R. David Simpson*

Do regulators overestimate the costs of regulation?

Abstract: It has occasionally been asserted that regulators typically overestimate the costs of the regulations they impose. A number of arguments have been proposed for why this might be the case. The most widely credited is that regulators fail sufficiently to appreciate the effects of innovation in reducing regulatory compliance costs. Most existing studies have found that regulators are more likely to over- than to underestimate costs. While it is difficult to develop summary statistics to aggregate the results of different studies of disparate industries, one such measure is the average of the ratio of *ex ante* estimates of compliance costs to *ex post* estimates of the same costs. This ratio is generally greater than one. In this paper I argue that neither the greater frequency of overestimates nor the fact that the average ratio of *ex ante* to *ex post* cost estimates is greater than one necessarily demonstrates that *ex ante* estimates are biased. There are several reasons to suppose that the distribution of compliance costs could be skewed, so that the median of the distribution would lie below the mean. It is not surprising, then, that most estimates would prove to be too high. Moreover, Jensen's inequality implies that the expected ratio of *ex ante* to *ex post* compliance costs would be greater than one. I propose a regression-based test of the bias of *ex ante* compliance cost estimates, and cannot reject the hypothesis that estimates are unbiased. Failure to reject a hypothesis with limited and noisy data should not, of course, be interpreted as a strong argument to accept the hypothesis. Rather, this paper argues for the generation of more and better information. Despite the existence of a number of papers reporting *ex ante* and *ex post* compliance cost estimates, it is surprisingly difficult to get a large sample with which to make such comparisons.

Keywords: compliance costs; *ex ante* estimation; *ex post* costs; innovation; skewed distributions.

DOI 10.1515/jbca-2014-0027

***Corresponding author: R. David Simpson,** National Center for Environmental Economics, United States Environmental Protection Agency, 1300 Constitution Avenue, NW, Washington, DC 20460, USA, e-mail: simpson.david@epa.gov

1 Introduction

Since the Reagan Administration regulatory agencies in the US have been required to perform cost-benefit analyses of high-profile regulations.[1] Many other nations have also instituted similar requirements for regulatory impact analyses (Radaelli, 2005). How accurate have such *ex ante* estimates of the costs and benefits of environmental, health, product, and other regulations proved to be?

With respect to the costs of regulatory compliance, available evidence seems to suggest that the answer is "not very." The cost estimates offered by regulators are generally higher than are *ex post* estimates of compliance costs.[2] A review of ten surveys, each of which reviews the results of a number of different case studies, finds that in each survey *ex ante* estimates of compliance costs exceed *ex post* estimates in a majority of instances. It is, admittedly, a risky venture to attempt to make comparisons across studies of different regulations from different regulators for different industries at different times, and sometimes in different places. If, however, one performs what may seem a natural test of the overall accuracy of regulatory cost estimates – averaging the ratio of *ex ante* to *ex post* cost estimates – she finds that the average ratio exceeds one, and often is considerably greater than one.

A number of hypotheses have been advanced to explain why *ex ante* cost estimates are often too high. Some emphasize that regulators do not have an incentive to conduct careful cost estimates: if it appears that a regulation will pass a cost-benefit test anyway, there is no real motivation to prepare a careful study, or, perhaps more importantly, to inflame opposition from the affected entities by venturing more controversial estimates. Other authors note that the "first draft" regulations for which compliance costs are predicted are often more stringent than those eventually passed and with which regulated entities must comply (or, to introduce another closely related hypothesis, with which they may not comply in practice). The explanation that has received the most attention and which seems to generate the most credence, however, is that regulators fail to account for innovation. As Lisa Heinzerling, former Associate Administrator in EPA's Office of Policy wrote (albeit in 2002, before coming to EPA), "Regulatory analysis is notorious for failing to take into adequate account the technological

1 In practice, this means regulations having an effect on the economy of $100 million or more per year, or designated as "significant" by the Office of Management and Budget (OMB).

2 I describe the studies in the following section. An exception in one respect is Hodges (1997), who focuses on the cost estimates offered by affected industries, rather than those prepared by government agencies. Not surprisingly, Hodges finds that such estimates are especially inflated.

https://doi.org/10.1515/jbca-2014-0027 Published online by Cambridge University Press

innovations that ultimately make many regulations cheaper to implement than regulators anticipate."

So, it seems that there are good reasons to suppose that regulators will overestimate the costs of compliance with environmental regulation, and compelling evidence that their cost estimates are biased.

Or are there?

In this paper I suggest that the evidence is not as clear-cut as it has seemed to some commentators. My main arguments are statistical. Neither of the procedures that have been employed to evaluate the accuracy of *ex ante* cost estimates in the existing studies provides a valid test to determine whether or not *ex ante* cost estimates are biased. The fact that most *ex ante* cost estimates exceed *ex post* estimates would only indicate a bias if it were reasonable to suppose that the distribution of *ex post* estimates were symmetric. I offer reasons to that they may not be.

The fact that the average ratio of *ex ante* to *ex post* estimates exceeds one is also not unexpected. Here the argument is very simple. Holding the numerator of a fraction fixed, a fraction is a convex function of its denominator. We would *always* expect the ratio of *ex ante* to *ex post* estimates to be greater than one if the *ex ante* estimates are unbiased– this is just an application of Jensen's inequality. If there is some probability that *ex post* costs are very low, then some of the ratios of *ex ante* to *ex post* estimates may explode.

