

# Structural Uncertainty and Pollution Control: Optimal Stringency with Unknown Pollution Sources

Richard T. Carson<sup>1</sup> · Jacob LaRiviere<sup>2</sup>

Accepted: 8 April 2017 / Published online: 4 May 2017  
© Springer Science+Business Media Dordrecht 2017

**Abstract** We relax the common assumption that regulators know the structural relationship between emissions and ambient air quality 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 unknown emission sources, can counterintuitively lead to increases in both optimal emissions and ambient pollution levels.

**Keywords** Information · Regulation · Externalities

**JEL Classification** Q58 · D81 · C11 · Q53

## 1 Introduction

Environmental regulations can have substantial economics impacts. As a result, there is considerable interest in knowing the precise relationship between emissions and ambient pollution levels in order to develop the optimal stringency pollution regulation. As long as the relationship between emissions and ambient pollution is known, any uncertainty reduces to

---

Thanks to Allen Basala, Roger Brode, Chris Costello, Gerardo Dominguez, David Evans, Meredith Fowlie, Ben Gillen, Ken Gillingham, Ted Groves, Stephen Holland, Charlie Kolstad, Mark Machina, Steve Polasky, Glenn Sheriff, Kerry Smith, Chris Timmins, Mark Thiemens, Roger Von Haefen, Christian Vossler, Choon Wang, seminar participants at Tennessee, Minnesota, TREE, participants of the 12th Occasional California Workshop on Environmental and Resource Economics, and two anonymous referees for very helpful comments and discussion.

---

✉ Jacob LaRiviere  
jlarivi1@utk.edu

<sup>1</sup> Department of Economics, UC San Diego, Lo Jolla, CA, USA

<sup>2</sup> Department of Economics, University of Tennessee, 525 Stokely Management Center, Knoxville, TN 37996-0550, USA

Weitzman (1974) 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 call “distributional uncertainty” to reflect the presence of stochasticity in the costs or benefits of emissions, 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. Examples include when the regulator is unaware of specific pollution sources or substantively misinformed about the magnitude of pollution coming from a source. We show that this form of uncertainty 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.<sup>1</sup> 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 “misapportioned”. 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; 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. China’s carbon emissions are notoriously challenging to track: the government announced in late 2015 they had been burning 17% more coal per year than previously announced.<sup>2</sup> In each case, and in many others, 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.<sup>3</sup>

Potential welfare gains from a better understanding of the structural relationship between emissions and ambient pollution seem to be largely ignored. For example, the 2010 EPA

<sup>1</sup> When the effect of harbor emissions on ambient sulfate levels was accurately estimated, emissions from ships were substantially reduced by plugging in to mainland electricity while in port.

<sup>2</sup> See New York Times article at: <http://tinyurl.com/gvodrvf>.

<sup>3</sup> Haagen-Smit (1952) famously discovered that automobile exhaust was largely responsible for smog in the L.A. air basin. Automobiles 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 finger printing 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.

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; 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 U.S. 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 delivering estimates for ozone and particulates 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 (McKee et al. 2004, 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 et al. 2012a, b).

Possible problems in translating emissions into ambient concentrations was occasionally mentioned in early economic work on pollution control (Spence and Weitzman 1978; Crandall 1981). The implications of imperfect information governing the relationship between emissions and ambient pollution, however, has received attention limited only to relatively narrow pollution control problems rather than pollution control problems more broadly. Examples include evaluating optimal pollution control when the effect of emitters on ambient pollution is stochastic (Horan et al. 1998; Hamilton and Requate 2012), when damages from emissions vary spatially (Muller 2011; Fowle and Muller 2013), when precisely which out of a set of possible non-point source emitters do in fact emit (Segerson 1988; Hansen 1998), and non-point source water pollution permit allowances (Horan and Shortle 2005). This literature considers structural uncertainty but do not explicitly consider cases where there is uncertainty over if the full set of all polluters are identified.

We build on this literature by investigating the more general issue of how allowing uncertainty in how known emission sources maps to observed ambient pollution levels impacts (1) expected damages from emissions and (2) the optimal stringency of pollution control.<sup>4</sup> In doing so we focus on air pollutants but the qualitative results generalize other forms of pollution, such as the water pollution examples discussed above. We do this in three parts, each of which is a distinct contribution. First, we develop a model of structural uncertainty. 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.<sup>5</sup> Second, we use the model to identify how resolving uncertainty through discovering new emission sources impacts optimal stringency of ambient pollution levels. This approach is motivated by discoveries of new sources in the physical sciences literature discussed above. Third, we investigate how explicitly accounting for the presence of structural uncertainty impacts

<sup>4</sup> Put 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.

<sup>5</sup> For an example of a well-executed reduced form model of damages see Lemoine and McJeon (2013).

optimal pollution stringency levels. This final approach seeks to be more normative, whereas the previous two are positive economic models.

## 2 Model

This section develops our theoretical model. Throughout we assume 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. We focus only on the stringency of pollution control in this paper.

