Structural Uncertainty and Optimal Pollution Control

Richard T. Carson∗ and Jacob LaRiviere†‡

February 2015

Abstract

We relax the economic literature’s usual assumption that the structural relationship between emissions and ambient levels of pollution is known with certainty. We find that uncertainty over this relationship can manifest as a unique form of multiplicative uncertainty in the marginal damages from emissions. We show how the optimal stringency of environmental regulation depends on this structural uncertainty. We also show how new information, like the discovery of previously unapportioned emission sources, can counterintuitively lead to increases in both optimal emissions and ambient pollution levels.

KEYWORDS: Information; Regulation; Externalities

JEL Codes: Q58, D81, C11, Q53

∗Department of Economics, UC San Diego.
†Corresponding author: Department of Economics, University of Tennessee. 525 Stokely Management Center, Knoxville, TN 37996-0550. jlarivi1@utk.edu.
‡Thanks to Allen Basala, Roger Brode, Chris Costello, Gerardo Dominguez, David Evans, Meredith Fowlie, Ben Gillen, Ken Gillingham, Ted Groves, Charlie Kolstad, Mark Machina, Steve Polasky, Glenn Sheriff, Kerry Smith, Chris Timmins, Mark Thiemens, Roger Von Haefen, Choon Wang, seminar participants at Tennessee, Minnesota, TREE, participants of the 12th Occasional California Workshop on Environmental and Resource Economics, two anonymous referees and Stephen Holland for very helpful comments and discussion.
1 Introduction

Environmental regulations can have substantial economics impacts (Hazilla and Kopp (1990) and Greenstone et al. (2011)). As a result, there is considerable interest in knowing the precise relationship between emissions and ambient pollution levels in order to develop optimal pollution regulations. As long as the relationship between emissions and ambient pollution is known, any uncertainty reduces to uncertainty in damages associated with a given level of ambient pollution or the costs borne by emitters for reducing emissions by a given amount. This type of uncertainty, which we define as “distributional uncertainty”, continues to receive considerable attention in the economics of pollution control.

In this paper, we examine the implications of a different kind of uncertainty we term “structural uncertainty”. Structural uncertainty occurs when the relationship between pollution emissions and ambient pollution levels is unknown. We show that this form of uncertainty, which has been little studied in the economics literature, gives rise to a unique set of modeling implications and regulatory issues.

A great deal of evidence in the physical science literature shows there is considerable uncertainty in the relationship between emissions and ambient concentrations of various pollutants. For example, it was recently discovered that up to 44% of sulfates in Los Angeles, the most studied air basin in the world, come from large ships burning bunker fuel in the Los Angeles-Long Beach harbors (Dominguez et al. (2008)). This far exceeded previous apportionments of ambient sulfate concentrations to harbor traffic. Even though sulfates were already known to come from ships, the level of contributions of their emissions to ambient pollution levels was very inaccurate. As a result, known sulfate emissions were “misappor-
tioned”. Recent work regarding the U.S. suggests that yearly anthropogenic emissions of methane, a powerful greenhouse gas (GHG), are 50-70% higher than previously thought (Miller et al. (2013)). Put another way, there are missing emissions inventories for methane. As a result, each ton of known methane was previously over-attributed to causing extant atmospheric methane concentrations. To that end, fugitive emissions from leaks in industrial processes of methane and Volatile Organic Compounds (VOCs) are thought to be both large and notoriously difficult to estimate (Chambers et al. (2008) and Wu et al. (2014)). Similarly, the relationship between mercury emissions and ambient mercury levels is subject to considerable uncertainty even though the health risks of mercury are now well-known. The uncertainty exists in large part since “it is difficult to obtain data [on mercury emissions] from national authorities” leading to uncertain emissions inventories (Pacyna et al. (2010)). Regardless, mercury emissions have been subject to recent regulatory action. In each case, and in many others especially in the developing world, ambient levels of pollution are known but the physical relationship between different emission sources and ambient pollution levels is unknown due to imprecision in attributing the set of total actual emissions to observed ambient pollution levels.\footnote{Haagen-Smit (1952) famously discovered that automobile exhaust was largely responsible for smog in the L.A. air basin. Automobile emissions were previously not recognized as a possible source of smog. In an example that clearly points to economic implications, Thiemens and Trogler (1991) discovered that about 30\% of nitrous oxide emissions, a precursor to ozone and a greenhouse gas, were unaccounted for in emissions inventories. Mounting a search for them by fingerprinting molecules, the EPA found that the use of adipic acid in the production process for nylon gave rise to large quantities of nitrous oxide. Upon learning of this discovery, nylon producers voluntarily modified their production process to greatly reduce nitrous oxide emissions at very little cost at a time when other emitters of nitrous oxide were undertaking high cost measures to abate emissions. Fu et al. (2012) shows precisely how volatile, variable emissions from disparate sources relate to urban ozone levels in China.}

Potential gains from a better understanding of the structural relationship between emissions and ambient pollution seem to be largely ignored. For example, the 2010 EPA budget
requested $842 million (less than 9% of the total agency budget) for improvement of science and technology. Of that amount, only a very small fraction is spent on improved modeling of the pollution output-ambient quality relationship, more accurately recording emissions or searching for new pollution sources. Such expenditures are orders of magnitude less than the direct costs of pollution regulation by firms, households and local governments nationwide (Gray (2002) and Gray and Shimshack (2011)). Therefore, it is possible there may be economic gains from basing pollution control regulations on a more precise understanding of the relationship between emissions and ambient pollution levels.

The EPA, and its counterparts in OECD countries and many developing countries, utilizes both an air pollutant emissions monitoring system and an ambient air quality monitoring system to measure air pollutants. Controlling for other factors such as temperature and wind, the two systems are linked to each other using an air dispersion model which takes emissions as an input affecting the stock of ambient pollution. The Community Multiscale Air Quality (CMAQ) model is one such model and its approach has been widely adopted in the atmospheric modeling literature (Kim and Ching (1999)). The potential for discord between the two monitoring systems is well known among air quality modelers (McKeen et al. (2007) and McKeen et al. (2009)) and difficulties in reconciling the two systems has taken on increasing importance as efforts are made to produce real time forecasts of ambient quality (Wang et al. (2011), Zhang, Bocquet, Mallet, Seigneur and Baklanov (2012a), and Zhang, Bocquet, Mallet, Seigneur and Baklanov (2012b)).

While possible problems in translating emissions into ambient concentrations was occasionally mentioned in early economic work on pollution control (Spence and Weitzman
(1978) and Crandall (1981)), the implications for optimal pollution control have received little attention in the economics literature. For example, economists have evaluated optimal pollution control when the effect of emitters on ambient pollution is stochastic (Horan et al. (1998) and Hamilton and Requate (2012)) or when damages from emissions vary spatially (Muller (2011) and Fowlie and Muller (2013)). In each case, though, the relationship between a particular emission source and expected ambient concentrations at a particular location is assumed to be known.

Our main contribution is that we address the fundamentally different problem of how allowing uncertainty between how known emission sources map to observed ambient pollution levels affects 1) expected damages from emissions and 2) the optimal stringency of pollution control.² A key aspect of our model is that the regulator must estimate the marginal contribution of observed emissions on ambient levels of pollution. In that sense, our model is akin to a structural model rather than a reduced form model of damages from emissions.³

We develop a simple theoretical model in which the regulator must estimate the structural relationship between emissions and ambient pollution concentration levels. We show that this type of structural uncertainty manifests as a type of multiplicative uncertainty in the distribution of pollution damages.⁴ One contribution of our paper, then, is that we provide a theoretical underpinning for multiplicative uncertainty, rather than additive uncertainty.

²Put another way, our problem is one about accurately apportioning known sources to ambient pollution levels. In that sense, one application of our paper is to consider how the level of a “pollution license” considered in Montgomery (1972) is affected by uncertainty in the relationship between how known emitters affect the ambient level of pollution.

³For an example of a well-executed reduced form model of damages see Lemoine and McJeon (2013).

⁴Multiplicative uncertainty was introduced into the pollution control literature by Malcomson (1978) as a potential issue with Weitzman (1974) explicit adoption of additive uncertainty as a modeling choice (see Weitzman (1978) for a reply). Since then, the possibility of multiplicative uncertainty has been noted occasionally and was studied in substantive detail by Hoel and Karp (2001).
in the pollution control literature. We show with a simple example that if some emitters are unknown or not discovered, then the regulator can over-attribute observed ambient pollution levels to known emitters. This over-attribution influences expected damages of emissions from known emitters and therefore both regulated emissions and ambient pollution levels. Our model is also applicable to stock pollutants if the rate of decay in the stock pollutant is known and total change in the ambient level of pollution in a given period is observed.

Two main policy implications follow from our model. First, we contrast how optimal pollution levels change when an additional source of emissions is discovered under three scenarios: full information, incomplete information over the set of emitters but no structural uncertainty, and incomplete information over the set of emitters in addition to structural uncertainty. In this final situation the regulator must estimate marginal emissions and we show those estimates will typically be biased upward. Contrary to standard intuition, the discovery of new emitters combined with structural uncertainty sometimes makes it optimal to increase the level of ambient pollution. Second, we show how the regulator’s problem changes if they explicitly account for structural uncertainty. When accounting for structural uncertainty the optimal pollution level is jointly determined by the costs of abatement on both the direct and indirect influence of the perceived degree of structural uncertainty. The model is able to give some simple comparative statics of the effect of accounting for structural uncertainty, though. For example, if the regulator believes that structural uncertainty is small, then they should increase the stringency of regulation on known emitters and vice-versa.

The remainder of this paper is as follows: Section two introduces and develops the
model while section three presents the main results. Section four offers discussion including
some approaches that might be implemented to reduce the adverse influence of structural
uncertainty, as well as some concluding remarks.

