Measuring the Stringency of Environmental Regulations

Claire Brunel* and Arik Levinson†

Introduction

The ability of researchers to answer questions about the effects of environmental policies often depends on measurements of the stringency of those policies. Did the U.S. Clean Air Act succeed in its goal of preventing individual U.S. states from lowering their environmental standards in a "race to the bottom" in order to attract investment (Portney 1990)? Should "harmonization" of regulatory stringency across European Union (EU) member countries be a prerequisite for further European unification (Bhagwati and Hudec 1996)? Are lax environmental regulations serving as a substitute for protectionist tariffs (Ederington and Minier 2003)? Are greenhouse gas emissions capped in one country simply emitted by nonparticipating countries instead? The jargon has proliferated along with the policy debates: pollution havens, industrial flight, environmental dumping, race-to-the-bottom, Not in My Back Yard, harmonization, and leakage. All of these policy concerns share a common feature: assessing them requires measuring the relative stringency of environmental regulations over time or across different states, countries, or industries. With this in mind, the purpose of this article is to identify the main challenges to measuring environmental regulatory stringency and evaluate existing measures of stringency based on how accurately they reflect stringency and how well they enable researchers to address these policy issues.

Perhaps because the consequences of environmental policies are so important, there is a vast literature on this topic. However, much of the literature reads as though the chief challenges to measuring the stringency of regulations are related to data collection, meaning that, if we gave the appropriate agencies adequate resources, they could simply collect the necessary information. But the challenges to using such information as measures of stringency for policy analysis actually derive from a much deeper set of conceptual (and econometric) problems: (1) multi-dimensionality—environmental regulations cannot easily be captured by a single measure of "stringency"; (2) simultaneity—regulations are meant to affect emission levels, but emissions levels can be a factor in determining the stringency of a regulation because, for

* American University; e-mail: brunel@american.edu.

† Georgetown University and NBER; e-mail: arik.levinson@georgetown.edu.

The authors are grateful to the OECD for helpful suggestions and financial support for this work and to Robert Elliott, Wayne Gray, Derek Kellenberg, and Dann Millimet for constructive insights.

Review of Environmental Economics and Policy, volume 10, issue 1, Winter 2016, pp. 47–67
doi:10.1093/reep/rev019
Advance Access published on January 21, 2016

example, jurisdictions with the most serious pollution problems may impose the most stringent regulations; (3) industrial composition—in places where the mix of industries is more pollution-intensive, the average business automatically faces more stringent regulations; and (4) capital vintage—regulatory standards are typically stricter for new sources of pollution, which may result in firms keeping older plants in operation longer, thus affecting the environment, the economy, and measures of regulatory stringency.

Because these conceptual problems are so challenging, a large number of studies have attempted to measure stringency using a variety of approaches. We divide these approaches into five categories: (1) private sector abatement costs, (2) direct assessments of the regulations themselves, (3) composite indexes, (4) measures based on pollution levels and changes or energy use, and (5) measures based on public sector expenditures or enforcement. We find that the concept of stringency is difficult to implement empirically and that the measures used thus far fall short in numerous ways.

This article is organized as follows. The next section describes the four conceptual obstacles (or challenges) to measuring and interpreting stringency measures. Then we discuss in more detail the five categories of approaches found in the literature, much of which focuses on the United States, and review their advantages and disadvantages. We present a summary and conclusions in the final section.

Conceptual Challenges to Measuring Environmental Regulatory Stringency

Most policy questions related to environmental regulation concern their effectiveness in achieving environmental goals or their consequences for economic outcomes. In such cases, the ideal measure of stringency would be a panel of data varying both across jurisdictions and within jurisdictions over time. Unfortunately, many available measures of stringency involve comparisons across countries or states in a single year, which further limits our ability to address the four challenges.

Multidimensionality

The "environment" is a complex multidimensional issue, and so are its associated regulations. Governments regulate various media—air, water, land—and pollutants into those media—sulfur dioxide, sewage, toxic chemicals, hazardous waste, and so on. Some regulations target households, whereas others target industries. Regulations set standards for total emissions, emissions concentrations, ambient environmental quality, and the technologies employed by producers. Finally, regulations are effective only if they are enforced.

The challenges of multidimensionality

This multidimensionality presents several challenges. The first and simplest is matching the regulation to the policy issue being addressed. If we are interested in whether regulations cause industrial flight from countries with strict regulations, then neither the lead content of automotive gasoline nor incentives to recycle household waste will directly affect industries' profitability in various locations. Some regulations target emissions, whereas others target ambient quality. For example, the U.S. Clean Air Act sets uniform national ambient air quality standards (NAAQSs). This means that in terms of ambient standards, all air quality control regions (typically set at the county level) face the same level of stringency. But in order to meet these standards, some counties need to impose costly emissions requirements, whereas others do not. To illustrate, in Los Angeles, where the local mountains trap air pollution over the city, the NAAQSs are costly to meet, whereas in Honolulu, where winds quickly blow air pollution out over the Pacific, the NAAQSs are easy to meet. If the regulation being considered is the local ambient standard, then both cities have equally stringent rules. But if the regulation is the control technologies that manufacturers are required to adopt, then Los Angeles is more stringent. Does this mean Los Angeles has more stringent regulations? Clearly, the answer depends on the context.

A second challenge is that complex regulations are not easy to compare. For example, the new U.S. standards for industrial boilers limit toxic emissions to 2.0–3.0 tons per year of mercury and 580,000 tons per year of sulfur dioxide. How do we determine which standard is more stringent? In 1987 the U.S. Environmental Protection Agency (EPA) set the NAAQSs for particulate matter at 150 micrograms per meter cubed ($\mu g/m^3$) of particles smaller than 10 micrometers in diameter (averaged over 24 hours), not to be exceeded more than once per year over 3 years. In 1997 the standard was changed to 65 $\mu g/m^3$ of particles smaller than 2.5 micrometers in diameter at the 98th percentile averaged over 3 years.[1] Which standard is more stringent? Here the answer is more complex because it depends on the compliance costs for industry, the health consequences for people, and the particular policy question being asked.

How has multidimensionality been addressed?

Thus far, researchers have dealt with the multidimensionality and complexity of regulations in one of two ways. First, some avoid the issue of multidimensionality altogether by focusing on one particular narrow environmental problem and using stringency measures that are directly comparable. For example, Berman and Bui (2001) study air pollution regulations as they affect oil refineries in Los Angeles. They use confidential plant-level data and a painstaking, line-by-line reading of the local air pollution regulations. This enables them to know the exact dates on which specific regulatory changes affected particular refineries, how those refineries responded, and what costs they incurred. Similarly, Levinson (1999b) narrows the multidimensionality by focusing on hazardous-waste disposal taxes, a single feature of state law that is easily measurable, comparable, and clearly targeted. The advantage of these focused approaches is clear: accuracy in identifying the appropriate regulations and comparability across regulations. The disadvantage of this approach is that the results may not be generalizable.