Each firm, industry or sector  $i$  chooses a level of emissions,  $x_i$ , to maximize profits  $\pi_i(x_i)$ . The indicator ‘ $i$ ’ could also index countries or anthropogenic versus geological sources. We assume that  $\pi_i(x_i)$  is initially increasing in  $x$  and twice differentiable with  $\pi_i''(x_i) < 0$ . Similar in spirit to [Horan and Shortle \(2005\)](#), each unregulated firm emits at a level  $x_i^*$  such that  $\pi_i'(x_i^*) = 0$ . Let the regulator know the exact form of  $\pi_i(x_i)$  for all  $i$  with certainty. There are  $N$  total emitters. Firm profit functions can be either homogeneous or heterogeneous (e.g.,  $\pi_i()$  or  $\pi()$ ). The steepness and level of marginal profits from emissions can vary by source.<sup>6</sup>

Emissions from firms contribute to the ambient pollution level,  $y$ , which decreases social welfare according to a strictly increasing damage function  $D(y)$ . We make two simplifying assumptions about ambient pollution and associated damages. First, the regulator knows the damage function  $D(y)$  perfectly. For exposition we assume that marginal damages are linear in  $y$ :  $D'(y) \propto cy$  although we don’t leverage it in a proof until our third Proposition. Second, the ambient level of pollution,  $y$ , is perfectly observable. We make this assumption throughout. Ambient pollution levels are actually only measured 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 do not investigate the implications of this source of uncertainty.

We assume ambient pollution is related to emissions according to the simple linear function  $y = f(\{x\}_{x=1}^N) = \sum_i x_i \beta_i + \epsilon$ , where  $\epsilon$  is an idiosyncratic term 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.<sup>7</sup> 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](#)).

A linear relationship between ambient pollution and emissions is appealing for simplicity but it is not innocuous. First, while linearity is a reasonable approximation 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,

<sup>6</sup> From a modeling 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.

<sup>7</sup> 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.

our ambient pollution  $y$  should be thought of as ambient pollution in a particular location.<sup>8</sup> 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 our model can be generalized, we restrict our analysis to the simplest case possible to clearly 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.,  $E[y] = \sum_i x_i \beta_i$ ). We initially assume all polluters' emissions have the same physical impact on ambient pollution levels (e.g.,  $\beta_i = \beta_j$ ), but later relax this assumption.<sup>9</sup>

If the marginal physical effects of emissions on ambient levels of pollution are known 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 (recall the expectations operator is needed since realizations of  $y$  are stochastic):

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

The  $N$  first order conditions for optimality are:

$$\pi'_i(x_i^*) = E[D'(y^*)\beta_i] = D'(y^*)\beta_i \quad \forall i = 1, \dots, N. \quad (2)$$

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 Eq. (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 Eq. (2). This represents the classic non-depletable externality problem.

## 2.1 Modeling Uncertainty Between Emissions and Ambient Pollution

We now relax the assumption that the marginal physical effect of known emitters on ambient pollution,  $\beta_i$ , is known and that all emitters are identified. We 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. 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 discuss the EPA's procedure for doing this in the "Appendix". To make our example as clear as possible, we initially assume the regulator completely

<sup>8</sup> Our model has implications for the related literature which uses "transfer coefficients" to improve efficiency of regulation (Muller 2011; Fowle and Muller 2013). However, a spatial dimensions is not needed to develop the intuition in our model.

<sup>9</sup> There are at least two common situations where this would not be the case that can be seen as extensions. First, there are many types of GHGs which contribute to climate change. For example, methane is a much more potent greenhouse gas than carbon dioxide, but ultimately cause damages via the same mechanism. Second, the same type of emissions from two different sources could impact ambient pollution at different rates. For example, particulate matter from two coal power plants, one close to a city and one far away, may both impact ambient PM 2.5 levels in a city, but have different transfer coefficients (Muller 2011; Fowle 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.

ignores that they observe an incomplete set of emitters. We also initially assume emitter homogeneity (e.g., for both  $\beta$  and  $\pi(\cdot)$ ).

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  $N \times 1$  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.<sup>10</sup> 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.<sup>11</sup> Introducing the  $\alpha$  parameter in this way is the key distinguishing factor in our approach relative to other structural uncertainty approaches such as those on water quality trading non-point source pollution (Segerson 1988; Horan and Shortle 2005) or transfer coefficients (Fowlie and Muller 2013).

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  $\tilde{X}$ . As before, assume the true data generating process is  $y_t = \sum_{i=1}^N x_{i,t} \beta_i + \epsilon_t$  but also assume  $\epsilon_t$  is a well-behaved normally distributed random error component and that emissions affect ambient pollution identically. For illustrative purpose, suppose the regulator estimates the coefficient vector  $\beta$  by regressing  $Y$  on  $\tilde{X}$  using Ordinary Least Squares (OLS). While regulators may not actually perform OLS to parameterize air dispersion models, we want to show very precisely in the simplest case how inferring transfer coefficients from an incomplete emissions inventory leads to biased estimates of transfer coefficients due to over-attribution of known emitters to ambient pollution levels.

Estimates from attributing known emissions to observed ambient pollution levels lead to upwardly biased estimates. Specifically, estimates take the form  $\hat{\beta} = \frac{Y' \tilde{X}}{\tilde{X}' \tilde{X}} = \frac{Y' X}{\alpha_0 X' X} = \frac{\beta}{\alpha_0}$ . 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.<sup>12</sup> Note that if the regulator knew  $\alpha_0$  with certainty, they could correct their biased estimate of  $\hat{\beta}$ , which we discuss in detail below.<sup>13</sup>

<sup>10</sup> 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,  $\alpha_0$ .

