting (IPW) estimators. The IPW estimator is derived from the fact that the ATE may be expressed as

$$\tau_{ATE} = \mathbb{E}\left[ \frac{Y \cdot T}{P(X)} - \frac{Y \cdot (1 - T)}{1 - P(X)} \right],$$

with the sample analogue given by

$$\hat{\tau}_{ATE} = \frac{1}{N} \sum_{i=1}^{N} \left[ \frac{Y_i T_i}{P(X_i)} - \frac{Y_i (1 - T_i)}{1 - P(X_i)} \right].$$ \hfill(14)

Similarly, the ATT and ATU are given by

$$\hat{\tau}_{ATT} = \frac{1}{\sum_{i=1}^{N} P(X_i)} \sum_{i=1}^{N} \hat{P}(X_i) \left[ \frac{Y_i T_i}{\hat{P}(X_i)} - \frac{Y_i (1 - T_i)}{1 - \hat{P}(X_i)} \right]$$ \hfill(15)

$$\hat{\tau}_{ATU} = \frac{1}{\sum_{i=1}^{N} [1 - \hat{P}(X_i)]} \sum_{i=1}^{N} [1 - \hat{P}(X_i)] \left[ \frac{Y_i T_i}{P(X_i)} - \frac{Y_i (1 - T_i)}{1 - P(X_i)} \right].$$ \hfill(16)
The estimator in (14) is the unnormalized estimator as the weights do not necessarily sum to unity (only in expectation). To circumvent this issue, a recently proposed alternative estimator, referred to as the normalized estimator, guarantees that the weights sum to one within the treatment and control groups. For the ATE, for instance, the estimator is given by

\[ \hat{\tau}_{ATE} = \left[ \sum_{i=1}^{N} \frac{Y_i T_i}{\hat{P}(X_i)} \right] \left/ \left[ \sum_{i=1}^{N} \frac{T_i}{\hat{P}(X_i)} \right] \right. - \left[ \sum_{i=1}^{N} \frac{Y_i (1 - T_i)}{1 - \hat{P}(X_i)} \right] \left/ \left[ \sum_{i=1}^{N} \frac{(1 - T_i)}{1 - \hat{P}(X_i)} \right] \right. \]

(17)

The literature provides evidence of the superiority of the normalized estimator. In practice, to avoid excessively large weights being given to particular observations with propensity scores close to the boundary, the sample is typically trimmed by including only observations with propensity scores away from zero and one (e.g., retaining observations with \( \hat{P}(X_i) \in [0.02, 0.98] \)).

The final set of estimators make use of both OLS and the propensity score and are known as doubly robust (DR) estimators. Estimators of this type specify functional forms for the potential outcomes and the propensity score model, but are consistent as long as either the outcomes or the propensity score is correctly specified. One such estimator combines weighting and regression by estimating (7) via weighted least squares, where the weights are \( [T_i/\hat{P}(X_i)] + [(1-T_i)/(1-\hat{P}(X_i))] \) for the ATE, \( T_i + [(1-T_i)\hat{P}(X_i)/(1-\hat{P}(X_i))] \) for the ATT, and \( T_i [1 - \hat{P}(X_i)]/\hat{P}(X_i)] + [(1 - T_i)] \) for the ATU. An alternative estimator is based on OLS estimation of the following specification

\[ Y_i = \alpha_0 + (\alpha_1 - \alpha_0)T_i + X_i \beta_0 + T_i X_i (\beta_1 - \beta_0) + \theta_0 (1 - T_i) \{1 - \hat{P}(X_i)\}^{-1} + \theta_1 T_i \{\hat{P}(X_i)\}^{-1} + \xi_i. \]

(18)

The DR estimate of the ATE is given by \( \hat{\tau}_{ATE} = \alpha_1 - \alpha_0 + \bar{X} (\beta_1 - \beta_0) + (1/N) \sum_i \hat{\theta}_1 \{\hat{P}(X_i)\}^{-1} - \hat{\theta}_0 \{1 - \hat{P}(X_i)\}^{-1} \). The ATT (ATU) is derived in a similar fashion except that \( \bar{X} \) is replaced by the average in the treatment (control) group and the summation is restricted to observations in the treatment (control) group.