Alternatively, some researchers have constructed composite indexes (or proxies) for environmental stringency. For example, Smarzynska and Wei (2004) use the number of international environmental treaties signed and the number of active environmental nongovernmental organizations (NGOs) as indicators of countries' environmental regulatory stringency. Cole and Elliott (2003) use an index based on a survey sent to each United Nations member country asking for details about its environmental policies, legislation, and enforcement. Kellenberg (2009) and Kalamova and Johnstone (2011) use the World Economic Forum (WEF) surveys,

[1]http://www.epa.gov/ttn/naaqs/standards/pm/s_pm_history.html.

which ask business executives in many countries about the stringency and enforcement of regulations imposed on their companies. Fredriksson and Millimet (2004) combine composite indexes, including pollution levels, resource endowments, and public and private abatement efforts.

These composite indexes are meant to compress multidimensional regulations into a single number, to apply broadly to entire economies, and to be inherently generalizable. The disadvantage is that, although these measures may successfully rank countries' stringency levels, they cannot assess their magnitude because they are ordinal rather than cardinal. Does signing twice as many treaties or having twice as many NGOs mean a country is twice as stringent? What does it mean that Germany's WEF index is above 6.5 whereas Argentina's is below 3.5? This means that, although studies using these indexes may be able to answer questions about the direction and statistical significance of the effect of regulations and describe outcomes when the indexes are relatively high or low, they cannot draw conclusions about how these indexes translate into real costs for regulated industries.

Simultaneity

Although the approaches for addressing multidimensionality have their disadvantages, simultaneity presents an even greater conceptual challenge to evaluating measures of stringency. This is because researchers seek to assess the effects of regulatory stringency on pollution, labor demand, trade, economic growth, and other environmental and economic variables, but each of these variables may be simultaneously contributing to regulatory stringency. That is, environmental and economic conditions may be influencing the stringency of the regulation even as researchers are trying to measure the opposite causality. In a widely cited example, Grossman and Krueger (1991, 27) find that in some cases U.S. "imports from Mexico appear to be lower [emphasis in original] in . . . sectors where US pollution abatement costs are relatively high," a counterintuitive result that they attribute to an unnamed omitted variable. A natural candidate for that omitted variable would be some source of comparative advantage that U.S. industries have that is correlated with pollution intensity, such as skilled labor, physical capital, or access to inexpensive energy. Researchers deal with this simultaneity of regulations in two closely related ways, through natural experiments and the use of instrumental variables.[2]

Natural experiments as a solution to simultaneity

Natural experiments are based on the assumption that some external force determines the stringency of regulations. The best example of this approach was first provided by McConnell and Schwab (1990) and Henderson (1996) and has since been followed by numerous researchers.[3] They used the U.S. Clean Air Act, which imposes uniform national ambient standards (the NAAQSs mentioned earlier) on every county in the United States. States are responsible for ensuring that air quality in all of their counties meets those standards. This means that a county whose air quality falls below the NAAQSs faces more-stringent standards,

[2]If the source of the simultaneity is time invariant, then the problem can also be addressed with panel data and fixed effects. But it is easy to imagine sources of simultaneity that change over time: demographic changes, technology trends, and so on. In that case, fixed effects cannot solve the problem.
[3]Numerous other researchers have since followed their example (see, for example, Greenstone [2002] and Chay, Dobkin, and Greenstone [2003]).

whereas a county whose air quality meets or exceeds the standards faces less-stringent regulations. Because these NAAQSs are set federally and apply nationally, they are not caused by or correlated with the economic activity in particular counties: in other words, they are plausibly exogenous. Researchers using this measure of stringency interpret the changes in local economic activity that follow changes in federal law as causal consequences of changes in regulatory stringency.

Similarly, Levinson (1999b) examined a 1992 U.S. Supreme Court ruling that prohibited states from charging higher fees for disposing of waste imported from other states, a practice that was common prior to 1992. Because the changes in fees following the ruling were not the result of individual states' decisions or disposal quantities, the changes in interstate shipments of waste that followed the 1992 ruling can be interpreted as the consequence of externally imposed fee changes.

The issue with using natural experiments to address the issue of simultaneity is that such experiments are few and far between. It is difficult to think of examples in which states or countries have been forced by outside circumstances to alter or adopt regulations with varying levels of stringency. Moreover, the few examples that do exist may not be representative of regulatory effects more generally.

Instrumental variables as a solution to simultaneity

A second way that researchers have tried to address the simultaneity problem is through statistical approximation of those experiments—that is, instrumental variables. The idea is to find a characteristic of jurisdictions or industries that is correlated with regulatory stringency but uncorrelated with the measure of economic activity being studied (except indirectly through its relationship to stringency) and then to use that correlated characteristic as a proxy for environmental regulatory stringency.

Unfortunately, as with natural experiments, there is a shortage of good examples of such instrumental variables. Xing and Kolstad (2002) use infant mortality and population density. Ederington and Minier (2003) use instruments motivated by political-economy theories: unionization rates, concentration ratios, and so on. Levinson and Taylor (2008) use the geographic distribution of manufacturing industries across U.S. states and the pollution abatement costs incurred by other industries in those states. Kellenberg (2009) uses lagged values of countries' corruption, income, urbanization, and education. Jug and Mirza (2005) use prior years' wages and investment in environmental equipment. Millimet and Roy (2012) provide an excellent review of all of these studies, as well as others, and they suggest two approaches of their own. The first is an instrument that relies on the fact that environmental regulatory stringency imposes higher costs on more pollution-intensive industries, whereas other local business conditions affect all industries equally. The second avoids the use of instruments altogether, instead exploiting assumptions about heteroskedasticity of the errors in the estimating equation.

There are several problems with using instrumental variables to address simultaneity. Most obviously, there are good reasons to question the underlying assumptions. Infant mortality may be a consequence of pollution and motivate stringent regulations; industry concentrations may affect regulatory stringency and be affected by those regulations. Moreover, the key assumption underlying instrumental variables techniques—that the variable being used as an instrument affects the outcome only through its effect on regulations—is not easily tested. Examples of instruments that stand up to this scrutiny—especially those that vary across jurisdictions and over time—are in short supply.

The next two obstacles to measuring regulatory stringency—industrial composition and capital vintage—can be thought of as special cases of simultaneity. But because they are central to evaluating stringency and to the research questions analysts want answered, it is useful to discuss them separately.

Industrial Composition

States and countries differ in their industrial composition. This means that some measures of stringency will have different values even for jurisdictions whose regulations are identical. To illustrate, consider two states: one producing cement, which creates pollution as a by-product, and one producing textiles, which causes much less pollution. Even if the two states have identical laws, the average manufacturer in the first state will incur more environmental costs than the average manufacturer in the second. Abatement costs, surveys of business executives, and pollution emitted per dollar of sales will vary across the two states for reasons having to do with industrial composition rather than regulatory stringency. Thus, depending on how we measure stringency, industrial composition could lead researchers to make incorrect inferences.

