Something in the Air: How Policy Affects Air Quality

Tristan Misko*

29 April 2022

Abstract

This paper addresses the extent to which national policy affects air quality in a cross-country comparison context. I construct a dataset from the ECOLEX database which uses the keyword coding ECOLEX provides to categorize policies across countries by their design and mechanisms. Tracking the implementation of policies across years, I estimate the effect of various policy categories in a time-series fixed effects regression at the country-year level to track the trends in pollution from road transport and electric power generation sources.

The second part of the paper addresses the extent to which markets respond to changes in air quality, specifically in the context of the United States Housing Market. I use wildfire smoke exposure, which varies quasi-randomly from county to county, as a source of exogenous variation in air quality to estimate a causal effect of air quality on housing prices.

Contents

*Acknowledgements: Thanks to Professor Clair Brown for her generous advising throughout the Honors Thesis process. David Card, Silvia Fregoni, Joaquin Fuenzalida, Thomas Tressel, and Matthew Swanson contributed ideas and asked incisive questions which benefitted this paper. Many ideas about policy analysis were formed in the context of work on the SSPI URAP team; I learned a lot from past and present members which has contributed to the writing of this paper.

1 Ambient Air Quality: A Pressing Concern

By the mid-twentieth century, ambient air pollution, especially in urban areas, became a serious public health concern and subsequently the focus of policy in many high-income countries (Rietze, 2001). A slate of pollutants including NOX, SOX, NMVOCs, $PM_{2.5}$, PM_{10}, and Carbon Monoxide are emitted by various consumer facing and industrial sources (OECD). For decades it has been known that ambient sulfur and nitrogen oxides produce acid rain (Likens et al., 1972), which can devastate freshwater ecosystems (Schindler, 1988). Recent research consistently shows that these pollutants large negative health and environmental consequences; particulate matter, especially $PM_{2.5}$, "pose[s] a hazard to public health even at low levels" (Feng et al., 2016). The threats that air pollutants pose to human and environmental well-being are grave, and the costs often accrue downstream of production processes, far from the original sources of pollution.

2 The Importance of Policy

Government policy has a uniquely important role in addressing sustainability since "business as usual" market mechanisms cannot effect the necessary changes with enough speed or impact to prevent catastrophic environmental outcomes. Individual economic agents like firms and households do not fully bear the costs of actions which degrade shared environmental resources like ambient air quality, so governments must design incentives and implement regulations to resolve externalities. Moreover, governments must intervene with forward-looking policy that overcomes the tendency for individual agents to act without considering the costs of their actions on the long-run well-being of current and future generations.

Policies set at the national and international level are particularly important in addressing the atmospheric pollution that leads to unhealthy ambient air quality and climate change. Atmospheric dispersal of pollutants means that air pollution cannot be geographically contained; its effects spread across regions and the globe. Hence, it is incumbent upon national governments, the largest unit of interpersonal and social organization, and intergovernmental organizations to implement policies that promote sustainability. For the most effective response, national governments must coordinate their policy responses with each other.

Given the importance of government policies in addressing air pollution, it is crucial that governments implement effective policies to make real progress on critical outcomes. But how should governments allocate scarce resources toward policies in this uncertain environment? The evaluation of policy efficacy is challenging. In a cross-country context and in complicated systems, the causal analysis approach to economic research often struggle to recover a significant, unbiased effect when the sample size is small (bounded by the number of countries) and when the number of omitted variables is large (the policy system is highly interconnected and complex). The difficulty in setting up well-controlled natural experiments to evaluate the efficacy of these policies cannot stand in the way of decisive policy action: the situation is too urgent. In this paper, I attempt to balance the concerns of empirical rigor and environmental urgency by using the best available empirical tools given the available data.

3 A Review of the Literature on Sustainability Policy Evaluation

This research began with my studying nitrogen management policies and their effects. A number of papers which focused on nitrogen policy were influential in structuring my policy analysis for air quality. Dalgaard et al. (2020) reviews in detail the nitrogen management policies which Denmark implemented in the period from 1985 to 2011. They analyze the time series responses of a set of outcome variables of interest to the implementation of policies targeted at those variables. The authors bring their specific institutional knowledge of Denmark agricultural policy to bear, hand coding and classifying policies according to a schema which distinguishes between input and output based policies; command and control, market based, and information/voluntary action polices; and geographically targeted policies and general regulations. The rich coding schema allows for the authors to distinguish between the effects of different kinds of policy at a granular level, providing useful feedback to policymakers.

One limitation of the paper is the limited geographic scope, which poses challenges for external validity and generalizability. The institutional knowledge that makes the paper so compelling is also a limitation for generalizability in the sense that their approach requires expert knowledge of the policies and institutions of a particular country in a particular sector. The challenge that motivated my work in this paper was to find a way to obtain a reasonably detailed policy dataset across many countries without having to hand code a large dataset.

Kanter et al. (2020) take a different from Dalgaard et al. by relying less on institutional knowledge and more on data mining. They gather nitrogen management policy data around the world from the ECOLEX database using keywords searches, collecting the country, year, name, keywords, and abstract of a policy. I follow their approach to policy data collection by collecting the same variables from ECOLEX that relate to air quality. One problem they encountered that I also address in my analysis is how to account for the number of policies passed by a country. A few countries pass many more policies than average, so it is difficult to tell the extent to which the number of policies passed represents a greater commitment to reducing pollution or reflects differences in the legislation or regulation procedures of the country. Kanter et al. take the number of policies passed as a “unit of analysis,” which is a step that brings up concern about the potential for institutional differences across countries: see my discussion below of the relationship between the average size or scope of a policy and the number of laws passed in a given country.

There are relatively few available cross-country analyses of the effects of air quality policy in general, although there are a number of papers which study the effect of policies in specific sectors. Kodjak (2015) provides an analysis of the fuel efficiency standards policies across the G20 countries, concluding that a focus on heavy-duty vehicle emissions is the most effective approach to reducing the emissions of vehicles. The OECD reports the Environmental Policy Stringency Index constructed by Botta and Koźluk (2014), which represents the best effort I have found at compiling a set of quantitative standards, taxes on emissions, and other policy variables. They aggregate these quantitative standards into a single score reflecting the degree of stringency of a country’s environmental policy regime. Unfortunately, the most recent data stop in 2012 and there appears to be no plan for updating the work.

Both of these papers represent attempts to understand a system of policies in terms of a few key quantitative metrics, which has a few important limitations. One concern is that policy metrics may not be exactly comparable across countries. For example, carbon tax policy can often be complex, with exemptions, deductions, rebates, and other nuances complicating the simple story that the nominal legal tax rate may tell. In both papers, the authors attempt to account for some of these differences across countries but acknowledge the complexity of the policy system. In addition, the work of tracking changes in the quantitative metrics by country is labor intensive and constrained by language in the cross country context, often limiting the potential sample size of this kind of analysis.

This paper pulls together an original dataset from the ECOLEX database in the spirit of Kanter et al. (2020) and matches these policy variables with a rich set of air pollution outcome variables and controls at the country-year level. Using the ECOLEX keyword approach provides less granularity and specificity than the quantitative metrics approach to measuring policy used by Kodjak (2015) and Botta and Koźluk (2014). In particular, the policy keywords do not furnish us with enough information to draw conclusions about how ambitious a particular policy is. What is lost in granularity is gained in the extensibility and generalizability of the analysis to a large sample of countries over a long period of time, however. By using the ECOLEX keyword data, the availability of outcome data becomes the binding constraint on the breadth analysis instead of the number of countries for which a labor-intensive policy parsing and coding can be conducted. I organize my analysis in a framework inspired by a "systems approach" to air quality and air pollution, conducting a detailed analysis of the effect of policy on road pollution and electricity power generation and showing how the approach can be generalized to other sources of air pollution.

4 Theory of Air Quality Policy: A Systems Approach

4.1 Setting Up the Structural Equation

Consider a set of n pollutants (e.g. $PM_{2.5}$, SO_2, CO, etc.) stored in the $n \times 1$ vector p and a set of d possible sources (e.g. vehicle emissions, construction, power generation, etc.) stored in the $d \times 1$ vector s. Let each component of s represent a proxy of the amount of a source behavior over a given period (e.g. average number of miles driven per year for vehicle emissions). The two vectors are related by an $n \times d$ matrix A, so that the equation $As = p$ holds. I assume that each source of pollution adds some linear contribution to the overall amount of each pollutant and that there are no interaction terms, so that each row of A encodes the relative weights on each source for a particular pollutant. Hence, the matrix equation above encodes a system of n equations of the form

$$p_j = A_{j1}s_1 + A_{j2}s_2 + \ldots + A_{jd}s_d.$$

Each coefficient A_{ji} corresponds to the rate at which increasing the source behavior increases the amount of pollutant in the air. For example, the coefficient on average distance driven per capita for pollutant j corresponds to the rate at which an increase in driving contributes to the amount of pollutant j observed in the atmosphere.

The linearity assumption seems reasonable on the surface: driving twice as much in a given period should cause approximately twice as much pollution, with a rough doubling of, say, the $PM_{2.5}$, SO_2, and CO emissions from vehicles. The simple story that the linearity assumption suggests may be leaving out important details. Perhaps an increase in the amount of driving increases the use of older, less efficient vehicles, so that increasing the amount of driving produces nonlinear increases in pollution. It is possible to extend this model by adding polynomial features to the vector s to account for such nonlinearities, but without compelling evidence of nonlinear behavior, the linearity should be sufficient if I assume that source variables vary somewhat smoothly.

There is a structural equation of the form $A_t s_t = p_t$ for each t over a series of annual observations for each country. The equation $A_t s_t = p_t$ has nd coefficients contained in the matrix A_t at each time t. Since I am assuming that the pollutants do not interact with one another, I treat the structural matrix equation as a set of n independent linear equations. The coefficients in A_t are *not* estimated via linear regression. Instead, I back out the coefficients contained in A_t from data which breaks down air pollution by source. For pollutant j, a emissions source i contributes a known share $A_{jit} s_{it}$ reported in the data, so dividing this share by s_i, also an observed value, allows us to estimate A_{ji}, which is not unobserved directly. Hence, the units of A_{jit} are quantity of pollutant p_{jt} per amount of source behavior s_{it}. The structural equation can be thought of as a useful form of organizing a basic accounting for the sources of emissions.

