

The RFF Regulatory Performance Initiative: What Have We Learned?

Richard D. Morgenstern

*This paper is one in a series of retrospective analyses from RFF's
Regulatory Performance Initiative. www.rff.org/regulatoryperformance.*

1616 P St. NW
Washington, DC 20036
202-328-5000 www.rff.org

The RFF Regulatory Performance Initiative: What Have We Learned?

Richard D. Morgenstern

Abstract

Retrospective analysis of federal environmental regulation aims to rigorously document the extent to which key policy objectives have been attained and at what cost. This paper reports the results of nine new case studies involving a total of 34 cost or benefit/effectiveness comparisons from a highly diverse set of environmentally oriented rules. Despite limitations of the case study approach and the non-representative nature of the sample, the results suggest somewhat of a tendency to overestimate both costs and benefits/effectiveness of regulation. Recommendations for revisions to future policy and analysis include: estimation of cost savings and distributional impacts from emissions trading, approaches to new technology development, use of up-to-date baseline information, analysis of prescriptive congressional mandates, federally mandated programs implemented by states, heterogeneity of costs and benefits, and treatment of uncertainty in energy prices and other macro factors. Recommendations also consider enhancements to the federal role in future retrospective analyses.

Key Words: retrospective analysis, regulatory review, benefit-cost analysis

JEL Classification Numbers: D61, H1, L51

© 2015 Resources for the Future. All rights reserved. No portion of this paper may be reproduced without permission of the authors.

Discussion papers are research materials circulated by their authors for purposes of information and discussion. They have not necessarily undergone formal peer review.

Contents

1. Introduction.....	1
2. Background	3
3. Review of Individual Papers	6
3.1. Cost Comparisons	9
3.2. Effectiveness Comparisons	11
4. Recommendations from the Case Studies.....	15
4.1. Direct Regulatory Improvements.....	15
4.2. Improving Future Retrospective Analyses.....	20
5. Conclusions.....	23
References	26
Appendix: Summaries of Case Studies	30

The RFF Regulatory Performance Initiative: What Have We Learned?

Richard D. Morgenstern*

1. Introduction

Federal regulation of the environment is designed to improve societal well-being, including protection of the air we breathe, the water we drink, the food we eat, and the quality of our natural environment, as well as the conservation of resources. Focusing on regulations that have been in operation for a sufficient period to allow for careful study of performance, retrospective analysis aims to rigorously document the extent to which these objectives have been attained and at what cost. Such evidence-based analysis can indicate which regulations and regulatory approaches work well and which do not. Most important, this process can support broad-scale learning about regulation and its impacts which, in turn can lead to improvements in new rules and the modifying of poorly performing ones.

By relying on outcome or ex post data, retrospective analysis differs from the now routine prospective or ex ante regulatory impact analysis (RIA) of the benefits or effectiveness and costs of major regulations, which, as one expert has noted, is developed at “the point when the least is known and any analysis must rest on many unverifiable and potentially controversial assumptions” (Greenstone 2009, 113). Retrospective analysis also differs from the qualitative look-backs periodically conducted by US executive branch agencies, which typically focus on paperwork reduction and what might be labeled “squeaky wheel” concerns about specific rules.¹

Unsurprisingly, and despite its obvious appeal, rigorous retrospective analysis of federal regulation is not commonly conducted either inside or outside of government. Such analysis can be both complicated and expensive to develop, and the incentives to do so are often limited.

* Richard D. Morgenstern, senior fellow, Resources for the Future, morgenstern@rff.org. The author acknowledges the contributions of all 21 authors or co-authors of the case studies summarized herein. Special thanks are due Arthur Fraas, Terry Davies, Randy Lutter and Winston Harrington for helpful comments on an earlier draft of this paper. Thanks are also due to Wayne Gray, Sebastian Houde, Albert McGartland, and William Pizer for cross-cutting comments on the case studies at an authors workshop held at RFF, December 5, 2014, and to Cary Coglianese, Susan Dudley, and Richard Schmalanese for comments at a Forum held at RFF on October 21, 2015. Generous financial support was provided to RFF by the Smith Richardson and Alfred P. Sloan Foundations.

¹Aldy (2014, 90) has identified 11 administration-wide regulatory look-backs over the period 1978–2012.

Beyond the difficulty of measuring outcomes in meaningful ways, there is the ever-present challenge of estimating what would have happened in the absence of the regulation. Would the status quo have continued, would conditions have worsened, or is it possible that the severity of the problem would have lessened due to technological advances or, perhaps, success of other regulations or voluntary programs independent of the specific rule in question? It is no exaggeration to say that developing a credible counterfactual or baseline analysis is one of the most demanding aspects of a retrospective study.

This paper reports on the results of nine new retrospective studies, covering a diverse set of regulations, carried out under the aegis of Resources for the Future's (RFF's) Regulatory Performance Initiative (RPI) over the period 2013–15.² The studies were conducted by independent researchers at RFF and other US institutions. Since a number of the individual studies cover multiple regulations or multiple requirements within individual regulations, mostly expressed in physical terms such as emissions or discharges, a total of 34 ex ante–ex post cost or benefit/effectiveness comparisons can be derived from the nine studies. The US Environmental Protection Agency (EPA) promulgated the preponderance of the regulations, although one or more rules issued by the US Department of Energy (DOE), Department of the Interior (DOI), and Federal Drug Administration (FDA) are also included.

The study authors were asked to conduct a thorough assessment of the rules' performance, not simply to update the RIA. While the ideal study would examine the monetary benefits on an ex post basis, that is not generally feasible. In fact, only two of the nine studies reported here actually estimated monetary benefits on a retrospective basis.³ Instead, the authors focused on observable changes in emissions, discharges, risks, energy usage, or other relevant metrics of effectiveness. They were also asked to distinguish these reductions from credible baselines, compare them with those originally forecasted in the RIA or other pre-regulatory assessment conducted by the issuing agency, and assess whether any observed discrepancies were due to conceptual flaws in the RIA or to factors that could not be reasonably predicted. On the cost side, the authors were asked a parallel set of questions: What are the true incremental

² RFF's Regulatory Performance Initiative aims to strengthen the measurement of the outcomes of federal regulations in the areas of safety, health, and environmental protection. The focus is on a series of studies that examine the actual realized costs and benefits of individual federal regulations, comparing observed outcomes with a range of credible baselines and ex ante estimates. See <http://www.rff.org/research/collection/regulatory-performance-initiative-rff>.

³ Those on EPA's Phase II acid rain rule and FDA's egg rule.

costs of the rule when measured on an ex post basis? How do they compare with the government's ex ante cost estimates?

In addition to the quantitative assessment of the regulation's performance, a further set of questions posed to the study authors involved the lessons learned from conducting the retrospective analysis. Here we look at two types of issues. First, when viewed through the rearview mirror, how could the regulation have been designed to be more efficient and/or achieve greater health/environmental/energy effectiveness? Based on information readily available to the agency at the time, could the RIA conducted prior to rule promulgation have been improved to yield a better understanding of the effectiveness and/or costs of the rule or of alternative options or regulatory refinements? Second, recognizing both the methodological difficulties of conducting retrospective studies and the often severe data limitations, are there steps the federal government can take to improve the quality of future retrospective analyses?

The plan for this paper is straightforward. After a brief background discussion in the next section, Section 3 introduces the nine individual studies, including a description of the methods and data sources employed. This section also presents the quantitative findings of the individual studies in terms of the relative accuracy—that is, the over- or underestimation of the effectiveness or benefits and the costs of the regulations vis-à-vis the ex ante estimates. Section 4 summarizes lessons learned from the case studies about regulation and regulatory analysis. It also presents the recommendations made by the study authors in terms of policy enhancements and improved approaches for conducting RIAs, as well as different steps the federal government could take to facilitate future retrospective analyses. Section 5 presents concluding observations.