<sup>11</sup> From a modeling perspective, this model is very similar to one with a set of point source polluters which can be regulated and monitored and a set of non-point source polluters which cannot be. The difference between our approach is the cause of partial regulation (unknown sources motivated by the physical sciences literature rather than known sources which can not be monitored). This matters for policy: paying for better technology to monitor non-point source polluters is different than paying for better scientific understanding of the pollution process.

<sup>12</sup> 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 et al. (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.

<sup>13</sup> 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 U.S. EPA estimates the level of “fugitive emissions” from sources that are difficult if not impossible to

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. The resulting optimality condition obtained from maximizing expected welfare is:

$$\max_{\{x\}_1^K} \quad \Sigma_{i=1}^K \pi_i(x_i) - E[D(y)|\alpha_0] = \Sigma_{i=1}^K \pi_i(x_i) - E[D(\Sigma_i x_{i,t} \hat{\beta}_i)|\alpha_0] \quad (3)$$

$$\pi'_i(x_i^*) = E[D'(y^*)\hat{\beta}] = D'(y^*)\hat{\beta} \quad \forall i = 1, \dots, K = \alpha_0 N. \quad (4)$$

where  $E[y^*] = \tilde{X}^* \hat{\beta}$ . Note that conditioning on  $\alpha_0$  indicates the regulator estimates  $\hat{\beta}_i$  for all identified emitters  $i = 1, \dots, K = \alpha_0 N$ . Given our homogeneity assumption,  $\hat{\beta}_i = \hat{\beta}_j = \hat{\beta}$ . Since  $\alpha_0 < 1$  and  $\pi''_i(\cdot) < 0$ , the  $\tilde{x}_i^*$  implicitly defined by the set of Eq. (4) must be lower than in the case above when  $\beta$  is estimated without upward bias.<sup>14</sup> The main 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.

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 and profits from emissions (via uncertainty in  $\beta$ ) rather than assumed uncertainty over costs of abatement themselves.

This structural uncertainty model can be extended to fit the stock pollutant model of [Newell and Pizer \(2003\)](#). Assume 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 stock pollutant decay rate. When  $\delta$  is known,  $y_t = S_t - (1 - \delta)S_{t-1}$ . In any given year, 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 we leave to future research.

Footnote 13 continued

measure such as particulates from traffic on roads, leaks from pipes, and small industrial processes. 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.

<sup>14</sup> 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  $d$ . 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  $\pi'_1(x_1^*) = E[d(y^*)2] = d(2x_1^*)2 = 4dx_1^*$ . Conversely, the full information first order condition, given that firms are homogeneous, would be  $\pi'_1(x_1^*) = d(2x_1^*)1$ . Since the profit function is concave, the known emitter is over regulated in the partial information case.

### 3 Analytical 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. The marginal benefit of emissions (e.g.,  $\pi_i(\cdot)$ ) and the marginal physical effects of emissions (e.g.,  $\beta_i$ ) are allowed to vary across firms.

#### 3.1 Discovering New Emission Sources

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.<sup>15</sup> This is similar to implications of the recent discovery that there were substantial unaccounted for inventories of fugitive methane emissions (Miller et al. 2013).

Figure 1 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.<sup>16</sup> Figure 1 holds ambient pollution levels,  $y$ , fixed in order to highlight how the rotation of the marginal damage curve is exclusively tied to changes in the estimated marginal physical effect of known emitters on ambient pollution levels,  $\hat{\beta}$ .

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 marginal abatement costs than that of a previously regulated firm:  $\frac{\partial \pi_i(x_{i,0}^*)}{\partial x} > \frac{\partial \pi_j(x_{un}^*)}{\partial x} = 0$ , where  $x_{un}^*$  is the emissions level of the unregulated emission source. An increase in the set of emitters increases the emissions of previously known emitters and decreases the emissions of a newly discovered polluter if the marginal profit of the unregulated source is less than the regulator’s expectations of their marginal damages.

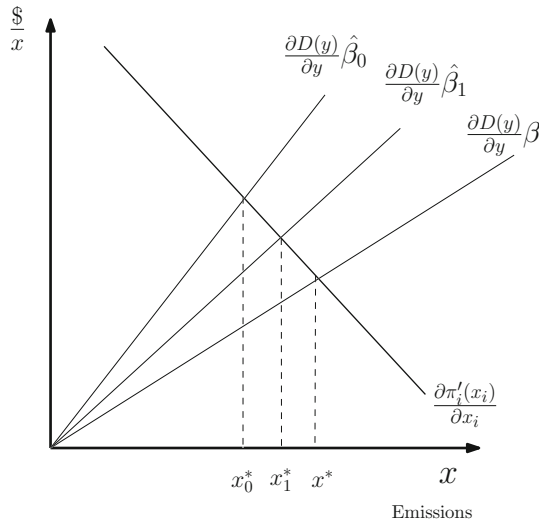
It is always weakly cheaper to achieve a given level of ambient pollution if a new emissions source is discovered. However, it may not be optimal to decrease ambient levels of pollution. Whether it is optimal or not for actual ambient pollution to decrease when a new emissions source is identified depends on what happens to expected ambient pollution levels when a new emitter is discovered (all proofs in the “Appendix”):