Over time, we observe variation in the quantities of pollutants in the air, which we see as variation in the components of p_t. We can use the structural equation to decompose the observed differences in a time series, tracking $p_{t+1} - p_t = A_{t+1}s_{t+1} - A_t s_t$. By tracking the evolution of A and s separately in time, we are able to untangle the extent to which observed changes in pollution levels over time are attributable to changes in the levels of source behavior (the difference $s_i^{(t+1)}$ - $s_i^{(t)}$) and which changes are attributable to changes in how much pollution the source behavior generates (the difference $A_{ji}^{(t+1)} - A_{ji}^{(t)}$).

4.2 The Interaction of Policy with the Structural Equation

The time dependent set of structural equations provides a rich framework for understanding the channels by which national policy acts. By thinking about the air pollution system in terms of its emissions sources s, their generation intensities A, and the outcomes p, there are then distinct mechanisms for policy to bring about changes in the amount of air pollution.

1. **Level of Source Behavior:** A policy might discourage certain high pollution behaviors by changing incentives, requiring permits, spreading information, or through other mechanisms. A gasoline tax is a canonical example of a policy which acts to reduce pollution from vehicle emissions by reducing the amount of driving (an example of a source behavior s) that occurs inside a country. This manifests as a change in s_i. Below, references to "source behavior" indicate the level of s_i.

2. **Pollution Generation Intensity of Source Behavior:** A policy might change the rate at which engaging in the source behavior contributes to air pollution. A policy which sets and enforces more stringent emissions standards for vehicles will decrease the amount of pollution due to driving without necessarily decreasing the amount of driving that occurs. This manifests as change in the coefficient A_{ji}, the generation intensity associated with source behavior i for pollutant j. Below, mentions

of "generation intensity," "air pollution generation intensity," or "generation intensity of source behavior" refer to the quantity A_{ji}, with j and i often indicated by context.

3. **Direct Capture of Pollutants:** This type of policy aims to decrease the stock of pollutants p in the air by acting directly on the pollutants in the environment. Since these policies are not widely implemented, I do not model them here, although in the Discussion below I outline a natural extension of the structural model should direct capture policies become more salient.

4.3 Modeling the Source Behavior and Generation Intensity

In the analysis that follows, I treat each s_i and each A_{ji} as a set of outcome variables whose variation we would like to explain in terms of policy and some control variables. Below, I take $j = \{\text{CO, NOX, NMVOC, PM}_{2.5}, \text{PM}_{10}, \text{SOX}\}$, with $n = 6$. Hence, for each source behavior i (below I consider a $d = 2$ case) there are seven regressions, one for the underlying source behavior s_i and one for each of the six A_{ji}.

The structural provides a roadmap for organizing and aggregating the outcome variables A and s, but it imposes no constraints on the way that A and s are modeled. Let s_{it} indicate the level of some source variable, for example total passenger kilometers driven, in country i in year t, and let A_{jit} denote the pollution generation intensity for pollutant j of the associated source in country i in year t.

We employ a fixed effects linear model with appropriate time-dependent controls x_{it} for each source variable as available along with a set of time-dependent dummies π_{it} which turn on when various policies are implemented. I discuss the particular models in extensive detail below. Writing out the models, I have

$$
\begin{aligned}
s_{it} - \overline{s}_i &= x_{it}^T \beta + \pi_{it}^T \delta + \epsilon, \\
A_{jit} - \overline{A}_{ji} &= x_{it}^T \beta + \pi_{it}^T \delta + \epsilon.
\end{aligned}
$$

There are a number of important subtleties associated with setting up these models. One choice concerns whether to tally the keywords or use a simple dummy. Below I try the model both ways and compare the results, preferring the tally approach. Another potential subtlety concerns correlations in pollution from drift between neighboring countries. The potential for spatial autocorrelation is real and requires some sophisticated modeling to overcome which is beyond the scope of this paper, so I assume that spillovers across borders are not important for this analysis and I proceed with caution.

4.4 Modularity of the Structural Equation

The ability to set up the structural model as described depends on the form of the available data and the particular definitions of the categories of pollution sources, since certain source variables are more suited toward the selection of a single proxy measure than others. For example, the variable "Pollution from Road Transit" has a fairly natural source proxy of "Total Passenger-Kilometers Travelled", allowing the division procedure detailed above to obtain an estimate of the associated coefficient in the matrix A. A variable like "Pollution from Industrial Processes" admits no easy proxy measure, however, because it is difficult to imagine a single variable that captures the notion of the level of industrial processing that occurs in a country for which there is data comparable across countries.

Fortunately, the structural model is modular in the sense the researcher can select which sources to model and which to leave as observed values. In this setting, the we have a structural model which looks something like $p = As + v$, where v is an $d \times 1$ vector of pollution amounts from unmodeled sources. Hence, the structural model allows for domain specific research to be contextualized in the larger air quality system as available instead of requiring the full modeling of the system, which would likely prove to be unmanageably complex. Below, I discuss the share of air pollution which is modeled by the sources I am modeling.

5 Policy Data by Country

The policy data are sourced from the ECOLEX database, which catalogs information on environmental law and policy across countries. The ECOLEX database contains the type of law (Legislation, Decision, Treaty), the name of the law, the year of implementation, and, most importantly, a keyword description of the law in English. The keywords are standardized across all variables, with 426 unique keywords available. I reproduce a typical entry in the database:

> **Country:** Argentina
> **Type:** Legislation
> **Name:** Decreto 3970/90 - Reglamentación de la Ley 5965.
> **Year:** 1990
> **Keywords:** Pollution control, Air quality/air pollution, Emissions, Environmental standards, Offences/penalties, Waste disposal, Effluent waste water/discharge, Sewerage, Freshwater quality/freshwater pollution, Water quality standards

The database sometimes contains further information like an abstract of the law, links to the full text, and further categorization information. I use these data when available to help determine which laws relate to the particular outcome variable of interest. The ECOLEX search bar does search across these fields, and I describe below I use this functionality to improve the sample of laws used in the analysis. Kralj et al. (2020) develop a rich framework for text analysis of the further information in the ECOLEX database via a document embedding procedure, which characterizes each document as a vector based on the frequency and placement of words used in the text. Their classification scheme could be used to extend the analysis performed here by developing a more nuanced scheme for classifying policies, but such an endeavor is beyond the scope of this paper.

5.1 Keywords as Selection Criteria and Policy Descriptors

Of the 426 unique keywords, a handful are particularly important for this analysis. ECOLEX is maintained in part by FAO, the Food and Agriculture Organization of the United Nations. The FAO provides detailed definitions of the keywords used in the ECOLEX dataset, which are reproduced for the most important keywords used in this analysis in Table 1 (FAOLEX).

The “air quality/air pollution” keyword signifies laws which contain text directly addressing air pollution. The presence of this keyword in the description of a law is taken to indicate that a policy addresses the generation intensity of a source behavior directly. It is possible that some policies may affect the air pollution generation intensity of a source behavior without explicitly mentioning a keyword. For example, a policy which promotes the adoption of electric vehicles through subsidies may not explicitly mention the effect that replacing internal combustion vehicles with electric vehicles has on the generation intensity of road transport, although certainly the effect exists. One limitation of the approach taken here is the difficulty of finding and accounting for such policies, as relevant search terminology must be included on a case-by-case basis.

Overall, the web-scraping approach taken here certainly risks excluding relevant policies because ultimately the selection process is entrusted to a computer program and not a human. The quality of the policy dataset could certainly be improved by a hand-coding procedure executed by experts in the policies of interest, but the automated procedure implemented here achieves a result approaching what could be done by hand-coding at a small fraction of the cost in time and effort.

5.2 Determining which Policies to Include in the Analysis

Determining the appropriate set of policies to include in the analysis is imperative. The ECOLEX database contains 158,838 policies across 237 countries and territories in the period from 1990 to 2019, which amounts to an average of over 22 policies per country per year. The scope of countries and territories included and the near-completeness of laws in the ECOLEX database ensure that the constraints on the analysis are imposed by the availability of outcome data and not by policy data.

For this analysis, the air pollution outcome data are sourced from OECD, so only policies implemented across the 38 OECD countries are considered. ECOLEX contains 64,475 laws passed across the OECD countries during the 1990 to 2019 period, for an average of 56.5 policies per country per year.

I use keywords and search terms to select only the laws which might be relevant for the outcome variable of choice. I discuss below why it is important to include only the relevant policies in order to estimate unbiased effects.

5.2.1 Road Transport Policies

The main step in the analysis is to filter via keyword and text search for policies relevant for driving. The filtering is done in two stages to obtain two distinct policy datasets for each source type. The first stage determines which policies relate to driving by using a text search, which searches across the law abstract and other metadata, for word fragments like ”automo*”, ”vehic*”, ”road”, ”highway”, ”truck”, ”car”. The laws which contain these patterns in their metadata are filtered further using ”Domain” and ”Primary Subject” Metadata from the ECOLEX database when it is available. These filters remove laws which mention vehicles in some capacity but which have a ”Primary Subject” listed as, for example, ”Livestock.” It is unlikely that laws which are primarily related to livestock have significant bearing on the country aggregated amount of driving or generation intensity, so these laws are removed from the analysis. At the end of this process, the filtered data contain laws which might plausibly affect the amount of driving in a country, so this data frame represents the policies which might plausibly affect the source behavior of road transport, i.e. the amount of driving.

The second stage of filtering produces a data frame which contains only the subset of the laws in the source behavior data frame which contain text which addresses air quality or air pollution directly, which is obtained by filtering out laws which do not contain the "air.quality.air.pollution" or "emissions" keywords. This choice of filter hinges on the reasoning that laws which do not have text addressing air pollution from the source are unlikely to be important determinants of the amount of air pollution released per unit activity from that source. The data frame which emerges from the second stage of filtering becomes the generation intensity policy data frame.

5.2.2 Power Generation Data

The same two stage process was applied to power generation data. The first stage determines which policies are relevant for the energy generation source behavior, primarily using the the "energy.conservation.energy.production" keyword. As it turns out, countries implement *many* policies to address energy, with 1778 implemented from 1990 to 2019 across the 38 OECD countries. This poses some problems for the analysis as mentioned in the Discussion Section below.