2. Background

Despite a growing interest in rigorous retrospective analysis, quality studies are generally few and far between. When retrospective assessments are undertaken, they often focus on input metrics such as inspections or audits rather than on outcomes such as effectiveness or net social benefits. In cases where retrospective analyses do adopt comprehensive welfare-oriented metrics, they often face problems of selection bias and disproportionately focus on costs rather than economic benefits or on standard metrics of environmental improvement such as changes in pollutant emissions or discharges. Broader assessments, such as the widely cited *Benefits and Costs of the Clean Air Act, 1990–2020* (EPA 2011), do not evaluate specific regulations and are often based on questionable baseline assumptions. These and other challenges have led some to question whether federal regulatory agencies can effectively play “moneyball” (Bridgeland and Orszag 2014) or whether outside actors need to perform this function, a topic we return to below.

Data availability, including credible baseline information on health and the environment, is a key issue in the conduct of rigorous retrospective analysis, particularly the type of micro-level information needed to draw credible conclusions about the performance of individual regulations. On the cost side, confidentiality and related concerns impede the conduct of ex post evaluations of many regulations. Economic incentive mechanisms can be an exception, as the transparency of market prices often provides a ready basis for analysis.

On the benefit side, some successes have been achieved in developing broad-scale, publicly available databases that allow assessment of effectiveness, such as improved food safety, reduced air emissions, or water discharges. Linking these gains to specific regulatory actions is rare. Even more infrequent, however, are retrospective evaluations that measure the ultimate objective—better health, enhanced recreational opportunities, and the like. Among federal regulatory agencies, the National Highway Traffic Safety Administration (NHTSA) has ready access to very rich data on auto safety, which it has used to conduct retrospective analysis of safety rules since the 1970s. For other agencies, the data limitations and, arguably, the institutional barriers are more severe. Although a number of academic studies have been conducted, they too face data limitations, and funding is often difficult to obtain.⁴ Several survey papers have also attempted to draw broader generalizations from the limited literature on retrospective analysis of environmental, health, and safety regulation.⁵

⁴ Examples of independent analysis of other agencies include Seong and Mendeloff's (2004) analysis of the mortality risk of specific workplace safety regulations issued by the Occupational Safety and Health Administration (OSHA), such as those related to construction scaffolding or electrical work practices. Thompson et al. (2002) focus on accidents in their analysis of the effectiveness of airbag regulation. In one of the first efforts to attribute air quality improvements to specific regulations—relying on available emissions and receptor data rather than the more standard physico-chemical air transport models—Harrington et al. (2012) find that Phase II of Title IV of the Clean Air Act Amendments reduced PM_{2.5} concentration in the eastern United States by an average of 8% compared with an explicit counterfactual. Here the authors focus on an outcome that is relatively easily measured (ambient pollutant concentration) rather than changes in health status, which are much more elusive. For additional examples, see Anderson and Sherwood (2002), Grosse et al. (2005), and Dale et al. (2009).

⁵ For example, Harrington et al. (2000) assembled the available retrospective cost estimates for safety, health, and environmental regulations and compared them with prospective estimates from the agencies' RIAs. This study identified only two dozen studies in both the published and the gray literature and focused on ex post evaluation of rules issued by OSHA and EPA. It found that total costs, calculated by RIAs and other government studies, tended to overstate observed costs, although the unit costs were generally accurate. One of the key reasons for the overstatement was that fewer safety or other benefits were achieved under the rules than originally anticipated. OMB examined this research in its 2005 and 2011 reports to Congress and in 2005 developed its own comparison of 47 prospective and retrospective case studies. The increased sample was largely driven by the addition of rules from other agencies: NHTSA, the DOE, and the Nuclear Regulatory Commission (NRC). OMB found that agencies had a tendency to overestimate benefits (generally defined as changes in emissions or other physical outcomes) more than

President Obama has emphasized the need for more ex post analysis, asserting in an executive order that federal agencies “must measure, and seek to improve, the actual results of regulatory requirements.”⁶ Cass Sunstein, then administrator of the Office of Management and Budget’s Office of Information and Regulatory Affairs (OMB’s OIRA), has emphasized the “culture of retrospective review and analysis,” noting that

to promote that culture, future regulations should be designed and written in ways that facilitate evaluation of their consequences and thus promote retrospective analyses. To the extent consistent with law, agencies should give careful consideration to how best to promote empirical testing of the effect of rules both in advance and retrospectively (2011b, 2).

A relatively bright spot in the field is the retrospective cost analyses recently carried out by EPA on four major regulations, although other agencies reportedly also have new retrospective studies under way.⁷

As noted, there are often many challenges in establishing a credible counterfactual for retrospective analysis. The existing literature, including the nine studies reported here, employ a range of methods, some based on well-defined treatment and control groups, using difference-in-differences regressions or other modeling techniques to quantify regulatory impacts, as well as simpler before-and-after analyses that try to capture the impacts.⁸

costs. Harrington (2006) assembled a still larger sample of 61 regulations and found that cost overestimates tended to predominate, although by a small margin. In reviewing the evidence, Simpson (2011) cannot reject the null hypothesis that the overestimates of costs are evidence of systematic bias. Lutter (2012) notes that the vast majority of all extant retrospective analyses of federal regulations provide insufficient information on benefits and costs to permit inferences about net benefits.

⁶ See EO 13610, May 10, 2012, Identifying and Reducing Regulatory Burdens, <http://www.whitehouse.gov/the-press-office/2012/05/10/executive-order-identifying-and-reducing-regulatory-burdens>, and also EO 13563, 76 FR 3821, January 18, 2011, Improving Regulation and Regulatory Review, http://www.reginfo.gov/public/jsp/Utilities/EO_13563.pdf.

⁷ The four rules studied by EPA are the Cluster Rule and those on arsenic, methyl bromide, and locomotives. Further information on the EPA retrospective studies is available at <http://www.epa.gov/regdarr/retrospective/documents/eparetroreviewprogressrpt-july2015.pdf>. Note that one of the new retrospective studies sponsored by the RPA is partially funded by the US DOE, to the Lawrence Berkeley Laboratory.

⁸ A before-and-after comparison is the basis of EPA’s 1997 *Report to Congress on the Benefits and Costs of the Clean Air Act*. In that analysis, EPA compares the estimated benefits of the act—in both physical and monetary terms—with the conditions existing in 1970, the year preceding enactment of the first federal air quality legislation. Despite initial criticism that such a before and after comparison ignored ongoing local efforts to improve urban air quality, EPA and its Advisory Council on Clean Air Act Compliance Analysis argued that any effort to develop an explicit counterfactual would be arbitrary, due in large part to the highly aggregate nature of the analysis.