In sum, the methods discussed here are potentially useful for answering a number of salient research questions concerning the causal effect of political institutions on environmental policy since institutional arrangements are typically measured as a binary treatment. However, since researchers must rely on observational data to investigate such questions, it is imperative that researchers understand the assumptions required for different estimators to provide unbiased estimates of different treatment effect parameters, as well as how to precisely interpret the parameters being estimated. In particular, it is paramount that researchers understand that estimators relying on the propensity score are not a panacea; such estimators still require unconfoundedness and thus do not address concerns over selection on unobservables. Moreover,
the apparent benefit of matching and weighting estimators – the fact that they are semiparametric in that no functional form for the potential outcomes need be assumed – is a bit of a mirage in that the functional forms assumed in Assumption (2) can be made arbitrarily flexible by incorporating higher order and interaction terms in $X$. Nonetheless, comparison of the propensity score distributions across the treatment and control groups is useful in highlighting failures of the common support, indicating that regression-based methods must extrapolate into empty regions in order to estimate at least some missing counterfactuals.

### 2.2.3 Selection on Unobservables Estimators

If one is concerned that Assumption (1) does not hold, the traditional response is to turn to instrumental variables (IV). However, researchers actually have at their disposal a wide array of estimation options from which to proceed. Even more so than among the estimators discussed in the Section 2.2.2, these different techniques rely on markedly different assumptions, entail different data requirements, and differ in what questions they can answer.

The first set of methods do not yield point estimates of the causal effect of the treatment in question. Instead, these techniques begin by obtaining an estimate of some average treatment effect parameter under Assumption (1) and then assess the robustness of this estimate to varying degrees of selection on unobservables. While the exact amount of selection on unobservables is unknown, if only a modest amount of selection on unobservables is needed to explain the average treatment effect estimated using a method requiring Assumption (1), the causal effect is said to not be robust. Conversely, if a large amount of selection on unobservables is needed to fully explain the average treatment effect obtained using a method requiring Assumption (1), and the researcher believes that such a large amount is not plausible, then the causal effect is said to be robust.

There are two important points to note. First, absent additional assumptions or data, the researcher cannot be certain of the extent (or even the existence) of any unobservables that invalidate the unconfoundedness assumption. Thus, these methods require either additional assumptions to pin down the amount of selection on unobservables (e.g., joint normality or equal amounts of selection on observables and unobservables) or an appeal to persuasion concerning the likely strength of any selection on unobservables. Second, even if the estimated effect obtained under unconfoundedness is determined to be robust, the magnitude of the causal effect is unknown. For example, if the impact of democracy on the probability of signing an international environmental agreement is estimated to be ten percent assuming no selection on unobservables and this is found to be robust to a large amount of selection on unobservables, one should not conclude that the average causal effect is ten percent. Rather, one can only conclude that it is not zero. An (unknown) portion of the ten percent effect may still be attributable to unobservables correlated
with both the probability of being democratic and entering into environmental agreements. Consequently, the trade-off when using these methods is that the assumptions are weaker than those (discussed below) required for point identification of some average treatment effect parameter, but one foregoes estimation of the magnitude of any robust causal effect.

In the interest of brevity, I do not discuss the details of empirical methods falling into this class. The most popular such method, however, is the so-called Rosenbaum bounds, applied after matching estimators.

The next set of estimators are applicable when the researcher has access to panel data and is willing to assume that the only unobservables responsible for the invalidation of Assumption (1) are invariant along one dimension of the panel. Typically, panel data involves a time dimension, but it is important to remember that this need not be the case. For simplicity, I will assume the usual data structure with a time dimension.