The second stage again selects for laws mention air quality directly, which creates a comparatively small sample of laws.

6 Air Pollution Data

6.1 Selection of Data

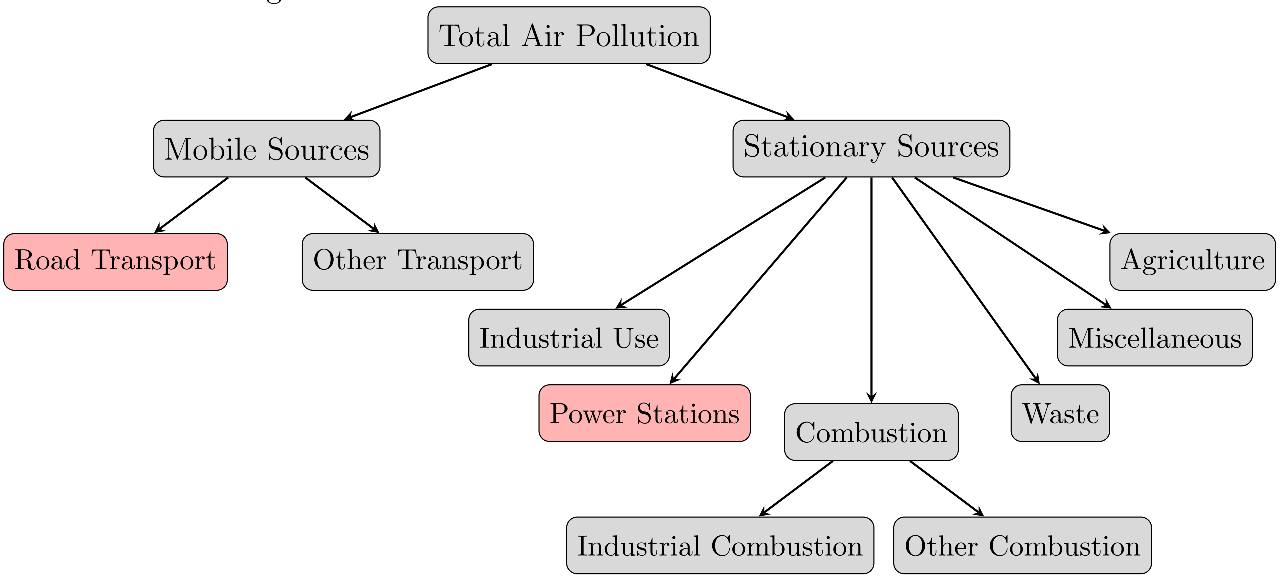
There are many potential sources of data on air pollution available. The OECD Statistics source was selected because it seems to be most suited to my analysis of sources I examined:

1. Data Reliability: OECD Statistics is a well-regarded data source for cross-country comparisons, offering clear documentation, sourcing, and descriptions of variables.

2. Level of Granularity in Time and Place: Many air quality data sources report daily values sourced in particular cities and localities which only go back a few years, which makes cross-country difficult because an aggregation scheme must be selected to analyze. The OECD source takes care of the aggregation, offering values from 1980 to 2019 for air quality at the country-year level, which is more suitable for a policy analysis than daily data.

3. Level of Granularity in Pollution Type: The OECD Pollution Dataset tracks six types of air pollution: Carbon Monoxide (CO), Non-Methane Volatile Organic Compounds (NMVOCs), Nitrogen Oxides (NOX), Coarse Particulates (PM10), Fine Particulates ($PM_{2.5}$), and Sulfur Oxides (SOX). These pollutants are the main pollutants policymakers track.

Figure 1: OECD Statistics Classification of Air Pollution



4. Level of Granularity in Pollution Source: The OECD categorizes air pollution according to the classification schema depicted in the chart in Figure 1. The OECD breaks down the amount of each of the six pollutants which comes from each of the nine source categories listed in the leaves of the tree below.

6.2 Determining which Source Behaviors to Model

Applying the discussion of the modularity of the structural equation above, I select two of the nine emissions source categories offered by OECD at the country-year level to model as a function of policy and some controls for this analysis. The categories modeled are **Road Transport** (summarized in Table 3) and **Power Stations** (summarized in Table 4) The categories are selected from the tree depicted in Figure 1. These two sources offer natural proxy variables for source behavior. I proxy Road Transport by Passenger-Kilometers per Capita and Power Stations by Total Energy Consumption per Capita. The other sources in this dataset do not readily admit the single variable proxy demanded by the structural equation, so I represent these as the unmodeled factors in the vector v.

This gives the following form for the structural equation:

$$\begin{bmatrix} p_{\text{CO}} \\ p_{\text{NMVOC}} \\ p_{\text{NOX}} \\ p_{\text{PM10}} \\ p_{\text{PM25}} \\ p_{\text{SOX}} \end{bmatrix} = \begin{bmatrix} A_{\text{CO, RT}} & A_{\text{CO, PS}} \\ A_{\text{NMVOC, RT}} & A_{\text{NMVOC, PS}} \\ A_{\text{NOX, RT}} & A_{\text{NOX, PS}} \\ A_{\text{PM10, RT}} & A_{\text{PM10, PS}} \\ A_{\text{PM25, RT}} & A_{\text{PM25, PS}} \\ A_{\text{SOX, RT}} & A_{\text{SOX, PS}} \end{bmatrix} \begin{bmatrix} s_{\text{RT}} \\ s_{\text{PS}} \end{bmatrix} + \begin{bmatrix} v_{\text{CO}} \\ v_{\text{NMVOC}} \\ v_{\text{NOX}} \\ v_{\text{PM10}} \\ v_{\text{PM25}} \\ v_{\text{SOX}} \end{bmatrix},$$

where s_{RT} denotes the total passenger kilometers travelled per year, s_{PS} denotes total energy consumption by country, and the quantities $A_{\text{i,RT}}s_{\text{RT}}$ and $A_{\text{i,PS}}s_{\text{PS}}$ are the observed totals in the OECD data for pollutant i.

This form of the structural equation motivates an analysis of the percentage of the air pollution being modeled. In Table 5, the sum of the amount of pollution from Road Transport and Power Generation sources is divided by the total amount of air pollution for each pollutant, and the mean across all available years is reported for each country. For example, the first entry under the P.CO column header in Table 5 shows that 48.55% of the overall carbon monoxide pollution in Australia is accounted for by Road Transport and Power Generation sources. The countries with *NA* values have a few missing values, so the group averaging methodology produced divide-by-zero errors. These countries are omitted in regressions for years in which there are *NA* values. The table shows that on average road transport and power generation sources account for a significant proportion of total CO (41% accounted for), NOX (56% accounted for), and SOX (35% accounted for). NMVOC, PM_{10}, and $PM_{2.5}$ have an average of around 15% of emissions accounted for by road transport and power generation, so the effect of policy is less important for these pollutants.

7 Source Behavior Data and Control Variables

7.1 Road Transport Data

The data on road transport present a number of complications. The OECD reports data on road passenger transport in Passenger Kilometers and data on goods transport in Tonne Kilometers. Ideally, the source behavior variable s_{RT} would be denominated in total vehicle-kilometers, but cross-country data on total vehicle kilometers travelled from the OECD (which is by far the most complete data set I could find) ranges only from 2010 to 2018 with many missing values. Leaving out heavy duty vehicles may be an important omission, however. As mentioned in the Kodjak (2015) paper, heavy duty vehicles involved in goods transport are some of the most important targets for pollution policy.

To deal with the data unavailability, imputations could be made to construct a proxy measure of vehicle-kilometers using the separate passenger and goods transport datasets. Using what vehicle-kilometers (VK_{it}) data is available as a regression output, estimates of weights for passenger-kilometers (PK_{it}) and tonne-kilometers (TK_{it}) are made using these more complete datasets via the model below:

$$VK_{it} = \alpha + \beta PK_{it} + \gamma TK_{it} + u_{it}.$$

The models were constructed with and without constants and with and without a population normalization.

The results are displayed in Table 6. Regressions (1) and (2) have no population normalization, yielding highly significant coefficients for both regressors. Accounting for country population in (3) and (4) gives an interesting result: the significance of tonne kilometers totally disappears, leaving only PK per capita as significant. This result justifies the use of Passenger Kilometers per Capita as a proxy for the source behavior variable in the base model. Model (4) could be used to generate a predicted vehicle-kilometers variable, but since the only the trends and not the levels of generation intensity and source behavior are important, this step is unnecessary.

7.2 Power Generation Data

The data on electricity generation by country are sourced from EIA. The analysis uses Total Electric Power Consumption[1] Total energy production for the United States also includes the production of biomass, geothermal, and solar energy not used for electricity generation. (TEPC) as a proxy for Total Net Electricity Generation, which is not available from the source. Here, net consumption differs from gross consumption in that "net consumption excludes the energy consumed by the generating units," e.g. the startup energy for powering on nuclear reactors. The EIA metadata describe the formula by which the TEPC is computed:

$$\begin{aligned}\text{Total Electric Power Consumption} &= \text{Total Net Electricity Generation} \\ &+ \text{Electricity Imports} - \text{Electricity Exports} \\ &- \text{Electricity Transmission and Distribution Losses.}\end{aligned}$$

Since imports and exports of electricity are small relative to total net electricity generation, TEPC is a viable proxy. Unfortunately, the transmission and distribution loss data are not available, so policy which improves grid efficiency will not have a visible effect in using this outcome data.

7.3 Control Variable Data

Other data from the OECD include GDP per capita, population, and rail transport in passenger-kilometers (a metric of public transit). Historical gas price data was sourced from the World Bank. All of these control variables were matched to the observations at the Country-Year Level.

[1]EIA also offers data on energy production, but this is not the correct outcome data for this analysis. The energy production series includes the production of petroleum, dry natural gas, and coal, which is measuring the amount of mining/extraction done, not the amount of energy used.