A major barrier to the conduct of these analyses is the difficulty of identifying a control group that is totally unaffected by the regulation. Even quite sophisticated studies can be limited in this regard. For example, a number of analyses have examined the effect of attainment status, with respect to National Ambient Air Quality Standards (NAAQS) on employment and other metrics of economic activity (e.g., Greenstone 2002; List et al. 2003, Greenstone, et. al. 2011). The authors identify the effects of regulation by comparing regulated plants in nonattainment counties for one or more pollutants with unregulated facilities in attainment areas.⁹ Although these analyses generally estimate some shift of economic activity from nonattainment to attainment counties, they do not conclusively demonstrate that there is an overall decline in aggregate economic activity due to environmental regulation. It is certainly plausible that plants in a county that changes from attainment to nonattainment for a particular NAAQS between one time period and the next could reduce employment, while plants in an attainment county gain. At the same time, some or all of the individuals who lost jobs in the nonattainment areas may have found work in attainment areas, and some or all of the capital and output losses may have been redeployed in such areas as well. However, because the potential gains arise at plants and locations that are not directly affected by regulation, studies based solely on attainment status cannot be used to calculate the net effects of a national regulation.¹⁰

3. Review of Individual Papers

The retrospective case studies in the RPI include a broad array of regulations, including both market-based and command-and-control oriented rules issued by four different federal agencies. Some of the cases have a clear health focus, while others are oriented to ecological or

⁹ More recent papers, which have expanded the attainment-nonattainment comparisons, also have limited application to national-level benefit–cost analyses. For example, Kahn and Mansur (2013) use a border-pair fixed effect model to compare employment by sectors at the county level for adjoining counties that vary by attainment status. Walker (2013) focuses on workers rather than plants as the unit of analysis, which permits estimation of actual costs related to regulatory job loss.

¹⁰ Furthermore, the costs associated with spatial reallocation of production or earnings losses as a potential consequence of regulation are not typically considered in national-level benefit–cost analyses. Even if they were included, the results from studies relying on a retrospective assessment of a specific regulatory change (e.g., the 1990 CAA Amendments) cannot be scaled in any obvious way for application in RIAs for new and potentially different regulations.

energy issues. Several cases involve the development of new technologies.. Specifically, they cover the following:¹¹

- EPA’s Phase II of the Acid Rain Program (ARP), focusing on the emissions trading component of the program, enacted under Title IV of the 1990 Clean Air Act Amendments (Chan et al. 2015);
- EPA’s Renewable Fuel Standard (RFS2), which mandates a major expansion of US biofuel consumption as stipulated in the Energy Independence and Security Act of 2007 (Lade et al. 2015);
- EPA’s Cluster Rule, the agency’s first integrated (air and water) multimedia regulation, which seeks to reduce toxic releases from the pulp and paper industry by establishing maximum achievable control technology (MACT) levels under Section 112 of the 1990 Clean Air Act Amendments and best available technology (BAT) standards, known as effluent guidelines, under the Clean Water Act of 1972 (Gray and Shadbegian 2015);
- EPA’s Air Toxics Program, establishing MACT standards for multiple industries (Fraas and Egorenkov 2015a);
- EPA’s Effluent Guidelines Program, setting BAT limits for multiple industries (Fraas and Egorenkov 2015b);
- EPA’s Clean Water State Revolving Fund (CWSRF), a federally supported state-level program that provides funding to states for publicly owned treatment works (POTWs) (Harrington and Malinovskaya 2015).
- FDA’s egg rule, which requires egg farmers to adopt control measures to limit possible *Salmonella* Enteritidis (SE) contamination (Lutter 2015);
- DOI’s designation of critical habitat (CH) on land cover change under the Endangered Species Act (Nelson et al. 2015); and

¹¹ Four of the studies (EPA’s Air Toxics and Effluent Guidelines Programs, EPA’s Clean Water State Revolving Fund (CWSRF), and FDA’s egg rule) were conducted by in-house RFF experts. The others were successful applicants under a competitive bidding process carried out at RFF in 2013 where selection was based on ‘...the originality and scientific merit of the proposed research, compatibility with the goals of the RFF RPI, and the demonstrated research competence of the applicants.’ For details see <http://www.rff.org/blog/2015/request-proposals-retrospective-studies-regulatory-performance>.

- DOE’s minimum efficiency performance standards (MEPS) for five appliance types—room air conditioners, refrigerators-freezers, dishwashers, clothes washers, and clothes dryers—established under the Energy Policy and Conservation Act of 1975 and subsequent amendments (Taylor et al. 2015).

Summaries of the nine studies, including descriptions of the specific regulations analyzed and the basic findings of the individual analyses, are presented in the appendix.¹²

The initial aim of the RPI was to avoid reliance on a “convenience sample” by examining entire classes of regulations, such as all MACT rules issued under Section 112 of the 1990 Clean Air Act Amendments or all BAT standards issued under EPA’s Effluent Guidelines Program. In the end, however, this proved to be infeasible because of unanticipated data limitations. The cases that come closest to this approach are the analysis of DOE’s MEPS, where the five regulations examined represent the full set of large household appliances subject to regulation prior to 2012, and the analysis of the CWSRF, which studied four individual state programs, those of Indiana, Iowa, Maryland, and Texas. The remaining cases mostly cover major or otherwise significant rules that represent an important part of the federal regulatory landscape. However, despite the relatively large number of cost-and-benefit or effectiveness comparisons developed in the RPI, we make no claim that the regulations examined are representative of all federal environmentally oriented rules.

For simplification, we label the ex ante–ex post comparisons as “accurate” where an RIA estimate falls roughly within $\pm 25\%$ of the ex post observation.¹³ Outside of these bounds, the results are deemed to be “underestimates” or “overestimates.” A summary of all 34 benefit/effectiveness and cost comparisons is displayed in Table 1. Although we find examples of accurate RIAs, as well as both underestimates and overestimates, and more than half the case comparisons involve accurate or underestimates of benefits, there is somewhat of a tendency among the studied regulations for overestimation of both costs and benefits/effectiveness, a conclusion supported by a number of earlier studies (see OMB 2005; Harrington 2006; Simpson 2011). Ideally, one could go further and actually quantify the extent, on average, that costs and benefits/effectiveness are overestimated. Unfortunately, the diversity of methods in the various

¹² The full set of papers is available at <http://www.rff.org/research/collection/regulatory-performance-initiative-rff>.

¹³ The convention of defining accuracy as $\pm 25\%$ was used in Harrington et al. (2000) and subsequently adopted by other authors, including OMB (2005).

case studies and the different baselines used present major challenges to estimating the average overestimation of either costs or benefits/effectiveness. Instead, we try to learn from the individual cases, focusing on the causes of the errors and potential for future improvements.

3.1. Cost Comparisons

Of the 12 regulations for which we are able to make ex ante–ex post cost comparisons, we see that 9 involve overestimates, 2 are underestimated, and 1 is accurate. Interestingly, all of the DOE MEPS of major household appliances count among the cost overestimates, as the authors demonstrate lower-than-expected purchase prices for all five products studied and greater-than-expected sales for one product (clothes washers).

The CWSRF, which accounts for the remaining four cost overestimates, differs from the other retrospective studies in that the cost comparison involves the size of the loans requested by municipalities in the four states versus the amount actually awarded by the administering agency (the state government). Although RIA-type pre-regulatory studies were not conducted for individual POTWs, one would expect that the typical loan request, a proxy for the ex ante cost estimate, would exceed the amount of the actual award, a proxy for the realized cost. What actually happened, however, was quite different. Across the four states, the case study authors report 126 loans that exceeded needs and only 109 cases of needs exceeding loans. At the same time, when considering the dollar difference between needs and loans, the reverse was true: total needs substantially exceeded total loans. For example, in Iowa 61 loans exceeded needs, but the sum of the excess was only \$144 million, and while needs exceeded loans in 37 cases, the total excess was nearly \$1.2 billion. Evidently, it was mostly the large cities that had requests larger than the loans. In small towns and rural areas, the reverse was true. The authors suggest that this result obtains because of a relative lack of sophistication among POTW managers in small towns, where few if any POTWs have full-time well-trained managers. The large dollar excess of needs versus loans in all four states is the primary basis for which the case study authors score as overestimates.¹⁴

Turning to the cost underestimates, the relevant metric for the emissions trading component of Phase II of the ARP is the difference between the costs of coal purchase and scrubbing under the program and a uniform performance standard that would have achieved the

¹⁴ Arguably, this interpretation is more of a judgment call than the other case comparisons, since the authors are assuming the loan amount reflect the realized cost of the projects.

same overall SO₂ reductions. To be clear, the entry for this rule in Table 1 should not be interpreted as saying that EPA underestimated the costs of complying with Title IV of the Clean Air Act Amendments. Rather, the focus is strictly on the trading component of the program. The authors calculate the savings from trading at \$250–\$300 million (1995 US\$) per year in Phase II of the program, less than half the amount estimated in an often-cited paper by Carlson et al. (2000), and an even smaller portion of the savings originally estimated in EPA’s RIA. Although the presence of institutional constraints on trading is well known, the magnitudes involved are surprisingly large.