This problem—that industrial composition differs across jurisdictions—poses especially serious challenges for interpreting measures of stringency that are based on pollution abatement costs. First, as discussed earlier, jurisdictions that are home to pollution-intensive industries will have relatively high average pollution abatement costs even if regulations are no more stringent than those in other jurisdictions. Second, if having a high concentration of polluting industries leads jurisdictions to enact more stringent regulations, the analysis becomes even more complicated because this simultaneity may cause some researchers to incorrectly conclude that environmental stringency actually attracts pollution-emitting industries or at least that stringency does not deter those industries. The solution to this problem, which we discuss in the "Capital Vintage" section, requires measuring a jurisdiction's pollution costs independently from its industrial composition.

Capital Vintage

The final obstacle to measuring regulatory stringency involves a feature that is common to many environmental regulations: they are "grandfathered" or "vintage-differentiated," which means they are stricter for new than for existing sources of pollution. One prominent example of grandfathered regulations is the U.S. Clean Air Act, which prescribes "new source performance standards" for large industrial sources of pollution that are new or significantly modified. Ironically, this can protect existing industries from potential competition, extend the profitable life of older equipment, and result in higher aggregate emissions.[4]

Depending on how we measure regulatory stringency, grandfathering could significantly bias those measurements. For example, suppose our measure is based on pollution abatement costs incurred. A stringent regulation that grandfathers existing sources may result in no new

[4]Buchanan and Tullock (1975) pointed this out long ago. See Stavins (2006) for a recent discussion.

development and low abatement costs. A less-stringent regulation or one that does not grandfather existing sources might result in more new development and higher abatement expenditures. Perversely, vintage-differentiated regulations can result in stringent jurisdictions appearing lax and vice versa. Alternatively, suppose our stringency measure is based on emissions, where low emissions are interpreted as the result of strict regulations. A strict vintage-differentiated regulation that deters new investment in cleaner production might be misinterpreted as a lack of stringency because emissions from existing production would remain high.

In the following sections, we describe the five categories of approaches researchers have developed to measure regulatory stringency and evaluate their success in overcoming the four obstacles. For each category, we also highlight some of the literature that uses that approach.[5]

Private-Sector Abatement Costs

Private sector abatement costs, reported in surveys of companies or indirectly calculated based on economic theory, constitute one of the most widely used measures of stringency.

Cost Surveys

The earliest and most comprehensive data on private-sector pollution abatement costs in the United States come from the U.S. Pollution Abatement Costs and Expenditures (PACE) survey.[6] Researchers have used the PACE data in various ways to address the problems of simultaneity, industrial composition, and capital vintage. For example, to address industrial composition, Levinson (1996) uses confidential plant-level data to regress each plant's abatement operating costs on other characteristics, including age of the facility and dummies for each industry and state. The age coefficient controls for capital vintage and the industry dummies account for industrial composition. A high state dummy coefficient indicates that manufacturers in that state that have similar observed characteristics in all other dimensions spend relatively more on pollution abatement.[7] Becker (2011) follows a similar approach, using plant-level data and controlling for capital vintage and industrial composition to create a county-level index of compliance costs.

Most researchers do not have access to confidential plant-level PACE data. Thus, Keller and Levinson (2002) construct an industry-adjusted, cost-based measure using published average annual PACE data by industry and state. They calculate the total pollution abatement costs per dollar of gross state product[8] and then compare these costs to the predicted abatement cost, which is simply a weighted average of the national pollution abatement costs for each of twenty

[5]For a more complete list of citations in each category, see the on-line Supplementary Materials.
[6]This survey was conducted annually by the US Census Bureau from 1973 until 1994, and then in 1999 and 2005.
[7]In Levinson (1996), the measure varies only across states, not over time. However, one could estimate the measure using annual cross-sections of plant-level pollution abatement cost data, and state dummies, year dummies, and the interactions between the two. The coefficients on the interaction terms would indicate whether a state became more or less stringent relative to the national trend.
[8]Formally, $S_{st} = P_{st}/Y_{st}$, where P_{st} is the pollution abatement cost in state s in year t, and Y_{st} is the gross state product in state s in year t.

two-digit Standard Industrial Classification industry codes.[9] Keller and Levinson's measure of stringency is the ratio of actual costs to predicted costs. When this ratio is greater than one, it means that pollution abatement costs are larger than would be expected given the state's industrial composition, leading Keller and Levinson to infer that in such cases the state's regulations are relatively stringent.

On the surface, the PACE survey data appear to be ideal for measuring stringency because the survey directly asks managers at industrial facilities how much their establishments spent on pollution abatement. However, one drawback of using PACE data is that they include all types of abatement costs, not simply those due to regulatory stringency. For example, states with the same regulatory stringency could have different abatement costs if the inputs to pollution abatement cost more in some states than others, such as if low-sulfur coal has to be shipped further or if environmental engineers are more highly paid. Thus, the PACE data should be viewed as measuring a jurisdiction's environmental costs overall, not just those related to regulatory stringency.

A more serious drawback of the PACE data is that the survey's central question—that is, How much did your plant spend on pollution abatement?—has become increasingly difficult for environmental managers to answer. Consider the instructions accompanying the 1994 PACE survey:

> For this survey, only expenditures with the primary purpose of protecting the environment are included. This survey does not . . . include expenditures that abate pollution when the primary purpose is to increase profits or cut costs, and the environmental protection is a side benefit.[10]

In the case of end-of-pipe technologies (e.g., scrubbers on smokestacks), the central question might be relatively easy to answer because such technologies entail modifying existing production processes in response to new regulations. However, the answer is much less clear when process or product changes have evolved in response to regulations that have existed for decades. For example, if a manufacturer installs capital equipment enabling it to begin using recycled materials, is this an environmental investment? Does it matter if doing so also increases profits? If an electricity generator switches from coal to natural gas and saves money partly because environmental regulations have made burning coal more costly and partly because natural gas prices have fallen, how much of that process change should be considered environmentally motivated? Just because a government agency asks survey respondents these questions does not mean the answers are meaningful to researchers.

Several studies have evaluated the results of the PACE survey. In a study of abatement costs in pulp and paper mills, Gray and Shadbegian (2003) find that the true abatement costs could be more than three times the reported costs. Joshi, Krishnan, and Lave (2001) interview accountants at fifty-five U.S. steel mills and conclude that for every \$1 of reported environmental costs, there are \$9–10 of unreported environmental costs because it is difficult to separately identify specific overhead and process changes as being implemented primarily for environmental purposes. Morgenstern, Pizer, and Shih (2001) further find that grandfathered regulations

[9] S_{st} is the actual abatement cost and $\hat{S}_{st}$ is the predicted abatement cost. The weights used to calculate $\hat{S}_{st}$ are the industries' shares of output in state s, Y_{sit}/Y_{it}.
[10] Current Industrial Reports, Pollution Abatement Costs and Expenditures: MA200(94)-1. Washington, DC: US Census Bureau, 1994.

mean that the reported abatement costs of existing facilities understate the costs faced by manufacturers that are expanding or opening new facilities. However, Morgenstern, Pizer, and Shih also find that surveys may overstate costs because, for example, there are complementarities between environmental and other objectives.