Resulting estimators are referred to as difference-in-differences (DD) estimators as they entail a comparison of the changes in outcomes across the treatment and control groups. The most frequent way of implement DD estimators is via fixed effects (FE) estimation. Application requires several assumptions.

**Assumption 4** Potential outcomes and latent treatment assignment are additively separable in observables and unobservables, with unobservables being decomposed into time invariant and time-varying components.

\[
Y_{it}(0) = \alpha_0 + \lambda_t + X_{it}\beta_0 + c_i + \varepsilon_{0it} \\
Y_{it}(1) = \alpha_1 + \lambda_t + X_{it}\beta_1 + c_i + \varepsilon_{1it} \\
T_{it} = \mathbb{I}[X_{it}\gamma + c_i + \eta_{it} > 0]
\]

where \(i = 1, \ldots, N, t = 1, \ldots, T, c\) is an observation-specific unobserved effect, \(\lambda\) is a period-specific unobserved effect, and \(\mathbb{I}[]\) is the indicator function. \(\varepsilon_0, \varepsilon_1, \eta \sim \mathcal{N}_3(0, \Sigma)\), where

\[
\Sigma = \begin{bmatrix}
\sigma_0^2 & \rho & 0 \\
\rho & \sigma_1^2 & 0 \\
0 & 0 & 1
\end{bmatrix}
\]

Implicit in Assumption (4) is the fact that unobservables captured by \(c\) have identical effects on both potential outcomes, and that both potential outcomes follow the same time trend (conditional on \(X\) and

---

1 More complex estimators, such as difference-in-difference-in-differences (DDD), are also utilized in certain contexts. It is important to note, however, that just because DDD contains an extra ‘D’ does not necessarily imply it relies on more palatable assumptions. Similarly, it is not necessarily the case that DD is always to be preferred over (pooled) OLS. The salient point is that each estimator relies on a different set of assumptions and it is incumbent upon the researcher to utilize the estimator that relies on the most plausible assumptions concerning a particular application.
Under Assumption (4), observed outcomes are given by

\[ Y_{it} = \alpha_0 + \lambda_t + (\alpha_1 - \alpha_0)T_{it} + X_{it}\beta_0 + T_{it}X_{it}(\beta_1 - \beta_0) + [c_i + \epsilon_{0it} + (\epsilon_{1it} - \epsilon_{0it})T_{it}] \]  

and the average treatment effect parameters continue to be reflected in (6). However, since the covariance between \( T_i \) and \( c_i \) is non-zero even conditional on \( X_i \), OLS estimation of (19) is biased and inconsistent.

To proceed, the model is transformed in order to eliminate \( c_i \) from the estimating equation. The fixed effects transformation entails subtracting \( Y_{it}(1/T) \sum_{t=1}^{T} Y_{it} \) from both sides. The FE transformation is equivalent to the Least Squares Dummy Variable (LSDV) approach that treats the \( c_i \) as parameters to be estimated. Equation (19) is estimated via OLS with the \( c_i \) replaced with \( N \) observation-specific dummy variables. Alternative transformations to the FE estimator are also available. First-differencing entails subtracting \( Y_{it-1} \) from each side of (19), while long-differencing entails subtracting \( Y_{it-s} \) from each side, where \( s \) is some integer greater than one.

Using the FE transformation, the model

\[ Y_{it} - \bar{Y}_i \equiv \tilde{Y}_{it} = \tilde{\lambda}_t + (\alpha_1 - \alpha_0)\tilde{T}_{it} + \tilde{X}_{it}\tilde{\beta}_0 + (\tilde{T}_{it}\tilde{X}_{it})(\beta_1 - \beta_0) + \tilde{\epsilon}_{it} \]  

is then estimated via OLS, where \( \tilde{\bullet} \) over a variable indicates deviations from observation-specific means, \( \lambda_t \) are estimated using period-specific dummy variables, and \( \tilde{\epsilon}_{it} \equiv \epsilon_{0it} + (\epsilon_{1it} - \epsilon_{0it})T_{it} \). For OLS estimation of (20) to yield unbiased estimates, the following two assumptions are needed.