It is worth emphasizing before suggesting a better procedure for evaluating the accuracy of *ex ante* cost estimates that any procedure involving comparisons between *ex ante*/*ex post* estimate pairs from different industries, regulations, time periods, countries, etc., must rest on heroic assumptions. For the sake of argument, however, let me proceed as follows. While both *ex ante* and *ex post* cost estimates are just that, "estimates," we have little prospect for drawing any conclusions regarding the accuracy of *ex ante* predictions if we do not suppose that the *ex post* estimates are at least somewhat informative. If we cannot measure the cost of a regulation even after the fact, any attempt to determine whether estimates made before the regulations were enacted were close to the mark would be futile. I will, then, maintain the hypothesis that the *ex post* cost estimate is an unbiased estimate of the "true" cost of the regulation. Treating this true cost as a random variable drawn from some probability distribution, we can, then, apply the principle of iterated expectations. The expectation formed now of the expectation that can be formed later, when more information will be available, is simply the unconditional expectation now.

So, if an *ex ante* estimate is unbiased we should be able to express it as the sum of the *ex post* realization plus an uncorrelated prediction error term. This hypothesis can be tested by regressing *ex ante* estimates on a constant term and

https://doi.org/10.1515/jbca-2014-0027 Published online by Cambridge University Press

the corresponding *ex post* estimates. The null hypothesis is that the intercept term of this regression will be zero and the slope one. I cannot reject this hypothesis in a sample of 18 *ex ante*/*ex post* compliance cost estimates.

What should we conclude from this exercise? First and most obviously, that existing studies do not establish that regulators generate biased estimates of costs. To be fair, it should be noted that this is not really the claim of most of the studies themselves so much as that of some second-hand summaries of their findings. It is, however, useful to be clear on this point. If policy makers were tempted to conclude that regulatory cost estimates are biased and should be revised downward so as to provide a more liberal benefit-cost test of proposed regulation, this would appear to be premature.

There is, however, a more important conclusion to be drawn from this exercise. The problem with existing *ex ante* cost estimates is not that they are *biased* so much as that they are *limited*. The reader may be surprised to learn that I could identify only 18 studies for which sufficient information was available to test the accuracy of *ex ante* cost estimates. So was I. I am agnostic on the question of whether regulators' *ex ante* cost estimates really are biased. With such a limited sample I certainly have not "proved" that they are. I am merely showing that we cannot reject the hypothesis that they are not.

While the conclusion that "more research is needed" is certainly hackneyed and, in many instances, self-serving, if ever it were justified, this would be an instance. This is not to say that previous authors have not been careful and diligent. They certainly have. But such fundamental questions as "what constitute costs?" have been answered in different ways by different authors. Greater methodological standardization would facilitate comparisons and conclusions of the type I have attempted to draw, and would provide better guidance for policy. Moreover, one finds on closer inspection that many existing studies do not record the kind of quantitative information that facilitates comparison. Authors of some studies have – often of necessity – confined themselves to qualitative assessments. One also finds that several surveys (such as this one) merely recombine existing studies rather than generating new data. More primary data collection and analysis would certainly be useful, as would greater consistency across studies in the definition of costs to be recorded.

This paper is presented in five sections, including this introduction. The next section reviews the literature on the accuracy of *ex ante* cost estimates. Following that, I offer arguments for why the measures reported in the existing literature – the frequency with which *ex ante* costs are overestimated and the ratio of *ex ante* to *ex post* cost estimates – do not necessarily shed light on the question of whether *ex ante* cost estimates are biased. In the fourth section I propose an alternative statistical test and report its results. A fifth section briefly presents conclusions.

https://doi.org/10.1515/jbca-2014-0027 Published online by Cambridge University Press

2 Previous studies

A number of researchers have studied the accuracy of *ex ante* estimates of the costs of environmental and other forms of regulation in the light of *ex post* estimates of such costs. In the interest of brevity I distinguish between studies of the disparity between *ex ante* and *ex post* estimates of costs and *surveys of studies of* such disparities, and focus on the latter. There are now quite a number of reports whose authors have taken as their data the results of earlier studies of particular regulations in particular industries and tried to evaluate the accuracy of such studies generally. As we will see below, one of the challenges of such undertakings is to define what it means for *ex ante* cost estimates to be "accurate." Existing studies generally report accuracy in terms either of the fraction of studies that overestimate costs, or in terms of the average ratio of *ex ante* to *ex post* cost estimates. Broadly speaking, existing studies find that overestimates are more common than underestimates, while the average ratio of *ex ante* to *ex post* estimates tends to be greater than one.

The first study of which I am aware devoted specifically to the consideration of the accuracy of *ex ante* projections of the costs of regulation was conducted for EPA by the consulting firm of Putnam, Hayes, and Bartlett and completed in 1980 (hereinafter, "PHB 1980"). The study compared EPA and industry *ex ante* estimates of required capital expenditures for five rules passed in the 1970's with actual capital expenditures. In four of five cases industry overestimated capital costs, while in three of five cases EPA overestimated capital costs for the period from 1974 to 1977. The PHB results are somewhat more ambiguous for a sixth case study, in which EPA and industry estimates of the effects of environmental regulations on new car prices were compared.

The next major study of the accuracy of cost projections was conducted in 1995 by the Office of Technology Assessment (OTA). OTA considered eight regulations in chemical, manufacturing, and service industries enacted between 1974 and 1989. In all cases in which numerical estimates were hazarded estimated costs exceeded actual costs. In two industries the OTA report suggests that costs may actually have been negative: in finding ways to reduce risks, producers may actually have identified processes that operate more efficiently. Such claims would substantiate Michael Porter's (1991) hypothesis, that firms that operate under tougher environmental regulation can actually be more competitive in world markets.