Table 1: A list of important ECOLEX Keywords

Air Quality/Air Pollution Air quality - The degree to which air is polluted; the type and maximum concentration of man-produced pollutants that should be permitted in the atmosphere. Air pollution -The general term alluding to the undesirable addition of substances (gases, liquids, or solid particles) to the atmosphere that are foreign to the natural atmosphere or are present in quantities exceeding natural concentrations.
Authorization/Permit: Authorization is the process of endowing or conferring a person with legal power or sanction to do something specific.
Basic Legislation: The prinicipal law within a particular sector or technical area.
Data Collection/Reporting: Data collection is the process of gathering and measuring information on targeted variables in an established system, which then enables one to answer relevant questions and evaluate outcomes. Reporting - the act of giving an official notification, for example about its accounts or activities.
Emissions: In the climate change context, emissions refer to the release of greenhouse gases and/or their precursors and aerosols into the atmosphere over a specified area and period of time.
Emissions Trading: Emissions trading (also known as cap and trade) is a market-based approach to controlling pollution by providing economic incentives for achieving reductions in the emissions of pollutants.
Energy Conservation/Energy Production: Energy conservation is the effort made to reduce the consumption of energy by using less of an energy service. This can be achieved either by using energy more efficiently or by reducing the amount of service used. Energy generation is the process of generating electric power from sources of primary energy.
Enforcement/Compliance: Policy implements or modifies a mechanism for the enforcement of compliance with standards.
Environmental Standards: Environmental standards are standards for materials, products and production processes to ensure that negative impacts on the environment are minimal or kept within certain limits.
Monitoring: Monitoring is a periodically recurring task already beginning in the planning stage of a project or programme. Monitoring allows results, processes and experiences to be documented and used as a basis to steer decision-making and learning processes. Monitoring is checking progress against plans. The data acquired through monitoring is used for evaluation.;
Nuclear Energy: Energy released by reactions within atomic nuclei, as in nuclear fission or fusion.
Pollution Control: Pollution control is a term used in environmental management. It means the control of emissions and effluents into air, water or soil. Without pollution control, the waste products from overconsumption, heating, agriculture, mining, manufacturing, transportation and other human activities, whether they accumulate or disperse, will degrade the environment.
Public Health: Public health means the health of individuals in the context of the wider health of the community.
Renewable Energy: Energy sources that are, within a short time frame relative to the Earth's natural cycles, sustainable, and include non-carbon technologies such as solar energy, hydropower, and wind, as well as carbon-neutral technologies such as biomass.
Standards: Provisions establishing requirements against which natural resources and products must conform. Standards may be "compulsory" or "recommended."
Subsidy/Incentive: Subsidies-Payment or benefit given to partially offset the cost of specific activities, such as the manufacture, production, or export of an article. Incentives - Generally refers to financial incentives (compensation, interest subsidies, production subsidies, tax exemption, tax reduction, grants, bonuses, rewards, loans with low interest rates, etc.) set by a government in order to support the management and development of a given sector.

Table 2: Count of Keywords Across the Four Policy Datasets

	Road Transport		Power Generation	
Keyword/Law Type	Source Behavior	Generation Intensity	Source Behavior	Generation Intensity
air.quality.air.pollution	110	110	26	26
authorization.permit	11	2	196	0
basic.legislation	4	3	107	8
bioenergy	2	2	112	8
biofuel	6	5	102	7
Constitution	0	0	0	0
data.collection.reporting	21	9	174	4
electric.vehicle	1	1	0	0
emissions	79	73	29	29
emissions.trading	2	2	3	3
energy.conservation.energy.production	11	4	1778	54
enforcement.compliance	17	15	162	4
environmental.standards	88	79	231	7
hydropower.generation	0	0	196	0
Legislation	23	13	443	26
monitoring	19	13	23	2
nuclear.energy	3	0	79	5
Policy	5	2	11	5
pollution.control	128	111	80	30
public.health	11	5	7	1
Regulation	127	95	1309	23
renewable.energy	8	6	588	37
standards	89	80	238	7
subsidy.incentive	5	2	276	7

Table 3: Mobile Source Air Pollution from Road Transit, 2000-2019, OECD Countries

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
YEAR	727	2009	5.696	2000	2005	2014	2019
CO	725	1,632.453	5,930.237	3.939	80.545	644.321	61,744.890
NMVOC	719	175.822	520.748	0.594	10.810	109.699	4,831.200
NOX	725	374.567	1,014.831	2.112	42.253	337.695	9,377.977
PM10	637	39.463	465.517	0.367	2.322	21.578	11,701.000
PM2-5	606	15.094	36.381	0.248	1.759	13.868	298.980
SOX	707	6.349	24.197	0.003	0.077	1.574	259.292

Table 4: Stationary Source Air Pollution from Power Generation, 1990-2019, OECD Countries

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
YEAR	1,075	2004	8.572	1990	1997	2012	2019
CO	1,064	55.304	176.699	0.001	2.564	34.511	1,318.000
NMVOC	1,040	3.161	7.062	0.000	0.293	2.979	56.111
NOX	1,064	220.961	685.327	0.002	10.507	205.896	6,044.674
PM10	887	21.427	80.398	0.000	0.597	8.514	1,109.000
PM2-5	845	13.698	55.219	0.000	0.411	4.903	563.540
SOX	1,064	460.267	1,547.191	0.001	5.171	268.573	14,432.650

Table 5: Mean Proportion of Total Air Pollution Accounted for from Road Transport and Power Generation Sources by Pollutant Type

	P.CO	P.NMVOC	P.NOX	P.PM10	P.PM2.5	P.SOX
Australia	0.4855	0.1839	0.3645	*NA*	*NA*	0.2395
Austria	0.1853	0.0936	0.6411	0.2162	0.3016	0.1306
Belgium	0.2207	0.1184	0.5513	0.2163	0.2333	0.1078
Canada	0.3956	0.0982	0.3924	0.0069	0.0209	0.2551
Switzerland	0.5141	0.1667	0.5913	0.2313	0.2463	0.0617
Chile	0.0869	0.0230	0.5721	0.1234	0.0430	0.2285
Costa Rica	0.7249	0.6076	0.8079	*NA*	*NA*	*NA*
Czech Republic	0.2289	0.1278	0.5844	0.1389	0.1354	0.5872
Germany	0.4037	0.1216	0.6257	0.2069	0.3117	0.4297
Denmark	0.4378	0.1394	0.4853	0.1257	0.1536	0.2853
Spain	0.2359	0.1039	0.5205	0.1506	0.1676	0.5247
Estonia	0.2270	0.1590	0.5781	0.4663	0.4068	0.8319
Finland	0.3063	0.1319	0.5324	0.2754	0.1550	0.4282
France	0.2085	0.1079	0.5714	0.1635	0.1974	0.2044
Greece	0.5393	0.3272	0.5917	0.2257	0.2332	0.6425
Hungary	0.3814	0.2151	0.5290	0.0874	0.0960	0.5013
Ireland	0.7165	0.0993	0.4919	0.1205	0.1910	0.4476
Iceland	0.1401	0.2098	0.1140	0.1792	0.2067	0.0009
Israel	0.9629	*NA*	0.8512	*NA*	*NA*	0.7822
Italy	0.3617	0.2134	0.5279	0.1807	0.1813	0.2345
Japan	0.3403	0.0983	0.3959	*NA*	*NA*	0.2302
Korea	0.6863	0.1114	0.5403	0.3183	*NA*	0.2954
Lithuania	0.2796	0.1194	0.5958	0.1653	0.2229	0.2010
Luxembourg	0.5486	0.1832	0.7637	0.4798	0.5034	0.0588
Latvia	0.2373	0.1347	0.4677	0.0697	0.0817	0.2251
Mexico	0.8143	0.4183	0.7515	0.1208	0.1387	0.4880
Netherlands	0.5515	0.1483	0.4891	0.2145	0.2361	0.2025
Norway	0.1995	0.0746	0.2461	0.1023	0.0857	0.0905
New Zealand	0.7239	0.5773	0.5848	*NA*	*NA*	0.2136
Poland	0.2802	0.1418	0.6084	0.1409	0.1603	0.6288
Portugal	0.3846	0.1761	0.6368	0.0921	0.1204	0.4271
Russia	0.7312	*NA*	0.7980	*NA*	*NA*	0.3548
Slovak Republic	0.2081	0.0945	0.5081	0.1246	0.1306	0.5504
Slovenia	0.2801	0.1295	0.6287	0.1035	0.0976	0.6087
Sweden	0.3579	0.1683	0.5138	0.4246	0.2476	0.1772
Turkey	0.3977	0.1402	0.7364	0.1254	0.1260	0.6062
United Kingdom	0.4175	0.1185	0.5295	0.1825	0.2056	0.4078
United States	0.5000	0.1997	0.5568	0.0360	0.1040	0.6932
Average	0.4132	0.1653	0.5599	0.1530	0.1510	0.3521

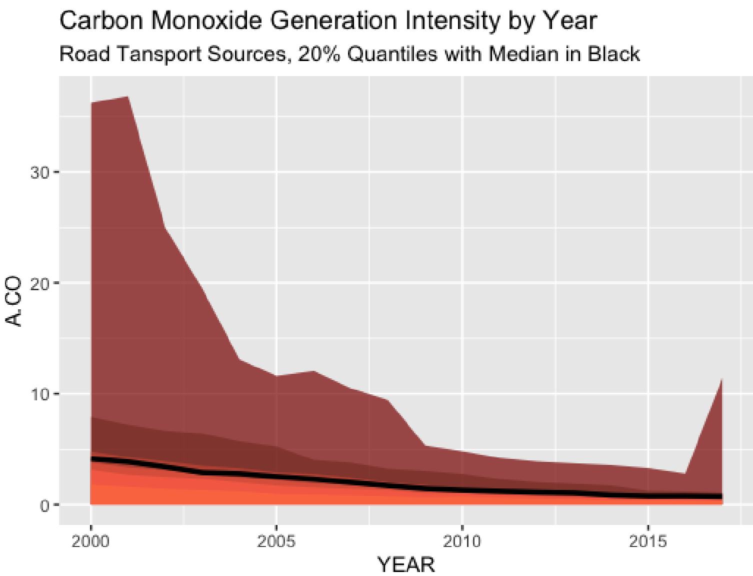
Table 6: Modeling Vehicle Kilometers with Tonne Kilometers and Passenger Kilometers