While EPA’s original analysis and Carlson et al. (2000) both assume that firms will achieve the least-cost solution to reducing emissions via fuel switching and that no additional scrubbers will be built after 1995, the retrospective analysis suggests that as a result of state-level political pressures and related factors, more scrubbers than necessary were actually installed. In total, almost one-fourth of the generating units were not using the least-cost compliance method. According to the retrospective analysis, additional annual cost savings of \$206 million (1995 US\$) could have been achieved had the program worked as assumed in the early studies.¹⁵

EPA’s Renewable Fuel Standard represents the other case of a cost underestimate from among our retrospective studies. In this case, the agency’s analysis missed the RIA estimate by quite a wide margin, as it failed to account for delayed development of an advanced biofuels industry, delayed investments in alternative fuel vehicles and fueling infrastructure necessary to consume more than 10% ethanol-gasoline blends, and the inherent uncertainty in the relative costs of oil and biofuels. Because of the specific volumetric requirements of the regulation, the failures to account for the delayed development of the new technology had major cost impacts, some of which may have been reduced by the adoption of a more flexible rule design in the first instance. As discussed in the next section, the case study authors identify a number of elements of the modeling and analysis that led to these errors and propose constructive paths for going forward.

The relatively accurate cost estimates developed by the DOI attempt to quantify potentially unmeasured costs of CH designation on land cover change under the Endangered

¹⁵ The authors also note that because of the installation of the additional scrubbers, the *global* least-cost solution was not achieved, as the cost per ton of SO₂ removed in the added scrubbers installed between 1996 and 2002 is considerably higher than the cost of purchasing an SO₂ allowance.

Species Act. Using recently published digital critical habitat maps and high-resolution geographic datasets, the authors do not find evidence that the average rate of change in developed and agricultural land is significantly affected by CH designation. Unsurprisingly, a number of important caveats apply here, including the finding that the average results are not robust for higher-value land, an issue we return to below.

3.2. Effectiveness Comparisons

Of the 22 regulations or regulatory requirements for which we are able to compare benefits/effectiveness estimates on an ex ante and ex post basis, we find that 10 involve overestimates, 6 are underestimates, and 6 are deemed to be relatively accurate. All five of DOE's MEPS are underestimates, as the case study authors find that DOE used overly cautious approaches to assessing the development of the new technologies.¹⁶ The other underestimate involved EPA's air toxics regulation of the printing/publishing industry. Note that this finding contrasts with the results from the other four air toxics regulations—pulp/paper, petroleum refining, pharmaceutical, and wood furniture industries—where effectiveness was overestimated. The benefits of emissions trading versus a uniform performance standard were also overestimated in Phase II of the Acid Rain Program, along with all three of the effectiveness measures used in the air toxics part of the Cluster Rule study (VOCs, toxics, and PM₁₀) and in FDA's egg rule.

Recall that emissions trading, per se, is the focus of the ARP Phase II case, rather than the overall benefits of the regulation. In fact, it has long been established that the benefits of the overall regulation were grossly *underestimated* due to the post-regulatory determination of the seriousness of the health damages associated with fine particles, an issue that was not formally considered at the time the initial regulation was issued in 1992. At the same time, and despite the substantial reductions in overall emissions, the case study authors calculate that health damages

¹⁶ Our reason for placing the DOE MEPS into the effectiveness underestimation category is twofold. First, the authors find that consumers purchased more efficient products than estimated for the one reviewed product (clothes washers) for which the RIA estimated the likely market share of more efficient products. In addition, the monthly sales-weighted average of the products sold in 2003–11 was better than the minimum standard for almost all data points in the study period. Second, the authors show that significant improvements occurred in the non-regulated aspects of the quality of products available on the market, as evidenced in Consumer Reports tests, which coincided with MEPS events. The authors also document post-purchase quality improvement over the longer period of time the products were regulated, as evidenced in annual large-scale surveys of the Consumer Reports readership regarding the incidence of major repairs. Of course, these views are not unanimous. For example, see <http://www.wikihow.com/Get-Rid-of-Mold-Smell-in-Front-Loader-Washing-Machine>.

under the ARP, which included trading, were greater than under a no-trade counterfactual that would have achieved the same overall emissions reductions, largely reflecting the transfer of allowances from power plants west of the Mississippi to facilities with higher exposed populations in the eastern United States. As further support for this result, the authors cite the finding by Henry et al. (2011) of a positive correlation between marginal abatement costs for SO₂ and marginal damages from SO₂ emissions. It should be emphasized that the negative impact of trading in this case is directly tied to a benefit category, namely the health impacts of fine particles, that for both legal and scientific reasons was not explicitly considered by EPA when the rule was issued. As discussed below, consideration of interregional trade constraints might be a desirable option to prevent this problem from occurring in future emissions trading rules.

The benefit overestimates in the egg rule were largely driven by deficiencies in FDA's pre-rule estimates of the effectiveness of the regulation in reducing illness and by revised estimates of the costs of medical treatment for *Salmonella* Enteritidis. As the author notes, a relatively effective program to reduce the spread of this pathogen was established somewhat earlier in the United Kingdom. In the case of the air toxics part of the Cluster Rule, the case study authors suggest that some of the overestimates may stem from inaccuracies in the pre-regulatory information available to EPA on plant-level baseline emissions of VOCs and PM₁₀, both of which were derived from the agency's National Emissions Inventory (NEI).

Two of the six cases where the pre-regulatory estimates of effectiveness were relatively accurate when compared with the ex post analysis involved chloroform emissions/discharges to air and water in the Cluster Rule. In both cases, the baseline emissions were drawn from EPA's Toxics Release Inventory (TRI), rather than the NEI. While the TRI also served as the source of emissions data for the regulatory metric labeled "air toxics," which, as noted, was overestimated in the pre-regulatory analysis, the authors suggest that the accuracy of baseline data may be a factor here as well.

Despite the absence of an RIA in the CWSRF, the case study authors have deemed the outcomes to be reasonably accurate compared with expectations. Although engineering estimates of secondary treatment at a POTW might suggest larger reductions in discharges than actually observed at the facilities, if one considers the improvements observed over the study period as part of a multiyear or even a multi-decadal process, then one would not expect to see the full gains achieved during the study period. Thus, despite the absence of a rigorous ex ante analysis, the authors' contention that the observed modest reductions in discharges are consistent with prior expectations seems eminently reasonable.

Taken together, we note a number of patterns in the ex ante versus ex post cost and performance comparisons. Errors in forecasting technology development are clearly an important issue in the case studies. While some might suggest a tendency to conservatism in forecasting technology development, in fact, the errors go both ways, as the DOE MEPS overestimated costs and underestimated effectiveness, while EPA's RFS2 underestimated costs.