2 Model

This section develops our theoretical model. We start with the simplest case: full infor-
mation, homogeneous emitters and a flow pollutant. We then gradually relax different
assumptions thereby adding increasing complexity to the model. Throughout this paper we
assume that the vector of optimal emission levels is achievable via a tax or tradable quota
system. We do not address the myriad issues with actual implementation of pollution control
mechanisms here, as we are focused only on the stringency of pollution control in this paper.

Our paper is motivated by the fact that any regulator must estimate the effect of emis-
sions on ambient levels of pollution. As a result, we develop our model in terms of emissions
and ambient pollution levels rather than abatement. In our model, we assume that each $i$
represents an emitter and that emitters benefit from emissions up to some point. It is also
useful to think of $i$ as indexing industries, sectors or any different set of emitters such as
different countries or anthropogenic versus geological sources. Damages to society, though,
stem from ambient levels of pollution.\footnote{Put another way, the benefit to a firm of the first unit of emissions is the same as the cost of its final unit of abatement. Similarly, the marginal cost to society of the first unit of ambient pollution is the marginal benefit to society of the last unit of ambient pollution abated.} For simplicity, we refer to emitters as firms throughout most of the paper.

We start with each firm $i$ choosing a level of emissions, $x_i$, to maximize profits $\pi_i(x_i)$.
We assume that \( \pi_i(x_i) \) is initially increasing in \( x \) and twice differentiable with \( \pi''_i(x_i) < 0 \). If unregulated, each firm emits at a level \( x_i^* \) such that \( \pi'_i(x_i^*) = 0 \). For simplicity, let the regulator know the exact form of \( \pi_i(x_i) \) for all \( i \) with certainty and there are \( N \) total emitters.

There are two characteristics of firms worth noting. First, firms can have either homogeneous or heterogeneous profit functions. The steepness and level of marginal profits from emissions could vary by source. Similarly, profits from emissions vary by sector or source.\(^6\) For simplicity and to highlight the implications of structural uncertainty, we start by assuming emitters are homogeneous. We relax the homogeneity assumption later.

Second, emissions could be correlated over time across firms or across sectors. For example, emissions from many pollutants are procyclical as economic activity increases during economic expansions (Heutel and Ruhm (2013)). Procyclical emissions amount to positively correlated emissions across firms or sectors over time. From a theoretical perspective, this amounts to the profits from emissions varying over time such that equilibrium emissions, \( x_i^* \), vary over time. We abstract from this intertemporal aspect of emissions behavior in the main text of our paper. We instead analyze the regulator’s problem in a static model. In an Appendix, we discuss the implications of explicitly incorporating time into the regulator’s problem. Like the profit function homogeneity assumption, the static model focuses the paper on new issues raised by allowing structural uncertainty.

Emissions from firms contribute to ambient levels of pollution, \( y \). Ambient pollution decreases social welfare according to a strictly increasing deterministic damage function \( D(y) \). We make three simplifying assumptions about ambient pollution and damages from

\(^6\)From a modelling perspective, geological emissions of methane, for example, have a very steep marginal benefit curve: it is very costly and could require geo-engineering to reduce such emissions.
ambient pollution. First, we assume the regulator knows the deterministic damage function $D(y)$ perfectly. Second, we assume the ambient level of pollution, $y$, is perfectly observable. While we assume that ambient levels of pollution are known with certainty throughout, this assumption is also incomplete if not incorrect. Ambient pollution levels are only known at discrete locations (e.g., monitoring stations) at discrete points in time (e.g., when readings are taken). Ambient levels of pollution between these spatio-temporal fixed points is itself uncertain. We leave this extension of our modeling approach to future work. Third, we initially assume that marginal damages are linear in $y$: $D'(y) \propto cy$. This last assumption is made only to build intuition and we relax it later.

We now make the very strong assumption that ambient pollution is related to emissions according to the simple linear function $y = f(\{x\}_i^{N}) = \Sigma_i x_i \beta_i + \epsilon$, where $\epsilon \sim N(0, \sigma^2)$ and $\beta_i$ reflects the marginal physical impact of each emitter’s contribution to, $y$, the ambient pollution level. The idiosyncratic term, $\epsilon$, explicitly models the typically assumed imperfections of any pollution dispersion model. The coefficient $\beta_i$ represents the contribution of different emitters to ambient pollution levels at a particular location. The existence and importance of these “transfer coefficients” is well-understood in the atmospheric chemistry literature (Fu et al. (2012)). This particular functional form is a generalization of a classical flow pollutant in which the total amount of ambient pollution levels is the sum of emissions.

A linear relationship between ambient pollution and emissions is appealing for simplicity of exposition but it does come with a cost. First, while linearity is a reasonable approximation

\footnote{Recent work has focused on the importance of stochastic environmental sinks, which could be modeled similarly, for the merits of policy instruments (Hamilton and Requate (2012)). We abstract from discussion of policy instruments here to focus on the role of structural uncertainty in the relationship between emissions and ambient pollution levels for pollution control stringency.}
in many cases, such as heavy metals and primary particulate matter, it is not a reasonable assumption in many other cases such as where the relationship with precursor emissions and ambient pollution is known to be non-linear. Second, since there is no explicit spatial dimension in this model, our ambient pollution $y$ should be thought of as ambient pollution in a particular location.\(^8\) Third, there is no temporal element in this initial flow pollutant model. We will later show that the important aspects of this model carry through directly to a stock pollutant when the stock’s decay rate is known. While the model can be extended to allow for all of these simplifications, we restrict analysis to the simplest model possible to show the first order effects of structural uncertainty on optimal pollution control.

Emissions from the $i$th polluter in the model, $x_i$, increases ambient pollution levels at some rate $\beta_i$ (e.g., $E[y] = \Sigma_i x_i \beta_i$). For simplicity, we start by assuming all polluters emit have the same physical impact on ambient pollution levels (e.g., $\beta_i = \beta_j$) although we later relax this assumption. There are at least two common situations where this would not be the case. First, there are many types of GHGs which all contribute to climate change. For example, methane is a much more potent greenhouse gas than carbon dioxide, but they both contribute to atmospheric GHGs which ultimately may cause damages via the same mechanism.\(^9\) Second, the physical effects from different sources can impact pollution levels at a particular location at different rates. These different rates are referred to as “transfer coefficients” in situations where emissions in one geographical region transfer to the stock of ambient pollution levels in another. There is a growing literature noting that using transfer coefficients to improve efficiency of regulation (Muller (2011) and Fowlie and Muller (2013)). However, a spatial dimensions is not needed to develop the intuition in our model.

\(^8\)This model clearly applies to the related literature with uses transfer coefficients to improve efficiency of regulation (Muller (2011) and Fowlie and Muller (2013)). However, a spatial dimensions is not needed to develop the intuition in our model.

\(^9\)As long as the physical conversion coefficients are known, this causes no issues. If the conversion coefficients are not known, it causes issues similar to those considered in this paper.
coefficients may improve efficiency of environmental regulation (Muller (2011) and Fowlie and Muller (2013)). Uncertainty over transfer coefficients can be seen as a special case where the translation of emissions into ambient pollution is not known with certainty. We directly address transfer coefficient heterogeneity in this paper. However, a simpler model assuming transfer coefficient homogeneity can provide the key insights of accounting for structural uncertainty.

In order to juxtapose our extension to the classical case, we start by assuming that the marginal physical effects of emissions on ambient levels of pollution are known. In this full information case, the social planner (e.g., regulator) sets the vector of emissions to maximize expected social welfare. As a result, the regulator maximizes the sum of all profits less expected damages from ambient levels of pollution:

\[
\max_{\{x\}} \sum_{i=1}^{N} \pi_i(x_i) - E[D(y)].
\]  

(1)

The \(N\) first order conditions for optimality are:

\[
\pi_i'(x_i^*) = E[D'(y^*)\beta_i] = D'(y^*)\beta_i \quad \forall \ i = 1, \ldots, N.
\]  

(2)

The first order conditions show the standard result that the change in profits from reducing a unit of emissions by the firm should equal the expected marginal damages caused by an increase in firm \(i\)'s contribution to the ambient pollution levels. Importantly, marginal damages of a firm's emissions are weighted by the contribution of their emissions to ambient pollution levels, \(\beta_i\). Given our assumptions that marginal damages are linear and the
relationship between emissions and ambient levels of pollution, $\beta_i$, is known we can drop the expectations operator in equation (2). Finally, due to our homogeneity assumption across profits, $\pi_i$, and the relationship between emissions and ambient levels, (e.g., $\beta_i = \beta_j$), the restricts all polluters to emit at the same level $x^*$ satisfying equation (2). This represents the classic non-depletable externality problem.\footnote{As a thought experiment, it is useful to think about how a firm’s optimal level of emissions changes if the marginal physical effect of that firm’s emissions is larger than another firm’s. For the purpose of this thought experiment, assume that two firms have different transfer coefficients $\beta_i > \beta_j$. In this case, for the same level of ambient pollution, $y$, the marginal damages from an additional unit of emissions from the more potent emitter is larger. Although $D'(y)$ is the same for all emitters since damages from ambient pollution levels do not vary within the model, the marginal physical effect of a polluter’s emissions on ambient pollution levels, $\beta_i$, causes the damages from firm $i$’s emissions to be greater than those from firm $j$. As a result, if the two firms were identical in every other way, the regulator would mandate a lower level of emissions for the more polluting firm so that the marginal benefit of emissions for that firm is equal to the marginal cost of emissions from that firm.}