**Assumption 5** The observable covariates, \( X \), are strictly exogenous. Formally,

\[ \mathbb{E}[\tilde{\epsilon}_{it}|Q_i, c_i] = 0 \quad \forall t = 1, ..., T \]

where \( Q_i = [T_{i1} X_{i1} T_{i1}X_{i1} \cdots T_{iT} X_{iT} T_{iT}X_{iT}] \).

**Assumption 6** The data are of full rank. Formally,

\[ \text{rank } \mathbb{E}[\tilde{Q}_i\tilde{Q}_i'] = K \]

where \( K \) is the number of covariates in \( Q \).

There are three important implications of Assumptions (4) – (6). First, strict exogeneity requires that the treatment, \( T \), and the covariates, \( X \), be uncorrelated with the time-varying errors in every time
period. This precludes situations where, say, current political institutions respond to past (unobserved) idiosyncratic shocks (so-called Ashenfelter’s Dip). Strict exogeneity is a strong assumption. Second, the rank condition rules out observable covariates that do not vary over time; it also requires that the treatment itself varies over time (which may be problematic for questions involving many political institutions). If one assumes that such time invariant covariates have identical effects on both potential outcomes (i.e., \( \beta_1 = \beta_0 \) for these observables), then this is not problematic for estimating the average treatment effect parameters in (6); these covariates simply drop out of (20). However, absent this assumption, then average treatment effect parameters are no longer identified. Finally, DD estimators identify the causal effect of the treatment from one-time changes in outcomes that occur at the time observations switch from the treatment to the control group (or vice versa). As such, it is imperative that the timing of the treatment be measured correctly and that the treatment does not alter environmental policy before (due to anticipatory effects) or after (due to lagged effects) the measured date of the treatment. In the context of analyzing political institutions, this point is particularly salient since institutional changes are not likely to be unexpected.

An alternative to regression-based DD estimation is DD matching and weighting estimators. Replacing Assumption (4), one needs a revised conditional independence assumption.

**Assumption 7** \( \Delta Y(0), \Delta Y(1) \perp T \mid X \), where \( X \) is set of pre-determined covariates and \( \Delta \) represents the change in a variable from an earlier time period.

Under Assumptions (3) and (7), matching and IPW estimators from Section 2.2.3 may be employed. The only difference is that now the outcome reflects the change in \( Y \) from an earlier time period. In most applications of DD matching or weighting, the data contain only two time periods, where no observations are treated in the first period and a subset of observations are treated in the second period. If the data do not conform to this setup, care must be taken to properly utilize DD matching or weighting.

The remaining estimators I will discuss are applicable to situations where panel data are unavailable or time-varying unobservables invalidate Assumption (7). The first technique is referred to as nonparametric bounds. This method abandons point estimation of the average causal effect and instead focuses on bounding the effect under minimal assumptions. The bounds are potentially informative if (i) they do not include zero (thus, the treatment has a definite impact even though the magnitude is not known), or (ii) they exclude certain values that are particularly interesting in context of the specific application. I briefly illustrate the idea of nonparametric bounds for the ATE below.
To proceed, $\tau_{ATE}$ is decomposed using the law of total probability

\[
\tau_{ATE} = \mathbb{E}[Y(1) - Y(0)]
\]

\[
= \{\mathbb{E}[Y(1)|T = 1] \Pr(T = 1) + \mathbb{E}[Y(1)|T = 0] \Pr(T = 0)\} \\
- \{\mathbb{E}[Y(0)|T = 1] \Pr(T = 1) + \mathbb{E}[Y(0)|T = 0] \Pr(T = 0)\}
\]

