Forest Policy and Economics 83 (2017) 142–145

Contents lists available at ScienceDirect

Forest Policy and Economics

journal homepage: www.elsevier.com/locate/forpol

Research trends: Policy impact evaluation: Future contributions from economics

Jing Lan[a], Runsheng Yin[b,c,*]

[a] College of Public Administration, Nanjing Agricultural University, Nanjing 210000, China
[b] College of Economics and Management, Zhejiang A & F University, Lin'an, 311300, China
[c] Department of Forestry, Michigan State University, East Lansing, MI 48824, USA

ARTICLE INFO

Keywords:
Impact evaluation
Forest policy
Difference in differences
Propensity score matching
Regression discontinuity
Instrumental variable method
Panel data models

ABSTRACT

Economic research has become more empirical, and much of the shift has been due to the broad applications of impact evaluation (or IE) methods for measuring the potential effects of policies or programs on outcomes of interest. The objective of this article is to introduce the essential IE methodology and literature to the audience of this journal and discuss how we can take advantage of these research developments to improve the quality of our own work and subsequent publications. Following an overview of the core IE concepts and methods, we will give an overview of the recent uses of these methods to address forest policy issues and outline the basic steps of sampling and data generation involved in an IE study.

1. Introduction

Economic research has become more empirical (Hamermesh, 2013), and much of the shift has had to do with the widespread applications of impact evaluation (or IE) methods for measuring the potential effects of policies (or programs) on outcomes of interest, especially at the micro-level (Athey and Imbens, 2017; Angrist and Pischke, 2017). As a result, evidence-based policymaking has been taken as a mantra of the day (Khandker et al., 2010). It appears, though, that this major trend in economics and the relevant econometric methods are still not well known or accessed by many forest policy analysts.[1] Therefore, the objective of this article is to introduce the essential methodology and literature of policy IE to the audience of this journal and discuss how we can take advantage of these research developments to improve the quality of our own work and subsequent publications.

The paper is organized as follows. An overview of the core IE concepts and methods will be given in the next two sections. Then, a summary of the recent applications of these methods to forest policy issues will be presented in Section 4. Section 5 will outline the basic steps of sampling and data generation involved in IE work. A few closing remarks will follow in Section 6.

2. The basic concepts

As summarized by Athey and Imbens (2017):

"*The gold standard for drawing inferences about the effect of a policy is a randomized controlled experiment. However, in many cases, experiments remain difficult to implement, for financial, political, or ethical reasons, or because the population of interest is too small… Thus, a large share of the empirical work in economics about policy questions relies on observational data—that is, data where policies were determined in a way other than through random assignment*" (p. 3).

Empirical work based on observational data, however, has faced challenges. Included in them are: First, how we can properly identify the counterfactual—what would have happened to the treated group of individuals (or participants) if a policy had not been adopted; second, how we can remove or at least control the confounding factor(s) that may induce correlation between the policy under evaluation and the outcome that is not indicative of what would happen if the policy could be adopted (Imbens and Rubin, 2015). Unfortunately, what would have happened to the treated group if a policy had not been adopted is not observable. It is natural to deal with this so-called "missing data problem" by creating a convincing and reasonable comparison (or control) group that has not been treated to determine the policy impact on the treated group (Khandker et al., 2010)[2]; thus, those who received treatment would have had outcomes very similar to those in the

* Corresponding author at: Department of Forestry, Michigan State University, East Lansing, MI 48824, USA.
E-mail addresses: lanjing@njau.edu.cn (J. Lan), yinr@msu.edu (R. Yin).

[1] As an editor of this journal, I have seen that a lot of the submitted manuscripts were not based on this framework and the relevant methods.

[2] A treatment is also referred to as an intervention or a manipulation in different disciplinary fields.

http://dx.doi.org/10.1016/j.forpol.2017.07.009
Received 17 July 2017; Accepted 19 July 2017
Available online 29 July 2017
1389-9341/

comparison group in the absence of treatment.

|So, it is presumed that only one of the two potential outcomes, Y_i, is observed for each sample unit i in the basic formulation of an IE model. That is:

$$Y_i = Y_i(W_i) = \begin{cases} Y_i(0) & \text{if } W_i = 0 \\ Y_i(1) & \text{if } W_i = 1 \end{cases}$$

where $W_i \in \{0,1\}$ indicates the treatment received. Then, the average treatment effect for the treated, or *ATT*, is

$$ATT = \frac{1}{N} \sum_{i|W_i=1} \{Y_i(1) - Y_i(0)\}$$

where $N = \sum_i W_i$ the number of treated units, for which we only observe $Y_i(1)$. Dealing with this missing data problem has led to several techniques for measuring the counterfactual effect, $Y_i(0)$, to determine the *ATT* (Imbens and Rubin, 2015). Notably, each of these techniques carries its own assumptions about the nature of potential selection bias in policy targeting and participation, and the assumptions are crucial to developing the appropriate model to determine causal effects (Khandker et al., 2010). Selection bias is the bias introduced by the selection of individuals, groups, or data for analysis in such a way that proper randomization is not achieved, with the result that the sample obtained is not representative of the population intended for analysis (Angrist and Pischke, 2008).

3. Analytic methods

The modeling techniques include difference in differences, propensity score matching, panel data models, and regression discontinuity.