Now assume that the regulator is uncertain about $\beta_i$ and must estimate it. This situation mirrors that of physical scientists in which $y$ is observed and $\beta_i$ is a parameter to be estimated. In this case, the regulator’s problem reduces to one with multiplicative uncertainty on the marginal cost of emissions. If the regulator’s estimate of marginal physical effects, $\hat{\beta}_i$, is an unbiased estimator of $\beta_i$, then expected marginal costs of a firm’s emissions equals actual marginal costs and the problem reduces to a standard treatment of multiplicative uncertainty. The main difference between this treatment and the seminal work on multiplicative uncertainty in Hoel and Karp (2001) is that our multiplicative uncertainty maps to damages from emissions (via uncertainty in $\beta$) rather than assumed uncertainty over costs of abatement themselves.\footnote{More specifically, uncertainty over costs of abatement would be represented in our model as uncertainty over increased profits due to emissions.}

Figure 1 shows how our structural uncertainty manifests as multiplicative uncertainty. The slope of the marginal cost of emissions curve for each firm is now subject to uncertainty.
The upper and lower rays from the origin represent upper and lower bounds of marginal damages from the $i^{th}$ emitter if $\beta_i$ is uncertain. If $E[\hat{\beta}_i] = \beta_i$, expected marginal damages equal actual marginal damages so long as marginal damages are linear.

$$\frac{\partial D(y)}{\partial y \beta} = \frac{\partial D(y)}{\partial y E[\hat{\beta}]}$$

Figure 1: Graphical Representation of Multiplicative Uncertainty

Finally, this structural uncertainty model can be extended to fit the stock pollutant model of Newell and Pizer (2003). Assume that yearly contributions to a stock of pollution, $S_t$ is $y_t$. Assume that $S_t$ is perfectly observed. As in Newell and Pizer (2003), the equation of motion is $S_t = (1 - \delta)S_{t-1} + y_t$ where $\delta$ is the decay rate of the stock pollutant. When $\delta$ is known, then $y_t = S_t - (1 - \delta)S_{t-1}$. In any given year, then, the set of known emitters contributes to the stock of ambient pollution as in the flow pollutant case since the decay rate is known (e.g., $S_t - (1 - \delta)S_{t-1} = \Sigma x_t \beta + \epsilon_t$). As a result, uncertainty over $\beta$ can be extended to stock pollutants so long as the net present value of damages is considered. If the decay rate $\delta$ is not known then the stock pollutant case presents a unique set of challenges that we leave to future research.
2.1 Partially Identified Sources of Emissions

We now again assume that $\beta_i$ is known but relax the assumption that all emitters are identified. This situation is called having “unapportioned sources” of emissions and is well studied for various pollutants in the atmospheric chemistry literature (Stone et al. (2008), Etope and Ciccioli (2009) and Miller et al. (2013)). We assume initially that the regulator knows the contribution of emissions to ambient levels of pollution, $\beta_i$, for the subset of known emitters. We briefly develop a simple model reflecting the regulator’s problem with a partially identified sets of emitters and known marginal physical effects. We then consider model dynamics if additional emitters are discovered and/or regulated. This gives us the opportunity to juxtapose our model of partially identified sets of emitters of known marginal physical effects against the same model in the presence of structural uncertainty.

Having unapportioned sources of emissions is akin to regulating only a subset of known emitters. There are several situations where only a subset of known emitters are regulated. Water pollution from industrial sources (rather than agricultural sources) and recent CO2 limits on power plants (rather than transportation fuel) are two prominent examples. This form of partial regulation has also received attention from the literature (Fowlie (2009) and Fowlie et al. (2012)). If marginal physical effects are known and only a subset of emitters are known, our model is qualitatively similar to those models in terms of optimal stringency of pollution regulation.\[12\]

A first best regulatory program equates marginal abatement costs across firms. In the

\[12\] From a modeling perspective, the two situations are identical but the difference between our approach is the cause of partial regulation, which matters for policy. For example, paying for better technology to monitor non-point source polluters is different than paying for better scientific understand of the pollution process.
simple case of homogeneous firms and full information (e.g., known transfer coefficients
and all polluters known) all firms would reduce emissions to the same level, satisfying the
equimarginal principal. If instead some firms are excluded from regulation because they
are unknown sources, their marginal cost of abatement is lower than regulated firms at the
margin. The best the regulator can do if some emitters are unknown is to regulate known
emitters at their “conditionally” optimal levels. In that situation, the social planner can
increase welfare by identifying an additional emitter and regulating them. This situation
is precisely what occurs in our model with known transfer coefficients but some unknown
sources.

More formally, assume as before that all emitters are identical and that the regulator can
identify only $\alpha_0 N = K$ emitters out of a total of $N$ total emission sources where $\alpha_0 \in (0, 1)$. It is convenient to think about $\alpha_0$ as the percent of emissions with identified sources given our
homogeneity assumption.\(^{13}\) Assume initially that the regulator knows the marginal physical
effect of each identified emitter on ambient pollution levels.

When only $K$ emitters are identified the social planner’s objective function and first order
conditions are:

$$\max_{\{x\}_i^K} \sum_{i=1}^{K} \pi_i(x_i) - E[D(y)]$$

$$\pi'_i(x^*_i) = E[D'(y^*)\beta_i] = D'(y^*)\beta_i \quad \forall \ i = 1, ..., K = \alpha_0 N. \quad (3)$$

In this situation the regulator can optimize using only $K$ rather than $N$ control variables
\(^{13}\)Identified emitters can refer to an entire industrial sector or geological source rather than individual polluting entities.
since there are only $K$ first order conditions. Since damages are strictly increasing in ambient pollution, all $K$ polluters are regulated more than if the regulator was able to identify all emitters.

Consider the dynamics of the social planner’s problem if an additional $K + 1$st emitter is exogenously discovered and that emitter’s marginal effect on ambient pollution levels can be regulated.\textsuperscript{14} At the time of discovery, the emitter will be producing at a level where the marginal profit of an additional unit of emissions is zero: $\pi'_{K+1}(x_{K+1}) = 0$. Since previously regulated firms are restricted to emitting at their conditionally optimal level, they have a higher marginal benefit of a unit of emissions on the margin. Since $D(y)$ is twice differentiable and increasing, welfare must increase when the newly discovered polluter is brought under regulation as emissions from the newly regulated firm can be decreased at a lower marginal cost than previously regulated firms. More generally, the discovery of a new emitter cannot decrease welfare as the regulator now has another degree of freedom to use in choosing optimal pollution levels.

This situation is shown explicitly in Figure 2. Assume that firm $i$ is a known emitter when a new emitter firm $j$ is discovered. Initially firm $i$ emits $x^*_{i,0}$ units of pollution and firm $j$ emits $x^*_{j,0}$ units of pollution. When discovered and regulated, firm $j$ is forced to reduce their emissions by $\Delta x^*_j$. As the newly discovered emitter is brought under regulation, ambient pollution levels will decline due to lower emissions from the newly regulated firm. This rotates the marginal damage associated with each of the previously known emitting firm’s emissions downward because there is less ambient pollution (e.g., ambient pollution

\textsuperscript{14}For simplicity, we do not model the discovery process here.
Figure 2: Optimal Firm Emissions Under Discovery of New Emitter

decreases by $\Delta x_j^* \beta_j$. Therefore, while the newly discovered firm suffers from lower profits, all other firms enjoy greater profits since their pollution increases by $\Delta x_i^*$. Ambient pollution levels fall thereby reducing both total and marginal damages and social welfare increases. Importantly, all changes in regulation and subsequent increases in welfare are predicated upon the regulator knowing $\beta_i$ and $\beta_j$ with certainty.

\footnote{If the damage function were perfectly linear, then this second order effect (due to decreased ambient pollution levels) would not occur. Put another way, if the marginal damage function were flat, then there would be no “rebound” effect like the one detailed here.}
2.2 Estimating Marginal Physical Effects with Incomplete Information

This subsection introduces the theory behind the main contribution of the paper. We identify the implications of relaxing the assumption that the marginal physical effect of known emitters on ambient pollution, $\beta_i$, is known and that all emitters are identified. We instead model the regulator’s task as both 1) estimating the marginal physical effect of known emitters and 2) using that parameter to inform the stringency of pollution regulation. We pay special attention to the dynamics of discovering new emission sources on each of these tasks. As a result, we add structural uncertainty to the preceding subsection.

In practice, the set of emitters that affect ambient pollution levels at a particular time and place are imperfectly observed. For both uniformly mixing pollutants like methane and non-uniformly pollutants like heavy metals the full set of emissions is unknown (Pacyna et al. (2010) and Miller et al. (2013)). Regulators sometimes, but not always, acknowledge these imperfections by allowing modelers to use estimates of “fugitive emissions” to account for emission inventories from unmeasurable sources.

In addition to having an incomplete set of emitters, the regulator must also estimate the marginal physical effect of emissions on ambient pollution levels from an incomplete set of emissions. The EPA and other industrialized countries acknowledge the imperfections of their theoretically driven air dispersion models in some cases. Implemented in 2006, “Response Surface Modeling” uses maximum likelihood statistical techniques to apportion known emissions to ambient pollution levels where atmospheric dispersion models are inac-
curate (EPA (2006) and Wang et al. (2011)).\textsuperscript{16} This amounts to the regulator estimating the contribution of each emitter to ambient levels of pollution, $\beta$, in our model with an incomplete set of emissions using statistical techniques.

We claim that estimating the relationship between emissions and ambient levels of pollution will often lead biased estimates of the marginal physical effect $\beta$. Assume, for example, that the regulator does not account for the fact that they only observe a subset of the true set of emitters when attributing known emissions to ambient pollution levels. Intuitively, the regulator will over-attribute known emitters to ambient levels of pollution, leading to upwardly biased estimates of the contribution of emissions to ambient pollution levels.