In 1997 Hart Hodges published a study of 12 environmental and workplace safety regulations initiated between the 1970s and 1990s (Hodges 1997; the results are also summarized in Goodstein & Hodges 1997). In each instance *ex ante* estimates of costs were greater than were costs recorded later; in 11 of 12 cases, *ex ante*

https://doi.org/10.1515/jbca-2014-0027 Published online by Cambridge University Press

cost estimates were more than double costs realized *ex post*. Hodges focuses on industry's rather than regulators' estimates of costs. Inasmuch as industry will, in general, have a powerful incentive to overstate costs, the discrepancies Hodges identifies are not surprising.

A very thorough comparison of *ex ante* to *ex post* estimates of costs was conducted in 2000 by Winston Harrington, Richard Morgenstern, and Peter Nelson. The researchers considered 28 regulations written by EPA, OSHA, and a handful of other regional and international regulators. A number of different industries were covered. *Ex ante* cost estimates were considered "accurate" if they were within ±25% of *ex post* values, and either too high or too low if they fell outside this range. By this standard total costs of regulation were overestimated in 14 instances, underestimated in only three, and deemed reasonably accurate in the remaining 11. Harrington et al. distinguish between *total* and *unit* costs of regulation (the numbers I have just reported are for "total" cost estimates). The latter refer to the costs per unit of output or the cost per plant. Total cost is per unit cost times output or number of plants affected. Harrington et al., find that unit costs tend to be overestimated as often as they are underestimated, in contrast to total cost estimates. I will discuss below some reasons for which this might be the case.

The next major retrospective study of the costs of regulation was completed in 2005 by the Office of Management and Budget (OMB, 2005). OMB reviewed 47 regulations initiated between 1976 and 1995. EPA issued 18 of the regulations in the OMB sample, the most of any of the five federal agencies included in the study.[3] As is generally the case with estimates of regulatory costs, the sample was determined by the availability of data, not by any attempt to generate a random cross-section of regulatory activity. The results of the OMB study are less striking than those of some other researchers. Of 40 regulations for which comparable *ex ante* and *ex post* data are available, 16 *ex ante* projections overestimated cost, 12 underestimated them, and 12 were approximately accurate. The OMB study was not completely independent of earlier work, however: for instance, nine of the studies in its sample were adopted from Harrington, Morgenstern, and Nelson (2000).

At least three studies have been conducted of the accuracy of *ex ante* cost measures in other countries (in addition, Harrington et al. 2000 includes three examples drawn from Singapore, Norway, and Canada among their 28 case studies). While such inquiries obviously consider costs generated under different legal and regulatory structures than prevail in the US, they may still be useful in interpreting general approaches to regulatory cost estimation. It might also

3 The others were the National Occupational Safety and Health Administration (13 regulations included), the National Highway Traffic Safety Administration (8), the Department of Energy (6) and the Nuclear Regulatory Commission (2).

https://doi.org/10.1515/jbca-2014-0027 Published online by Cambridge University Press

be noted in passing that international standards for the analysis of regulatory impacts have become more similar over time, with the UK (MacLeod, Moran, Dominic, Lago, Harrington, & Morgenstern, 2006) and the European Union adopting such requirements.[4] A study conducted by the Stockholm Environmental Institute considered the cost estimates presented by industry in regulatory negotiations, and found them to be consistently higher than *ex post* realizations of actual costs (Bailey, Haq & Goudson, 2002).

MacLeod et al. (2006) performed a similar analysis of *ex ante* costs in UK rulemaking. The authors of this study adopted the same ±25% standard as used in Harrington et al. (2000), and found that by this standard the costs of five of eight regulations considered were overestimated, those of two regulations were underestimated, and those of one were approximately on target.

In 2006 Oosterhuis published estimates of *ex ante* and *ex post* costs of regulation with five EU environmental regulations. They report that in four instances *ex ante* cost estimates exceeded *ex post* costs by a factor of two or more, while the *ex ante* and *ex post* estimates were roughly the same in the fifth case.[5] Oosterhuis also reports on an earlier study of costs of compliance with Dutch environmental regulations of the first Dutch National Environmental Policy Plan of 1988, as predicted *ex ante* by Jantzen (1989) and later estimated *ex post* by RIVM (2000). These Dutch studies were, by the standards of the field, unusually accurate. While the costs of five of the eight regulations considered were overestimated, only one *ex ante* estimate was as much as twice its *ex post* realization, and in aggregate the total *ex ante* estimate of slightly over €12 billion was only 13% higher than the *ex post* realization. Oosterhuis (2006) credits this unusually accurate performance to the existence of relatively good statistics and studies in the Netherlands.

I will conclude this section with summaries of two studies that considered the accuracy of *ex ante* cost predictions for specific consumer products. Anderson and Sherwood (2002) compare cost estimates for EPA mobile source rules. These include six fuel-quality regulations and eleven vehicle emission standards. In most instances Anderson and Sherwood found that *ex ante* estimates of price increases induced by regulation were greater than actual price changes observed. They also found, however, that EPA estimates tended to be closer to actual price changes than were industry estimates.

4 See Radaelli 2005, however, who notes that "regulatory impact assessments" may still differ significantly from one jurisdiction to another.