**Proposition 1** *An increase in the set of known emitters may decrease the expected level of ambient pollution if 1) marginal physical effects are sufficiently biased upward and 2) the set of new emitters is sufficiently small.*

<sup>15</sup> 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  $\tilde{X} = \alpha_1 X$  where  $\alpha_1 > \alpha_0$ . As a result, the estimated coefficient is lower than what is was before for previously known emitters:  $E[\hat{\beta}_1] = \frac{y'X}{\alpha_1 X'X} = \frac{\beta}{\alpha_1} < \frac{\beta}{\alpha_0}$ . Given that  $\alpha_1 > \alpha_0$ , the first order effect on  $x_i^*$  implied by the new optimality condition in Eq. (4) for  $i = 1, \dots, \alpha_1 N$  is that previously known emitters’ optimal emissions levels will be higher than they were before.

<sup>16</sup> In Fig. 2 and subsequent figures, we assume  $D(\cdot)$  is a convex function. Therefore, marginal damages are upward sloping. However, even if damages were linear (Muller and Mendelsohn 2009; Newell and Pizer 2003) in pollution (e.g., marginal damage curves are flat), discovering new emitters ( $\alpha$  increasing) shifts the marginal damage curve down causing emissions of previously known sources to increase.





**Fig. 1** Optimal firm emissions under different  $\hat{\beta}$

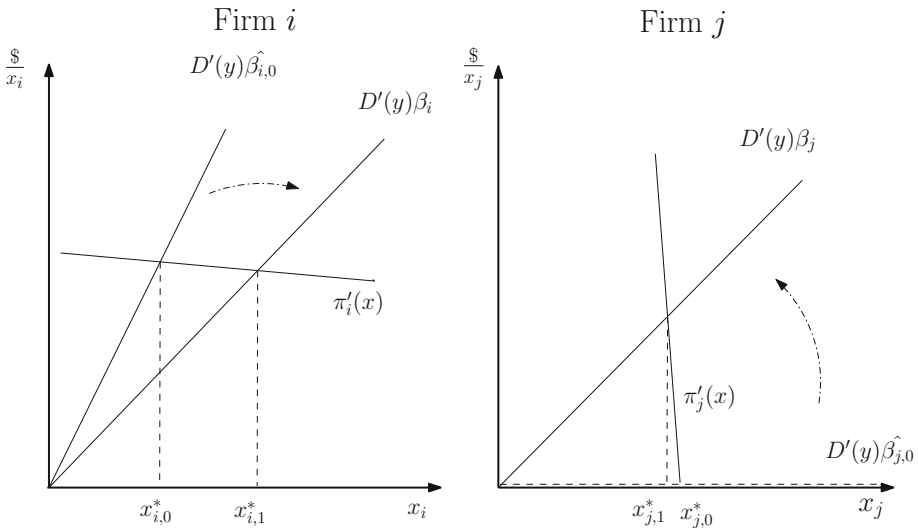
Intuition suggests that discovering new emitters leads expected emissions to increase. Proposition 1 says that once new emitters are discovered, expected ambient pollution levels could be lower than before. The level of expected ambient pollution can change once the new emitters are discovered because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

Proposition 1 results from non-linearity 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. If marginal physical effects are biased downward; then expected emissions always increase.

While Proposition 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 two Propositions:

**Proposition 2** *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 3** *An increase in the set of known emitters can decrease the optimal level of ambient pollution if benefits of emissions for newly discovered emissions is sufficiently large (e.g., decreasing emissions is sufficiently costly).*



**Fig. 2** Example of discovery of new emissions sources

Proposition 2 relates the nature of the damage function and to updated estimates of the marginal effect of emissions on ambient levels of pollution. 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.<sup>17</sup> 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.

Proposition 3 identifies the impact of discovering additional emitters on optimal 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 known emitters in each period. Third, each additional known emitter is now regulated. Fourth, the level of emissions from the previously regulated firms can be adjusted. The first and third effects act to decrease ambient pollution levels while the second acts to increase them. The final effect is the response of the regulator: in response to effects one through three the regulator could increase, decrease or leave regulated emission levels unchanged. Since there are effects inducing the regulator to both increase and decrease regulated emissions, it is not clear whether previously identified emitters will be regulated more or less stringently.

<sup>17</sup> 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  $(1 - \delta)\beta$ . Over time, the regulator could account for this discrepancy in order to partially correct for biased estimates.

The key intuition behind Proposition 3 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 Fig. 2. This situation arises when a new source of emissions is discovered that cannot be changed for any plausible cost, as is the case when new emissions are part of the earth's natural processes. Such discoveries were made by [Keppler et al. \(2006\)](#) and [Etiopo and Ciccioli \(2009\)](#) when they found that both plants and the ocean floor are major sources of potent greenhouse gases.<sup>18</sup> Alternatively, if new emissions sources can be cheaply reduced, such a discovery would result in lower optimal total ambient pollution levels. 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 2 and 3 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.