We now show with an example how estimating $\beta_i$ from an incomplete set of emissions, and not accounting for only observing an incomplete set of emissions, leads to the over-attribute described in the previous paragraph. To make our example as clear as possible, we assume the regulator completely ignores that they observe an incomplete set of emitters. We discuss relaxing this assumption in detail below. We also maintain both the homogeneity of emitters assumption (across both $\beta$ and $\pi(\cdot)$) and linearity in the relationship between emissions and ambient pollution levels assumption (e.g., $y = \sum_i x_i \beta_i + \epsilon$) to clearly show how

\textsuperscript{16}This situation is likely the case for many pollutants: in the United States, the EPA implicitly assumes that their models correctly identify the relationship between the history of emissions and current ambient pollution levels. For example, in order to perform forecasts of the effect of new pollution sources on ambient pollution levels, the EPA typically calibrates its air quality models to the ambient pollution readings at any given monitoring station. The estimated or calibrated parameters are then to forecast the effect of reducing emissions from existing sources or the effect of additional emissions on ambient air quality from siting a new point source at a particular location. This problem may be getting worse rather than better over time: as the number of regulated air and water pollutants increases, the need for additional EPA air and water dispersion modelers also increases but it is unclear if funding for these modelers increases as well. Conversations with two different economists and two different air modelers at the EPA suggest that there is a bottleneck at the EPA with respect to their capacity to develop high quality emission and pollution models despite the best efforts of the atmospheric chemists and modelers currently on staff. In one conversation, a modeler noted that after the EPA began regulating a new pollutant not a single new modeler was hired.
uncertainty in the emissions inventory affects transfer coefficient estimates.

Formally, assume that the regulator has identified only a subset of $K = \alpha_0 N$ out of $N$ total emitters where $\alpha_0 < 1$. There are two helpful ways of formalizing “$\alpha_0$” in this example. First is as the percentage of identified emitters. A second way of representing the percentage of known emitters, $\alpha_0$, in this situation is as a $Nx1$ vector of ones and zeros. If the $i$th emitter is known, then the $i$th place define a vector $\alpha_0$ will have a one. Alternatively, if the $j$th emitter is unknown then the $j$th place in the vector $\alpha_0$ will have a zero.\footnote{In this way, the sum of the number of ones in the vector $\alpha_0$ divided by the length of the vector is the percentage of known emissions, $\alpha_0$.} Since we assume that all emitters are identical except for being identified or unidentified, we can interpret observed emissions as the sum of the elements in the true vector of emissions, $x$, multiplied by $\alpha_0$ being the total amount of observed emissions without loss of generality.

In this framework, there are two ways to think about what the regulator observes. Using vector notation, the regulator only observes $\tilde{x} = \sum_{i=1}^{N} \alpha_{0,i} x_i$ where $\alpha_{0,i}$ is the $i$th element of the vector $\alpha_0$. Using scalar notation a leveraging the homogeneity assumptions, if all emissions contribute equally to ambient pollution levels we can also have the regulator observe $\tilde{x} = \alpha_0 x$ where $\alpha_0 \in (0,1)$ and $x$ is the sum of all firms’ emissions. Thus, since all emitters are homogeneous except for being identified or unidentified, then, we can interpret observed emissions as $\alpha_0$ percent of the sum of total emissions observed by the regulator without loss of generality. Recall, though, that the regulator does not know $\alpha_0$. For much of this section, we use the scalar notation for notational ease.

The regulator’s task in this subsection is estimating marginal physical impacts, $\beta$, of each emitters’ contribution to ambient pollution levels (call the estimate $\hat{\beta}$). We consider
the simplified strategy of estimating $\hat{\beta}$ by ordinary least squares (OLS) as an example below. We don’t claim that regulators actually perform OLS to parameterize air dispersion models. However, the OLS example provides the needed intuition for how over-attribution can occur given regulators’ reliance on statistical procedures to parameterize their models as in Response Surface Modeling used in practice (EPA (2006)).

To allow the regulator form an estimate of $\beta$ by OLS, we need to add a temporal element to our model. Assume that the regulator observes a series of observations for emissions and ambient pollution levels: $\{\tilde{x}_t, y_t\}_{t=1}^T$. Stack ambient pollution levels in a vector $Y$ and observed emissions for each time period $t$ in a vector $\bar{X}$. Recalling the true data generating process is $y_t = \sum_{i=1}^{N} x_{i,t} \beta_i + \epsilon_t$ where $\epsilon_t$ is a well-behaved random error component and that emissions affect ambient pollution identically, the estimated coefficient vector if $Y$ is regressed on $\bar{X}$ takes the form $\hat{\beta} = \frac{Y'\bar{X}}{\alpha_0 X'X} = \frac{X'X}{\alpha_0}$. Because $\alpha_0 \in (0,1)$ due to incomplete information of the regulator, the estimate of marginal physical effects of known pollutants, $\hat{\beta}$, is upwardly biased in this case.\footnote{This type of upward bias of the OLS estimator is a special case of multiplicative measurement error. While somewhat common in the epidemiology literature, we are not aware of any similar measurement error present in the economics literature. In our model $\alpha_0$ is an unknown constant between zero and one. Zhang, Liu, Dong, Holovati, Letcher and McGann (2012) shows that the asymptotic distribution of $\hat{\beta}$ given multiplicative measurement error of this form can be expressed as $\sqrt{N}(\hat{\beta} - \frac{\beta}{\alpha_0}) \sim N(0, \sigma^2 \alpha_0^2 (X'X)^{-1})$. Hence, OLS will overestimate the marginal physical effect and underestimate its variance of the estimator. Intuitively, there is over-attribution of perfectly observed ambient pollution levels to a subset of the universe of emitters in the physical model being used in the regulation of pollution.}

Note that if the regulator knew $\alpha_0$ with certainty, they could correct their biased estimate of $\hat{\beta}$. That scenario would give rise to the situation presented in the previous subsection in which the regulator could correctly estimate the marginal physical effect of known emissions for only a subset of emitters.

We don’t imply that regulators actually perform OLS estimation to attribute observed emissions to ambient pollution levels. We do, though, want to show very precisely how
this situation leads to biased estimates of transfer coefficients due to over-attribution of known emitters to ambient pollution levels. While this type of over-attribution and bias in the estimating procedure is misspecified, researchers and regulators continually discover unobserved emissions and mis-attributed emissions for both uniformly and non-uniformly mixing pollutants (Pacyna et al. (2010) and Miller et al. (2013)). As a result, this example highlights a relevant policy concern.

Consider a regulator who uses the estimated marginal effect of known emitters on ambient pollution from the misspecified model above, \( \hat{\beta} \), to set emissions levels as a function of expected ambient pollution levels. Maintaining the linear marginal damages assumption for simplicity and recalling that all emitters are homogeneous, the resulting optimality condition obtained from maximizing expected welfare is:

\[
\max_{\{x\}^K_i} \Sigma_{i=1}^K \pi_i(x_i) - E[D(y) | \alpha_0] = \Sigma_{i=1}^K \pi_i(x_i) - E[D(\Sigma_i x_{i,t} \hat{\beta}_i) | \alpha_0] \quad (4)
\]

\[
\pi'_i(x^*_i) = E[D'(y^*) \hat{\beta} = D'(y^*)\hat{\beta} \quad \forall \ i = 1, ..., K = \alpha_0 N. \quad (5)
\]

where \( E[y^*] = \bar{X}^* \hat{\beta} \). Note that conditioning on \( \alpha_0 \) indicates the regulator estimates \( \hat{\beta}_i \) for all identified emitters \( i = 1, ..., K = \alpha_0 N \). Given our homogeneity assumption, we allow \( \hat{\beta}_i = \hat{\beta}_j = \hat{\beta} \). Since \( \alpha_0 < 1 \) and \( \pi''(\cdot) < 0 \), the \( \bar{x}^*_i \) implicitly defined by the set of equations (5) must be lower than in the case above when \( \beta \) is estimated without upward bias.\(^{19}\) The main

\(^{19}\)Consider the following example: assume there are two homogeneous polluters but only one is known, the true \( \beta = 1 \) but that only one polluter is identified. Lastly, assume marginal damages increase at a constant rate \( d \) for simplicity. In this case the regulator would form an estimate of marginal emissions of \( \hat{\beta} = 2 \). As a result the first order condition for the regulator is \( \pi'_i(x^*_i) = E[d(y^*)2] = d(2x^*_i)2 = 4dx^*_i \). Conversely, the full information first order condition, given that firms are homogeneous, would be \( \pi'_i(x^*_i) = d(2x^*_i)1 \). Since the profit function is concave, it is clear that the known emitter is over regulated in the partial information case.
implication for individual firms is that known sources are over-regulated in this misspecified model relative to when marginal damages of known emissions are known. The reason is that perceived marginal damages of known emitters are higher than their actual damages.

In sum, we see that structural uncertainty in the relationship between emissions and ambient pollution levels manifests as multiplicative uncertainty. It is realistic that the regulator both 1) observes only a subset of total emissions and 2) must attribute and/or estimate the effect of emissions from known sources to ambient pollution levels (e.g., estimate transfer coefficients). This process can lead to upwardly biased estimates of the relationship between known emitters and ambient pollution levels. Upwardly biased estimates imply that known emitters are over-regulated relative to the case when transfer coefficients are unbiasedly estimated. While this section has presented our modeling framework, the main contribution considers the regulator’s response to discovering new emitters.

Before proceeding to our main results, we note that it is also possible that the contribution of the subset of known emissions are underestimated by the regulator in some cases. The clearest example is for local air pollutants like particulate matter. For particulate matter the EPA estimates the level of “fugitive emissions” from sources that are difficult if not impossible to measure such as particulates from traffic on roads, leaks from pipes, small industrial processes, etc. These emissions are important to account for because they contribute to ambient pollution just as emissions from well-identified and measurable point sources.