About the time the United States stopped collecting the first wave of PACE data (the mid-1990s), Canada and the EU began. Pasurka (2008) reviews these efforts, including Canada's Survey of Environmental Protection Expenditures (SEPE) (Statistics Canada 1995, 2000) and the joint OECD/Eurostat Questionnaire on Environmental Protection Expenditure and Revenues (Eurostat 2005). He finds that it is sometimes difficult to compare surveys across countries. For example, Germany's data focus on end-of-pipe expenditures from 1996 to 2002, whereas other countries' surveys include all abatement costs. The U.S. survey includes capital depreciation, whereas Canada's does not. Countries' surveys also differ in their industry classification systems and industry coverage, making it difficult to account for differences in industrial composition.

Although these abatement cost surveys clearly have drawbacks, it is important to note that, in aggregate terms, reported abatement costs vary over time and across industries and jurisdictions in ways that are consistent with intuition. Moreover, the industries we would expect to have high abatement costs (e.g. primary metals, paper) top the list; countries and U.S. states we would expect to have low abatement costs (e.g. Finland, Sweden, New Mexico, Delaware) rank toward the bottom; and changes in pollution regulations appear to be reflected in reported abatement costs (Becker 2005). Thus, we are not arguing that these surveys contain no useful information. Rather, we are suggesting that survey responses be treated with caution and recognized as speculative answers to increasingly difficult and abstract questions, questions that go beyond regulatory costs and are not necessarily applicable to new plants.

Shadow Costs

Some studies have avoided the drawbacks of cost surveys by relying on economic theory and firms' choices to indirectly calculate the "shadow price" of pollution. Van Soest, List, and Jeppesen (2005, 1155) define the shadow price of an input as "the potential reduction in expenditures on other variable inputs that can be achieved by using an additional unit of the input under consideration (while maintaining the level of output)." For simplicity, let us assume that a firm has two inputs—emissions and some "generic" input X (e.g., labor)—and that the firm seeks to maintain a certain level of output.[11] When there is no regulation, the price of emissions is low (or even zero), and thus profit-maximizing firms will choose to use relatively more emissions and less of the other input. When the price of emissions is higher, perhaps because of stringent regulations, the firm will choose lower emissions.

The key to this approach is that, with the exception of the price of emissions, all of the prices and quantities can be found in government statistical tables and reports.[12] This means that if firms are profit maximizing and we know the output, the amounts of all inputs used, emissions, and the prices of all of the other inputs, then we can calculate the implicit—or "shadow"—

[11]Economists consider pollution as an "input" to production, even though it physically emerges from smokestacks or wastewater pipes because it is an activity undertaken in order to generate the main product of the firm. Some readers may find it simpler to think of emissions as "waste disposal services."
[12]However, the level of detail will depend on data availability, which varies by pollutant and country.

price of emissions. This price will be higher in jurisdictions where the cost of abatement is higher. Like the cost survey approach, the shadow price approach interprets higher costs as being a result of more stringent regulation.

The shadow cost approach to measuring environmental regulatory stringency has a number of advantages. First, it converts a multidimensional environmental regulation into a single cardinal measure of costs. Second, it controls for both capital vintage and industrial composition. Finally, it can be estimated across countries, industries, years, and pollutants. Of course, the shadow price approach also has drawbacks. Shadow prices will depend in part on the functional forms chosen for cost functions or production functions and on the set of other inputs used in their estimation, which might depend on which inputs have readily available price and quantity data. Like the abatement expenditures from cost surveys, the expenditures measured by shadow prices are not necessarily the result of regulatory stringency.

Direct Assessments of Regulation

An alternative to using the costs imposed by a regulation to measure its stringency is to use a direct assessment of the regulation itself. Such regulation-based measures face two main challenges: multidimensionality and simultaneity. Thus, most studies that use this approach ask very narrow questions about particular pollutants and try to avoid the simultaneity problem by using natural experiments or instrumental variables.

Addressing Simultaneity through Natural Experiments

One widely used natural experiment relies on the U.S. Clean Air Act because its NAAQSs address both the simultaneity and the multidimensionality problems. The NAAQSs set a maximum allowable ambient concentration level for six common air pollutants and thus can be viewed as a general measure of multidimensional stringency. In addition, because these standards are set federally and apply to every U.S. county, they can be considered exogenous to any one county's economic or environmental conditions. However, using county attainment status as a measure of stringency has two drawbacks. First, it is difficult to translate NAAQSs attainment into a cardinal measure of stringency, which makes it difficult to assess magnitudes or draw general conclusions about the consequences of regulations. Second, economic consequences (e.g., the effect on foreign direct investment) can only be identified as the outcomes of regulation in stringent nonattainment counties relative to the outcomes in less-stringent attainment counties. In some cases, regulations in nonattainment counties can have spillover effects in attainment counties—for example, if the regulation causes investment to shift from nonattainment to attainment counties. These spillover effects imply that this approach can only identify the difference between the two sets of counties, not the overall effect of the regulations on investment. Thus, researchers must be careful not to double-count the regulations' effect by interpreting the total difference in outcomes as the effect of the regulations in nonattainment counties.

Using NAAQSs, Henderson (1996) and Becker and Henderson (2000) show that more stringent regulations in nonattainment counties improve air quality and reduce the number of new polluting manufacturing plants locating there. More recent studies that use this approach examine industry location and employment (Greenstone 2002), housing prices (Chay and Greenstone 2005), pollution (Greenstone 2004), new manufacturing plants (List, McHone, and Millimet 2004), and mortality (Chay, Dobkin, and Greenstone 2003). Although there has been extensive research on the effects of U.S. air quality standards (NAAQSs), the results cannot necessarily be extrapolated to standards imposed for other pollutants or by other countries.

Targeted Approaches to Regulation-Based Measures

An alternative regulation-based measure uses a specific regulation as an indicator of overall environmental regulatory stringency. One example is using the maximum allowable level of lead per gallon of gasoline as a proxy for a country's overall level of environmental concern. Although this measure only applies directly to the transportation sector, Damania, Fredriksson, and List (2003) show that gasoline lead levels are correlated with three other measures of regulatory stringency (i.e., public environmental R&D expenditures as a proportion of GDP, per capita membership in environmental organizations, and an index derived from Dasgupta et al. [2001]). Cole, Elliott, and Fredriksson (2006) and Cole and Fredriksson (2009) use this measure to study the relationship between foreign direct investment and environmental regulations.