\[
= \{g_1 - \mathbb{E}[Y(0)|T = 1]\} \Pr(T = 1) + \{\mathbb{E}[Y(1)|T = 0] - g_0\} \Pr(T = 0)
\]

where $\Pr(T = 1)$, $\Pr(T = 0)$, and $g_j$, $j = 0, 1$, are all estimable from the data. Thus, only $\mathbb{E}[Y(0)|T = 1]$ and $\mathbb{E}[Y(1)|T = 0]$ are unknown. The intuition behind this approach, then, is to bound the feasible values of $\tau_{ATE}$ by replacing these missing components with their maximum and minimum values. For example, if the outcome, $Y$, is binary, then the upper bound is obtained by setting $\mathbb{E}[Y(0)|T = 1] = 0$ and $\mathbb{E}[Y(1)|T = 0] = 1$; the lower bound is found by flipping the values of each expectation. Intuitively, the upper bound is found by painting the treatment in the most favorable light: assume all treated observations would have experienced $Y = 0$ in the absence of the treatment and all control observations would have experienced $Y = 1$ with the treatment. The lower bound assumes the reverse, thereby painting the treatment in the least favorable light.

While interesting, this approach may be of limited use absent further assumptions since the bounds just described necessarily include zero. Thus, one will never be able to rule out the possibility of no treatment effect. Nonetheless, this approach may still be of use if the goal is to obtain a maximum possible benefit of a particular treatment in order to conduct a cost-benefit or other welfare-type analysis.

However, the bounding approach does not stop here. The researcher may invoke (hopefully) modest assumptions in order to tighten (reduce the width of) the bounds and potentially exclude zero (and additional extreme values). While omitting the details, examples of such assumptions include level-set restrictions (whereby the bounds are estimated on different subsets of the data and then $\tau_{ATE}$ is assumed to be constant across these subgroups), monotone treatment selection (whereby both potential outcomes are assumed to be at least as high in expectation for the treatment group as the control group), monotone treatment response (whereby the expected outcome under the treatment is assumed to be at least as great as the expected outcome absent the treatment), and monotone instrumental variable (whereby expected potential outcomes are assumed to be monotonically increasing in the value of some variable). Often combinations of these assumptions are invoked in order to assess their power in narrowing the bounds.

IV is the final estimation technique I will discuss. IV is the most frequently method utilized by applied researchers concerned with the presence of unobservables correlated with both treatment assignment and potential outcomes conditional on $X$. While the advantage of IV relative to nonparametric bounds is that
it yields a point estimate rather than an interval of possible values, it requires several strong assumptions and careful interpretation. To illustrate, consider a binary instrument. Let $T_i(Z)$ denote the potential treatment assignment of observation $i$ under instrument realization $Z$, $Z \in \{0, 1\}$. Under the following assumptions

**Assumption 8** $Y(1), Y(0), T(1), T(0) \perp Z \mid X$

**Assumption 9** $\Pr(T_i(1) = 1 \mid X) > \Pr(T_i(0) = 1 \mid X)$

**Assumption 10** $\Pr(T_i(1) \geq T_i(0) \mid X) = 1$,

IV estimates a new causal parameter referred to as the local average treatment effect (LATE). The LATE is expressed formally as $E[\tau_i \mid T_i(1) > T_i(0)]$. In words, this represents the expected treatment effect for an observation randomly selected from the subpopulation of so-called compliers, defined as observations whose treatment assignment is determined by the instrument.