5 Oosterhuis actually considers six environmental directives, addressing large combustion plants, integrated pollution prevention and control, ozone control, ozone depleting substances packaging, and nitrates, but are unable to develop *ex ante* compliance cost estimation numbers for the packaging directive.

https://doi.org/10.1515/jbca-2014-0027 Published online by Cambridge University Press

Dale, Antinori, McNeil, McMahon, and Fujita (2009) considered the costs associated with the Department of Energy's efficiency regulations on consumer appliances such as air conditioners, refrigerators, and washing machines. This study illustrates the challenges inherent in developing estimates for the costs of regulation for consumer goods. Dale et al. derived their *ex post* cost estimates using hedonic regressions to tease out the separate effects of scale, general technological progress, and more competitive behavior from those of the energy efficiency regulations themselves. Having isolated these effects, the authors found, as have the other studies, that *ex ante* cost estimates generally exceed those developed *ex post*.

3 What does the literature show, and what does it mean?

The studies I have reviewed uniformly find that regulators overestimate the costs of regulatory compliance more often than they underestimate them, and that the ratio of *ex ante* to *ex post* compliance cost estimates is, on average, considerably greater than one. While this might seem at first blush to establish that regulators' *ex ante* estimates of the costs of compliance are biased upward, this assertion does not actually withstand closer scrutiny. I consider the two types of evidence in turn, and show that neither necessarily reveals a bias in estimates.

3.1 Overestimation is more common

One of the most robust findings in the existing literature comparing *ex ante* to *ex post* estimates of costs is that the former generally exceed the latter. I am aware of no study of in which more *ex ante* cost estimates were lower than *ex post* estimates, as opposed to higher, and in many a substantial majority of *ex ante* estimates were higher than the corresponding *ex post* estimates. Can we then conclude that *ex ante* cost estimates are generally biased upward?

The answer would appear to be "No," at least not on the basis of this simple observation alone. Consider a skewed distribution whose median lies below its mean. An unbiased estimate of the mean of the distribution would then, by construction, be greater than most observations likely to be drawn from the distribution. In our context, if *ex ante* cost estimates are unbiased estimates of the mean of *ex post* costs, realizations of the latter will be more likely to be lower than the *ex ante* estimate than they are to be higher.

https://doi.org/10.1515/jbca-2014-0027 Published online by Cambridge University Press

Do we have any reason to suppose that the distribution of costs would not be symmetric? There are several. First, *total* costs are often estimated by multiplying an estimate of *unit* costs by the *number* of units affected by the regulation (Harrington et al., 2000).[6] "Units" might, in this case, refer to units of production, plants, or units of emissions. The analyst is typically uncertain as to the change in unit costs. While it may seem that regulators would know to whom their proposed rules would apply, large errors are sometimes made in estimating the extent of a regulation's reach as well. When EPA estimated costs under its enhanced automobile inspection and maintenance program, for example, analysts assumed that 56 million cars would be covered by the program. Only four states actually implemented the program, however (Harrington et al., 2000; the authors also provide examples of inaccurately estimated quantities in other EPA and OSHA rules). It should also be recognized that costs are often estimated based not on current levels of production, but rather, on anticipated future levels. Prediction errors can, then, also affect the accuracy of cost estimates. EPA assumed a faster rate of growth in electricity demand than actually occurred, an error which Harrington et al. (2000) blame in part for the Agency's overestimation of the costs of its NO_x rules.

How do uncertainty about both the increase in unit costs and volume of production affected result in a skewed distribution of the estimate of total costs? Heuristically, if there is a small probability that unit costs will be large, and a small probability that the number of units affected will be large, there is a *very* small probability that the cost of regulation will be *very* large. Under such conditions the distribution of total costs will have a long right tail, and hence, be asymmetric (recall that "costs" are, by assumption, positive, and so the left tail should not extend beyond zero). A simple example illustrates this point. Suppose that both the increase in unit costs and the number of affected units are distributed independently and uniformly on the interval [0, 1] (we can always make the supports the same by choice of units of measurement). Then it is easily demonstrated that the product of these two symmetrically distributed random variables is distributed asymmetrically on the interval [0, 1] with probability distribution function $-\ln\theta$. This function has mean ¼ and median of approximately 0.187, and about three-fifths of observations are less than the mean.

6 The reader may wonder what the rationale is for adopting this multiplicative approach. It is a reasonable question; it need not generally be the case that costs would naturally be decomposed into these two factors. The fact is, however, that analysts often do estimate total costs by multiplying an estimate of the cost of compliance times the number of facilities required to comply, or the cost per unit of emissions reduced times the number of units of emissions reduced. Since this is the way it is often done, it is reasonable to ask what the implications are of doing it this way.

https://doi.org/10.1515/jbca-2014-0027 Published online by Cambridge University Press

Another reason to suppose that the distribution of costs is asymmetric is because the mathematical forms that give rise to such costs may be asymmetrically distributed. Consider another simple example. Suppose the production function is of the constant-returns-to-scale Cobb-Douglas form

$$f(x, e)=x^{\alpha}e^{1-\alpha}, \tag{1}$$

where x is the quantity of purchased inputs employed in production and e the amount of effluent discharged. Suppose the prices of output and inputs are not affected by the regulation, so the only cost of restricting e is the loss of profit. If p is the price of output and w the price of the input, x, it can be shown that

$$\pi=(1-\alpha)p^{\frac{1}{1-\alpha}}\left(\frac{\alpha}{w}\right)^{\frac{\alpha}{1-\alpha}}e, \tag{2}$$

and so when allowed pollution is reduced by de, firm profit will decline by $(1-\alpha)p^{\frac{1}{1-\alpha}}\left(\frac{\alpha}{w}\right)^{\frac{\alpha}{1-\alpha}}$ per unit of emissions reduction required.