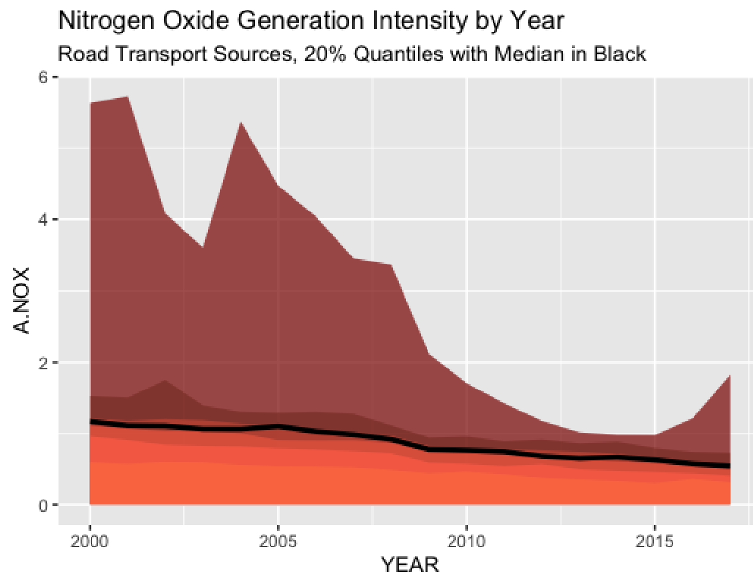
	Dependent variable:			
	VehicleKilometers		VKperCap	
	(1)	(2)	(3)	(4)
TonneKilometers	0.656***	0.596***		
	(0.062)	(0.059)		
PassengerKilometers	0.486***	0.523***		
	(0.027)	(0.026)		
TKperCap			−0.090*	0.013
			(0.049)	(0.052)
PKperCap			0.674***	0.802***
			(0.023)	(0.036)
Constant		−40,368.530***		−2.018***
		(6,584.419)		(0.446)
Observations	276	276	276	276
R^2	0.993	0.992	0.914	0.652
Adjusted R^2	0.992	0.992	0.913	0.649
Note:		*p<0.1; **p<0.05; ***p<0.01		

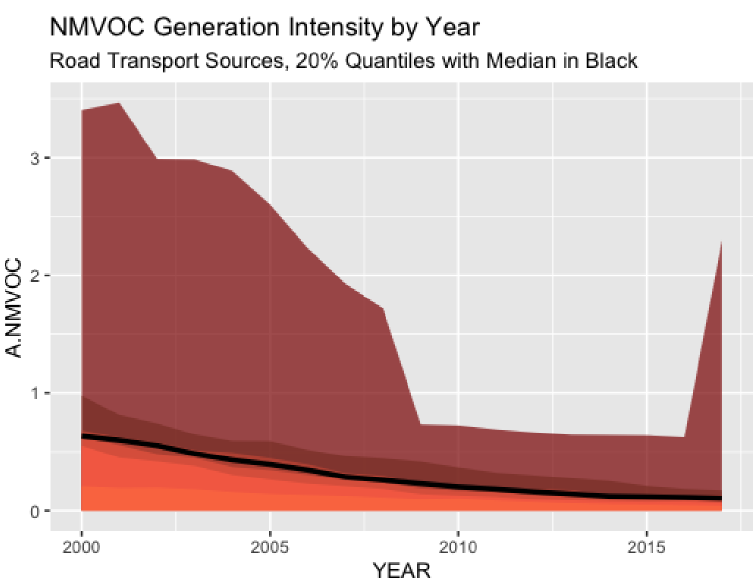
Table 7: Availability of Road Data by Country in Combined Dataset (Policy, Outcome, and Control Variables Included)

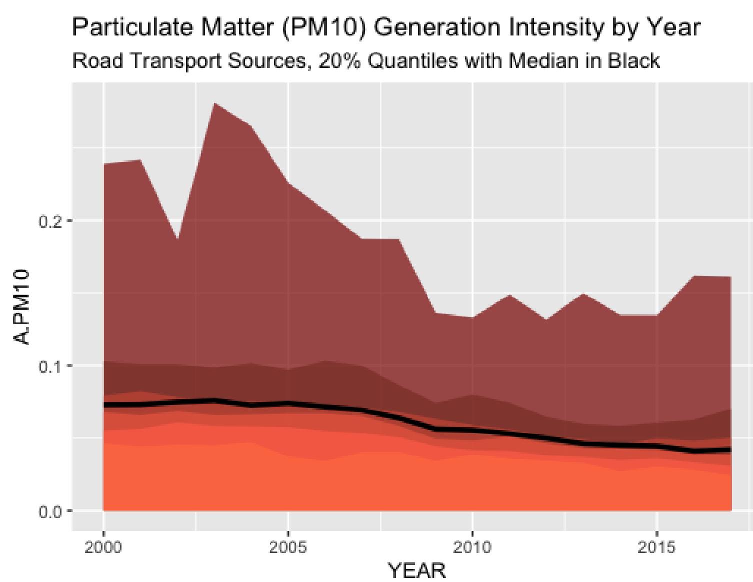
Country	**N Years with Data**	**Country**	**N Years with Data**
Australia	19	Netherlands	9
Belgium	17	New Zealand	12
Canada	8	Norway	20
Czech Republic	20	Poland	20
Denmark	17	Portugal	9
Finland	20	Russia	13
France	20	Slovak Republic	20
Germany	20	Slovenia	11
Greece	9	Spain	20
Hungary	20	Sweden	20
Iceland	20	Switzerland	20
Italy	20	Turkey	20
Japan	20	United Kingdom	20
Korea	18	United States	18
Lithuania	17		

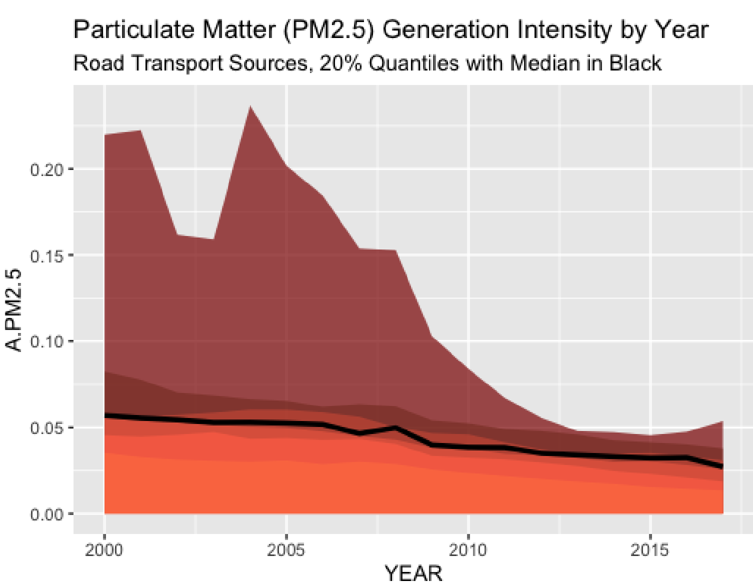
Total Observations: 438 country-year observations across 29 of the 38 OECD countries from 2000 to 2019.


Carbon Monoxide Generation Intensity by Year
Road Transport Sources, 20% Quantiles with Median in Black
A.CO
0
10
20
30
2000
2005
2010
2015
YEAR


Nitrogen Oxide Generation Intensity by Year
Road Transport Sources, 20% Quantiles with Median in Black
A.NOX
0
2
4
6
2000
2005
2010
2015
YEAR


NMVOC Generation Intensity by Year
Road Transport Sources, 20% Quantiles with Median in Black
A.NMVOC
0
1
2
3
2000
2005
2010
2015
YEAR


Particulate Matter (PM10) Generation Intensity by Year
Road Transport Sources, 20% Quantiles with Median in Black
A.PM10
0.0
0.1
0.2
2000
2005
2010
2015
YEAR


Particulate Matter (PM2.5) Generation Intensity by Year
Road Transport Sources, 20% Quantiles with Median in Black
A.PM2.5
0.00
0.05
0.10
0.15
0.20
2000
2005
2010
2015
YEAR

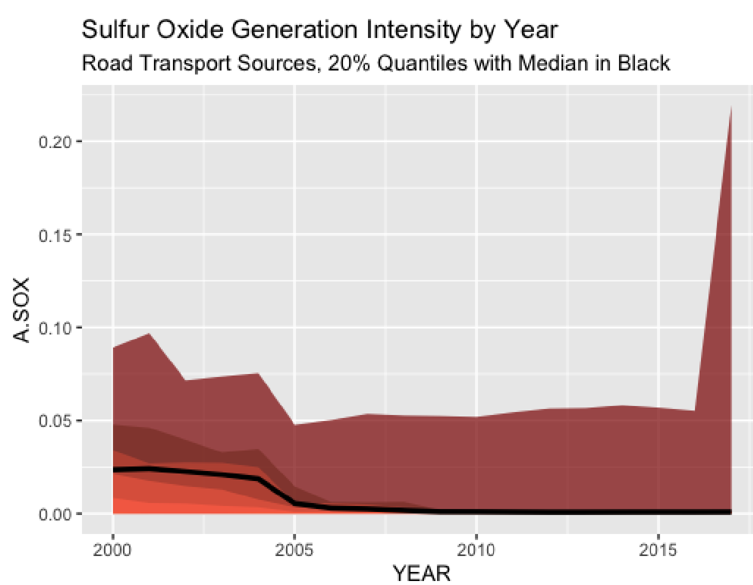
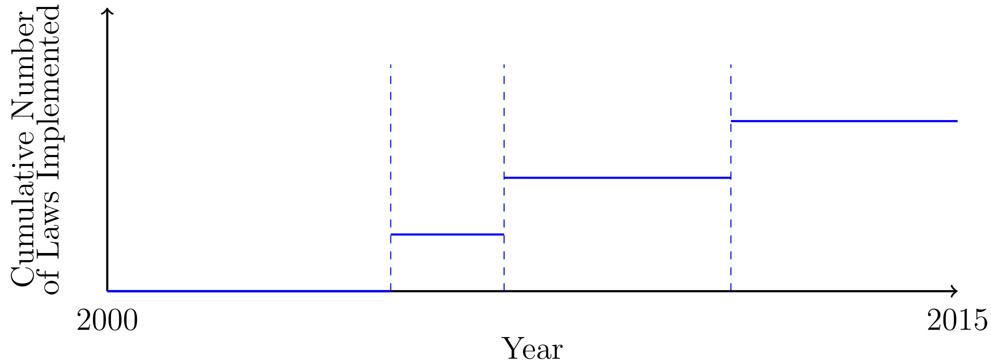

Sulfur Oxide Generation Intensity by Year
Road Transport Sources, 20% Quantiles with Median in Black
A.SOX
0.00
0.05
0.10
0.15
0.20
2000
2005
2010
2015
YEAR

Figure 2: Example Structure of the Policy "Cumulative Dummies," with Policy Implementation Dates Indicated by Steps in the Graph



8 Statistical Models

I now turn to estimating the fixed effects models described above. I estimate four different models for each outcome variable of interest. All models include year fixed effects but not country fixed effects, as including country fixed effects produced overcontrolled models with R^2 values over 0.90. Year fixed effects can be thought of as accounting for technological progress and other unobserved factors which affect all countries in the sample differently at different times.