It is possible that the EPA could overestimate such “fugitive emissions” and also observe all emitters. In our model, it amounts to adding an unobserved emitter (motivated by unobserved fugitive emissions) to the set of known emitters such that $\alpha_0 > 1$. In that
case, estimates of known emissions on ambient pollution levels would be biased downward (e.g., \( \hat{\beta} = \frac{\beta}{\alpha_0} < \beta \)). Indeed, one possible interpretation of our model is that it is a model of uncertain fugitive emissions: uncertainty in fugitive emissions manifests as structural uncertainty in the relationship between emissions and ambient pollution levels. We mainly focus on the case of incomplete emissions sources below but are careful to discuss this alternative interpretation when appropriate.

3 Results

This section presents analytical results from the theoretical model when the regulator must estimate marginal physical effects of known emissions from an incomplete set of emitters. As motivated above, for most of this section we therefore assume that the regulator forms upwardly biased estimates of marginal physical effects of known emissions on ambient levels of pollution. We perform comparative statics over the optimal level of regulated emissions and ambient levels of pollution when new emitters are discovered or estimated transfer coefficients are revised toward their true levels. In this section we separately allow the marginal benefit of emissions (e.g., \( \pi_i(\cdot) \)) and the marginal physical effects of emissions (e.g., \( \beta_i \)) to vary across firms. Lastly, we discuss the implications of a regulator having downwardly biased estimates of the marginal physical effects of known emissions on ambient levels of pollution at the end of the section.

There are two situations we consider in this section. First, if a scientific discovery reveals more emission sources than were previously known, then there are more emissions to attribute to ambient pollution levels (e.g., \( \alpha_0 \) increases to \( \alpha_1 \)). As a result the marginal damage
associated with emissions from each emitter that was identified before the scientific discovery is lower since the marginal contribution of their emissions to ambient levels of pollution is lower. Put another way, $\beta_i$ is revised down when the vector of transfer coefficients are re-estimated since observed emissions are now larger than they were before.\textsuperscript{20} This is similar to implications of the recent discovery that there were very significant unaccounted for inventories of fugitive methane emissions (Miller et al. (2013)). Second, a scientific discovery could re-order the relative importance of known emitters. For example, better air dispersion models or better measurement technology could more accurately attribute different emission sources to ambient pollution levels in the absence of new emission discoveries. This is similar to the recent attribution of ambient sulfate levels in the LA air basin to ships in the LA harbor (Domínguez et al. (2008)). Even though sulfates were already known to come from ships, the level of contributions of their emissions to ambient pollution levels was considered negligible so that other emissions sources we over-attributed to ambient pollution levels.

Figure 3 shows the intuition for this result graphically. As the transfer coefficient of known emitters becomes less upward biased (e.g., $\hat{\beta}_0 > \hat{\beta}_1 > \beta$), the curve representing perceived marginal damage of emissions rotates down.\textsuperscript{21} Figure 3 holds ambient pollution levels, $y$, fixed in order to highlight how the rotation of the marginal damage curve is exclusively

\textsuperscript{20}In the context of the OLS estimation example above, the scientific discovery implies that the percentage of observed emissions is larger than it was before. If $\alpha_0$ is the old percentage of identified emissions and $\alpha_1$ is the new percentage, then $\bar{X} = \alpha_1 X$ where $\alpha_1 > \alpha_0$. As a result, the estimated coefficient is lower than what is was before for previously known emitters: $E[\hat{\beta}_1] = \frac{y'X}{\alpha_1'X} = \frac{\beta}{\alpha_1} < \frac{\beta}{\alpha_0}$. Given that $\alpha_1 > \alpha_0$, the first order effect on $x_i^*$ implied by the new optimality condition in equation (5) for $i = 1, ..., \alpha_1 N$ is that previously known emitters’ optimal emissions levels will be higher than they were before.

\textsuperscript{21}In Figure 2 and subsequent figures, we assume $D(\cdot)$ is a convex function. Therefore, marginal damages are upward sloping. In certain situations, damages may be linear (Muller and Mendelsohn (2009) and Newell and Pizer (2003)). However, even if damages were linear in pollution (e.g., marginal damage curves are flat), discovering new emitters ($\alpha$ increasing) would shift down the marginal damage curve causing emissions of previously known sources to increase.
tied to the change in the estimated marginal physical effect of known emitters on ambient pollution levels, $\hat{\beta}$. The first order effect is that profits increase for a firm who was emitting at level $x_0$ and who now emits at level $x_1$.

\[
\frac{\partial D(y)}{\partial y} \hat{\beta}_0 \quad \frac{\partial D(y)}{\partial y} \hat{\beta}_1 \\
\frac{\partial \pi_i'(x_i)}{\partial x_i} \\
\frac{\partial \pi_i'(x_i)}{\partial x_i}
\]

Figure 3: Optimal Firm Emissions Under Different $\hat{\beta}$.

An important result from discovering new emissions sources is that it is always strictly cheaper to achieve a given level of ambient pollution so long as at least one newly discovered emitter has lower abatement costs than that of a previously regulated firm: $\frac{\partial \pi_i(x_{i,0}^*)}{\partial x_i} > \frac{\partial \pi_i(x_{un}^*)}{\partial x_i} = 0$, where $x_{un}^*$ is the emissions level of the unregulated emission source. Put another way, an increase in the set of emitters will increase the emissions of previously known emitters and decrease the emissions of a newly discovered polluter so long as the marginal profit of the unregulated source is less than their perceived marginal damages.

It is always cheaper to achieve a given level of ambient pollution if a new emissions source is discovered. However, it may not be optimal to decrease ambient levels of pollution.
Whether it is optimal or not for actual ambient pollution to decrease when a new emissions source is identified depends on what happens to expected ambient pollution levels when a new emitter is discovered. The effect on expected emissions depends critically on both the estimation of marginal physical impacts of known emitters on ambient pollution levels, \( \hat{\beta}_{i,0} \), and the regulated emission levels of known emitters, \( x_{i,0}^* \). The intuition behind the effects of newly discovered sources of emissions on expected ambient levels of pollution can be formalized as:

**Lemma 1:** An increase in the set of known emitters from may increase or decrease the expected level of ambient pollution if marginal physical effects are biased upward.

Lemma 1 says that once new emitters are discovered, expected ambient pollution levels could be higher or lower than before. The level of expected ambient pollution can change once the new emitters are discovered because 1) now their emissions are counted in inventories, 2) marginal physical effects of previously identified emitters were biased upward and 3) now the regulator must revise their estimates of marginal physical effects of known emitters and newly discovered emitters. The new expected ambient level of pollution is a function of the interaction these three parameters. As a result, expected ambient levels of pollution can change even before the regulator decides to restrict emissions from the newly identified emitter.

Lemma 1, then, results from nonlinearity in revisions of marginal physical effects from previously identified emitters. Sometimes the change in the estimated physical effects from previously known emitters dominates the additional contribution from newly discovered sources and vice-versa. As a result, expected ambient pollution levels can increase or decrease
as new sources of emissions are discovered. We note that if marginal physical effects are biased downward; in that case expected emissions always increase.

While Lemma 1 deals with changes in expected pollution, there is the larger question of how newly discovered sources or transfer coefficient revisions affect optimal expected pollution levels. If expected ambient pollution falls upon the discovery of new emitters, marginal expected damages also fall, so long as the damage function is convex leading to the following Propositions:

**Proposition 1**: If expected ambient pollution decreases upon the discovery of new emitters, then previously known emitters’ optimal emissions levels strictly increase. Conversely, if expected ambient pollution increases, previously known emitters’ optimal emissions levels strictly decrease.

**Proposition 2**: An increase in the set of known emitters can increase or decrease the optimal level of ambient pollution.

Proposition 1 relates the nature of the damage function and to updated estimates of the marginal effect of emissions on ambient levels of pollution. If the joint effect of changes in damages from expected ambient levels of pollution and updated estimates for a particular emitter is strong enough (e.g., if expected marginal damages attributable to a specific emitter increase), then that emitter should be regulated more intensely. Importantly, expected damages from ambient pollution levels are a function of three things: 1) the nature of the damage function, 2) the emission levels of known emitters, and 3) the estimated contribution of those known emissions on ambient pollution levels. The emission levels of known polluters
and the estimated contribution of those known emissions on ambient pollution levels form expected ambient pollution levels.\textsuperscript{22}

Proposition 2 identifies the impact of discovering additional emitters on expected ambient levels of pollution. There are four effects which all occur simultaneously when a new emitter is discovered which all affect whether expected ambient pollution levels increase or decrease upon discovery of additional emitters. First, the estimates of marginal physical effects of emissions on ambient pollution levels decrease. Second, there are now more emitters in each period. Third, each additional emitter is now regulated. Fourth, the level of emissions from the previously regulated firms can be adjusted. The first and third effects act to decreases ambient pollution levels while the second acts to increase them. The final effect is the policy instrument of the regulator of interest. Since there are effects inducing the regulator to both increase and decrease their emissions, it is not clear that previously identified emitter’s will be regulated more or less stringently.