This reduction of $d\pi$ will represent the cost of the regulation if prices may be treated as constant. If we treat α, p, and w as unknown random variables with independent symmetric distributions, the resulting function $d\pi$ is likely to be asymmetrically distributed. This is not surprising, as the central limit theorem applied to the product, rather than the sum, of independent random variables, implies that the product will be lognormally distributed. A histogram for one

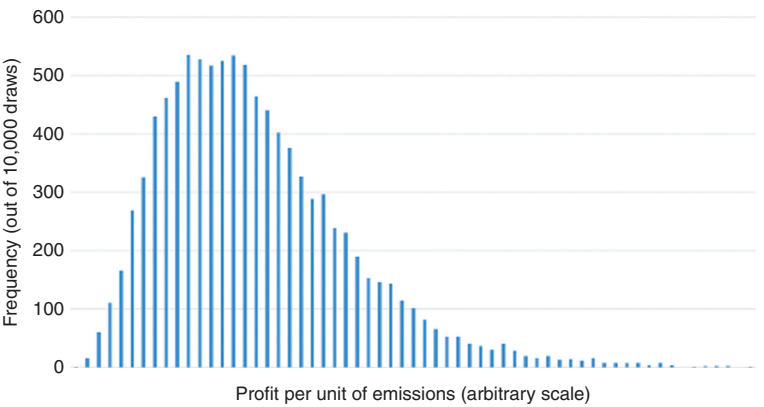

Figure 1 An example in which the distribution of profit per unit of emissions is skewed.

https://doi.org/10.1515/jbca-2014-0027 Published online by Cambridge University Press

such distribution is presented in Figure 1. To generate it, I assumed that the parameter α in equation (2) is distributed normally with mean 0.75, and standard deviation 0.05, and that both p and w are distributed normally with mean 1 and standard deviation 0.10. I then generated 10,000 realizations for α, p, and w and computed the cost per unit of emissions reduction required. The units in which profits are measured on the horizontal axis, as they would vary with the units in which prices are specified.

The expectation of the resultant distribution is 0.120, the median 0.108, and 57.3% of observations are less than the mean. In other words, if the regulator understood the uncertainties she were facing in forming her *ex ante* estimate of costs, and formed an unbiased estimate, she would still expect to overestimate costs in about three out of five instances.

The reader may object that I have done nothing more than showing that functional forms *may* exist under which the distribution of costs is asymmetrically distributed. This is true. Recall, however, that all I am trying to show is that the fact that most cost estimates prove to be too high does not necessarily imply that those estimates were biased. It suffices for my purposes simply to generate plausible examples. Similarly, the reader may object that simply supposing that regulations require a reduction in emissions by a fixed amount is not a very realistic depiction of real-world regulations that may impose restrictions on emissions per unit produced, or prescribe technologies. This is also true. But again, I am just giving simple examples in which a plausible depiction of a production function and a schematic treatment of enhanced regulatory stringency demonstrate my point.

A third reason for supposing that the distribution of costs might be skewed may arise from the attributes of innovation. The story that is often told as to why costs tend to be overestimated is that the people recording estimates tend to discount the possibility of innovation. They do not reflect the high likelihood that much more cost-effective ways of complying with regulation will be identified. It could well be, however, that while very cost-effective strategies are, in fact, identified *most* of the time, spectacularly costly exceptions could draw the mean cost of regulation considerably higher than the median.[7]

7 One might point to some possible instances of such events. While it is difficult to compare the circumstances characterizing the Renewable Fuel Standard to those of EPA regulations of emissions, some authors have suggested that provisions such as the cellulosic ethanol mandate were based on rosy assumptions concerning the evolution of technology (see NRC, 2011). The costs of meeting the standard were higher than expected, as the technology for producing cellulosic ethanol did not evolve as rapidly as was hoped.

https://doi.org/10.1515/jbca-2014-0027 Published online by Cambridge University Press

3.2 The average ratio of *ex ante* to *ex post* costs exceeds one

The fact that *most ex ante* estimates of costs are higher than are the corresponding *ex post* estimates does not necessarily imply that such estimates are biased. In order to make the determination of whether or not estimates are biased, we would need to know whether there are occasional spectacular exceptions in which *ex ante* costs severely *under*estimated actual costs.[8]

There are certainly instances in which the costs of rules have been underestimated. The problem, though, is that we never have repeated samples from the same distribution. There is, in each case, *one* rule whose costs are estimated, yielding one *ex ante* estimate and one *ex post* estimate for each case. It would require some truly heroic assumptions to say that the fact that costs were *under*estimated on a regulation affecting electric utilities, for example, somehow "offsets" the fact that costs were overestimated on rules affecting automobiles or appliances.

Let us suppose, however, that one were prepared to make such heroic assumptions. What might we infer from the observation that the average ratio of *ex ante* to *ex post* costs is typically greater than one?

Again, the frustrating answer may be "not much." To see why, consider a very simple example. Suppose that a rule is being contemplated, and that with 50% probability costs will be 50, while with 50% probability costs will be 150. In expectation, then, costs would be 100. This would be an unbiased *ex ante* estimate. On average, however, the *ratio* of predicted *ex ante* to realized *ex post* costs would be $\frac{1}{2} \cdot 100/50 + \frac{1}{2} \cdot 100/150 = 1\ 1/3$. Note that the issue here is *not* the symmetry of the distribution of realized costs. It is simply Jensens's Inequality. Fixing the numerator, a fraction is a convex function of its denominator.