I illustrate the form of the policy variables used in the analysis in Figure 2. As policies containing a specific keyword are implemented, the "cumulative dummy" for that keyword counts up, creating a step-like function. This gives all policies equal weight so that the regression can estimate the average effect of a policy containing the keyword while using the time structure to compare multiple pre- and post-implementation periods at a time.

At its core, each model consists of an outcome variable modeled by an appropriate selection of the cumulative policy variables, a set of appropriate control variables, and the year fixed effects. With this structure in mind, the regression tables give the details of exactly which specification is being used.

The regressions in Table 8 to Table 21 report the coefficients and fixed effects robust standard errors.

Table 8: The Effect of Policy on Road Transport Air Pollution Source Behavior (Passenger-Kilometers per Capita)

	Dependent variable:			
	PKperCap			
	(1)	(2)	(3)	(4)
Legislation	−0.693***	−0.294***		
	(0.122)	(0.075)		
Regulation	−0.099	0.079		
	(0.088)	(0.051)		
Policy	2.809***	4.212***		
	(0.661)	(0.545)		
authorization.permit			−1.989***	−1.766***
			(0.444)	(0.339)
basic.legislation			1.373	1.137**
			(0.902)	(0.552)
data.collection.reporting			2.173***	−0.111
			(0.741)	(0.522)
electric.vehicle			6.174***	5.035***
			(1.096)	(0.751)
emissions			−0.060	−0.626***
			(0.242)	(0.167)
emissions.trading			−1.953	−3.632***
			(1.320)	(0.787)
enforcement.compliance			1.286***	−0.221
			(0.318)	(0.214)
environmental.standards			−0.534	−0.709***
			(0.346)	(0.183)
pollution.control			−0.263	1.041***
			(0.466)	(0.286)
public.health			−2.801***	−0.668**
			(0.890)	(0.308)
subsidy.incentive			0.180	−0.391
			(0.773)	(0.450)
gas.price.usd.per.liter	−4.388***	−3.310***	−3.879***	−2.763***
	(0.784)	(0.325)	(0.692)	(0.341)
RAILperCAP	0.269	−1.348***	0.054	−1.287***
	(0.264)	(0.144)	(0.238)	(0.155)
Lagged2YInfraSpend	−2.488***	−0.308	−1.310***	−0.212
	(0.493)	(0.356)	(0.494)	(0.369)
GDPperCAP		0.0003***		0.0003***
		(0.00001)		(0.00001)
Observations	414	414	414	414
R^2	0.184	0.670	0.327	0.709
Adjusted R^2	0.136	0.650	0.272	0.684

Table 9: The Effect of Policy on Road Transport Air Pollution CO Generating Intensity

	Dependent variable:			
	A.CO			
	(1)	(2)	(3)	(4)
Legislation	40.704**	54.549***		
	(16.471)	(18.651)		
Regulation	−24.609***	−20.327***		
	(7.458)	(7.789)		
authorization.permit			−185.814	−210.106
			(116.350)	(132.194)
basic.legislation			47.657	40.828
			(77.677)	(94.056)
data.collection.reporting			−142.996***	−189.504***
			(47.265)	(54.225)
electric.vehicle			58.955	96.669*
			(48.222)	(55.217)
enforcement.compliance			−81.868**	−111.565***
			(36.773)	(42.414)
environmental.standards			−49.364**	−32.536
			(24.659)	(29.157)
pollution.control			40.494*	42.639*
			(21.995)	(24.812)
public.health			196.514**	309.520***
			(77.399)	(114.145)
subsidy.incentive			182.796	109.648
			(125.719)	(145.940)
gas.price.usd.per.liter	−762.507***	−732.254***	−772.918***	−732.767***
	(130.872)	(116.597)	(132.061)	(113.135)
RAILperCAP	−124.405***	−164.932***	−147.758***	−208.237***
	(30.321)	(39.478)	(37.238)	(50.237)
Lagged2YInfraSpend	−287.414***	−235.477***	−323.779***	−274.206***
	(53.852)	(48.329)	(61.676)	(55.232)
GDPperCAP		0.006**		0.009**
		(0.003)		(0.003)
Observations	414	414	414	414
R^2	0.291	0.307	0.304	0.332
Adjusted R^2	0.251	0.266	0.251	0.279

Note: *p<0.1; **p<0.05; ***p<0.01

Table 10: The Effect of Policy on Road Transport Air Pollution NMVOC Generating Intensity

	Dependent variable:			
	A.NMVOC			
	(1)	(2)	(3)	(4)
Legislation	4.210***	5.167***		
	(1.475)	(1.614)		
Regulation	−2.428***	−2.132***		
	(0.676)	(0.713)		
authorization.permit			−32.360***	−34.242***
			(10.469)	(11.668)
basic.legislation			13.645*	13.116
			(6.997)	(8.188)
data.collection.reporting			−14.877***	−18.481***
			(4.336)	(4.815)
electric.vehicle			8.340**	11.263**
			(4.170)	(4.684)
enforcement.compliance			−8.774***	−11.075***
			(3.309)	(3.694)
environmental.standards			−3.330	−2.026
			(2.270)	(2.619)
pollution.control			2.966	3.133
			(1.997)	(2.214)
public.health			18.505***	27.263***
			(6.795)	(9.475)
subsidy.incentive			5.992	0.324
			(11.117)	(12.890)
gas.price.usd.per.liter	−64.563***	−62.471***	−65.838***	−62.727***
	(10.751)	(9.671)	(10.752)	(9.259)
RAILperCAP	−11.155***	−13.957***	−12.936***	−17.623***
	(2.645)	(3.189)	(3.260)	(4.098)
Lagged2YInfraSpend	−23.415***	−19.824***	−27.784***	−23.942***
	(4.570)	(4.322)	(5.205)	(4.851)
GDPperCAP		0.0004*		0.001**
		(0.0002)		(0.0003)
Observations	414	414	414	414
R^2	0.303	0.314	0.320	0.343
Adjusted R^2	0.264	0.274	0.269	0.292
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 11: The Effect of Policy on Road Transport Air Pollution NOX Generating Intensity

	Dependent variable:			
	A.NOX			
	(1)	(2)	(3)	(4)
Legislation	16.045***	17.116***		
	(3.710)	(4.010)		
Regulation	−5.436***	−5.105***		
	(1.658)	(1.759)		
authorization.permit			−104.480***	−107.338***
			(25.494)	(27.525)
basic.legislation			52.078***	51.274***
			(16.834)	(18.741)
data.collection.reporting			−53.688***	−59.160***
			(14.970)	(15.259)
electric.vehicle			4.285	8.723
			(10.850)	(12.609)
enforcement.compliance			−16.343*	−19.837**
			(8.868)	(9.379)
environmental.standards			−15.669**	−13.689*
			(7.414)	(8.303)
pollution.control			13.502**	13.755**
			(6.571)	(6.993)
public.health			45.206***	58.502***
			(16.756)	(22.453)
subsidy.incentive			−27.286	−35.892
			(26.917)	(29.844)
gas.price.usd.per.liter	−119.669***	−117.329***	−122.982***	−118.258***
	(26.175)	(24.038)	(26.078)	(23.215)
RAILperCAP	−21.840***	−24.974***	−25.647***	−32.762***
	(6.365)	(7.247)	(7.596)	(9.318)
Lagged2YInfraSpend	−52.240***	−48.224***	−62.325***	−56.492***
	(10.726)	(9.900)	(12.353)	(11.392)
GDPperCAP		0.0005		0.001*
		(0.001)		(0.001)
Observations	414	414	414	414
R^2	0.230	0.233	0.262	0.273
Adjusted R^2	0.187	0.188	0.206	0.217
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 12: The Effect of Policy on Road Transport Air Pollution PM_{10} Generating Intensity

	Dependent variable:			
	A.PM10			
	(1)	(2)	(3)	(4)
Legislation	0.769***	0.916***		
	(0.214)	(0.226)		
Regulation	−0.371***	−0.335***		
	(0.091)	(0.094)		
authorization.permit			−8.623***	−9.656***
			(1.297)	(1.330)
basic.legislation			5.396***	6.204***
			(0.828)	(0.824)
data.collection.reporting			−2.570***	−3.018***
			(0.749)	(0.772)
electric.vehicle			−0.257	0.0005
			(0.661)	(0.708)
enforcement.compliance			−0.706	−0.948**
			(0.440)	(0.468)
environmental.standards			−0.635*	−0.498
			(0.356)	(0.396)
pollution.control			0.445	0.460
			(0.314)	(0.339)
public.health			0.744	0.980
			(0.759)	(0.850)
subsidy.incentive			−3.084***	−4.121***
			(0.919)	(0.975)
gas.price.usd.per.liter	−5.854***	−5.448***	−5.960***	−5.511***
	(1.244)	(1.083)	(1.281)	(1.107)
RAILperCAP	−2.206***	−2.937***	−2.340***	−3.184***
	(0.445)	(0.490)	(0.541)	(0.563)
Lagged2YInfraSpend	−2.926***	−2.463***	−3.512***	−3.109***
	(0.462)	(0.441)	(0.523)	(0.495)
GDPperCAP		0.0001**		0.0001***
		(0.00002)		(0.00003)
Observations	377	377	377	377
R^2	0.301	0.319	0.328	0.356
Adjusted R^2	0.257	0.275	0.272	0.301
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 13: The Effect of Policy on Road Transport Air Pollution $PM_{2.5}$ Generating Intensity