A second pattern relates to basic analytic errors that should not occur in a quality RIA, especially for a transformative regulation like the RFS2. The authors note that beyond the technology forecasting errors, for example, the RFS2 also failed to consider variability in oil and biofuel prices, as well as potential delays in infrastructure investments. There was also a single-minded focus on the long-term analysis without regard to short- and medium-term "start-up" problems that could occur. Going forward, the authors argue, in effect, that for future regulations of this type the RIA should assess what 'could' happen rather than what 'would' happen.

A further theme of the cases relates to the heterogeneity of impacts along several dimensions. The authors of the paper on CH designation under the US Endangered Species Act (ESA) note the possibility of adverse impacts of CH designation for high-value land, even though these impacts do not show up in their analysis of average impacts across all land types. The Acid Rain Program case study authors state that failure to account for differences in population density across regions can result in perverse impacts from emissions trading. Similarly, the Cluster Rule study authors point to the importance of cost heterogeneity among pulp and paper plants. The egg rule study author suggests that the different levels of management sophistication across egg farms might have argued for a different regulatory approach for smaller enterprises. All these examples indicate the need to go beyond consideration of the average impacts and also examine potential outliers. As we discuss in the next section, additional ex ante analysis could shed light on all these complexities prior to rule promulgation.

Finally, we note the importance of institutional constraints that are sometimes not adequately accounted for in RIAs. This is most apparent in the ARP, where state-level political factors limited the attainment of the forecast least-cost solution by a considerable amount. Institutional constraints may also have been a factor in the cost underestimates of the RFS regulation, as the production ramp-ups were constrained in various ways not anticipated in the pre-regulatory analysis.

Table 1. Summary of Ex Ante–Ex Post Comparisons

RULE	METHODS	ESTIMATED COSTS			ESTIMATED BENEFITS/EFFECTIVENESS		
		Accurate	Over	Under	Accurate	Over	Under
Acid Rain Program Phase II trading	Model estimation, simulation			x		x	
Cluster Rule: overall							
air: chloroform	Diff-in-diff, propensity score matching				x		
water: chloroform					x		
air: VOCs						x	
air: toxics						x	
air: PM ₁₀						x	
Pulp/paper effluent guideline	diff-in-diff estimation, time series analysis (OLS)					x	
Air toxics							
printing/publishing	diff-in-diff estimation, time series analysis (OLS)						x
pulp/paper						x	
petroleum refining						x	
pharmaceutical						x	
wood furniture						x	
Clean Water State Revolving Fund							
Indiana	event study/time series analysis		x		x		
Iowa			x		x		
Maryland			x		x		
Texas			x		x		
Renewable Fuel Standard (RFS2)	credit prices proxies for cost various stat. methods			x			
Egg rule: <i>Salmonella</i>	diff-in-diff estimation					x	
Critical habitat designation under the ESA	propensity score matching	x					
MEPS							
room air conditioners	time trends, OLS multiple datasets		x				x
refrigerators			x				x
dishwashers			x				x
clothes washers			x				x
clothes dryers			x				x

4. Recommendations from the Case Studies

This section summarizes the key recommendations made by the case study authors. For ease of presentation, these recommendations are divided into two categories: (i) those aimed directly at improving the effectiveness and efficiency of regulation, which can include improving the RIA to better understand the costs and benefits/effectiveness of the regulation, as well as the scoping out of additional regulatory options; and (ii) those relating to the conduct of future retrospective analyses, focusing on specific enhancements to the federal role.

4.1. Direct Regulatory Improvements

One of the strongest recommendations overall is from the authors of the ARP Phase II case study, which finds that as a result of state-level constraints, the emissions trading regime failed to deliver the full amount of expected cost savings and, further, that it resulted in a net *increase* in health damages compared with a uniform performance standard, as emissions moved from lower to higher population density areas. The authors' proposal is that future emissions trading regulations should drop the default assumption that trading will achieve the least-cost outcome and, instead, treat cost savings based on the least-cost solution as a lower bound estimate. Even though this recommendation is based on a single study, the rigor of the analysis provides great weight to the proposal. A second recommendation from the same authors focuses on the benefits side. Here the authors suggest that the pre-regulatory analysis should specifically consider the health and environmental impacts of the spatial redistribution of emissions associated with trading in future regulations.

Both of these recommendations represent a significant departure from current EPA practices, and both could lead to important changes in the design and stringency of future emissions trading programs. Dropping the least-cost default assumption would clearly result in higher expected costs of the program. All else equal, higher expected costs might lead to the adoption of less stringent regulations. Explicit consideration of the spatial redistribution of emissions associated with emissions trading could potentially provide support for zoned trading or other geographic-related mechanisms that valued some emissions reductions more than others versus the current standard practice of 1:1 trading. This last change might alter both the costs and benefits of trading programs.

Another interesting set of recommendations emerges from the case study on the Renewable Fuel Standard, which, the authors say, overlooked several fundamental factors essential to conducting a quality RIA. These problems proved to be especially important because

the regulation contains such specific volumetric requirements. The authors recommend that the agency implement a series of changes: (i) simplify future RIAs, focusing on the big picture rather than attempting to predict outcomes for every fuel type in every region; (ii) avoid large-scale engineering-type models, which lack transparency and tend to shift focus away from the most critical aspects of the regulation toward less important details; (iii) focus on short- to medium-term compliance scenarios, not just long-run outcomes; (iv) consider “worst case” compliance scenarios in an effort to anticipate the effects of delays in technology development or major investments and to “stress test” how regulatory costs respond; (v) recognize the inherent uncertainty of predicting the relative prices of oil and biofuels by constructing a number of plausible price outcomes and specifying a range of costs of the regulation rather than relying on a single price trajectory; and (vi) periodically update the analysis to assess whether the program is keeping pace with its intended goals.

Collectively, adherence to these recommendations would have likely unmasked the inherent uncertainties in achieving the congressionally mandated objectives for new biofuel technologies. Such an analysis, had it been conducted prior to promulgation of the rule, might have lent support to more modest regulatory objectives, proactive measures to relax the infrastructure constraints, or changes in the design of the rule, perhaps by the adoption of a safety valve in the form of a compliance credit price ceiling/floor or other mechanism that would have allowed greater flexibility in the face of delays in technology development or related factors. As it turns out, the schedule has been revised several times already since the initial rule promulgation, although not without the imposition of significant costs on both producers and consumers.

As noted earlier, there is a degree of overlap between the Cluster Rule case study, imposing technology standards for the pulp and paper industry, and the case studies on both the air and water toxics rules, which included the pulp and paper industry plus several others. While all three papers rely on TRI data, they differ somewhat in terms of the time periods and specific facilities covered. In addition, the papers use somewhat different measures of toxic pollutants. The Cluster Rule analysis for the pulp and paper industry relies strictly on chloroform concentrations in the water as the toxics metric, while the study of water toxics in multiple industries is based on a broader set of TRI toxics. Both papers recognize that over the lengthy course of rule development, EPA had to modify the estimates of the expected pollution reductions and the costs of the rules to reflect ongoing changes in the industry. One of the principal differences among the papers is the relative emphasis placed on prior EPA rulemaking in driving the early reductions. The authors of the air and water case studies attribute a large

portion of the early reductions to prior EPA rulemakings under subsection 304(l) of the Clean Water Act, while the Cluster Rule case study authors see them as driven by a number of voluntary initiatives, evolving dioxin standards in many states (perhaps triggered by EPA's 304(l) rule), and anticipatory actions to the Cluster Rule itself.