Both of these Propositions rely on the assumption that the regulator under counts emissions. If they instead over count emissions then (e.g., account for too much fugitive dust) then the results reverse. These two Propositions are explicit about this with their key assumption: if  $\alpha$  is assumed to be greater than one the converse of these results would be true. As a result, each Proposition has an associated Corollary which we omit here in the interest of conserving space. If the regulator over counts fugitive emissions the opposite would be true and known emitters would be under-regulated relative to the full information case.

### 3.2 Accounting for Uncertain Emission Inventories

Here we consider the case when the regulator accounts for the possibility they may observe only a subset of emitters.<sup>19</sup> We primarily assume that all emitters are homogeneous in most, but not all, of this section to show the first order effects of accounting for structural uncertainty.

Assume the regulator knows that  $\alpha \leq 1$ , but only has prior beliefs over the true fraction of identified emitters at any point in time:  $\alpha \sim \mu(\alpha)$ . By acknowledging imperfect information, 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  $(1 - \alpha)$ .

There are two implications of allowing for priors  $\mu(\alpha)$  over the fraction of identified emissions. First, it impacts the stock of total expected emissions inventories. Second, it impacts estimates of marginal physical impacts of emissions on ambient pollution levels. Continuing with the OLS intuition from above, conditional on a particular percentage of known emitters,  $\alpha$ , the regulator can correct for the bias in marginal physical impacts by

<sup>18</sup> 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.

<sup>19</sup> 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.

estimating  $\hat{\beta}$  if they multiply by  $\alpha$ :  $\alpha\hat{\beta} = \alpha\frac{\beta}{\alpha} = \beta$ . Put another way, it is possible that conditional on knowing  $\alpha$ , it is possible that  $E[\hat{\beta}|\alpha] - \beta < \hat{\beta} - \beta$  if  $\hat{\beta}$  is the estimated effect when not accounting for bias; in this way, conditional on knowing the level of incomplete information it is possible to reduce or eliminate bias in an estimated transfer coefficient. In this way, priors  $\mu(\alpha)$  imply a distribution over physical impacts as a function of the model used to estimate marginal physical impacts by atmospheric chemists.

Maintaining homogeneity of all emitters for simplicity, take  $\tilde{x}$  to denote the actual emissions of each of the  $\alpha N$  identified emitters and  $x_{un}^*$  to represent unidentified emitters. When accounting for structural uncertainty, the regulator can account for the contributions of the unknown emitters in contributing to ambient levels,  $(1 - \alpha)Nx_{un}^*\beta$  because they condition on  $\alpha$ . However, 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 when accounting for structural uncertainty is:

$$\max_{\{\tilde{x}\}^{\alpha N}} \int (\alpha N)\pi_i(\tilde{x}) + (1 - \alpha)N\pi_i(x_{un}^*) - D(\alpha N\tilde{x}\beta + N(1 - \alpha)x_{un}^*\beta)\mu(\alpha)d\alpha. \tag{5}$$

Note that the regulator has only  $\alpha N$  control variables since only  $\alpha N$  emitters are known with certainty. The resultant first order conditions for the known bias case is therefore:

$$\begin{aligned} \pi_i'(\tilde{x}) &= \int D'(\alpha N\tilde{x}\beta + N(1 - \alpha)x_{un}^*\beta)\beta\mu(\alpha)d\alpha \quad \forall i = 1, \dots, \alpha N \\ &= E[D'(Y)\beta|\mu(\alpha)] \end{aligned} \tag{6}$$

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. In the first order condition, Eq. (6), we can see precisely where multiplicative uncertainty enters the regulators problem since the estimated transfer coefficient,  $\beta$ , is a function of priors,  $\mu(\alpha)$ . The expectations operator now is evaluated over  $\beta$ .

Now consider the 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 *correctly* identified in both cases, akin to Sect. 2.1 above:

**Proposition 4** *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 with Bayesian priors relative to when it is not accounted for.*

Proposition 4 stems from the regulator acknowledging their policy will have less influence on actual ambient pollution levels than a regulator with naive beliefs. Regulators in this case know they have biased estimates of marginal physical effects due to the possibility of unknown unapportioned sources of emissions. They balance the loss in profits of known emitters with the social welfare gains of lower expected ambient pollution.

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. Because  $\alpha$  can impact the estimate of marginal physical effects of emissions

on ambient levels of pollution non-linearly (e.g., the OLS example has that  $\hat{\beta} = \frac{\beta}{\alpha_0}$ ), deriving analytical results are more challenging. We therefore make the strong assumption that the slope of marginal damages with respect to ambient pollution levels is linearly increasing. Even with this restrictive assumption we are unable to determine whether currently known emitters should always be regulated more or less stringently:

**Proposition 5** *For linearly increasing marginal damages and upwardly biased estimates, it is not optimal to regulate known emitters more stringently when accounting for structural uncertainty if expected marginal damages when accounting for structural uncertainty are lower than when not accounting for structural uncertainty.*

There are three key pieces of intuition for Proposition 5. 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 lower 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. Even 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 5 only relies on estimates conditioning on  $\mu(\alpha)$  being less biased than when not.

The indeterminate result from Proposition 5 is unsatisfying but there are still policy lessons embedded in it. For example, if the regulator is fairly certain they have identified all emitters then it is more likely that accounting for structural uncertainty leads to increased stringency on known emitters. The reason is the non-linearity in the bias when naively estimating marginal emissions. In the 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.