Other research has used specific regulations to study narrow policy questions related to those regulations. For example, Berman and Bui (2001) examine the regulations that apply to particular petroleum refineries in Los Angeles. Refineries not subject to those rules—because they use different technologies or are located elsewhere—are used as a comparison group. Other examples include McConnell and Schwab (1990), who use states' standards for the maximum amount of volatile organic compounds in automobile paint, and Levinson (1999a), who uses indicator variables for whether or not U.S. states' toxic air pollution rules grandfathered existing sources of pollution. Hascic and Johnstone (2011) examine the effect of fuel taxes and fuel efficiency standards on innovation aimed at alternative fuel vehicles. However, all of these studies have the same limitation: the results cannot be generalized to other conditions and outcomes.

Composite Indexes

This leads to the next category of measures: composite indexes, which are meant to serve as comprehensive indicators of a jurisdiction's overall environmental regulatory stringency. Although composite indexes are aimed at solving the multidimensionality issue, they are vulnerable to criticism for being arbitrary. Moreover, their magnitudes can be difficult to interpret. Some of the earliest attempts to quantify regulatory stringency were based on simple indexes constructed from counts of regulations, environmental NGOs, international treaties signed, and similar easily enumerated characteristics. For example, researchers have used the voting records of U.S. states' congressional delegations (Gray 1997) and counts of the number of statutes each state has from a list of fifty common environmental laws (Levinson 1996).

Indexes Based on Surveys of Government Officials

A large number of composite indexes have been used to conduct cross-country comparisons. Among the earliest efforts was a 1976 survey sent to 145 countries by the United Nations

Conference on Trade and Development (UNCTAD), asking government officials about their environmental policies. The UN ranked the overall responses of the 40 countries that responded on a seven-point scale (Tobey 1990; Walter and Ugelow 1979).

Cross-country indexes have come a long way since the UNCTAD survey. For example, Dasgupta et al. (2001) randomly selected thirty-one of the 145 national environmental reports that were prepared in advance of the first UN Conference on Environment and Development (UNCED) in 1992. They assessed the answers to twenty-five questions as they applied to four media (air, water, land, and wildlife) and five economic sectors, resulting in 500 separate scores for each country. This was done separately for five different environmental dimensions: awareness, policies, legislation, control mechanisms, and implementation. Moreover, the ranking can be compiled separately for each media, industry, or environmental dimension. Not surprisingly, this composite index is correlated with economic development, with Switzerland ranking first and Mozambique ranking last. Damania, Fredriksson, and List (2003) use the Dasgupta approach, along with several other approaches, to try to calculate the causal relationship between trade policy and environmental policy. Raspiller and Riedinger (2008) use the Dasgupta approach to estimate the effect of regulations on the location decisions of French multinational firms.

There are two problems with this approach to measuring environmental regulatory stringency. First, the UNCED survey was one single cross-section for 1990. Thus, it is not possible to construct a panel of data and to include country fixed effects to control for unobserved country characteristics correlated with both regulatory stringency and economic activity.[13] Second, although Dasgupta et al. were sensible and methodical in constructing their index, different researchers might have constructed a different index with the same data and thus arrived at different country rankings. We are not aware of any studies that conduct this robustness check—that is examining whether alternative indexes constructed with the same data deliver similar conclusions.

Indexes Based on Surveys of Business Managers

Rather than surveying government officials about their countries' environmental regulatory stringency, a number of studies have used surveys of business managers. Among the most widely used surveys is the WEF Executive Opinion Survey. Kellenberg (2009) examines a question that has been asked consistently since 2000: How do regulatory stringency and enforcement affect economic growth? Wagner and Timmins (2009) find that outbound foreign direct investment from Germany is strongly negatively associated with high values of the WEF index for one particular industry, chemical manufacturing, suggesting that there is less German investment in chemical manufacturing in more stringent countries. Kalamova and Johnstone (2011) find a more broad-based effect that is relatively small, nonlinear, and diminishes above a certain threshold of stringency.

As Kalamova and Johnstone emphasize, this measure of stringency is based not on hard data on costs or regulations, but rather on perceptions. Although perceptions may be correlated with

[13]Although Raspiller and Riedinger (2008) use firm-level data and include country-level fixed effects, those fixed effects capture unobserved characteristics of the set of firms that choose to locate in a given country, not unobserved characteristics of that country. Damania, Fredriksson, and List (2003) try to convert the Dasgupta index into a panel by predicting its values in 1990 using country characteristics that vary over time.

regulatory stringency, Kalamova and Johnstone cannot say whether the nonlinear relationship they find between stringency and investment reveals a true nonlinear relationship or a spurious nonlinear relationship between actual and perceived stringency. Moreover, the perception of regulations may be determined even more simultaneously by economic activity than by the actual regulations. We know, for example, that public support for environmental policies falls when unemployment rates rise (Kahn and Kotchen 2011). If economic downturns increase perceived stringency, then researchers may misinterpret this result as indicating that actual stringency reduces investment.

Another problem with surveys of executives concerns industrial composition. As Albornoz et al. (2009) note, firms that come from more pollution-intensive industries are more likely to report that they incur environmental costs.

Indexes Based on Counts of Regulation

In the same way that the Dasgupta index improves on the UNCTAD index from the early 1970s, indexes based on counts of regulations have also become more sophisticated. Smarzynska and Wei (2004) count whether or not a country signed or ratified one of four international environmental treaties, along with the number of environmental NGOs in the country. Johnstone, Hascic, and Popp (2010) examine the degree to which policies supporting renewable energy have spurred technological innovation. They create a list of policies, including tax incentives, investment subsidies, differentiated tariffs, voluntary programs, quotas, and tradable certificates, and count the number of these policies implemented in each of twenty-five countries from 1978 to 2003. Although not technically a measure of regulatory stringency, such a measure could be constructed with that alternative goal in mind.

Other Indexes

Finally, Cole, Elliott, and Okubo (2010) use an index constructed by the Japanese Research Institute of Economy, Trade and Industry to measure regulatory stringency. The index calculates a weighted average of 303 four-digit manufacturing industries within each of forty-one two-digit sectors that are governed by 3,000 broad industrial regulations. The index represents the share of each sector that is regulated. Unfortunately, as with the other indexes, we do not know whether slight deviations in its construction might lead to large differences in its rankings and conclusions. And, like all such indexes, this one conveys little sense of magnitude because regulations are weighed equally regardless of the burden they impose.

Emissions, Pollution, or Energy Use

Some studies have turned the entire concept of measuring stringency on its head and used emissions, ambient pollution, or energy use as measures of stringency. On the surface, this approach seems counterintuitive. The regulations whose stringency is to be measured are designed to reduce pollution; so does pollution indicate regulatory stringency or laxity? The answer depends on the context.