The key intuition behind Proposition 2 is from allowing the marginal benefits of emissions to vary by source (e.g. $\pi_i \neq \pi_j$). If the marginal benefit of emissions of the newly discovered polluter is large (e.g., $\pi'_j(x_j)$ steep), then the increase in emissions from the previously identified source ($x^*_{i,1} - x^*_{i,0}$) could be smaller than the decrease in emissions from the previously unidentified source ($x^*_{j,0} - x^*_{j,1}$) for a given revision in the marginal impact of emissions on ambient pollution levels. As a result, an increase in ambient pollution may be optimal, as shown in Figure 4. This situation would arise when a new source of emissions is discovered that cannot be changed for any plausible cost, as is the case when new emissions

\textsuperscript{22}One important feature of this model is that expected ambient levels of pollution, even when emitters are regulated, might not equal actual ambient pollution levels due to biased estimates of the physical effect of known emitters. We discuss this shortly.
are part of the earth’s natural processes. Such discoveries were made by Keppler et al. (2006) and Etiope and Ciccioli (2009) when they found that both plants and the ocean floor are major sources of potent greenhouse gas.\textsuperscript{23} Alternatively, if new emissions sources can be cheaply reduced, then the discovery of such new emissions source would result in optimal total ambient pollution levels being lowered. Recent recognition of the greenhouse gas potential of atmospheric black carbon from stoves in India is one such example (Ramanathan and Carmichael (2008)).

Taken together, Propositions 1 and 2 imply that if the resource regulator has misspecified the relationship between emissions and ambient pollution levels, then known emitters are often over-regulated relative to the full information case. Hence, known emitters in this scenario may have an incentive to learn more about the true nature of the polluting process.

\textsuperscript{23}Put another way, if a regulator sets a limit on total known emissions in order to meet an ambient quality standard and discovers later that some emissions are from very high cost sources such as geological processes, then the expected cost of reaching the ambient pollution target will increase.
Note that it is possible that the change in marginal damages is precisely offset by the drop in emissions from the newly regulated firm. The direction of the change of a particular firm, however, is determined exclusively by the relative slopes of the marginal benefit function of emissions by previously known versus newly discovered emitters. It is possible, for example, in contrast to the case shown in Figure 3, that the opposite effect is true. If so, the discovery of a new high marginal benefit emitter means that previously known emitters should abate more if the change in the estimate of previously known emissions’ ambient pollution is not sufficiently large.

A theoretical drawback of the approach in this section is the lack of dynamic consistency in the relationship between changes in emissions caused by the regulation of emissions and ambient levels of pollution. For example, assume that through a change in some policy instrument, a regulator reduces known emissions from $x_i$ to a level $\delta x_i$ where $\delta \in (0, 1)$. In this case, expected ambient pollution will fall by $(1 - \delta)\hat{\beta}_i$ but actual ambient pollution will fall by only $(1 - \delta)\hat{\beta}$. Over time, the regulator could account for this discrepancy in order to partially correct for biased estimates. In essence, the regulation of pollutants amounts to a natural experiment that can be used to identify parameter estimates themselves. This drawback and more proactive possibilities are addressed in the next section.

3.1 Accounting for Structural Uncertainty

An incomplete set of emissions can lead to changes in the stringency of environmental regulation. It can also lead to uncertain and possibly biased estimates of the marginal physical effect of known emissions on ambient pollution levels since total pollution must be attributed
to a subset of emitters. Such “transfer coefficient” uncertainty manifests as multiplicative uncertainty. This subsection considers the case when the regulator accounts for the possibility they may observe only a subset of emitters.24 Throughout this section we assume all emitters are homogeneous to show the first order effects of accounting for structural uncertainty.25

Consider a case in which the regulator knows that only a subset of polluters are identified but does not know the fraction of known emitters relative to actual emitters, \( \alpha \). Put another way, the regulator knows that \( \alpha \leq 1 \), but only has prior beliefs over what the true fraction of identified emitters is at any point in time: \( \alpha \sim \mu(\alpha) \). In this case of “known bias” the regulator knows their estimates of identified emissions’ marginal effect on ambient pollution levels are upwardly biased and knows there is a fraction of unknown emitters of size \((1 - \alpha)\).26 Maintaining homogeneity of all emitters for simplicity, take \( \tilde{x}_i \) to denote the actual emissions of each of the \( \alpha N \) known emitters. Similarly, since unknown emitters are unregulated by definition, take \( x^*_{un} \) to denote the actual emissions of the \((1 - \alpha)N\) unknown

---

24 Thus far we have focused on how changes in the percentage of known emissions affect estimates of known emissions on ambient pollution levels. One reason for this is that it is impossible for the regulator to know with certainty what the new percentage of emissions, \( \alpha_1 \), or the old percentage of emissions, \( \alpha_0 \), actually are. Instead, the discovery only reveals the amount of new emissions discovered as a percentage of old emissions. Thus, the discovery informs the regulator of the relative change in the composition of emissions but not the level of known emissions directly.

25 While the heterogeneous emitters case is more realistic, as a first step in this line of inquiry we are trying to develop intuition for how accounting for structural uncertainty relates to the two simple homogeneous emissions cases presented above. Further, we also do not address the comparative statics of relative effects of discovering additional emitters while accounting for structural uncertainty. This is a choice made to keep focus on how accounting for structural uncertainty changes the first order question of optimal emission levels conditional on an information set.

26 An alternative form of misspecification has the regulator accounting for the possibility of unknown emitters but not correcting the estimate of the marginal physical impact. In that situation, optimal policy is fully determined by the shape of the damage function: the regulator should regulate known emitters more when setting expected marginal benefit equals expected marginal cost if the damage function is convex. The opposite is true when the damage function is concave. A proof of this result is available from the authors upon request.
emitters. Conditioning on a particular percentage of known emitters, $\alpha$, we argue that the regulator can correct for the bias induced by estimating $\hat{\beta}$ from an incomplete set of emissions. Consider the simple OLS example above: conditional on $\alpha$, the regulator can form an unbiased estimate of the marginal physical effect of known emitters by multiplying $\hat{\beta}$ by $\alpha$: $\alpha \hat{\beta} = \alpha \frac{\beta}{\alpha} = \beta$. We don’t claim that regulators and atmospheric chemists actually use OLS estimation and correct transfer coefficient estimates in this way. We argue rather that conditional on knowing the level of incomplete information, it is possible to reduce or eliminate bias in an estimated transfer coefficient. Put another way, it is possible that conditional on knowing $\alpha$, it is possible that $E[\hat{\beta}|\alpha] = \beta$.27

Conditional on a particular $\alpha$, then, expected ambient pollution is:

$$E[y|\alpha] = \alpha N \bar{x} \beta + (1 - \alpha) N x_{un}^* \beta$$

(6)

In equation (6) the regulator fully accounts for the contributions of the unknown emitters in contributing to ambient levels, $(1 - \alpha) N x_{un}^* \beta$, because they condition on $\alpha$. Of course, at any point in time the regulator does not know the true fraction of identified emitters but instead has prior beliefs $\mu(\alpha)$ over the distribution of $\alpha$. As a result, the objective function of the regulator is:

$$\max_{\{\bar{x}\}_i^N} \int \left( (\alpha N) \pi_i(\bar{x}) + (1 - \alpha) N \pi_i(x_{un}^*) - D(\alpha N \bar{x} \beta + N(1 - \alpha) x_{un}^* \beta) \mu(\alpha) \right) d\alpha. \quad (7)$$

27Homogeneity of all emitters is equivalent to assuming $\pi_i(x_i) = \pi_j(x_j)$ and $\beta_i = \beta_j$ for all $i$ and $j$ in the set of $N$ emitters.

28Note that in our proof of Proposition 4, we make the less stringent assumption of $E[\hat{\beta}|\alpha] - \beta < \hat{\beta} - \beta$ if $\hat{\beta}$ is the estimated effect when not accounting for bias.
Note that the regulator has only $\alpha N$ control variables since only $\alpha N$ emitters are known with certainty.\footnote{Given that the regulator accounts for the possibility of an incomplete emissions inventory, an alternative objective function of the regulator worth noting is one that would put extra weight on errors in their regulatory actions. The implication of the regulator penalizing errors is that there is an additional benefit associated with increases in profits of regulated firms. Similar intuition holds if the regulator asymmetrically weights the interests of different political constituencies such as environmental groups or industries.} The resultant first order conditions for the known bias case is therefore:

$$
\pi_i'(\tilde{x}) = \int D'(\alpha N \tilde{x} \beta + N(1 - \alpha)x^*_{un} \beta) \beta \mu(\alpha) d\alpha \quad \forall \ i = 1, \ldots, \alpha N
$$

$$
= E[D'(Y) \beta | \mu(\alpha)] \tag{8}
$$

In this circumstance, the regulator accounts for the known bias in both calculating the expected level of ambient pollution and in correcting their estimates of the marginal physical effect of emissions on ambient pollution, making this management regime dynamically consistent. No elements of the unknown portion of ambient pollution, $(1 - \alpha)Nx^*_{un} \beta$, are known to the regulator. However, if $E[\alpha | \mu(\alpha)] < 1$, then there must be more total emissions than the set of observed emissions.

We now show our two main proofs for this subsection. First consider the relatively simple case of the regulator setting optimal equilibrium ambient pollution and emission levels when accounting for structural uncertainty versus the case when only a subset of emitters are known but the marginal effect of all emitters is \textit{correctly} identified in both cases, akin to section 2.1 above:

\textit{Proposition 3: If only a subset of homogeneous emitters is accounted for by the regulator but marginal emissions are estimated unbiasedly, it is optimal to over-regulate known emitters when accounting for structural uncertainty.}
Intuitively, Proposition 3 stems from the regulator acknowledging their policy will have less influence on actual ambient pollution levels than a regulator with naive beliefs acknowledges. Regulators in this case know they have biased estimates of marginal physical effects due to the possibility of unknown unapportioned sources of emissions. As a result, they balance the loss in profits of known emitters with the social welfare gains of lower expected ambient pollution. When uncertainty in the relationship between emissions and ambient pollution levels is systematically accounted for, optimal emission levels depends jointly on the nature of the damage function, the profit function and beliefs about the proportion of known emitters.