Prior to examining the LATE, examine Assumptions (8) – (10). Assumption (8) ensures that the instrument is randomly assigned conditional on $X$. In particular, $Z$ is conditionally independent of potential treatment assignments and $Z$ represents a valid exclusion restriction in that it is conditionally independent of potential outcomes. Assumption (9) guarantees that $Z$ is correlated with treatment assignment; the probability of receiving the treatment is greater when $Z = 1$ than when $Z = 0$. Finally, Assumption (10) is referred to as the monotonicity assumption. It rules out the possibility of so-called defiers by assuming that the instrument only manipulates behavior in one direction; there are no observations who would receive the treatment when $Z = 0$ but not receive the treatment when $Z = 1$. These are strong and, except for Assumption (9), untestable assumptions. Finally, note that actual estimation requires either a functional form assumption like Assumption (2) or a common support condition, $\Pr(Z = 1 \mid X) \in (0, 1)$. However, in light of the selection on unobservables, an appropriate functional form assumption for IV estimation relaxes the assumption of independence between $\varepsilon_0$, $\varepsilon_1$, and $\eta$. The instrument, $Z$, also now appears in the equation for treatment assignment.

Despite the popularity of IV estimators among applied researchers, there are many details that are often overlooked in practice. First, consider the LATE in more detail. While the LATE is a well-defined parameter, it is unusual in that the subpopulation of compliers is not known and potentially varies from instrument to instrument. Specifically, it is not possible to distinguish compliers in the data from so-called always-takers and never-takers. Always-takers are observations that receive the treatment regardless of the value of the instrument; i.e., $T(1) = T(0) = 1$. Never-takers are observations that do not receive the treatment regardless of the value of the instrument; i.e., $T(1) = T(0) = 0$. Compliers are defined as those
with $T(1) = 1$ and $T(0) = 0$. Thus, observations with $Z = 1$ and $T = 1$ ($Z = 0$ and $T = 0$) represent a mix of compliers and always-takers (never-takers).

In light of this, researchers must carefully consider who the complier population represents, as well as be clear on the internal versus external validity of the LATE. Specifically, while the LATE does provide a consistent estimate under the preceding assumption of the average treatment effect for the subpopulation of compliers associated with the particular instrument (internal validity), there is no guarantee that this parameters provides any guidance concerning the expected treatment effect in other subpopulations (external validity). Moreover, this distinction also means that researchers must be cautious comparing IV estimates based on alternative instruments as these do not necessarily estimate the same parameter.

Second, a lengthy literature exists documenting the pitfalls associated with invalid instruments. In particular, the poor performance of IV estimators based on weak instruments – Assumption (9) holds, but only weakly – is now well known. As a result, testing the strength of the first-stage relationship between $Z$ and $T$ and overidentifying restrictions (if the $Z$ includes more than one variable) is crucial. Utilizing weak instrument robust inference is preferable if strong identification is not feasible. Third, a number of IV estimators are available in practice, including Two-Stage Least Squares, Generalized Method of Moments, Limited Information Maximum Likelihood, Jackknife IV, among others. There are some useful results in the literature suggesting the use of particular estimators in different contexts. Fourth, while researchers are typically concerned solely with the causal effect of $T$ on $Y$, it is imperative that the decision over which variables to include in $X$ not be taken lightly; $X$ must continue to be pre-determined. If any of the elements in $X$ are correlated with unobservables impacting potential outcomes and correlated with $Z$, then IV estimators will no longer provide a consistent estimate of the LATE.

Finally, while valid instruments are often difficult to find, procedures have recently been proposed to partially address situations of imperfect or only plausibly exogenous instruments. Moreover, there are many atypical means of achieving identification without utilizing an instrument as usually conceived. Fuzzy regression discontinuity designs, higher moments estimators, and identification through heteroskedasticity are three examples.

### 2.2.4 Measurement Error

Before concluding this section, it is crucial to spend a moment pondering an often overlooked issue: measurement error. In the context of political institutions and environmental policy, measurement should be a concern. While many institutions may be clearly measured, others are less clear. For example, even the characterization of countries as democratic or not entails subjective choices, especially considering the timing of any changes within a country. Environmental policy outcomes are also fraught with measure-
ment difficulties. Typically, proxies are used in place of difficult-to-measure outcomes. Finally, many of the controls, $X$, included in analyses reflect attributes that are notoriously mismeasured. In light of this, measurement error may enter the empirical methods discussed here via four routes: the outcome, treatment assignment, the covariates, and/or the instrument.