Emissions and Pollution

Some studies have taken high levels of pollution as evidence that regulations are relatively lax. For example, Xing and Kolstad (2002) use aggregate U.S. sulphur dioxide emissions in this way. Others use high pollution as evidence that regulations are stringent, arguing that governments will be forced to tighten regulations to deal with the problem. McConnell and Schwab (1990) use the degree to which a U.S. county was out of compliance with national standards as a proxy for the stringency of the regulations the state would have to impose to meet those standards.

Several research projects have used reductions in emissions as indicators of stringency. Smarzynska and Wei (2004) used declines in carbon dioxide, lead, and water pollution as a share of GDP. In a classic study, Gollop and Roberts (1983) examined fifty-six U.S. electric utilities from 1973 to 1979 using a measure of stringency that is based on the difference between actual observed emissions and an engineering estimate of what the utility's unconstrained emissions rate would have been (i.e., without any regulation). These two studies represent extremes of aggregation and disaggregation. Smarzynska and Wei use aggregate country-wide emissions reductions, which could be due to regulatory changes but could also result from changes in industrial composition or factor prices caused by other variables. Gollop and Roberts (1983) use emissions reductions below unconstrained levels for one particular industry in the United States. Their approach would be difficult to extrapolate to other industries or countries because legal standards and unconstrained emissions would have to be identified for each case.

Energy Use

The last approach in this category involves using energy consumption as a proxy for regulatory stringency. For example, Cole and Elliot (2003) use countries' energy consumption divided by GDP in 1980 and the change in that variable from 1980 to 1995 to rank countries and create an index of environmental regulatory stringency.[14] Harris, Konya, and Matyas (2003) expand on this energy index, using two measures of energy (final consumption and primary supply) scaled by population and two alternative measures of GDP (based on purchasing power parity and exchange rates).

However, these indexes make it difficult to determine whether the measure of stringency is largely the result of changes in energy use or levels of energy use. Both changes and levels could vary across countries for reasons other than environmental regulatory stringency, including energy prices, industrial composition, and trade liberalization. Moreover, if environmental regulations drive up the price of energy, energy expenditures may not decline as a share of GDP even if energy use has also declined.

Public Sector Expenditures or Enforcement

The final category involves using public sector environmental efforts—either expenditure or enforcement or both—as a measure of stringency. For example, Gray (1997) uses U.S. state budgets for environmental and natural resources, which has the advantage of including

[14]This energy-based measure is highly correlated with the Dasgupta index.

enforcement (an important component of stringency). However, this approach should be used with caution because some public-sector expenditures reduce private-sector costs (e.g., tax incentives and subsidies for clean-up). Do such public expenditures increase or reduce stringency? The answer depends on the policy in question.

Shadbegian and Gray (2012) and Levinson (1996) get around this issue by focusing specifically on public enforcement efforts. The major drawback of enforcement measures, however, is that they are at best remote indicators of stringency. State laws can be made stringent either through frequent inspections by well-staffed environmental agencies even if the punishment for violations is not steep or through infrequent inspections by understaffed agencies if the punishment for violations is sufficiently onerous.

Pearce and Palmer (2001) combine the public-expenditure and private-cost approaches by asking whether the burden of environmental regulations has shifted over time from the public sector to the private sector. However, as they point out, the division of the burden between private and public is not always obvious.

In general, public-sector effort has not been widely used to measure regulatory stringency, most likely because its shortcomings tend to outweigh its advantages and because, as more emissions and cost data have become available, the need to use public-sector data as a proxy has declined.

Comparison of the Measures

We have shown that existing measures of environmental stringency vary depending on what they cover and how well they tackle the four conceptual obstacles. Because of these differences, each approach results in its own ranking of stringency across jurisdictions. One way to evaluate how well the approaches measure stringency is to compare them with one another. To illustrate how different the rankings are, we examine correlations between measures of stringency across U.S. states and across countries. But does a low correlation necessarily imply that the measures are failing to accurately capture stringency? And what does this mean for researchers?

The correlations among eight recent measures of the stringency of U.S. state environmental regulations show that the measures are remarkably uncorrelated (see Appendix table 1).[15] The highest correlations are among the composite indexes. This is not surprising because these indexes share many components. These composite indexes are also correlated with the League of Conservations Voters' evaluation of the voting record of state congressional delegations but not with the share of the population in each state living in counties with air quality that exceeds federal standards or with pollution costs. Pollution abatement costs adjusted for states' industrial compositions are not strongly positively correlated with any of the other indexes. We also find that some measures vary within a narrow range across states, whereas others vary considerably from state to state.

We also examined the correlations among four similar measures of stringency for twenty-four countries (see Appendix table 2). These correlations are low as well. In fact, the differences in indexes are even greater across countries than across U.S. states. The highest correlation is

[15]More details about the indexes discussed here can be found in the online Supplementary Materials. Note that List and Co (2000) find similarly low correlations across states' abatement expenditures, regulatory budgets, and a composite regulatory index.

only 0.52, and the relative swings in one index can be as much as fifteen times larger than the relative swings in another index.

One interpretation of these low correlations is that the information in the indexes is not consistent across measures and that even if one index captures a comprehensive measure of stringency, the others are failing. A more positive interpretation is that regulatory stringency is multidimensional and each index captures different aspects of that stringency (List and Co 2000). For example, there is no particular reason to expect that the "42/7 Wall Street" ranking of a state's environmental policy would be correlated with the share of its population in nonattainment counties (see Appendix table 1). And there is no reason to expect that business executives' ranking of a country's environmental policy would be correlated with the company's abatement costs, which may have more to do with industrial composition (see Appendix table 2). Whether or not these results concerning correlations undermine the credibility of any one index or reinforce the idea that regulatory stringency is multidimensional, the low correlations suggest that researchers should avoid putting too much faith in results that rely on only one measure of stringency.

Summary and Conclusions

This article has shown that the obstacles to evaluating environmental regulatory stringency are not simply issues of data collection. Rather they derive from the deeper conceptual challenges of multidimensionality, simultaneity, industrial composition, and capital vintage, which should be kept in mind when evaluating any proposed measure of stringency. We divided the approaches to measuring stringency into five broad categories and have argued that each has both strengths and weaknesses.

What would an ideal measure of environmental regulatory stringency look like? It would be relatively easy to calculate, using data that governments already collect or that they should collect as part of efforts to achieve other policy objectives (e.g., pollution control). Such a measure would be available annually in order to facilitate panel data models that address some sources of simultaneity. It would be cardinal, enabling assessment of magnitudes, and either available for various pollutants and media or combinable into a single overarching measure of multidimensional stringency. It would be theoretically related to the costs facilities incur when they abate pollution, but it would not be determined by industrial composition.