Applying this model and the insights from Proposition 5 is relatively straightforward, but the correct functional relationships between pollution and damages, the general functional form on profits, and the non-linearities induced by biased estimation will all be context specific. This naturally lends itself to simulation based exercises to size the magnitude of the problem in different contexts. One example is bounding exercises where efficiency losses from not accounting for structural uncertainty are going to be largest versus smallest. This allows regulators to focus effort on use cases which have the highest expected value in terms of increased regulatory efficiency.

## 4 Discussion

While we have focused on upwardly biased transfer coefficient estimates over counting fugitive emissions could lead to downwardly biased estimates. In that situation, our results

would be reversed. It is unclear if fugitive emissions are over or under counted in general. In the case of methane, for example, fugitive emissions were under counted (Miller et al. 2013). In other contexts like particulate matter, though, the opposite could be true. This possibility highlights the complexity associated with the general problem of structural uncertainty in the relationship between emissions and ambient levels of pollution. Complicating matters further, the regulator might face punitive legal or political costs in over or under regulation of regulated firms if they acknowledge scientific uncertainty.

There are also welfare implications of accounting for structural uncertainty in the pollution process. Increasing the accuracy of information assumptions in any economic model should, in expectation, increase total welfare of the solution to the regulator's problem. Indeed, the change in welfare is equal to the social value of information. We note, though, that the social value of information is a function to the regulator's response to it; in this way the social value of the discovery of new emission sources is a function of whether the regulator explicitly accounts for structural uncertainty.

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 passively react to scientific discoveries that alter the understanding of the pollution process. Our work can 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. In particular, extensions of this model could include optimal experimentation to learn about structural uncertainty faced by the regulator.

## Appendix 1

**Proposition 1** *An increase in the set of known emitters may decrease the expected level of ambient pollution if (1) marginal physical effects are sufficiently biased upward and (2) the set of new emitters is sufficiently small.*

*Proof* Let a subset of emitters  $K = \alpha_0 N$  initially be used to estimate marginal effects of emissions on ambient pollution levels where  $\alpha_0 < 1$ . 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 Eq. (4).

Given the homogeneity assumption, the new class of unregulated emissions that were emitting at a level  $\tilde{x}_{un}^*$  where  $\tilde{x}_{un}^* > \tilde{x}_0^*$ . 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  $E[y|\alpha_1] = (\alpha_1 - \alpha_0)N \tilde{x}_{un}^* \hat{\beta}_1 + \alpha_0 N \tilde{x}_0^* \hat{\beta}_1$ .

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

$$\begin{aligned} \text{sign}[E[y|\alpha_1] - E[y^*|\alpha_0]] &= \text{sign}[(\alpha_1 - \alpha_0)N\tilde{x}_{un}^*\hat{\beta}_1 + \alpha_0N\tilde{x}_0^*\hat{\beta}_1 - \alpha_0N\tilde{x}_0^*\hat{\beta}_0] \\ &= \text{sign}[\alpha_0\tilde{x}_0^*(\hat{\beta}_1 - \hat{\beta}_0) + \tilde{x}_{un}^*\hat{\beta}_1(\alpha_1 - \alpha_0)]. \end{aligned} \tag{7}$$

There are two terms in Eq. (7). 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 second term dominates expected ambient pollution increases. However, if the first term dominates the second, expected ambient pollution decreases giving the desired result. □

**Proposition 2** *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* To determine the impact of new emission sources on known emitter’s optimal pollution levels, Proposition 1 shows we must address both increases and decreases in expected ambient pollution levels. 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 Eq. (4), is smaller than the marginal expected damages before the discovery:

$$E[D'(y|\alpha_1)]\hat{\beta}_1 < E[D'(y^*|\alpha_0)]\hat{\beta}_0. \tag{8}$$

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

$$\pi'_i(\tilde{x}_i^*|\alpha_0) = D'(E[y^*|\alpha_0])\hat{\beta}_0 > D'(E[y|\alpha_1])\hat{\beta}_1. \tag{9}$$

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

**Proposition 3** *An increase in the set of known emitters can decrease the optimal level of ambient pollution if benefits of emissions for newly discovered emissions is sufficiently large (e.g., decreasing emissions is sufficiently costly).*

*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:

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

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  $\hat{\beta} > 1$ .

Leveraging the linear marginal damage assumption to drop the expectations operator, the regulator will let the following single first order condition set emitter’s  $i$ ’s:

$$\pi'_i(x_{i0}^*) = D'(x_{i0}^*)\hat{\beta}. \tag{10}$$

Note that emitter  $j$  will emit at the level defined by  $\pi'_j(x_{j0}^*) = 0$  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

$$\begin{aligned} \pi'_i(x_{i0}^* + \Delta x_i) &= D'(x_{i0}^* + \Delta x_i + x_{j0}^* - \Delta x_j) \\ \pi'_j(x_{j0}^* - \Delta x_j) &= D'(x_{i0}^* + \Delta x_i + x_{j0}^* - \Delta x_j) \end{aligned} \tag{11}$$

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  $\pi_j(\cdot)$ .