In terms of the outcome, classical measurement error – mean zero, independent and identically normally distributed – leads only to a loss in efficiency. However, classical measurement error is a strong assumption. There are many examples of measurement error in outcomes – at the individual or the aggregate level – being mean-reverting (i.e., negatively correlated with the true outcome) or correlated with $T$ or $X$. This can introduce substantial bias in the estimation of treatment effects.

In terms of treatment assignment and the covariates, even classical measurement error in a regression context leads to biased estimates of the coefficients on the mismeasured variables. More importantly, perhaps, even if only $X$ is measured with error, this will also bias the estimate of the average treatment effect parameters if $T$ and $X$ are correlated. Nonclassical measurement only reinforces these results. Furthermore, note that measurement error in treatment assignment must be nonclassical; with a binary (or, in fact, any bounded) variable, measurement error must be negatively correlated with the true value.

While much less work has explored the effect of measurement error on some of the non-regression estimators discussed here (e.g., matching and weighting estimators), measurement error in $X$ will also bias these estimators (due to the inconsistency of the estimated propensity score model). Finally, to my knowledge, there exists no work on measurement error in an instrument. However, this should lead to biased estimates of the average treatment effects given the inconsistency of the first-stage estimates.

3 Spatial Interactions and Environmental Policy

3.1 Background

Political decisions are not made in a vacuum. Not only might environmental policy outcomes depend on factors internal to a political unit, such as the institutional arrangements discussed in Section 2, but external factors may also play a crucial role. These external factors may be decomposed into two groups: (i) exogenous attributes of other political units, and (ii) endogenous policy choices of other political units. Much of the theoretical and empirical research on political interactions focuses on the latter; I do the same. Not only are investigations into the role of political interactions useful for understanding the determination of policy, but they also shed light on the debate concerning environmental federalism. The theoretical literature concerning environmental federalism emphasizes the trade-off concerning local versus centralized control over environmental policy. Local control allows for the tailoring of policy to local preferences,
whereas centralized policy tends to emphasize uniformity across space. However, local control may lead to inefficiencies due to the failure to internalize externalities (e.g., transboundary pollution) or excessive competition for resources (the so-called race-to-the-bottom). Thus, understanding the magnitude of political interactions is crucial for measuring the relative merits of decentralized environmental policymaking.

There are several theoretical reasons why one might expect environmental policy in one location to depend on the policy choices of other locations. In the spillover model, locations are intertwined due to transboundary pollution. As such, the welfare in one jurisdiction depends not only on its own policy choice, but also on the choices of other jurisdictions as these external decisions affect the amount of transboundary pollution. Thus, policy choices become linked across jurisdictions. The model of yardstick competition can also be thought of as a spillover model. Here, policymakers are evaluated through a comparison of outcomes in the electorate’s own jurisdiction to outcomes in other jurisdictions. Such comparisons generate informational spillovers across jurisdictions and create an alternative explanation for policy linkages across political units. Finally, in the resource competition model, jurisdictions compete for mobile resources through the setting of policy. If these mobile resources (e.g., capital or labor) impact welfare (either positively or negatively), then the welfare of one jurisdiction depends not just on its own policy choice, but also on the choices of competing jurisdictions. Again, policy choices become linked across jurisdictions.

3.2 Empirical Methods

As the preceding background discussion makes clear, researchers seeking to empirically assess the importance of political interactions in the determination of environmental policy are interested in the causal effect of external environmental policy on policy choices. The empirical difficulty in such analyses is that, absent strong assumptions, the joint determination of policy across political units generates a simultaneity problem. Fortunately, the literature concerned with spatial statistical and econometric methods has experienced similar growth to the literature on treatment effects.