Most of the measures we have discussed here fall far short of this ideal. This is because the regulations themselves are too complex and dissimilar across countries to create consistent measures of stringency, except in narrow case studies that are not generalizable. It is hard to imagine surveys of business executives or government officials meeting the conditions of the ideal measure. Although numerous, composite indexes are rarely created as panels, are not theoretically grounded in costs, and are typically impossible to disaggregate by pollutant or media. Public-sector efforts, measured by expenditures or enforcement, are conducted differently in every state and country and fail to capture key aspects of stringency. Pollution abatement cost surveys shift the burden to private-sector managers by asking them to answer difficult conceptual questions they may be incapable of answering. The challenges of capturing environmental stringency in a single, easily computable, cardinal measure for multiple pollutants over multiple years suggest that the characteristics of the "ideal" measure may depend on the policy issue being studied.

We hope that the discussion here has provided readers with an understanding of both the importance of measuring environmental regulatory stringency and the serious challenges facing researchers and policymakers seeking to undertake this task.

Appendix Table 1 Correlations among various measures of the stringency of U.S. state policy

	Nonattainment counties (1)	24/7 Wall Street (2)	Greenopia (3)	Forbes (4)	League of Conservation Voters (5)	PACE ÷ value added (6)	PACE adjusted 3-digit (7)	PACE adjusted 4-digit (8)	Coefficient of variation[a] (9)
Nonattainment counties	1.00								0.92
24/7 Wall Street	−0.21	1.00							0.57
Greenopia	0.11	0.56	1.00						0.57
Forbes	0.27	0.51	0.81	1.00					0.26
League of Conservation Voters	0.31	0.24	0.60	0.68	1.00				0.50
PACE ÷ value added	−0.22	−0.22	−0.56	−0.57	−0.50	1.00			0.58
PACE by 3-digit NAICS	−0.25	0.11	−0.35	−0.37	−0.44	0.68	1.00		0.40
PACE by 4-digit NAICS	−0.04	0.24	−0.08	−0.05	−0.04	0.10	0.61	1.00	0.55

Sources: (1) Share of state population living in nonattainment counties based on EPA nonattainment designation (and county-level U.S. Census data from 2010; (2) 24/7 Wall Street 2010 ranking of US states based on the environmental problems in each state and how effectively these problems are addressed; (3) Greenopia 2011 State Sustainability Index based on air quality, water quality, recycling rate, number of LEED buildings, green business density, per capita water consumption, per capita energy consumption, per capita emissions, per capita waste generation, and several renewable energy statistics; (4) Forbes 2007 America's Greenest States ranking based on six equally weighted categories: carbon footprint, air quality, water quality, hazardous waste management, policy initiatives and energy consumption; (5) LCV 2010 Congressional delegation environment scorecard based on the voting record of all members on Congress on the most important environmental legislation considered that year; (6) U.S. Bureau of Economic Analysis 2005 PACE data divided by value added (U.S. Census Bureau 2008); (7) 2005 PACE data adjusted for states' industrial compositions at the level of three-digit NAICS codes; (8) 2005 PACE data adjusted for states' industrial compositions at the level of four-digit NAICS codes.

[a]The coefficient of variation is a standardized measure of the dispersion of each index, defined as the ratio of the variance over the mean.

Appendix Table 2 Correlations among measures of countries' policies

	Environmental performance (1)	Abatement costs (2)	WEF survey (3)	Energy intensity (4)	Coefficient of variation[a] (5)
Environmental performance	1.00				0.11
Abatement costs	0.26	1.00			1.72
WEF survey	0.36	−0.04	1.00		0.14
Energy intensity	0.44	0.38	0.52	1.00	0.47

Sources: (1) 2010 Environmental Performance Index, which ranks countries based on changes in their environmental performance and is produced by the Yale Center for Environmental Law, CIESIN of Columbia University, the World Economic Forum, and the Joint Research Center of the European Commission; (2) Pollution abatement costs from the U.S. Bureau of Economic Analysis (2005), Eurostat Environmental Expenditure Database (2005), and Statcan (2006); (3) 2011 World Economic Forum Global Competitiveness Executive Opinion Survey ranking of environmental regulatory stringency by CEOs; (4) Energy use from the World Bank World Development Indicators divided by GDP from the IMF World Economic Outlook for 2010.

[a]The coefficient of variation is a standardized measure of the dispersion of each index, defined as the ratio of the variance over the mean.

References

Albornoz, Facundo, Matthew Cole, Robert Elliott, and Marco Ercolani. 2009. In search of environmental spillovers. *World Economy* 32 (1): 136–63.

Becker, R. A. 2005. Air pollution abatement costs under the clean air act: evidence from the PACE survey, *Journal of Environmental Economics and Management* 50: 144–69.

———. 2011. Spatial heterogeneity in environmental compliance costs. *Land Economics* 87 (1): 28–44.

Becker, R. A., and J. V. Henderson. 2000. Effects of air quality regulations on polluting industries. *Journal of Political Economy* 108: 379–421.

Berman, E., and L. Bui. 2001. Environmental regulation and productivity: Evidence from oil refineries. *Review of Economics and Statistics* 83 (3): 498–510.

Bhagwati, J., and R. Hudec, eds. 1996. *Fair trade and harmonization: Prerequisite for free trade?* Cambridge, MA: MIT Press.

Buchanan, J. M., and G. Tullock. 1975. Polluters' profits and political response: Direct controls versus taxes. *American Economic Review* 65 (1): 139–47.

Chay, K., C. Dobkin, and M. Greenstone. 2003. The Clean Air Act of 1970 and adult mortality. *Journal of Risk and Uncertainty* 27: 279–300.

Chay, K., and M. Greenstone. 2005. Does air quality matter? Evidence from the housing market. *Journal of Political Economy* 113: 376–424.

Cole, M., and R. J. R. Elliot. 2003. Do environmental regulations influence trade patterns? Testing old and new trade theories. *World Economy* 26 (8): 1163–86.

Cole, M., R. J. R. Elliot, and P. Fredriksson. 2006. Endogenous pollution havens: Does FDI influence environmental regulations? *Scandinavian Journal of Economics* 108 (1) 157–78

Cole, M. A., R. J. R. Elliott, and T. Okubo. 2010. Trade, environmental regulation and industrial mobility: An industry-level study of Japan. *Ecological Economics* 69 (10): 1995–2002.

Cole, M. A., and P. G. Fredriksson. 2009. Institutionalized pollution havens. *Ecological Economics* 68 (4): 925–1274.

Damania, R., P. G. Fredriksson, and J. A. List. 2003. Trade liberalization, corruption, and environmental policy formation: Theory and evidence. *Journal of Environmental Economics and Management* 46: 490–512.

Dasgupta, S., D. Wheeler, S. Roy, and A. Mody. 2001. Environmental regulation and development: A cross-country empirical analysis. *Oxford Development Studies* 29 (2): 173–87.

Ederington, J., and J. Minier. 2003. Is environmental policy a secondary trade barrier? An empirical analysis. *Canadian Journal of Economics* 36 (1): 137–54.