Difference in differences (or DID) is an often-used IE method, in which outcomes are observed for both the treated group ($W_i = 1$) and the control group ($W_i = 0$) for a pre-treatment period t_1 and a post-treatment period t_2. So,

$$y_{it} = \beta_0 + \beta_1 d2 + \beta_2 dT + \beta_3 d2dT + u_{it}$$

where y_{it} is outcome of interest, $dT = 1$ for treatment group ($W_i = 1$) and 0 otherwise, $d2 = 1$ for the post-treatment period (t_2) and 0 otherwise, $d2dT$ is the interaction term between $d2$ and dT, u_{it} is random error, and $\beta_1, \beta_2, \beta_3$ are parameters to be estimated. Specifically, β_0 is the average outcome for the control group prior to the treatment (t_1), β_1 captures changes in y caused by aggregate factors even in the absence of the treatment, and β_2 captures possible differences in outcomes between treatment and control groups prior to the treatment (t_1). The coefficient for the *ATT* is β_3, which is defined as

$$\widehat{\beta_3} = (\overline{y_{i2}}(1) - \overline{y_{i1}}(1) - (\overline{y_{i2}}(0) - \overline{y_{i1}}(0))$$

where $\overline{y_{i2}}(1)$ and $\overline{y_{i1}}(1)$ are the mean outcome for the treatment group in period t_2 and t_1, and $\overline{y_{i2}}(0)$ and $\overline{y_{i1}}(0)$ are the corresponding values for the control group.

In addition to clear-cut cases of treatment vs. control groups, we encounter situations of different treatment types/intensities in reality. The above DID estimator remains applicable in the latter case so long as a proper indicator of the treatment types/intensities can be formulated and the confounding factors taken into account (Finkelstein, 2007). One key assumption of the DID estimator is that the average outcomes for the treatment and control groups would follow parallel trends in the absence of the treatment (Khandker et al., 2010). If it does not hold, then the *ATT* estimate will be biased. Thus, alternative methods have been developed (Heckman et al., 1997).

Propensity score matching (PSM) strives to construct a statistically determined comparison group based on a model of the probability of participating in the treatment, using observed characteristics. Participants are matched on the basis of this probability, or propensity score, to nonparticipants. The *ATT* is then calculated as the mean difference in outcomes across the two groups. In addition to DID, there are several specific methods to determine the mean difference numerically (Khandker et al., 2010).

The validity of PSM assumes that unobserved factors do not affect participation and that there is sizable common support in propensity scores across the two groups (Imbens and Rubin, 2015). Note that although PSM can be deployed even if the analyst has only cross-sectional data—observations at a single point of time and use of a simple matching method other than PSM—this type of data and method, especially in a small sample, will make it less likely to identify a truly meaningful counterfactual and thus a reliable estimate of the policy impact.

Panel data models (PDMs) can be employed to explore the advantages of panel data by running the following model:

$$y_{it} = \lambda_t + \tau w_{it} + \gamma x_{it} + c_i + u_{it}, \quad t = 1, 2, \ldots, T$$

where $w_{it} = 1$ if unit i is treated at time t and 0 otherwise, x_{it} is a vector of covariates, c_i is an unobserved individual-level effect, λ_t accounts for aggregate time effects, and u_{it} are the idiosyncratic errors (Wooldridge, 2009). Coefficient τ is the average treatment effect. This model can be estimated with a fixed effects or first differencing method, provided that the treatment indicator w_{it} is exogenous, conditional on unobserved heterogeneity (Woodridge, 2010). If w_{it} is correlated with unit-specific trends, a correlated random trend model can be used.

The availability of panel data also allows for selection bias on unobserved characteristics to vary with time. Here, selection bias on unobserved characteristics can be corrected by using an instrumental variable (IV) method to find a variable (or instrument) that is correlated with participation but not so with unobserved characteristics affecting the outcome (Wooldridge, 2009). This instrument is then used to predict participation. Again, a proper indicator of the types/intensities of participation can be used as a treatment proxy in this setting, which allows the analyst to deal with issues like lag time and spatiotemporal correlations as well.

Regression discontinuity (RD) methods exploit discontinuities in policy rules (e.g., eligibility requirements) and thus incentives or ability to receive a discrete treatment. Participants and nonparticipants are compared in a close neighborhood around the eligibility cutoff. The boundary demarcation of a protected area and the public assistance to local community members below the official poverty line are just two of the many familiar examples.

To model the effect of a particular policy on individual outcomes y_i through an RD method, one needs a variable S_i that determines participation eligibility with an eligibility cutoff of s_*. The estimating equation is $y_i = \beta S_i + \varepsilon_i$, where individuals with $s_i \leq s_*$, for example, receive the treatment, and individuals with $s_i > s_*$ are not eligible to participate. Individuals in a narrow band above and below s_* need to be "comparable" in that they would be expected to achieve similar outcomes prior to the treatment.

4. Forest literature overview

There exists a long list of practical cases for which we can apply IE methods in measuring the potential impacts of forest policies. They include the establishment of a protected areas (PAs), the growth of ecotourism in a particular place, the retirement and restoration of degraded cropland, the execution of a firefighting plan, the implementation of a biofuel initiative, the devolution of forest tenure, and the adoption of an REDD + program….[3] In fact, IE methods have been applied in forestry contexts. Andam et al. (2008) stated that conventional methods of evaluating the effectiveness of PAs can be biased because protection is not randomly assigned and because protection can induce deforestation spillovers to neighboring forests. By adopting

[3] REDD + refers to reducing CO_2 emissions from deforestation and forest degradation and increasing carbon stocks by enhancing forest regeneration and regrowth.