To determine how discovering new emission source  $j$  impacts optimal ambient pollution, Proposition 1 and Proposition 2 show we must analyze three cases: no change in emissions after discovery of  $j$ , an increase in emissions and finally an 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 Eq. (11). 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 ( $x_{j0}^* - \Delta x_j$ ) entering the damages function. Hence, emissions could remain unchanged after the discover of an additional emitter.

Now augment the profit function for emitter  $j$  such that  $\tilde{\pi}'_j(x_{j0}^*) = 0$  at the same point as previous but that  $\tilde{\pi}''_j(\cdot) = -\tilde{k}_j$  where  $\tilde{k}_j > k_j$ . By assumption,  $\tilde{\pi}'_j(x_{j0}^* - \Delta x_j) > \pi'_j(x_{j0}^* - \Delta x_j)$ . 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. Note that to arrive at a decrease in emissions let  $\tilde{k}_j < k_j$  and the converse is true. This completes the proof.  $\square$

**Proposition 4** *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 with Bayesian priors relative to when it is not accounted for.*

*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, Eq. (6), can be rewritten as  $\pi'(\tilde{x}) = E[D'(\alpha N\tilde{x}\beta + (1 - \alpha)Nx_{un}^*\beta)\beta|\mu(\alpha)]$ . The term  $(1 - \alpha)Nx_{un}^*$  will be positive for non-degenerate beliefs. Now assume that  $\tilde{x} = x^*$ . Because the term  $(1 - \alpha)Nx_{un}^*$  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,  $\tilde{x}$ , should be reduced to less than emissions of emitters in the full information case,  $x^*$ , to satisfy the first order condition (6), thereby completing the proof.  $\square$



**Proposition 5** For linearly increasing marginal damages and upwardly biased estimates, it is not optimal to regulate known emitters more stringently when accounting for structural uncertainty with Bayesian priors relative to when it is not accounted for if expected marginal damages when accounting for structural uncertainty are lower than when not.

*Proof* Assume that the marginal damages from ambient pollution increase at a constant rate  $c$ . The first order condition when accounting for structural uncertainty, Eq. (6), can be rewritten as  $\pi'(\tilde{x}) = E[c(\alpha N \tilde{x} \hat{\beta} + (1 - \alpha) N x_{un}^* \hat{\beta}) \hat{\beta} | \mu(\alpha)]$ . Assuming that that bias can be corrected (e.g.,  $E[\hat{\beta} | \alpha] = \beta$ ) and simplifying terms, we can further reduce the expression to  $\pi'(\tilde{x}) = E[c(\alpha N \tilde{x} \beta + (1 - \alpha) N x_{un}^* \beta) \beta | \mu(\alpha)]$ . The term  $(1 - \alpha) N x_{un}^* \beta$  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  $\pi'(x^*) = E[c(\alpha N x^* \hat{\beta}) \hat{\beta}]$  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:

$$E[c(\alpha N x^* \hat{\beta}) \hat{\beta}] - E[c(\alpha N \tilde{x} \beta + (1 - \alpha) N x_{un}^* \beta) \beta | \mu(\alpha)] \quad (12)$$

Taking advantage of the linearity and  $\tilde{x} = x^*$  assumptions, Eq. (12) 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 incorrectly 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.  $\square$

## Appendix 2

### *Regulator Techniques to Account for Incomplete Emission Inventories*

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; 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 U.S. EPA and its counterparts in 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 inaccurate (U.S. EPA 2006; Wang et al. 2011).

This situation is likely the case for many pollutants: in the United States, the U.S. EPA implicitly assumes that their models correctly identify the relationship between historical 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 U.S. 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 U.S. 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 U.S. EPA suggest that there is a bottleneck at the U.S. 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.

We believe 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 a 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.

## References

- Chambers AK, Strosher M, Wootton T, Moncrieff J, McCready P (2008) Direct measurement of fugitive emissions of hydrocarbons from a refinery. *J Air Waste Manag Assoc* 58(8):1047–1056
- Crandall R (1981) Controlling industrial pollution. Brookings, Washington
- Dominguez G, Jackson T, Brothers L, Barnett B, Nguyen B, Thiemens MH (2008) Discovery and measurement of an isotopically distinct source of sulfate in Earth's atmosphere. *Proc Natl Acad Sci* 105(35):12769–12773
- EPA: (2006), Technical support document for the proposed pm naaqs rule, *Office of Air Quality Planning and Standards* February
- Etiopie G, Ciccioli P (2009) Earth's degassing: a missing ethane and propane source. *Science* 323(5913):478–482
- Fowlie M, Muller N (2013) Market-based emissions regulation when damages vary across sources: What are the gains from differentiation?, *NBER Working Paper 18801*
- Fu JS, Dong X, Gao Y, Wong DC, Lam YF (2012) Sensitivity and linearity analysis of ozone in east asia: the effects of domestic emission and intercontinental transport. *J Air Waste Manag Assoc* 62(9):1102–1114
- Gray W (2002) Economic costs and consequences of environmental regulation. Ashgate Publications, Farnham
- Gray W, Shimshack J (2011) The effectiveness of environmental monitoring and enforcement: a review of the empirical evidence. *Rev Environ Econ Policy* 5(1):3–24
- Haagen-Smit A (1952) Chemistry and physiology of los angeles smog. *Ind Eng Chem* 44(6):1342–1346
- Hamilton S, Requate T (2012) Emissions standards and ambient environmental quality standards with stochastic environmental services. *J Environ Econ Manag* 64:377–389
- Hansen LG (1998) A damage based tax mechanism for regulation of non-point emissions. *Environ Resour Econ* 12(1):99–112
- Hoel M, Karp L (2001) Taxes and quotas for a stock pollutant with multiplicative uncertainty. *J Public Econ* 82(4):91–114
- Horan RD, Shortle JS (2005) When two wrongs make a right: second-best point-nonpoint trading ratios. *Am J Agric Econ* 87(2):340–352
- Horan T, Shortle J, Abler D (1998) Ambient taxes when polluters have multiple choices. *J Environ Econ Manag* 36(2):186–199
- Keppeler F, Hamilton J, Brab M, Rockmann T (2006) Methane emissions from terrestrial plants under aerobic conditions. *Nature* 439:187–191
- Kim D, Ching J (1999) *Executive summary of science algorithms of the EPA Models-3 community multiscale air quality (CMAQ) modeling system*, EPA

