Journal of Environmental Management 119 (2013) 220–229

Contents lists available at SciVerse ScienceDirect

Journal of Environmental Management

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

# Controlling for exogenous environmental variables when using data envelopment analysis for regional environmental assessments

Alexander J. Macpherson[a,1], Peter P. Principe[a,*], Yang Shao[b]

[a] Office of Research and Development (E243-05), United States Environmental Protection Agency, 109 T.W. Alexander Drive, Research Triangle Park, NC 27711, USA
[b] Department of Geography, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA

ARTICLE INFO

*Article history:*
Received 28 June 2012
Received in revised form
11 December 2012
Accepted 21 December 2012
Available online 12 March 2013

*Keywords:*
Data envelopment analysis
Environmental management
Natural resources
Exogenous variables
Environmental planning

ABSTRACT

Researchers are increasingly using data envelopment analysis (DEA) to examine the efficiency of environmental policies and resource allocations. An assumption of the basic DEA model is that decision-makers operate within homogeneous environments. But, this assumption is not valid when environmental performance is influenced by variables beyond managerial control. Understanding the influence of these variables is important to distinguish between characterizing environmental conditions and identifying opportunities to improve environmental performance. While environmental assessments often focus on characterizing conditions, the point of using DEA is to identify opportunities to improve environmental performance and thereby prevent (or rectify) an inefficient allocation of resources. We examine the role of exogenous variables such as climate, hydrology, and topography in producing environmental impacts such as deposition, runoff, invasive species, and forest fragmentation within the United States Mid-Atlantic region. We apply a four-stage procedure to adjust environmental impacts in a DEA model that seeks to minimize environmental impacts while obtaining given levels of socioeconomic outcomes. The approach creates a performance index that bundles multiple indicators while adjusting for variables that are outside management control, offering numerous advantages for environmental assessment.



## 1. Introduction

The environmental research community is increasingly interested in applying data envelopment analysis (DEA) to examine environmental problems (Zhou et al., 2008). DEA offers several advantages for examining environmental problems. DEA relies upon a flexible, nonparametric specification of production relationships, making no assumptions about the functional relationships among inputs and outputs. This flexibility is advantageous for environmental assessments when the relationships among model variables are not well understood. Some DEA models are invariant to the units used in the analysis, which is desirable when the environmental impacts of interest are measured in different units, or at different scales, that cannot be easily transformed into comparable quantities. The incomparability of units and scales has been shown to be problematic for the multicriteria decision tools that are often employed in environmental decisionmaking, because different conversion methods yield different values, which could lead to different management recommendations (Steele et al., 2009). Another common approach to solving a multi-objective problem that shares the problem of incomparability of units and scales is the aggregation of multiple outputs into a single objective function or index by assigning weights or relative scores to the outputs. While intuitively appealing, without complete preference information over multiple outputs, the analysts constructing the indices will inevitably be obliged to rely heavily upon expert, administrative, or political judgments, all of which will come freighted with errors or biases that will affect the relative values of the multiple outputs and, consequently, the value of the indices.

An important assumption of the DEA model is that decision-makers operate within homogenous environments. However, this assumption does not hold in environmental systems where the production of environmental benefits or negative impacts is influenced by variables beyond managerial control. For example, all else being equal, an area with higher precipitation is likely to see higher nitrogen wet deposition. Similarly, all else being equal, greater topographic variation within an area may lead to higher

* Corresponding author. Tel.: +1 919 541 1422; fax: +1 919 541 1138.
*E-mail addresses:* macpherson.alex@epa.gov (A.J. Macpherson), principe.peter@epa.gov (P.P. Principe), yshao@vt.edu (Y. Shao).
[1] Current address: Office of Air and Radiation (C439-02), United States Environmental Protection Agency, 109 T.W. Alexander Drive, Research Triangle Park, NC 27711, USA.

0301-4797/$ – see front matter Published by Elsevier Ltd.
http://dx.doi.org/10.1016/j.jenvman.2012.12.044

levels of erosion or stream runoff. Again, all else being equal, environmental managers in areas with higher precipitation or greater topographic variation may find it more difficult to mitigate impacts from deposition or runoff within their management unit, solely because of the precipitation and topography that were beyond their managerial control.

This example leads to a crucial point that in our experience is typically overlooked in environmental assessments: a distinction must be made between the overall condition of a management unit and the relative opportunity to improve the environmental performance of the management unit. Environmental assessments tend to focus on the former, which could lead to an inefficient allocation of resources directed toward environmental improvement. A more accurate accounting of variables beyond managerial control and how these variables affect performance will be critical to shift managerial attention from condition to opportunities for improvement.