4 Evaluating the accuracy of *ex ante* cost estimates: An alternative approach

Both *ex ante* and *ex post* estimates of costs are just that – estimates of random variables whose true values remain unknown. In the case of *ex ante* estimates,

8 The 2005 OMB study raises another interesting point that is worth considering in passing: we can never observe the accuracy of cost estimates for rules that were never issued. An anonymous referee has noted that when I test to see if I can reject the hypothesis that *ex ante* cost estimates are unbiased, I can only perform that test on a sample that may already have been selected to exclude *ex ante* cost estimates that were high enough to dissuade the regulator from enacting the rule.

https://doi.org/10.1515/jbca-2014-0027 Published online by Cambridge University Press

the rule under contemplation has not yet been enacted. In the case of *ex post* estimates, we cannot completely and accurately observe all affected entities' costs of compliance. Having said this, however, it seems reasonable to suppose that *ex post* estimates are giving us at least a somewhat informative signal of what the "real" costs of regulation are. Let us, then, refer to an *ex post* cost estimate as θ_1, and suppose that it is an unbiased estimate of the "real" cost, which I will designate as θ: $\theta_1=E(\theta)$. *Ex post* cost may not be exactly equal to "real" cost, as it may be measured with error.

Now suppose the *ex ante* estimate is θ_0. If the *ex ante* estimate is unbiased, then by iterated expectations, $\theta_0=E(\theta_1)=E(\theta)$. Note that for each regulation we consider we precisely observe one *ex ante* estimate – the estimate is what was reported in the study. We also observe one and only one *ex post* estimate. However, the *ex post* estimate we happen to observe could be drawn from anywhere along a spectrum of possible values. Whatever the "real" cost of the regulation might be is drawn from the support of the distribution of values that cost could take on. Moreover, *ex post* cost might also be observed with some error.

Everything I have said so far concerns the *one* estimate of *ex ante* costs and the *one* estimate of *ex post* costs we have for each regulation. We still have the problem of making inferences regarding the accuracy of *ex ante* cost estimates for many *different* regulations. Under the assumptions I have made so far, however, I can say that for any unbiased estimate of costs of regulation *i* proposed for industry *j*, I can say

$$\theta_0^{ij}=\theta_1^{ij}+\varepsilon^{ij}, \tag{3}$$

where ε is a measurement error with mean zero. While the magnitudes of the θ's might differ considerably across different regulatory cost estimates, equation (3) suggests that if I run the regression

$$\theta_0^{ij}=\alpha+\beta\theta_1^{ij}+\varepsilon^{ij}, \tag{4}$$

I should find that the estimated value of α should be equal to zero, while the estimated value of β should be equal to one.

I have estimated equation (4) using as data *ex ante* and *ex post* cost estimates reported in Harrington, et al. (2000); specific cost data were found in an earlier working paper, Harrington et al. 1999), MacLeod et al. (2006), and Oosterhuis (2006); this study includes both original case studies conducted by the authors and summaries of eight other case studies in which *ex ante* estimates were developed by Jantzen (1989) and *ex post* estimates reported by RIVM (2000).

As detailed descriptions of the data from these studies is included with each, I will not repeat such descriptions here. I might, however, note, in passing

https://doi.org/10.1515/jbca-2014-0027 Published online by Cambridge University Press

that I was unable to employ nearly as many data points as might be inferred from the numbers of cases considered in the studies. It is rather surprising when one consults the actual studies that clear, consistent, quantitative statements concerning both *ex ante* and *ex post* costs are more the exception than the rule. Harrington et al., for example, cite 28 cases. I use only seven. The others were eliminated for want of quantitative data (either in Harrington et al., 2000, or the working paper on which it was based, Harrington et al., 1999)[9], or because the authors reported only unit-cost estimates which may not be comparable with aggregate estimates (this, incidentally, is why I have not included any cost estimates from Anderson & Sherwood, 2002, or Dale et al., 2009). Similarly, it was possible to derive comparable numbers for *ex ante* and *ex post* costs for only three instances in the MacLeod et al. (2006) report, and Oosterhuis (2006) proved useful only inasmuch as we adopted figures that it reported from Jantzen (1989) and RIVM (2000). I did not consider studies such as OMB (2005), which compiles estimates from other sources (relying heavily, for example, on Harrington et al., 2000), or Hodges 1999, which reports *industry*, rather than regulators', estimates of costs.

I decided on a sample of 18 regulations (see Table 1). Six are from the US, one from Canada, eight from the Netherlands, and three from the UK. Regrettably – and surprisingly – only one US EPA regulation has clear quantitative estimates of both *ex ante* and *ex post* costs corresponding to total (as opposed to unit) effects. Performing the regression indicated in (4) yields the following results:

$$\begin{matrix}\theta_0 = 0.197 + 0.940\theta_1 \\ (0.124)\quad (0.083)\end{matrix} \tag{5}$$

R^2=0.889 (standard errors in parentheses).

I cannot reject the hypothesis that the intercept is zero and slope one, i.e., that *ex ante* estimates of the costs of regulation are unbiased.