Notwithstanding the different interpretations of the motivations for early reductions, which cannot be easily resolved after almost two decades have passed, what can be learned from the technical analyses about the technology-based regulation of air and water toxics for the pulp and paper industry, and possibly the other industries as well? Defining the baseline is crucial for measuring both benefits and costs of any rule or regulation. For regulations, such as the Cluster Rule, that involve a particularly long development process, developing an accurate baseline is especially challenging, as documented in all three case studies. In fact, during the development of the Cluster Rule, EPA updated its calculations of benefits and costs to reflect the changing situation between proposal and promulgation. The fact that the resulting emissions/discharges were still overestimated suggests that the agency's baseline revisions did not adequately account for the realized reductions prior to rule promulgation. Thus, going forward, the clearest recommendation is for the agency to make an even stronger effort to acquire up-to-date baseline information in their pre-regulatory analyses, including both the emissions/discharge characteristics of the plants and the technologies and abatement equipment employed. Further, all three cases argue that the pre-regulatory analyses should put more effort into identifying the range of expected compliance methods and potential heterogeneity across plants.

Both the use of more up-to-date performance data and more detail on compliance options would reduce the tendency to make inaccurate claims about the expected benefits and costs of the regulation and possibly influence stringency levels. For the pulp and paper regulations, such data would have resulted in smaller estimates of both benefits and costs. However, because the MACT rule is a technology-based standard keyed to the top 15% performing plants in the industry, it probably would not have influenced the overall stringency of the rule in this case.

A second issue arising from the study of the Cluster Rule, as well as the broader analyses of the air and water toxics regulations, involves the treatment of industry laggards. The authors of all three cases argue that even though the direct impact of a rule may have been mostly to clean up the laggards, its issuance may have been important to avoid undermining the agency's ability to set future regulations. That is, if the laggards were not brought up to appropriate BAT or MACT levels, as defined in the Clean Water and Clean Air Acts, respectively, a future rule would have to confront a less controlled set of sources in determining the BAT or MACT—both of which give considerable weight to economic achievability. Others might argue that a rule

addressing the laggards was necessary in order to level the playing field for competitors. At the same time, the benefits of reducing the discharges from the still modestly controlled facilities may not justify the costs involved, because of the location of the mills or other factors. EPA's analysis does not present the benefit and cost estimates in a way that allows further judgment on the merits of setting tighter limits for these plants. A more complete economic analysis would be needed to address this issue. Regardless of the results of such an assessment, however, revisions to both the Clean Water and Clean Air Acts would likely be required to tailor these technology-based rules in this manner.

A third issue of interest arising from the Cluster Rule concerns the value of integrated, multimedia rulemaking as opposed to the standard medium-by-medium approach. One of EPA's goals in adopting a multimedia approach was to reduce compliance costs, and the authors cite evidence that actual costs were lower than the agency's pre-regulatory estimates. At the same time, they note that the inability to observe the counterfactual of separate air and water regulations makes it impossible to judge the extent of cost savings, if any, properly attributable to the multimedia approach. As a general matter, the authors argue that RIAs should put more emphasis on identifying the range of expected compliance methods and potential heterogeneity across plants. This would allow retrospective analyses to identify changes in compliance methods and technical advances that might help explain observed differences between predicted and actual effectiveness and costs, thereby improving future ex ante estimates. Thus, while not undermining the potential cost savings potential of multimedia regulation, the authors remain agnostic about the likely economic gains from such an approach.

Turning to the analysis of the potential for unquantified costs of CH designation on land cover change under the ESA, the authors of this case study find, as noted, that the average rate of change across all developed and agricultural land is not significantly affected by such designation, although that conclusion is not robust for higher-value land. Going forward, they recommend that pre-regulatory analyses need to continue to develop disaggregated cost estimates into categories of land cover, land management, and administrative/consulting. Pre-regulatory studies also need to more carefully distinguish incremental costs of CH from coextensive costs of listing endangered species. Finally, in light of the nonrobust findings for higher-value land, the authors recommend that future CH designations, along with the corresponding RIAs, should be sensitive to land management impacts, especially for higher-value land.

The finding that CH designation does not impact land development is certainly not a consensus view among experts, and it is possible that it may not stand up to subsequent analysis.

Nonetheless, the authors deserve enormous credit for using new mapping data to quantify a difficult-to-study impact. The recommendation that future RIAs should be especially sensitive to any land management impacts of CH designation, especially for higher-value land, seems quite sound and reflects the uncertainties in these findings.

Regarding the case study on FDA's egg rule, the author reports, as noted, overestimates of the benefits. Going forward, he makes two recommendations of note. First, he suggests that FDA collect information via demonstration programs or some other means that is better able to assess both the potential effectiveness of future rules and the societal gains (benefits) likely to be realized from such actions. These further assessments should probably distinguish farms according to their management sophistication, perhaps based on the size of the operations. Second, the author calls for changes in the presentation of information to the public, noting the importance of using specific health outcomes and other risk measures that matter most to people, as well as the value of identifying comparison pathogens, so that individuals can better gauge the potential severity of the problem. Taken together, these recommendations would likely result in more efficient and, arguably, more effective regulation, along with an improved public understanding of the issues.

Unlike most other studies in the RPI, the study of the CWSRF was not an analysis of a specific regulation, but a study of a federal program providing loans to POTWs that were already subject to water quality regulations. Only some of the facilities seeking loans were funded in the sample period (2008–12), providing an instructive contrast in performance: between 2007 and 2014, the authors observed a modest but unmistakable difference in effluent quality, which largely improved in the plants receiving loans but was trendless in those that did not.

However, the authors also observed another finding from the data: those POTWs receiving loans tended to have better effluent quality in 2007, before 2008–12 loans were distributed. At first blush this outcome seems odd, given that pollutant discharge into receiving waters typically has rising marginal costs. That is, the greatest improvement in environmental quality for a given investment can usually be achieved by targeting the worst performers first. The matter is complex, however, and other considerations may prevail. POTWs also have strong economies of scale, and this might lead state authorities to continue to favor plants receiving loans in an earlier period. Then, too, the authorities might wish to use the revenues to reward previous good performance, a criterion that would also tend to aid plants with superior effluent quality. Going forward, as the case study authors suggest, it seems reasonable for state authorities to revisit the loan selection criteria as a means of achieving greater environmental improvements from the available resources.

The authors of the MEPS case study demonstrate that the standards result in the purchase of energy-efficient appliances at lower-than-expected prices. Moreover, the quality indicators the authors assess suggest that consumers of energy-efficient appliances are not being made worse off in nonregulated aspects of product performance and, indeed, appear to be made better off in several dimensions of product quality. In the case of clothes washers, this is particularly salient, as it aligns with an important court decision regarding alleged mold growth in clothes washers.¹⁷

At the same time, the authors do suggest some changes to the RIA process that could improve the understanding of the proposed standards and generate more accurate regulatory cost predictions. First, they suggest that the interpretation of the statutory requirement for “technologically feasible” standards be changed from current practice. Specifically, they are concerned that potential efficiency-enhancing design options are eliminated from consideration at the outset of the multiyear rulemaking process if these options are not currently incorporated in a commercial product or working prototype. In essence, they argue that the current practice works to enshrine the outcomes of previous iterations of model design in regulatory expectations of future generations of products. While the authors’ proposal might present implementation challenges to regulators, it does hold the promise of greater efficiency gains. Second, the case study authors argue for more extensive analyses to account for the likelihood of feature correlation and cross-functionality, including consideration of nonenergy product quality improvements that can occur with regulation. Overall, they believe these changes would support more cost-effective standards in future regulations.