This paper examines the role of several exogenous variables (*e.g.*, climate, hydrology, and topography) in altering environmental impacts such as deposition, runoff, the presence of invasive species, and forest fragmentation within 134 watersheds in the Mid-Atlantic region of the United States. Without adequately controlling for these exogenous variables, the efficiency estimates generated by the DEA model will be potentially biased as inefficiencies are assumed to be fully attributable to managerial decisions. When the effects of uncontrollable variables are not controlled, the managerial efficiency of units operating in adverse environments will be underestimated (Yang and Pollitt, 2009). Conversely, the managerial efficiency of units operating in favorable environments will be overestimated. These biased estimates can potentially lead to an inefficient allocation of resources if decisions are informed by the efficiency analysis results. For example, if the resource allocation strategy is to direct scarce resources toward improving environmental management within units where managerial performance is low, it is important to have an unbiased assessment and ranking of the managerial component of environmental performance.

A four-stage approach based on Hof et al. (2004) is applied to remove the influence of the exogenous variables and the bias they introduce (Table 1). While Hof et al. (2004) presented results adjusted for exogenous variables, the authors did not compare the adjusted results with the unadjusted results. In this paper, we provide a detailed comparison of adjusted and unadjusted results to better draw out the importance of considering the role of exogenous variables in performing environmental assessments.

Using the four-stage approach, a case study is presented using data collected by the U.S. Environmental Protection Agency's Regional Vulnerability Assessment (ReVA) program as part of its Mid-Atlantic Assessment (Smith et al., 2004). While more recent datasets may be available, we chose to use this particular dataset for this study as it is contains information on multiple environmental media, as well as socioeconomic information, estimated across a large, heterogeneous region. As is typical in regional environmental assessments, the socioeconomic and environmental data used in the ReVA assessment were drawn from a wide variety of sources and are from a variety of time periods. The ReVA program estimated each variable for 134 Mid-Atlantic region watersheds at the 8-digit Hydrological Unit Code (HUC) scale. HUCs are units within a standardized hierarchical system of hydrologic units identified by the U.S. Geological Survey, where the unit boundaries are delineated so surface waters within the boundaries drain to a single outlet (Seaber et al., 1987). Watersheds delineated at the 8-digit HUC scale represent sub-basins and average about 3100 km$^2$ within our study region. The Mid-Atlantic region spans ten states across five Level II Ecoregions (Secretariat of the Commission for Environmental Cooperation, 1997) (Fig. 1):

(1) Atlantic Highlands;
(2) Southeastern USA Plains (abbreviated to Southeastern Plains);
(3) Mixed Wood Plains;
(4) Mississippi Alluvial and Southeast USA Coastal Plains (Coastal Plains); and,
(5) Ozark, Ouachita-Appalachian Forests (Appalachian Forests).

The results of our analysis indicate that the proposed approach affords insights into which exogenous variables influence condition assessments and the magnitude of their influence. The proposed approach also enables environmental managers to more effectively compare the environmental performance of spatial units to identify opportunities for maximizing the efficiency of investments for environmental improvement.

The case study presented here contributes to a growing literature on environmental performance. While there exists a substantial efficiency analysis literature accounting for the effect of variables exogenous to decisionmakers (sometimes referred to as "operating environment") on the economic performance of economic entities such as banks, airports, hospitals, and nursing homes, for example, there is a more limited literature describing how to evaluate the effects of exogenous *environmental* variables. This literature has been largely focused on the agricultural sector where farm and regional agricultural productivity is affected by farm environmental attributes (Asmild and Hougaard, 2006; Barnes, 2006; Binam et al., 2003; Gómez-Limón et al., 2012; Otsuki et al., 2002; Picazo-Tadeo et al., 2011; Wossink and Denaux, 2006). Hof et al. (2004) is the most comprehensive example of evaluating and controlling for exogenous environmental variables in the efficiency literature and, with the exception of our current study, is the only broad-scale application we are aware of that models multiple environmental media, as well as landscape and demographic factors.

**Table 1**
The four stages of the proposed process.

| Stage | Action |
|---|---|
| 1 | Regression models are used to estimate the influence of the exogenous variables on environmental impacts. |
| 2 | The influence of the exogenous variables is removed by adjusting environment impacts according to the results of the regression models, based on an approach used in Hof et al. (2004). |
| 3 | DEA is used with socioeconomic variables as outputs and the adjusted environmental impacts as inputs. |
| 4 | The total radial and non-radial input slacks estimated in Stage 3 are regressed against the exogenous variables using Tobit regression models to verify the removal of the bias introduced by the exogenous variables. |

## 2. Material and methods

### 2.1. Data

The input, output, and exogenous variables used in this case study are summarized in Table 2. The interaction of income, population, and environmental quality has been the subject of significant research over time. Additionally, balancing rapid population growth, quality of life, and environmental quality are important objectives for environmental managers and planners at local and regional scales. As such, we incorporate socioeconomic measures—per capita income and population density—as outputs. From a planning perspective, being able to maximize population density and per capita income with a minimum of environmental impact is a desirable objective. Fortunately, the dataset used for the case study included these two variables for the watersheds analyzed.