	Dependent variable:			
	A.PM2.5			
	(1)	(2)	(3)	(4)
Legislation	0.416***	0.467***		
	(0.159)	(0.165)		
Regulation	−0.233***	−0.212***		
	(0.069)	(0.072)		
authorization.permit				
			(0.000)	(0.000)
basic.legislation			0.921*	1.058**
			(0.477)	(0.485)
data.collection.reporting			−2.597***	−2.827***
			(0.642)	(0.648)
electric.vehicle			−0.162	−0.028
			(0.492)	(0.544)
enforcement.compliance			−0.678*	−0.803**
			(0.365)	(0.382)
environmental.standards			−0.613**	−0.542*
			(0.303)	(0.329)
pollution.control			0.515*	0.522*
			(0.271)	(0.285)
public.health			0.397	0.519
			(0.677)	(0.752)
subsidy.incentive			1.361**	1.100*
			(0.557)	(0.577)
gas.price.usd.per.liter	−3.595***	−3.394***	−3.531***	−3.302***
	(0.995)	(0.884)	(1.015)	(0.899)
RAILperCAP	−1.784***	−2.161***	−1.942***	−2.377***
	(0.377)	(0.409)	(0.448)	(0.468)
Lagged2YInfraSpend	−2.144***	−1.905***	−2.562***	−2.349***
	(0.370)	(0.363)	(0.432)	(0.420)
GDPperCAP		0.00003		0.00004*
		(0.00002)		(0.00002)
Observations	369	369	369	369
R^2	0.255	0.263	0.278	0.290
Adjusted R^2	0.207	0.214	0.218	0.230
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 14: The Effect of Policy on Road Transport Air Pollution SOX Generating Intensity

	Dependent variable:			
	A.SOX			
	(1)	(2)	(3)	(4)
Legislation	−0.030	0.045		
	(0.076)	(0.093)		
Regulation	−0.037*	−0.014		
	(0.021)	(0.021)		
authorization.permit			−2.125***	−2.237***
			(0.410)	(0.418)
basic.legislation			1.038***	1.007***
			(0.248)	(0.240)
data.collection.reporting			−0.059	−0.274
			(0.144)	(0.182)
electric.vehicle			0.457***	0.631***
			(0.154)	(0.191)
enforcement.compliance			−0.149	−0.286*
			(0.112)	(0.149)
environmental.standards			−0.092	−0.014
			(0.065)	(0.081)
pollution.control			0.065	0.075
			(0.059)	(0.069)
public.health			−0.177	0.344
			(0.305)	(0.426)
subsidy.incentive			−1.166**	−1.504***
			(0.479)	(0.445)
gas.price.usd.per.liter	−2.442***	−2.278***	−2.518***	−2.332***
	(0.584)	(0.524)	(0.600)	(0.527)
RAILperCAP	0.032	−0.187	0.066	−0.213
	(0.142)	(0.195)	(0.175)	(0.250)
Lagged2YInfraSpend	−0.231	0.051	−0.270	−0.041
	(0.257)	(0.200)	(0.290)	(0.238)
GDPperCAP		0.00003***		0.00004***
		(0.00001)		(0.00001)
Observations	414	414	414	414
R^2	0.210	0.243	0.214	0.254
Adjusted R^2	0.166	0.198	0.154	0.196
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 15: The Effect of Policy on Power Generation Air Pollution Source Behavior (Energy Consumption per Capita)

	Dependent variable:			
	s			
	(1)	(2)	(3)	(4)
Legislation	−0.002***			
	(0.0002)			
Regulation	0.00001			
	(0.0001)			
Policy	−0.023***			
	(0.005)			
bioenergy		0.006*		
		(0.003)		
biofuel		−0.004		
		(0.003)		
hydropower.generation		−0.006***		
		(0.001)		
nuclear.energy		0.004**		
		(0.002)		
renewable.energy		−0.002***		
		(0.001)		
authorization.permit			0.003*	
			(0.001)	(0.000)
basic.legislation			−0.005***	−0.004
			(0.001)	(0.005)
data.collection.reporting			0.009***	0.061***
			(0.001)	(0.012)
emissions			0.009**	0.004
			(0.004)	(0.005)
emissions.trading			0.017	−0.051***
			(0.014)	(0.008)
enforcement.compliance			0.001	−0.018**
			(0.001)	(0.007)
environmental.standards			−0.008***	0.058***
			(0.001)	(0.006)
pollution.control			−0.024***	−0.043***
			(0.003)	(0.007)
public.health			−0.029***	−0.025
			(0.006)	(0.023)
subsidy.incentive			−0.001	0.0002
			(0.001)	(0.016)
GDPperCAP	0.004***	0.004***	0.005***	0.005***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Observations	978	978	978	978
R^2	0.416	0.426	0.498	0.438
Adjusted R^2	0.396	0.404	0.476	0.414
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 16: The Effect of Policy on Power Generation CO Generation Intensity

	Dependent variable:			
	A.CO			
	(1)	(2)	(3)	(4)
Legislation	−7.127***			
	(2.038)			
Regulation	−1.674***			
	(0.578)			
Policy	−210.164***			
	(55.494)			
bioenergy		−109.640***		
		(24.883)		
biofuel		75.109***		
		(21.739)		
hydropower.generation		−42.348***		
		(9.030)		
nuclear.energy		−31.689***		
		(5.868)		
renewable.energy		5.775		
		(3.766)		
authorization.permit			−70.113***	
			(16.910)	(0.000)
basic.legislation			−31.274***	13.227
			(5.681)	(40.622)
data.collection.reporting			35.062***	−295.145***
			(12.565)	(84.440)
emissions			57.560**	16.494
			(23.501)	(34.590)
emissions.trading			−269.053***	−346.416***
			(84.372)	(64.311)
enforcement.compliance			11.668**	165.897***
			(5.007)	(53.384)
environmental.standards			−51.152***	115.484***
			(12.339)	(36.247)
pollution.control			−123.249***	−66.728*
			(27.711)	(34.188)
public.health			−221.944***	304.052**
			(52.165)	(134.735)
subsidy.incentive			79.807***	59.358
			(17.021)	(79.257)
GDPperCAP	−9.240***	−9.349***	−13.734***	−9.283***
	(1.893)	(1.829)	(2.819)	(2.253)
Observations	967	967	967	967
R^2	0.057	0.069	0.098	0.049
Adjusted R^2	0.024	0.034	0.059	0.009
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 17: The Effect of Policy on Power Generation NMVOC Generation Intensity

	Dependent variable:			
	A.NMVOC			
	(1)	(2)	(3)	(4)
Legislation	−0.211**			
	(0.083)			
Regulation	0.170***			
	(0.030)			
Policy	2.993			
	(7.510)			
bioenergy		−0.149		
		(1.420)		
biofuel		0.547		
		(1.503)		
hydropower.generation		−0.846		
		(0.529)		
nuclear.energy		−0.911**		
		(0.440)		
renewable.energy		0.501**		
		(0.206)		
authorization.permit			0.617	
			(0.499)	(0.000)
basic.legislation			−0.564***	4.446*
			(0.175)	(2.395)
data.collection.reporting			−1.612**	−23.393***
			(0.716)	(7.091)
emissions			1.572	2.491
			(1.382)	(2.400)
emissions.trading			−26.913***	−23.061***
			(6.207)	(5.939)
enforcement.compliance			0.606**	10.626***
			(0.277)	(2.723)
environmental.standards			1.280*	14.905***
			(0.715)	(3.738)
pollution.control			−2.621***	−9.521***
			(0.868)	(2.074)
public.health			−3.971**	−1.752
			(1.927)	(6.067)
subsidy.incentive			−0.150	7.914
			(0.514)	(9.943)
GDPperCAP	−0.074	−0.103*	0.008	−0.028
	(0.051)	(0.053)	(0.061)	(0.066)
Observations	943	943	943	943
R^2	0.031	0.037	0.063	0.063
Adjusted R^2	−0.004	−0.00004	0.022	0.022
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 18: The Effect of Policy on Power Generation NOX Generation Intensity

	Dependent variable:			
	A.NOX			
	(1)	(2)	(3)	(4)
Legislation	−23.499***			
	(4.091)			
Regulation	2.091*			
	(1.165)			
Policy	−74.327			
	(155.071)			
bioenergy		−315.412***		
		(64.445)		
biofuel		102.051*		
		(57.912)		
hydropower.generation		−132.745***		
		(17.875)		
nuclear.energy		−76.093***		
		(14.820)		
renewable.energy		58.099***		
		(9.277)		
authorization.permit			−12.716	
			(32.522)	(0.000)
basic.legislation			−80.036***	204.309
			(11.857)	(135.794)
data.collection.reporting			−0.425	−1,609.860***
			(29.823)	(431.354)
emissions			342.811***	344.667***
			(83.348)	(132.081)
emissions.trading			−1,034.875***	−1,148.763***
			(203.882)	(355.675)
enforcement.compliance			7.026	85.465
			(12.626)	(163.759)
environmental.standards			−55.377**	825.165***
			(25.020)	(152.353)
pollution.control			−156.701***	−467.872***
			(58.725)	(113.218)
public.health			−461.079***	1,267.067***
			(96.835)	(388.522)
subsidy.incentive			46.400*	−110.788
			(23.830)	(271.651)
GDPperCAP	−10.243**	−12.897***	−11.717**	−7.260
	(4.123)	(4.008)	(5.417)	(5.007)
Observations	967	967	967	967
R^2	0.029	0.049	0.043	0.035
Adjusted R^2	−0.005	0.013	0.002	−0.005
Note:			*p<0.1; **p<0.05; ***p<0.01	