4.2. Improving Future Retrospective Analyses

Retrospective analysis is a relatively new topic for regulatory agencies. As the case studies demonstrate, there are numerous barriers to conducting such analyses. Other than general discussions in various OMB *Reports to Congress on the Costs and Benefits of Federal Regulation*, a recent executive order encouraging agencies to conduct such analyses, and the limited agency-level work previously cited, the federal government has not been a major driver of rigorous retrospective analysis of individual regulations. While the RPI was designed to

¹⁷ In *Whirlpool Corp. Front-Loading Washer Products Liability Litigation No. 1:08-WP-65000 in the U.S. District Ct. Northern District of Ohio*, the jury’s decision generally supported the firm’s contention that “the vast majority” of consumers of the models cited in the class did not complain about performance issues and “exonerated” the company’s design. See Frankel (2014).

develop additional experience with retrospective analysis, it is probably not a coincidence that this work has been carried out by independent researchers at RFF and various academic institutions, with funding from nongovernment sources. While we discuss below alternative institutional settings for conducting such studies, here we turn to the case study authors' proposals to enhance the federal role in promoting future retrospective studies.

The strongest recommendations for changes in federal government procedures are made by Fraas and Egorenkov, the authors of the case studies on both the air and water toxics regulations. They propose three specific measures, two to be carried out by OMB and one by regulatory agencies. Mirroring a recent recommendation by Aldy (2014) and, subsequently, by the Administrative Conference of the United States (2014), Fraas and Egorenkov suggest that OMB formalize a process for the retrospective evaluation of important rules under the authority of Executive Order (EO) 13563. To implement this recommendation, OIRA could issue guidance identifying factors agencies should consider in selecting rules for regulatory review, the kinds of measurable outcomes to be targeted in the analysis, the associated data requirements, the preferred methods to be used, and the time frame to be evaluated.¹⁸ Such a guidance document would supplement the long-established OMB guidance on the conduct of ex ante analysis for RIAs, codified in OMB Circular A-4.

The authors also offer suggestions on how OMB could provide a streamlined process for data collection under the Paperwork Reduction Act (PRA). To obtain OMB approval under the PRA, agencies would have to show that data collection has "practical utility" and is the least burdensome way of obtaining the information. The PRA also requires the agency to go through a public comment process.

OMB has previously moved to streamline its PRA review by establishing a generic clearance process for specific types of information collection. In a memo titled "Facilitating Scientific Research by Streamlining the Paperwork Reduction Act Process," OMB has outlined

¹⁸ OIRA has already provided general guidance along these lines. OIRA (2011) and Sunstein (2011b) include advice on the implementation of the retrospective review provision of EO 13563. OIRA (2011) identifies the following topics as areas that agencies should address in conducting retrospective review: public participation, prioritization, analysis of costs and benefits, and coordination with other forms of mandated retrospective analysis and review. Sunstein (2011b) also recommends that to promote a consistent culture of retrospective review, "future regulations should be designed and written in ways that facilitate evaluation of their consequences and thus promote retrospective analyses. To the extent consistent with law, agencies should give careful consideration to how best to promote empirical testing of the effects of rules both in advance and retrospectively." See also Sunstein (2011a).

options and strategies for agencies to use to streamline the process of getting approval of research-related information collection. Fraas and Egorenkov recommend that OIRA explicitly provide this streamlined PRA process for data collection for retrospective studies.

Turning to actions by the regulatory agencies, the authors recommend that OMB require agencies to provide in the final rule for the collection of data on the control measures adopted and the costs incurred to comply with the rule, as well as production characteristics for at least a representative sample of plants. One way of doing this would be to require the reporting of this information to the regulatory agency as part of the final rule. However, the collection of such information could be burdensome, even if restricted to a representative sample of covered plants, so the regulatory agency would need to be strategic in its selection of retrospective data to be collected, possibly coordinating with existing data collection efforts by the US Census Bureau or other relevant agencies.¹⁹

Several other case study authors also weigh in on these issues. The author of the egg rule calls for better cooperation among the multiple agencies involved in the regulation and analysis of food safety issues, including adequate sampling of both large- and small-scale producers' practices and costs, as well as public disclosure of that data. Similarly, the authors of the CWSRF case study note the relatively lax data collection and reporting requirements imposed on states as a condition to be eligible for federal subsidies or loans. Potentially valuable datasets are compromised by such failures as a lack of uniform reporting requirements or lack of audits to correct missing or obviously incorrect data. These and other problems could be avoided by closer federal oversight of datasets created by the states or localities with federal funds.²⁰

¹⁹ In their analysis of the air toxics rule, Fraas and Egorenkov recommend that EPA provide for the collection of data on control measures adopted and costs to comply with the rules, emissions data, and plant production characteristics for at least a representative sample of plants.

²⁰ In working with the federal and state datasets related to the CWSRF and local water quality management, including the Clean Water Benefits Reporting (CBR) system, the Clean Watersheds Needs Survey (CWNS), and the Discharge Monitoring Report (DMR), a federal pollutant discharge dataset consisting of assembled state data, the authors note the difficulty of merging datasets owing to oversights as simple as the failure to maintain consistent facility ID numbers across the datasets. They are further dismayed by the frequent absence of complete information in the DMR, even in data fields that would be required to ensure that plants are in compliance with regulations, including basic operating data such as biochemical oxygen demand and wastewater flow.

5. Conclusions

Rigorous retrospective analysis of federal regulations can provide the foundation for improving new regulations, including the conduct of future RIAs. Such analysis also establishes the basis for a better understanding of the impacts of existing regulations, revealing opportunities for strengthening some and modifying others. Further, it may suggest new opportunities and methods for future retrospective studies. While a single project, even one based on a diverse set of case studies, cannot possibly fill the enormous knowledge gaps in this area, a number of observations do emerge from this work.

Of course a long list of caveats apply to any conclusions drawn from a limited selection of case studies, especially those based on a small and non-representative sample. The results presented here support earlier findings that agencies have somewhat of a tendency to overstate both the costs and the benefits/effectiveness of regulations. This does not mean that costs and benefits are *always* overstated. In fact, slightly more than half the case comparisons reported here involve accurate or underestimates of benefits. On the cost side, the overestimates are principally driven by DOE's MEPS. No evidence is offered on whether costs are overstated more than benefits/effectiveness, or vice versa.

While some may see the analysis produced here as an indictment of regulation, that is far from the case. Rather, this undertaking provides a framework for assessing the outcomes of federal regulations. Since the goal here is to learn from these early efforts at retrospective analysis, we highlight a few specific findings from individual cases that may have application to the future universe of environmental regulation:

- *The realized cost savings from emissions trading may fall short of the idealized textbook case of a least-cost outcome.* While the notion of a shortfall not a surprise to experts in the field, the size of the shortfall reported here is somewhat surprising. Going forward, to avoid overstating the cost savings from emissions trading, the recommendation is that the least-cost textbook assumptions be treated as an upper bound rather than as a best estimate.
- *Emissions trading may lead to a spatial rearrangement of emissions that reduces overall benefits compared with a uniform performance standard with the same aggregate emissions.* This result, too, has long been seen as a possibility in the expert community, albeit not a well-documented finding. The recommendation here is that regulators should explicitly consider this possibility when developing new regulations and, where appropriate, introduce geographic-based trading designs that limit the adverse outcomes.
- *Highly prescriptive congressional mandate warrant the same quality of analysis as other major regulations.* Highly prescriptive legislative mandates, especially involving large-

scale changes in economic activity, could involve costs or benefits that may become apparent in a regulatory setting via a well-done RIA. Under extant executive orders, a regulatory agency has no less of an obligation to conduct rigorous RIAs for such prescriptive regulations than for others. The recommendation here is for agencies to conduct, and for OMB to demand, such analyses. *Up-to-date baseline information is essential for sound regulation.* Several case studies presented here reveal the absence of current baseline information that led to significant errors in the estimation of both the expected costs and benefits/effectiveness of new regulations. Greater effort on the part of the agencies is clearly needed to obtain more current information on abatement equipment in place and actual environmental performance.