The empirical model used to test for policy interactions between jurisdictions – assuming a linear approximation to the underlying policy reaction function – takes the general form

\[ Y_i = \delta \sum_{j \neq i} \omega_{ij} Y_j + X_i \beta + \varepsilon_i \]  

(21)

where \( Y_i \) is a measure of environmental policy in jurisdiction \( i \), \( \omega_{ij} \) is the weight assigned to jurisdiction \( j \) by \( i \), \( Y_j \) is the measure of environmental policy in \( j \), \( \delta \) is the parameter of interest, \( X_i \) is a vector of covariates, and \( \varepsilon_i \) is the error term. If \( \delta \neq 0 \), then this indicates the presence of policy spillovers.

There are two issues to be addressed when estimating an equation such as (21). The first issue is
the choice of weights, \( \omega \), which must be specified by the researcher. As such, the choice is often ad hoc. Common examples are to define the weights such that they are declining in the distance between \( i \) and \( j \), or simply to give positive weight to contiguous neighbors. However, depending on the context, the relevant spillovers or resource competition may entail certain jurisdictions regardless of distance. The arbitrary nature of the choice of weights has several implications. First, researchers ought to assess the robustness of the empirical results to alternative weighting schemes. Second, any weighting scheme that differs from the ‘true’ weighting scheme induces measurement error in the so-called spatial lag. Thus, the estimate of \( \delta \) will most likely suffer from attenuation bias. This means that estimates of \( \delta \) close to zero do not necessarily imply the absence of political interactions; rather, one could simply be ‘looking in the wrong place’ due to a poor choice of weighting schemes. Similarly, if estimates of \( \delta \) are statistically significant for certain weighting schemes, then one can be confident that political interactions do exist. In fact, other weighting schemes may even yield stronger evidence of interactions.

The second issue of concern in the estimation of (21) pertains to the potential endogeneity of the environmental policy of other states. In the presence of political interactions, absent some strong assumptions, jurisdictions are choosing their environmental policies simultaneously and incorporating their expectations about the decisions of other political units into their own decision-making process. This may give rise to concerns about the direction of causation. In addition, there may be unobservable determinants of environmental policy that are correlated with the decisions of multiple jurisdictions (known as spatial error dependence). To address these concerns, the model in (21) is either estimated by IV or maximum likelihood (ML).

To apply IV, typically instruments of the type \( \sum_{j \neq i} \omega_{ij} X_j \) are utilized, where the same weighting scheme is utilized for the creation of the instruments as in (21). IV yields consistent estimates of \( \delta \), even if \( \varepsilon \) is spatially correlated, as long as these instruments are valid. Alternatively, ML estimation may be used. To proceed, the model in (21) may be written in matrix form as

\[
Y = \delta W Y + X \beta + \varepsilon
\]  

(22)

where \( Y \) and \( \varepsilon \) are \( N \times 1 \) vectors, \( W \) is a \( N \times N \) weight matrix (with zeros along the diagonal and representative element \( \omega_{ij} \)), and \( X \) is an \( N \times K \) matrix. With spatial error dependence, \( \varepsilon = \rho W \varepsilon + u \) where \( u \) is a \( N \times 1 \) vector and \( \rho \) is the spatial error correlation. To derive the likelihood function, note that (22) may be solved for \( Y \), yielding

\[
Y = (I - \delta W)^{-1} X \beta + (I - \delta W)^{-1} (I - \rho W)^{-1} u
\]  

(23)
where $I$ is an $N \times N$ identity matrix. While complex, (23) may be estimated using standard nonlinear optimization methods.

4 Conclusion

This survey provides a brief introduction to some of the empirical methods useful in addressing research questions within the political economy arena of environmental policy. As discussed, many empirical questions addressed in this area pertain to the causal effect of a binary variable reflecting institutional arrangements on environmental policy, or the causal effect of environmental policy in neighboring jurisdictions on own environmental policy. In the former case, the literature on estimating treatment effects is profoundly helpful. In the latter case, the spatial econometric literature is relied upon.

Further Reading