Consider the regulator setting optimal ambient pollution and emission levels when accounting for structural uncertainty versus the case when they are not accounting for structural uncertainty (e.g., only a subset of emitters are known and the marginal effect of all emitters is estimated with bias) akin to section 2.2 above. This is a more difficult situation to analyze theoretically because $\alpha$ can enter non-linearly in the estimate of marginal physical effects of emissions on ambient levels of pollution. In the simple OLS example above, for example, estimated coefficients are $\hat{\beta} = \frac{\beta}{\alpha_0}$. In an effort to obtain a readily interpretable result, we make the strong assumption that the slope of marginal damages with respect to ambient pollution levels is linearly increasing. However, even with this restrictive assumption we are unable to determine whether currently known emitters should be regulated more or less stringently:

**Proposition 4:** For linearly increasing marginal damages and upwardly biased estimates, it is not always optimal to regulate known emitters more stringently when accounting for
There are three key pieces of intuition for Proposition 4. First, if the set of unknown emitters is believed to be large, then there is a large quantity of emissions that the regulator believes they cannot control. This is the direct effect of structural uncertainty on the regulator’s problem. This gives the regulator incentive to increase regulatory burden on known emitters when they account for structural uncertainty since unapportioned emission sources increase the marginal cost of an additional unit of pollution from known emitters. Second, estimated marginal physical effects of emissions from known emitters are always higher when the regulator accounts for structural uncertainty relative to when they do not. This gives the regulator incentive to reduce the regulatory burden on known emitters when they account for structural uncertainty. This is the indirect effect of uncertainty on the regulator’s problem. Third, if the marginal profit from emissions is small, then the importance of the first two effects is magnified since, from a social welfare perspective, it is costly to significantly alter emissions of known polluters. For linearly increasing marginal damages, then, we cannot say whether known emitters should always be regulated more stringently when accounting for structural uncertainty. Lastly, while full bias correction is feasible conditional on priors $\mu(\alpha)$ is feasible, the proof of Proposition 4 only relies on estimates conditioning on $\mu(\alpha)$ being less biased than when not.

The indeterminant result from Proposition 4 is somewhat unsatisfying. It is unexpected, though, as even with the restrictive linearity assumption on increases in marginal damages, the general functional form on profits and the non-linearities induced by biased estimation make definitive analytic statements challenging. Generally speaking, however, Proposition
4 does offer insight for comparative statics. For example, if the regulator is fairly certain they have identified all emitters (e.g., $E[\alpha] >> 0$) then it is more likely that accounting for structural uncertainty will lead to increased stringency on known emitters. The reason is the non-linearity in the bias when naively estimating marginal emissions. In the simple OLS case, for example, as $\alpha$ approaches one, the marginal change in $\alpha$ is constant but the marginal change in $\frac{1}{\alpha}$ gets smaller. The converse is true for lower levels of $\alpha$. Similarly if marginal profits are very steep, then $x^* \approx x^*_{un}$, and accounting for structural uncertainty will lead to decreased stringency on known emitters.

4 Discussion

We introduce structural uncertainty to the relationship between emissions and ambient levels of pollution. Allowing for structural uncertainty leads to a theoretical foundation of multiplicative uncertainty over marginal damages of pollution. Our model shows that estimates of marginal damages from known emitters will often be biased upward if this structural uncertainty is not explicitly accounted for by regulators. As a result, known emitters can be over-regulated relative to when the regulator conditions on the possibility of structural uncertainty.

Our analysis puts an emphasis on whether the marginal impacts of emissions on ambient pollution levels are known or whether they are unknown. While we argue marginal physical impacts will often be uncertain when estimated from an incomplete set of emissions, structural uncertainty can exist even when all emitters are known. Perhaps the clearest example of this phenomenon is the recent discovery of large ships in the Los Angeles Harbor con-
tributing 10%-40% of sulfates in the Los Angeles air basin, a far larger level than previously thought (Dominguez et al. (2008)). When the effect of harbor emissions on ambient sulfate levels was accurately estimated, the ships began to plug-in to mainland electricity while in port. The process by which previously regulated sulfate emitters were affected by the sulfate reductions from the ships, however, is unclear.

Our work suggests there may be substantial efficiency gains to investing resources to help resolve and reduce structural uncertainty since we show it affects the optimal stringency of environmental regulation. Currently, budget allocations suggest that pollution control agencies react to scientific discoveries that alter the understanding of the pollution process, but do so in a passive manner. We provide a framework for evaluating the potential gains from actively putting resources into reducing scientific uncertainties surrounding the pollution process. There is, therefore, an interesting political economy question as to the incentives faced by actors involved in the pollution control process. Our work can also be used as the starting point to formally estimate the expected gains from R&D programs aimed at quantifying the total amount of known emissions, identifying unapportioned sources of a particular pollutant and reducing structural uncertainty. While greenhouse gases are an obvious candidate, a much smaller system with a well-understood physical model for translating emissions into ambient pollution levels might be a better test case. Lastly, our results are primarily existence results. Therefore, there is an opportunity for additional research to provide uniqueness results defined by specific economic parameters and physical relationships.
References


Proofs

Lemma 1: An increase in the set of known emitters from may increase or decrease the expected level of ambient pollution if marginal physical effects are biased upward.

PROOF: Let a subset of emitters $K = \alpha_0 N$ initially be used to estimate marginal effects of emissions on ambient pollution levels where $\alpha_0 < 1$. For simplicity, assume all emitters are homogeneous except that some are known and some are unknown. Let estimated transfer coefficients conditional on $\alpha_0$ be $\hat{\beta}_0$ and assume that $\hat{\beta}_0 > \beta$.

Expected ambient pollution, conditional on regulation, is $E[y^*|\alpha_0] = \tilde{X}^* \hat{\beta}_0 = \alpha_0 N \tilde{x}_0^* \hat{\beta}_0$ where $\tilde{x}_0^*$ is the regulated level of emissions of known emitters conditional on $\alpha_0$ and $\hat{\beta}_0$ given by the first order conditions in equation (5).

Assume there is a discovery of a new class of unregulated emissions that were emitting at a level $\tilde{x}_{un}^*$ where $\tilde{x}_{un}^* > \tilde{x}_{un}'$. Assume there are $(\alpha_1 - \alpha_0) N$ emitters discovered (e.g., $\alpha_1 > \alpha_0$). Taken together, this implies that total discovered emissions could be either larger or smaller than the existing set of emissions. Lastly, assume that the new estimated transfer coefficient is $\hat{\beta}_1 < \hat{\beta}_0$ so that transfer coefficient estimates are revised downward.

The new expected level of ambient pollution before any changes in emissions of regulated and unregulated emitters is $E[y|\alpha_1] = (\alpha_1 - \alpha_0) N \tilde{x}_{un}^* \hat{\beta}_1 + \alpha_0 N \tilde{x}_0^* \hat{\beta}_0$.

The sign of the difference between expected emissions ex ante and ex post of discovery can be expressed as:

$$\text{sign}[E[y^*|\alpha_1] - E[y|\alpha_0]] = \text{sign}[(\alpha_1 - \alpha_0) N \tilde{x}_{un}^* \hat{\beta}_1 + \alpha_0 N \tilde{x}_0^* \hat{\beta}_0 - \alpha_0 N \tilde{x}_0^* \hat{\beta}_0] = \text{sign}[\alpha_0 \tilde{x}_0^*(\hat{\beta}_1 - \hat{\beta}_0) + \tilde{x}_{un}^* \hat{\beta}_1 (\alpha_1 - \alpha_0)].$$ (9)

There are two terms in equation (9). The first term describes the decrease in expected emissions from previously identified emitters when the regulator revises their transfer coefficient estimate downward. This term is negative since transfer coefficients are initially biased upward by assumption. The second term is the additional emissions expected from the newly discovered emitters. If the first term dominates the second, expected emissions decreases. Conversely if the second term dominates expected emissions increases giving the desired result.

Proposition 1: If expected ambient pollution decreases upon the discovery of new emitters, then previously known emitters’ optimal emission levels strictly increase. Conversely, if expected ambient pollution increases, previously known emitters’ optimal emission levels strictly decrease.

PROOF: Let initial estimated transfer coefficients be $\hat{\beta}_0$ and the estimated transfer coefficient after new emissions are discovered be $\hat{\beta}_1 < \hat{\beta}_0$. Assume that $\alpha_0$ increases to $\alpha_1$ causing $E[y|\alpha_1] < E[y^*|\alpha_0]$ and damages are convex, $E[D'(y)|\alpha_1] < E[D'(y^*)|\alpha_0]$. By construction, the marginal expected damages attributable to previously known emitters in the first order
condition after the discovery of new emitters, shown in equation (5), is smaller than the marginal expected damages before the discovery:

\[ E[D'(y|\alpha_1)]\hat{\beta}_1 < E[D'(y^*|\alpha_0)]\hat{\beta}_0. \]  

(10)

The previous level of regulated emissions, therefore, cannot be optimal:

\[ \pi'_i(\hat{x}_i^*|\alpha_0) = D'(E[y^*|\alpha_0])\hat{\beta}_0 > D'(E[y|\alpha_1])\hat{\beta}_1. \]  

(11)

Concavity of the profit function dictates that the marginal benefit of emissions is too high and it must be that \( \hat{x}_i^*|\alpha_1 > \hat{x}_i^*|\alpha_0 \). By inspection, the converse is also true giving the desired result.

**Proposition 2:** An increase in the set of known emitters can increase or decrease the optimal level of ambient pollution.