Table 19: The Effect of Policy on Power Generation PM_{10} Generation Intensity

	Dependent variable:			
	A.PM10			
	(1)	(2)	(3)	(4)
Legislation	−2.330***			
	(0.689)			
Regulation	−0.067			
	(0.264)			
Policy	−48.660*			
	(29.043)			
bioenergy		−17.504*		
		(10.046)		
biofuel		5.618		
		(10.844)		
hydropower.generation		−15.602***		
		(3.177)		
nuclear.energy		−8.648***		
		(2.052)		
renewable.energy		3.277**		
		(1.667)		
authorization.permit			−14.845**	
			(6.717)	(0.000)
basic.legislation			−8.502***	−19.609*
			(1.520)	(10.552)
data.collection.reporting			−7.773	17.840
			(4.865)	(54.204)
emissions			23.532**	−5.835
			(9.970)	(15.365)
emissions.trading			−94.749**	−49.641
			(40.559)	(39.130)
enforcement.compliance			4.566***	18.053
			(1.554)	(16.476)
environmental.standards			8.237	−26.433
			(5.599)	(24.101)
pollution.control			−18.167	−10.174
			(12.190)	(16.910)
public.health			−45.392	71.498
			(27.947)	(64.030)
subsidy.incentive			9.064	34.922
			(5.557)	(29.326)
GDPperCAP	−2.757***	−2.889***	−3.246***	−2.768**
	(0.898)	(0.849)	(1.172)	(1.083)
Observations	797	797	797	797
R^2	0.053	0.062	0.063	0.049
Adjusted R^2	0.012	0.019	0.014	0.0003

Note: *p<0.1; **p<0.05; ***p<0.01

Table 20: The Effect of Policy on Power Generation $PM_{2.5}$ Generation Intensity

	Dependent variable:			
	A.PM2.5			
	(1)	(2)	(3)	(4)
Legislation	−1.552***			
	(0.367)			
Regulation	0.126			
	(0.104)			
Policy	−12.404			
	(8.960)			
bioenergy		−4.490		
		(5.602)		
biofuel		−0.944		
		(7.061)		
hydropower.generation		−9.154***		
		(1.639)		
nuclear.energy		−5.498***		
		(1.222)		
renewable.energy		2.235**		
		(1.104)		
authorization.permit			−10.555***	
			(3.299)	(0.000)
basic.legislation			−5.612***	−14.010**
			(1.057)	(6.415)
data.collection.reporting			−6.130*	−13.520
			(3.428)	(29.422)
emissions			21.152***	2.529
			(6.657)	(8.064)
emissions.trading			−62.732***	−44.235**
			(16.968)	(21.059)
enforcement.compliance			3.108***	13.344
			(1.071)	(9.529)
environmental.standards			11.846***	−3.607
			(4.013)	(12.843)
pollution.control			−11.994**	−12.682
			(5.522)	(10.393)
public.health			−26.306***	32.886**
			(9.683)	(15.054)
subsidy.incentive			2.431	17.966
			(2.295)	(11.669)
GDPperCAP	−0.850*	−0.985**	−0.976	−0.790
	(0.471)	(0.445)	(0.605)	(0.622)
Observations	767	767	767	767
R^2	0.035	0.045	0.051	0.033
Adjusted R^2	−0.008	−0.001	−0.001	−0.019
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 21: The Effect of Policy on Power Generation SOX Generation Intensity

	Dependent variable:			
	A.SOX			
	(1)	(2)	(3)	(4)
Legislation	−59.052***			
	(10.560)			
Regulation	7.134*			
	(4.087)			
Policy	830.481			
	(746.904)			
bioenergy		−949.943***		
		(250.269)		
biofuel		324.297*		
		(189.090)		
hydropower.generation		−246.723***		
		(45.782)		
nuclear.energy		−202.426***		
		(45.914)		
renewable.energy		144.182***		
		(28.210)		
authorization.permit			307.406**	
			(125.779)	(0.000)
basic.legislation			−210.356***	1,315.638**
			(40.056)	(600.079)
data.collection.reporting			−240.204**	−4,624.204***
			(104.692)	(1,391.793)
emissions			1,088.794***	1,699.239***
			(285.527)	(562.261)
emissions.trading			−2,090.965***	−2,024.989***
			(609.607)	(744.557)
enforcement.compliance			19.607	−1,627.602**
			(32.175)	(777.403)
environmental.standards			−289.614***	143.222
			(73.267)	(391.404)
pollution.control			−184.384	−1,339.885***
			(139.557)	(403.256)
public.health			−597.627**	8,064.535***
			(268.815)	(2,065.562)
subsidy.incentive			46.320	−1,024.470
			(75.692)	(1,192.211)
GDPperCAP	−57.124***	−66.169***	−59.069***	−46.522***
	(12.125)	(11.965)	(15.129)	(14.102)
Data	PGSB	PGSB	PGSB	PGGI
Observations	967	967	967	967
R^2	0.069	0.083	0.092	0.080
Adjusted R^2	0.036	0.048	0.052	0.042
Note:			*p<0.1; **p<0.05; ***p<0.01	

9 Discussion of Results

I organize the discussion in two separate subsections for the Road Transport and Power Generation. I subdivide these subsections into discussions of regressions for the source behavior outcome variables and regressions for the generation intensity outcome variables.

Each discussion section begins with a short description of the units of the variables, then proceeds into a discussion of the coefficients on the controls as a robustness check. Then the discussion turns to the policy variables themselves. The main focus of the discussion is on regressions (3) and (4), which offer a much more granular description of policies than regressions (1) and (2). Regressions (1) and (2) are included because they present a coarse-grained distinctions among policies to check the robustness of the methodology in addition to providing some insight into the behavior of the controls in the regressions. Comparison of the results of (1) and (2) to (3) and (4)

9.1 Road Transport

9.1.1 Source Behavior (Regression Table 8)

PKperCAP, the outcome variable in this regression, is denominated in units of millions of passenger kilometers per thousand people so that the coefficients are easily readable numbers. Rail passenger kilometers per capita is denominated in , and GDPperCAP

In the regressions for Passenger-Kilometers per Capita, the level of gas prices and rail passenger kilometers per capita are expected to have negative effects on the amount of driving which occurs in a given country, while lagged road infrastructure spending (as a percentage of GDP) and GDP per capita are expected to have positive effects on the amount of driving.

Gas prices exhibit a significant negative relationship in regressions (1), (2), (3), and (4), with large effect sizes relative to the estimates for policy implementation. The coefficient on gas price represents the expected decrease in passenger kilometers for a one USD per liter increase in gas prices, which represents a fairly large change in gas prices and therefore provides a basis for comparison of the effect size of policy. On average, the effect of an average policy is expected to be substantially smaller than the effect of a one USD increase in gas prices, which is generally what is observed in the regression results in (3) and (4).

The coefficients on rail per capita are not significant in regressions (1) and (3) but are negative and highly significant when the GDP per Capita control is added in regressions (2) and (4).

9.1.2 Generating Intensity (Regression Tables 9-14)

9.2 Power Generation

9.2.1 Source Behavior

9.2.2 Generating Intensities

9.3 The Importance of Restrictive Selection Criteria

Ultimately, the policy data are used in a time series regression to model the pollution level as a linear function of some time-dependent cumulative dummies and (see below for a detailed description of the setup). To avoid constructing an overdetermined model, it is important that the the average number of policies per year per country be small enough to untangle the effects of individual policies and policy types. In particular, if the number of policies per country per year is greater than one, then it is impossible for the regression to untangle the effects of the individual policies, especially if the effect of policy is not immediate.

Ideally, the number of relevant policies per country per year should be much smaller than one to be confident that the system does not suffer from overdetermination. This metric will serve as a necessary but not sufficient condition for obtaining good estimates of the effect of policy from this methodology and is reported for each dataset described below.

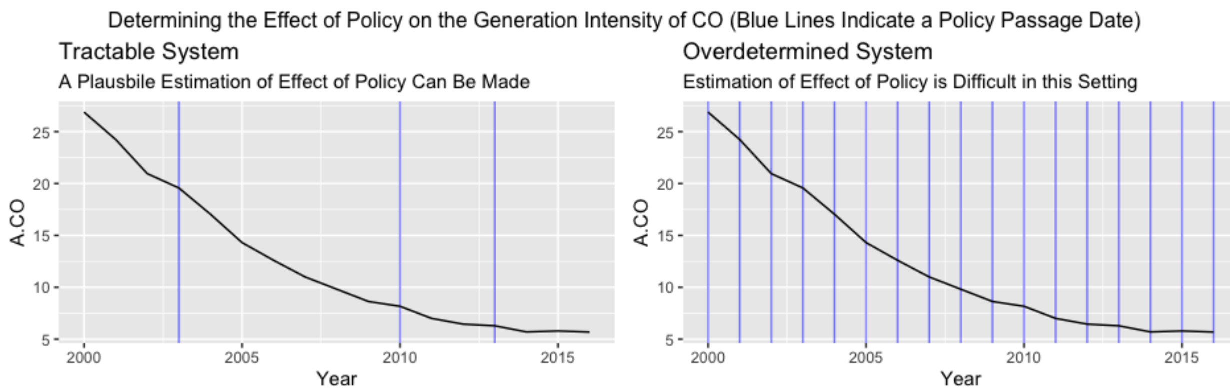


Figure 3: A Toy Example of the Importance of Restrictive Selection Criteria

Hence, a key step in the analysis is to determine which of the 64,475 policies implemented across the OECD from 1990 to 2019 might plausibly affect the outcome variables of interest. Including policies which have no plausible relationship to the outcome variable clutters the analysis and contaminates the estimated effect size by "soaking up" some of the effect. Keyword selection combined with the search function of the ECOLEX database are used to determine which policies are relevant for a particular outcome variable. The subsections below describe in detail the methodology for selecting the relevant policies for each outcome variable.

The reverse problem of omitting a relevant policy is also important. If a country implements a few policies during the period but one is omitted, then the other policies which are included in the analysis will soak up the effect of the omitted policy, inflating the estimated effect size of the included policies. Thus, the

A pattern emerged when selecting the relevant policies to include in the analysis: determining which policies might affect the level of source behavior poses a much greater challenge than determining those which affect the generation intensity. Policies which affect the generation intensity of a particular source behavior generally must mention the source behavior by name in order to affect the outcome, while source behavior levels can be affected through many complex mechanisms.

9.4 Power Generation Omitted Variables Bias

It is clear from the regression tables for Power Generation that the models are under-controlled, with R^2 values often less than 0.05 and all less than 0.10. The low R^2 indicates that the model is not explaining much of the variance in outcomes at all, suggesting that some other factors may be at play which are unaccounted for. The only obvious and widely available control for the amount of energy consumed I could find was GDP per capita. Adding other controls may help provide better estimates

9.5 Geography

Geographic features are an important factor that remains unaccounted for in this analysis. The geography of a country certainly contributes to the amount of driving which occurs in a country. Perhaps population density could be added as control

Further research could improve this analysis by controlling for the salient geographic factors