- Lemoine D, McJeon H (2013) Trapped between two tails: trading off scientific uncertainties via climate targets. *Environ Res Lett* 8(3):34019–34028
- McKeen S, Chung S, Wilczak J, Grell G, Djalalova I, Peckham S, Gong W, Bouchet V, Moffet R, Tang Y, Carmichael G, Mathur R, Yu S (2007) Evaluation of several pm<sub>2.5</sub> forecast models using data collected during the icartt/neaqs 2004 field study. *J Geophys Res Atmos*. doi:10.1029/2006JD007608
- McKeen S, Grell G, Peckham S, Wilczak J, Djalalova I, Hsie E-Y, Frost G, Peischl J, Schwarz J, Spackman R, Holloway J, de Gouw J, Warneke C, Gong W, Bouchet V, Gaudreault S, Racine J, McHenry J, McQueen J, Lee P, Tang Y, Carmichael GR, Mathur R (2009) An evaluation of real-time air quality forecasts and their urban emissions over eastern texas during the summer of 2006 second texas air quality study field study. *J Geophys Res Atmos*. doi:10.1029/2008JD011697
- Miller S, Wofsy S, Michalak A, Kort E, Andrews A, Biraud S, Dlugokencky E, Eluszkiewicz J, Fischer M, Janssens-Maenhout G, Miller B, Miller J, Montzka S, Nehrkorn T, Sweeney C (2013) Anthropogenic emissions of methane in the united states. *Proc Natl Acad Sci* 110(50):20018–20022
- Montgomery D (1972) Markets in licenses and efficient pollution control programs. *J Econ Theory* 5(3):395–418
- Muller N (2011) Linking policy to statistical uncertainty in air pollution damages, *The B.E. Press Journal of Economic Analysis and Policy*, 11(1), Contributions, Article 32
- Muller N, Mendelsohn R (2009) Efficient pollution regulation: getting the prices right. *Am Econ Rev* 99(5):1714–1739
- Newell R, Pizer W (2003) Regulating stock externalities under uncertainty. *J Environ Econ Manag* 45(2):416–432
- Pacyna E, Pacyna J, Sundeth K, Munthe J, Kindbom K, Wilson S, Steenhuisen F, Maxson P (2010) Global emission of mercury to the atmosphere from anthropogenic sources in 2005 and projections to 2020. *Atmos Environ* 44:2487–2499
- Ramanathan V, Carmichael G (2008) Global and regional climate changes due to black carbon. *Nat Geosci* 1(4):221–227
- Segerson K (1988) Uncertainty and incentives for nonpoint pollution control. *J Environ Econ Manag* 15(1):87–98
- Spence A, Weitzman M (1978) Approaches to controlling air pollution. In: Friedlaender AF (ed) *Regulating strategies for pollution control*. MIT Press, Cambridge
- Thiemens M, Trogler W (1991) Nylon production: an unknown source of atmospheric nitrous oxide. *Science* 251(4996):932–934
- Wang S, Xing J, Jang C, Zhu Y, Fu JS, Hao J (2011) Impact assessment of ammonia emissions on inorganic aerosols in east china using response surface modeling technique. *Environ Sci Technol* 45(21):9293–9300
- Weitzman M (1974) Prices versus quantities. *Rev Econ Stud* 41:477–491
- Wu C-F, gang Wu T, Hashmonay RA, Chang S-Y, Wu Y-S, Chao C-P, Hsu C-P, Chase MJ, Kagann RH (2014) Measurement of fugitive volatile organic compound emissions from a petrochemical tank farm using open-path fourier transform infrared spectrometry. *Atmos Environ* 82:335–342
- Zhang P, Liu J, Dong J, Holovati J, Letcher B, McGann L (2012) A bayesian adjustment for multiplicative measurement errors for a calibration problem with application to a stem cell study. *Biometrics* 68(1):268–274
- Zhang Y, Bocquet M, Mallet V, Seigneur C, Baklanov A (2012a) Real-time air quality forecasting, part i: history, techniques, and current status. *Atmos Environ* 60:632–655
- Zhang Y, Bocquet M, Mallet V, Seigneur C, Baklanov A (2012b) Real-time air quality forecasting, part ii: state of the science, current research needs, and future prospects. *Atmos Environ* 60:656–676