**PROOF:** The objective function of the regulator in this model is to maximize expected social welfare. If the regulator is fully informed, their maximization problem is:

\[ \max_{\pi_i} E \left[ \sum_{i=1}^{N} \pi_i(x_i) - D(y) \right], \]

where \( y = \sum x_i^*\beta_i + \epsilon \) and \( x \) is the \( N \times 1 \) vector of emissions. Differentiating this equation with respect to the \( N \) control variables gives the \( N \) first order conditions (\( E[\pi_i(x_i^*)] = E[D'(y^*)\beta_i] \) \( \forall i \)) and implicitly define the set of optimal emission levels \( \{x^*\} \).

Take a simplified case where there are only two emitters \( i \) and \( j \). Without loss of generality, assume that \( \beta_i = \beta_j = \beta = 1 \). Assume that both \( \pi_i(\cdot) \) and \( \pi_j(\cdot) \) are strictly concave and that \( \pi_i''(\cdot) = -k_i \) and \( \pi_j''(\cdot) = -k_j \) so that \( k \) indexes the magnitude of the second derivative for each emitter’s profit function. Assume that the regulator only knows of emitter \( i \) and has formed an upwardly biased estimate the marginal physical effect of emitter \( i \) of \( \hat{\beta} > 1 \). The regulator will let the following single first order condition set emitter’s \( i \)'s:

\[ \pi'_i(x_{i0}^*) = D'(x_{i0}^*)\hat{\beta}. \]  

(12)

Note that emitter \( j \) will emit at the level defined by \( \pi'_i(x_{j0}^*) = 0 \) since they are unidentified and therefore unregulated. As a result, before being regulated emitter \( j \)'s total emissions are \( x_{i0}^* + x_{j0}^* \).

Assume that after identifying emitter \( j \), the regulator correctly identifies marginal physical effects \( \beta = 1 \). The new optimal levels of regulated emissions are given by

\[ \pi'_i(x_{i0}^* + \Delta x_i) = D'(x_{i0}^* + \Delta x_i + x_{j0}^* - \Delta x_j) \]

\[ \pi'_j(x_{j0}^* - \Delta x_j) = D'(x_{i0}^* + \Delta x_i + x_{j0}^* - \Delta x_j) \]  

(13)

where \( \Delta x_i \) and \( \Delta x_j \) represent the increase and decrease in emissions after emitter \( j \) is identified and \( x_{i0}^* + \Delta x_i + x_{j0}^* - \Delta x_j \) the associated level of emissions. Since \( D'(x_{i0}^* + \Delta x_i + x_{j0}^* - \Delta x_j) > 0 \) there must be a decrease in emitter \( j \)'s emissions due to the concavity of \( \pi_j(\cdot) \).
We now proceed by analyzing three cases: no change in emissions after discovery of \( j \), an increase in emissions and finally a decrease in emissions. Assume first that \( \Delta x_i - \Delta x_j = 0 \). In this case there is no change in total emissions after \( j \) is discovered and the regulator sets emissions according to equations (13). This would occur if \( \hat{\beta} \) was sufficiently high so that when the marginal damage curve rotates down, the increase in \( i \)'s emissions exactly offset the decrease in \( j \)'s. This would occur despite emitter \( j \)'s emissions \( (x^*_j - \Delta x_j) \) entering the damages function. Hence, emissions could remain unchanged after the discovery of an additional emitter.

Now augment the profit function for emitter \( j \) such that \( \tilde{\pi}'(x^*_j) = 0 \) at the same point as previous but that \( \tilde{\pi}''(\cdot) = -\tilde{k}_j \) where \( \tilde{k}_j > k_j \). By assumption, \( \tilde{\pi}'(x^*_j - \Delta x_j) > \pi'(x^*_j - \Delta x_j) \). This cannot be an equilibrium. As a result, emissions from \( j \) must increase and emissions from \( i \) must decrease thereby leading to an increase in emissions after the discovery of emitter \( j \), correctly revising estimates of marginal physical effects of emissions and subsequently revising regulated emission levels. To arrive at a decrease in emissions let \( \tilde{k}_j < k_j \) and the converse is true. This completes the proof.

**Proposition 3:** If only a subset of homogeneous emitters is accounted for by the regulator but marginal emissions are estimated unbiasedly, it is optimal to over-regulate known emitters when accounting for structural uncertainty.

**Proof:** The first order condition in the full information case is \( \pi'(x^*) = D'(Nx^*\beta) \) which implicitly defines the optimal level of emissions for all emitters. The first order condition when accounting for structural uncertainty, equation (8), can be rewritten as \( \pi'({\tilde{x}}) = E[D'(\alpha N\tilde{x}\hat{\beta} + (1 - \alpha)Nx^*_{un}\hat{\beta})\beta|\mu(\alpha)] \). The term \((1 - \alpha)Nx^*_{un}\beta \) will be positive for non-degenerate beliefs. Now assume that \( \tilde{x} = x^* \). Because the term \((1 - \alpha)Nx^*_{un}\beta \) will be positive and the profit function, \( \pi'(\cdot) \) is strictly concave, it must be the case that \( x^*_{un} > \tilde{x} \), therefore this cannot be an equilibrium. By concavity of the profit function, allowed emissions of known emitters, \( x^* \), should be reduced to less than emissions of emitters in the full information case, \( x^* \), to satisfy the first order condition (8), thereby completing the proof.

**Proposition 4:** For linearly increasing marginal damages and upwardly biased estimates, it is not always optimal to regulate known emitters more stringently when accounting for structural uncertainty relative to when it is not accounted for.

**Proof:** Assume that the marginal damages from ambient pollution increase at a constant rate \( d \). The first order condition when accounting for structural uncertainty, equation (8), can be rewritten as \( \pi'({\tilde{x}}) = E[d(\alpha N\tilde{x}\hat{\beta} + (1 - \alpha)Nx^*_{un}\hat{\beta})\beta|\mu(\alpha)] \). Assuming that that bias can be corrected (e.g., \( E[\hat{\beta}|\alpha] = \beta \)) and simplifying terms, we can further reduce the expression to \( \pi'({\tilde{x}}) = E[d(\alpha N\tilde{x}\hat{\beta} + (1 - \alpha)Nx^*_{un}\hat{\beta})\beta|\mu(\alpha)] \). The term \((1 - \alpha)Nx^*_{un}\beta \) will be positive for non-degenerate beliefs. Take \( \tilde{x} \) to be the optimal level of emissions of identified emitters when accounting for structural uncertainty.
When structural uncertainty is not accounted for assume that $\hat{\beta} > \beta$, the first order condition for the $\alpha N$ known emitters is $\pi'(x^*) = E[d(\alpha N x^* \hat{\beta}) \hat{\beta}]$ which implicitly defines the optimal level of emissions, $x^*$ for all known emitters.

Further, temporarily assume that $\tilde{x} = x^*$ so that $\pi'(\tilde{x}) = \pi'(x^*)$. We can compare the expected marginal damages across regulatory regimes by signing the expression:

$$E[d(\alpha N x^* \hat{\beta}) \hat{\beta}] - E[d(\alpha N \tilde{x} \beta + (1 - \alpha) N x_{un}^* \beta) \beta | \mu(\alpha)]$$  \hspace{1cm} (14)

Taking advantage of the linearity and $\tilde{x} = x^*$ assumptions, equation (14) simplifies to $x^* \alpha (\hat{\beta}^2 - \beta^2) - (1 - \alpha) x_{un}^* \beta^2$. The first component of this expression is positive since $\hat{\beta}^2 - \beta^2 > 0$ by assumption. Note further that positivity requires only that the uncorrectly estimated coefficient, $\hat{\beta}$, be larger than the corrected coefficient, $\beta$. The second component of this expression is negative since $(1 - \alpha) x_{un}^* \beta^2 > 0$. As a result, the sign of this expression is determined by the relative magnitude of the bias in $\beta$, the relative size of identified versus unidentified emitters and the relative slope of the profit function (e.g., the difference between $\tilde{x}$ and $x_{un}^*$). If the expression is positive, the known emitters must be less intensely regulated when not accounting for bias. The converse is also true. This completes the proof.
Estimating Marginal Physical Effects

For simplicity of exposition, one key assumption in the above analysis is that although some emitters are identified and some are not, all emitters have perfectly correlated emissions. More precisely, if emissions from a known emissions source increase by $\phi\%$, then we assume that all emissions from unknown emitters increase by $\phi\%$ as well (e.g., $\text{cov}(x_i, x_j) = 1$ for all $i, j$). We assume that this relationship holds even if an emitter $i$ is regulated and emitter $j$ is not.

We acknowledge that perfect correlation is a very strong assumption. Insofar as emissions are procyclical, though, it is reasonable that emissions are positively correlated. We make it to clearly show how uncertainty in the relationship between emissions and ambient pollution manifests for the regulator.

An alternative situation arises when unknown emissions are imperfectly correlated with known emitters. One intuitive way to understand how imperfect correlation affects estimated marginal effects is by setting up the problem as an omitted variable bias regression. If $X$ is the set of known emitters and $Z$ the set of unknown emitters, assume the true data generating process for ambient pollution is $Y = X\beta + Z\delta + \epsilon$. If instead of estimating the true data generating process, the regulator estimates $Y = X\beta + \epsilon$, the estimated coefficient is: $E[\hat{\beta}|X] = \beta + \frac{X'Z}{X'X}\delta$. The second part of this expression is the well-known bias due to omitted variables. If the covariance of $Z$ and $X$ is negative, then $\hat{\beta} < \beta$, as in the fugitive dust case considered in the main text.

In either case, that of overestimated fugitive dust and negatively correlated unknown emitters, the estimated marginal physical effect of known emitters on ambient pollution levels is biased downward. As a result, the expected marginal damages associated with known emitters will be too low and lead to under-regulation of known sources. While this is a possibility, the major feature of the model remains that an uncertain relationship between emissions and ambient pollution manifests as multiplicative uncertainty related to uncertain and biased marginal physical effects of emissions on ambient pollution levels.