9 For example, the authors write of the phase-out of lead from gasoline that "There has not been a retrospective analysis of the rule's costs but evidence indicates that EPA's analysts correctly forecast the costs or even overestimated them." While this judgment allowed Harrington et al. to classify this rule among those for which *ex ante* costs were estimated with reasonable accuracy, it does not allow me to employ the observation in my quantitative procedures. Harrington et al. (2000), like some other studies, does not provide explicit quantitative evidence even for those studies it determines meet its less than/within/more than 25% of *ex post* cost standard. They simply include a table with a "↑," "↓," or "↔" symbol to show if costs were over-, under-, or approximately accurately estimated, respectively.

https://doi.org/10.1515/jbca-2014-0027 Published online by Cambridge University Press

Table 1 Studies used for regression analysis.

Jurisdiction	Rule	*Ex ante* cost estimate (millions of US dollars)	*Ex post* cost estimate (millions of US dollars)	Source
Ontario	Ontario water	58	51	Harrington et al. (2000)
US (OSHA)	Vinyl Chloride	1000	253	Harrington et al. (2000)
US (OSHA)	Cotton Dust	280	83	Harrington et al. (2000)
US (OSHA)	Occupational Lead	224	20	Harrington et al. (2000)
US (OSHA)	Formaldehyde	11	6	Harrington et al. (2000)
US (EPA)	SO2 Phase I	764	779	Harrington et al. (2000)
US (OSHA)	Ethylene oxide	24	25	Harrington et al. (2000)
UK	Control of Major Accidents Hazards	155	416	MacLeod et al. (2006)
UK	Food Safety (General Food Hygiene/ Butchers' Shops)	5	25	MacLeod et al. (2006)
England	The Welfare of Farmed Animals	3	3	MacLeod et al. (2006)
Netherlands	Acidification	2620	1248	Jantzen 1989
Netherlands	Climate change	617	839	Jantzen 1989
Netherlands	Eutrophication	1471	814	Jantzen 1989
Netherlands	Hazardous Substances	3465	2738	Jantzen 1989
Netherlands	Waste Management	4848	5443	Jantzen 1989
Netherlands	Soil sanitation	914	881	Jantzen 1989
Netherlands	Disturbance	923	763	Jantzen 1989
Netherlands	Other	1939	2140	Jantzen 1989

I suggested above that *ex post* costs are likely measured with some error. If this is the case, the "iron law of econometrics" (the phrase comes from Hausman, 2001) holds that the coefficient of the mismeasured variable will be biased toward zero. I might investigate this possibility by just reversing the positions of *ex ante*, θ_0, and *ex post*, θ_1, costs in the regression equation, (4). Doing so, I find

$$\underset{(0.132)\quad(0.084)}{\theta_1 = -0.092 + 0.948\theta_0} \tag{6}$$

Again, standards errors are in parentheses. The R^2 is, of course, exactly the same in (6) as it is in (5), since we are fitting the same data. What *is* different, though, is that if (6) comes from inverting (5), we should expect the coefficient estimated for θ_1 in (6) to be the inverse of that estimated for θ_0 in (5). In fact, (6)

https://doi.org/10.1515/jbca-2014-0027 Published online by Cambridge University Press

implies a slope coefficient of 1/0.948=1.06, which is higher than the estimate of 0.940 in (5). This is consistent with the idea that *ex post* costs, θ_1, are a noisy measure of "real" costs.[10] This observation just underscores my main point: without more and better data, we cannot hope to come to any very firm conclusions regarding the accuracy of regulatory cost estimates.

It would be foolish to try to make too much of these results. Among other potential problems, it is more reasonable to regard the eighteen before-and-after estimates of costs I have used as a convenience sample than as any sort of random draw from the entire universe of cost estimates. In fact, one might suggest that the fact that I do have good before-and-after estimates for these eighteen rules is evidence that they were more carefully analyzed than were the many other rules that have been mentioned in studies comparing *ex ante* to *ex post* estimates of costs.[11] Moreover, the observation that I cannot reject the hypothesis that *ex ante* cost estimates are biased does not imply that such estimates are "good". There is still considerable variation in the sample, as evidence by the fact that the ratio of *ex ante* to *ex post* estimates ranges from 0.207 to 11.2. If nothing else, it would appear that cost estimates can be a long way off, in either direction. Again, the reader should remember the statistical proposition that we are testing is whether we can reject a null hypothesis. I would certainly not claim to have established that *ex ante* cost estimates are unbiased. All I am saying is that we need more and better data before we conclude that they are not.

5 Conclusions

The observation that concluded the previous section provides a good segue to a few brief concluding remarks. While much of my analysis is intended to suggest that it would be premature to draw certain conclusions from the evidence we have concerning the accuracy of *ex ante* cost estimates, my real point is that we are sorely in need of better evidence with which to make such judgments.

10 It might reasonably be suggested that the estimates I have reported will be inefficient, as we might reasonably expect considerable heteroskedasticity: the variance of random errors vary with the cost of the rules. I also transformed the regressions I have reported by weighting by the *ex ante* estimates of each, and found again that I could not reject the hypothesis that *ex ante* estimates were unbiased.

11 Some authors have noted that estimates reported in such studies might not have been chosen at random (see, e.g., Hahn & Tetlock, 2008). High-profile regulations, rules for which *ex ante* predictions were spectacularly inaccurate, or instances illustrating economists' favorite hobby horses (e.g., those allowing allowance trading) might all be more likely to be considered.

https://doi.org/10.1515/jbca-2014-0027 Published online by Cambridge University Press

While the authors of existing studies have labored diligently to gather evidence, the evidence remains limited. Moreover, different studies have assembled different data in different ways. While I have tried to compare studies that report similar measures of costs, discrepancies remain between studies as to, e.g., how to include capital investments and variable costs, time periods, discounting, etc. My results can only be considered suggestive at best.