Eurostat. 2005. *Environmental protection expenditure by industry in the European community*. Office for Official Publications of the European Communities, Luxembourg.

Fredriksson, P., and D. Millimet. 2004. Electoral rules and environmental policy. *Economics Letters* 84 (2): 237–44.

Gollop, F., and J. Roberts. 1983. Environmental regulations and productivity growth: The case of fossil-fueled electric power generation. *Journal of Political Economy* 91 (4): 654–74.

Gray, W. B. 1997. Manufacturing plant location: Does state pollution regulation matter? NBER Working Paper Series, No. 5880, NBER, Cambridge, Massachusetts.

Gray, W. B. and R. J. Shadbegian. 2003. Plant vintage, technology and environmental regulation. *Journal of Environmental Economics and Management* 43: 384–402.

Greenstone, M. 2002. The impact of environmental regulation on industrial activity: Evidence from the 1970 and 1977 Clean Air Acts Amendments and the Census of Manufacturers. *Journal of Political Economy* 110 (6): 1175–1219.

———. 2004. Did the Clean Air Act cause the remarkable decline in sulfur dioxide concentrations? *Journal of Environmental Economics and Management* 47: 585–611.

Grossman, G., and A. Krueger. 1991. Environmental impacts of a North American Free Trade Agreement. NBER Working Paper Series, No. 3914, NBER, Cambridge, Massachusetts.

Harris, M., L. Konya, and L. Matyas. 2003. Modelling the impact of environmental regulations on bilateral trade flows: OECD, 1990–1996. *World Economy* 25 (3) 387–405.

Hascic, I., and N. Johnstone. 2011. Innovation in electric and hybrid vehicle technologies: The role of prices, standards and R&D, in *Invention and Transfer of Environmental Technologies*, ed. OED, 85–125. OECD Publishing, Paris.

Henderson, V. 1996. Effects of air quality regulation. *American Economic Review* 86: 789–813.

Johnstone, N., I. Hascic, and D. Popp. 2010. Renewable energy policies and technological innovation: Evidence based on patent counts. *Environmental and Resource Economics* 45 (1): 133–55.

Joshi, S., R. Krishnan, and L. Lave. 2001. Estimating the hidden costs of environmental regulation. *Accounting Review* 76 (2): 171–98.

Jug, J., and D. Mirza. 2005. Environmental regulations in gravity equations: Evidence from Europe. *World Economy* 28 (11): 1591–1615.

Kahn, M., and M. Kotchen. 2011. Business cycle effects on concern about climate change: The chilling effect of recession. *Climate Change Economics* 2 (3): 257–73. OECD Publishing, Paris.

Kalamova, M., and N. Johnstone. 2011. Environmental policy stringency and foreign direct investment. OECD Environment Working Papers, No. 33.

Kellenberg, D. 2009. An empirical investigation of the pollution haven effect with strategic environment and trade policy. *Journal of International Economics* 78 (2): 242–55.

Keller, W., and A. Levinson. 2002. Pollution abatement costs and foreign direct investment inflows to U.S. *states. Review of Economics and Statistics* 84 (4): 691–703.

Levinson, A. 1996. Environmental regulations and manufacturers' location choices: Evidence from the Census of Manufactures. *Journal of Public Economics* 62 (1–2): 5–29.

———. 1999a. Grandfather regulations, new source bias, and state air toxics regulations. *Ecological Economics* 28: 299–311.

———. 1999b. State taxes and interstate hazardous waste shipments. *American Economic Review* 89 (3): 666–77.

Levinson, A., and M. Taylor. 2008. Unmasking the pollution haven effect. *International Economic Review* 49 (1): 232–54.

List, J. A., and C. Y. Co. 2000. The Effects of Environmental Regulations on Foreign Direct Investment. *Journal of Environmental Economics and Management* 40: 1–20.

List, J. A., W. W. McHone, and D. L. Millimet. 2004. Effect of environmental regulation on foreign and domestic plant births: Is there a home field advantage. *Journal of Urban Economics* 56: 303–26.

Millimet, D. L., and J. Roy. 2012. Three new tests of the pollution haven hypothesis when environmental regulation is endogenous. IZA Discussion Paper No. 5911. Institute for the Study of Labor, Bonn.

McConnell, V., and R. Schwab. 1990. The impact of environmental regulation on industry location decisions: The motor vehicle industry. *Land Economics* 66 (1): 67–81.

Morgenstern, R., W. Pizer, and J. Shih. 2001. The cost of environmental protection. *Review of Economics and Statistics* 83 (4): 732–38.

Pasurka, C. 2008. Perspectives on pollution abatement and competitiveness: Theory, data, and analyses. *Review of Environmental Economics and Policy* 2 (2): 194–218.

Pearce, D., and C. Palmer. 2001. Public and private spending for environmental protection: A cross-country policy analysis. *Fiscal Studies* 22 (4): 403–56.

Portney, Paul R., ed. 1990. *Public policies for environmental protection.* Washington, DC: Resources for the Future.

Raspiller, S., and N. Riedinger. 2008. Do environmental regulation influence the location behavior of French firms? *Land Economics* 84 (3): 382–95

Shadbegian, R. J., and W. B. Gray. 2012. Spatial patterns in regulatory enforcement: local tests of environmental justice. In The political economy of environmental justice, ed. H. S. Banzhaf. Stanford, CA: Stanford University Press. p. 225.

Smarzynska, B. J., and S. J. Wei. 2004. "Pollution havens and foreign direct investment: Dirty secret or popular myth? *B.E. Journal of Economic Analysis and Policy* 3 (2): 8.

Statistics Canada, Environment Accounts and Statistics Division. 1995. Environmental Protection Expenditures in the Business Sector. Statistics Canada, Ottawa.

———. 2000. Environmental Protection Expenditures in the Business Sector. Statistics Canada, Ottawa.

Stavins, R. 2006. Vintage-differentiated environmental regulation. *Stanford Environmental Law Journal* 25 (1): 29–63.

Tobey, J. 1990. The effects of domestic environmental policies on patterns of world trade: An empirical test. *KYKLOS* 43 (2): 191–209.

U.S. Census Bureau. 2008. *Pollution abatement costs and expenditures: 2005.* Washington, DC: U.S. Government Printing Office.

Van Soest, D., J. List, and T. Jeppesen. 2005. Shadow prices, environmental stringency, and international competitiveness. *European Economic Review.* 50 (5): 1151–67.

Wagner, U., and C. Timmins. 2009. Agglomeration effects in foreign direct investment and the pollution haven hypothesis. *Environmental and Resource Economics* 43 (2): 231–56.

Walter, I., and J. Ugelow. 1979. Environmental policies in developing countries. *Ambio* 8: 102–9.

Xing, Y., and C. D. Kolstad. 2002. Do lax environmental regulations attract foreign investment? *Environmental and Resource Economics* 21: 1–22.