Moreover, as other authors have suggested, conducting retrospective studies of the accuracy of *ex ante* cost estimates remains something of an orphan activity (see, e.g., Hahn & Tetlock, 2008). It is understandable that regulators would put a higher priority on predicting the effects of prospective regulations than they would on evaluating the accuracy of their predictions of regulations that have already been promulgated. It is also understandable that those who have ventured predictions in the past would be reluctant to revisit them: the best possible outcome for them would be that they would be shown to have done their job competently, while the alternative is that their best efforts would be found lacking. Be that as it may, however, it would certainly be useful to high-level decision makers to know how reliable the information they are receiving is – or at least, how reliable it has been in the past. Ultimately, this information might show *why* different studies have over- or underestimated costs, and whether the prospect for technological innovation is, in fact, underappreciated.

Acknowledgments: I thank Jennifer Bowen, Cynthia Morgan, Carl Pasurka, Ronald Shadbegian, Nathalie Simon, William Wheeler and Ann Wolverton for comments on related work, Winston Harrington for a helpful discussion, and Robert Hahn for comments on an earlier draft. Both the style and the substance of the paper have benefited from a number of helpful suggestions from two anonymous reviewers. I remain responsible for any errors.

References

Anderson, J. F., & Sherwood, T. (2002). Comparison of EPA and Other Estimates of Mobile Source Rule Costs to Actual Price Changes. Paper presented at the SAE Government Industry Meeting, Washington, DC, May 14, 2002.

Bailey, Peter D., Haq, Gary, & Goudson, Andy. (2002). Mind the gap! Comparing *ex ante* and *ex post* assessments of the costs of complying with environmental regulation. *European Environment*, *12*(5), 245–256.

Dale, Larry, Antinori, Camille, McNeil, Michael, McMahon, James E., & Fujita, K. Sydny. (2009). Retrospective evaluation of appliance price trends. *Energy Policy, 37*(2), 597–605.

Goodstein, Eban, & Hodges, Hart. (1997). Polluted data: Overestimating environmental costs. *American Prospect, 8*(35), 64–69.

https://doi.org/10.1515/jbca-2014-0027 Published online by Cambridge University Press

Hahn, Robert, & Tetlock, Paul. (2008). Has economic analysis improved regulatory decisions? *Journal of Economic Perspectives, 22*(1), 67–84.

Harrington, Winston, Morgenstern, Richard D., & Nelson, Peter. (2000). On the accuracy of regulatory cost estimates. *Journal of Policy Analysis and Management, 19*(2), 297–322.

Hausman, Jerry A. (2001). Mismeasured variables in econometric analysis: problems from the right and problems from the left. *Journal of Economic Perspectives, 15*(4), 57–67.

Heinzerling, Lisa. (2002). Markets for arsenic. *Georgetown Law Review, 90*, 2311–2339.

Hodges, Hart. (1997). Falling prices: Cost of complying with environmental regulations almost always less than advertised. Economic Policy Institute Briefing Paper. Washington, DC: Economic Policy Institute.

Jantzen, J. (1989), *Costs of Environmental Management, 1988 – 2010, 3 policy scenarios.* Report for the Ministry of VROM, The Hague, 25 May 1989.

MacLeod, Michael, Moran, Dominic, Aresti, Manuel Lago, Harrington, Winston, & Morgenstern, Richard. (2006). Comparing the *ex ante* and *ex post* costs of complying with regulatory change. Final report to DEFRA. London: Department for Environment, Farms, and Rural Affairs.

National Research Council of the United States National Academies (NRC). (2011). *Renewable Fuel Standards: Potential Economic and Environmental Effects of US Biofuel Policy*. Washington, DC: National Academies Press.

Office of Management and Budget, Office of Information and Regulatory Affairs (OMB). (2005). Validating Regulatory Analysis: 2005 Report to Congress on the Costs and Benefits of Federal Regulations and Unfunded Mandates on State, Local, and Tribal Entities. Washington DC: Government Printing Office.

Office of Technology Assessment (OTA). (1995). Gauging Control Technology and Regulatory Impacts in Occupational Safety and Health: An Appraisal of OSHA's Analytic Approach. OTA-ENV-635. Washington DC: Government Printing Office.

Oosterhuis, Frans (editor). (2006). Ex-post estimates of costs to business of EU environmental legislation. Report to European Commission DG Environment. Amsterdam: Institute for Environmental Studies, Vrije Universiteit.

Porter, Michael. (1991). America's green strategy. *Scientific American, 264*(4), 96.

Putnam, Hayes, & Bartlett (PHB). (1980). "Comparisons of Estimated and Actual Pollution Control Capital Expenditures for Selected Industries," Report Prepared for the Office of Planning and Evaluation, U.S. Environmental Protection Agency, Cambridge, Mass., (mimeo).

Radaelli, Claudio. (2005). *What Does Regulatory Impact Assessment Mean in Europe?* Washington: AEI-Brookings Joint Center for Regulatory Studies.

RIVM (2000), *Techno 2000; Modellering van de daling van eenheidskosten van technologieën in de tijd.* Rapportnummer 773008003, Bilthoven, April 2000.

Article note: All opinions expressed here are the author's and do not necessarily represent those of the United States Environmental Protection Agency.

https://doi.org/10.1515/jbca-2014-0027 Published online by Cambridge University Press