- *Regulations based on new, noncommercial technologies require particularly rigorous assessments.* Forecasting technological progress over five years or even shorter periods can be extremely challenging. In cases where new regulations are particularly sensitive to the performance and costs of still-evolving technologies, extra care should be taken to ensure the use of credible methods. Evidence unearthed here suggests it is especially important for regulators to avoid early screening out of promising options, and equally important to be rigorous in considering difficult-to-estimate costs of options that are included, along with the timing and feasibility of introducing these products in the marketplace. Some of these problems can be at least partially addressed by improved pre-regulatory analysis, including greater use of uncertainty analysis.
- *Federal programs operated by states should be evaluated retrospectively to assess both their effectiveness and their efficiency.* Evidence presented here suggests that federal funds allocated by states may not be fully achieving their intended purposes. Specifically, there is considerable potential in at least one major program to reallocate funds to obtain better environmental results. While there may be a management rationale for the observed allocations, the burden is on the program managers to provide further evidence on this issue.
- *Potential heterogeneity of costs and benefits needs to be further considered by regulators.* An important theme in several of the case studies is that the focus by regulators on the average costs or benefits/effectiveness often misses important differences among regulated facilities. While it is clearly not possible for a regulatory agency to consider all possible impacts, the case study results suggest that current practices in pre-regulatory studies are inadequate in some cases. At a minimum, greater agency attention to the potential heterogeneity of impacts across groups of firms or population subgroups seems warranted.
- *For some rules, the potential uncertainties are so great that an RIA should seek to illustrate a range of possible outcomes rather than focus on narrow point estimates.* At a minimum, the inherent uncertainties in world energy prices and other macro variables

should be considered when they could have a major influence on regulatory outcomes. Alternative regulatory designs that allow for flexibility in regulatory requirements in the face of extreme situations should be considered in some cases.

Finally, we turn to the issue of future retrospective studies of federal regulation. Especially at this very early stage of policy development on the issue, OMB can and should take a leadership role by formalizing a process for retrospective analysis to be conducted for important rules. Specifically, OMB should issue guidance to agencies on how to conduct retrospective analyses, akin to the guidance issued for pre-regulatory analyses. Further, OMB should work with regulatory agencies to identify priority regulations to be subjected to retrospective analysis, and to include a plan with data collection requirements in the final rules. OMB should also take steps to reduce the barriers to data collection for retrospective studies imposed by the Paperwork Reduction Act.

Regarding the actual conduct of future retrospective studies, multiple paths seem best. Regulatory agencies should clearly step up their game, by mandating more data collection as part of the regulatory process, by conducting more high quality internal and extramural analyses. At the same time, regulatory agencies are clearly limited institutionally in their ability to conduct rigorous retrospective studies of their own regulations. Thus, independent researchers have an important role to play, both in ensuring that the tough questions are asked and in advancing the use of new data and methods. The lack of funding for retrospective assessments, both inside and outside of government, is clearly a barrier to further progress. Funding for the enhanced pre-regulatory analyses recommended here is also an issue. As agencies face growing budget stringency, the challenges will likely increase. There is no simple answer here. Agencies have always faced challenges in balancing concerns about the improved effectiveness and efficiency of new regulations along with a host of other concerns. Various proposals have been made for an independent agency, perhaps an arm of Congress, to take on part of the role for retrospective analysis, although those approaches have problems of their own.²¹ In the end, a hybrid approach, with both government and independently supported studies, is probably the most practical way forward.

²¹S 1472, introduced by Senators Klobuchar and Collins in the 113th Congress but never enacted, would have created a division within the Congressional Budget Office to perform certain types of regulatory analysis, including retrospective analysis.

References

- Administrative Conference of the United States. 2014. “Retrospective Review of Agency Rules,” Recommendation number 2014-5, adopted December 4, 2014; published December 9, 2014. <https://www.acus.gov/research-projects/retrospective-review-agency-rules>.
- Aldy, Joseph E. 2014. “Learning from Experience: An Assessment of the Retrospective Reviews of Agency Rules and the Evidence for Improving the Design and Implementation of Regulatory Policy.” Report to the Administrative Conference of the United States. <http://www.acus.gov/sites/default/files/documents/Aldy%20Retro%20Review%20Draft%2011-17-2014.pdf>.
- Anderson, J. F., and T. Sherwood. 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.
- Bridgeland, John, and Peter Orszag. 2013. “Can Government Play Moneyball?” *Atlantic*, June 19, 2013.
- Carlson, C., Dallas Burtraw, Maureen Cropper, and Karen Palmer. 2000. “Sulfur Dioxide Control by Electric Utilities: What Are the Gains from Trade?” *Journal of Political Economy* 108 (6).
- Chan, H. Ron, B. Andrew Chupp, Maureen L. Cropper, and Nicholas Z. Muller. 2015. “The Net Benefits of the Acid Rain Program: What Can We Learn from the Grand Policy Experiment?” Discussion paper 15-25. Washington, DC: Resources for the Future.
- Dale, L., C. Antinori, M. McNeil, J. E. McMahon, and K. S. Fujita. 2009. “Retrospective Evaluation of Appliance Price Trends.” *Energy Policy* 37: 597–605.
- EPA (US Environmental Protection Agency). 2011. *The Benefits and Costs of the Clean Air Act, 1990–2020*. Washington, DC: EPA. http://www.epa.gov/air/sect812/feb11/fullreport_rev_a.pdf.
- Fraas, Arthur G., and Alexander Egorenkov. 2015a. “A Retrospective Study of EPA’s Air Toxics Program under the Revised Section 112 Requirements of the Clean Air Act.” Discussion paper 15-23. Washington, DC: Resources for the Future.

- . 2015b. “A Retrospective Study of EPA’s Rules Setting Best Available Technology Limits for Toxic Discharges to Water under the Clean Water Act.” Discussion paper 15-41. Washington, DC: Resources for the Future.
- Frankel, A. 2014. “Why Whirlpool’s Win in Moldy Washer Trial Vindicates Consumer Class Actions.” Reuters. <http://blogs.reuters.com/alison-frankel/2014/10/30/why-whirlpools-win-in-moldy-washer-trial-vindicates-consumer-class-actions/>
- Gray, Wayne B., and Ronald Shadbegian. 2015. “Multimedia Pollution Regulation and Environmental Performance: EPA’s Cluster Rule.” Discussion paper 15-26. Washington, DC: Resources for the Future.
- Greenstone, Michael. 2002. “The Impacts of Environmental Regulations on Industrial Activity: Evidence from the 1970 and 1977 Clean Air Act Amendments and the Census of Manufactures.” *Journal of Political Economy* 110 (6): 1175–1219.
- . 2009. “Toward a Culture of Persistent Regulatory Experimentation and Evaluation.” In *New Perspectives on Regulation*, edited by David Moss and John Cisternino, 111–26. Cambridge, MA: Tobin Project.
- Greenstone, Michael, John A. List, and Chad Syverson. 2011. “The Effects of Environmental Regulation on the Competitiveness of U.S. Manufacturing.” Paper no. CES-WP-11-03. Cambridge, MA: US Census Bureau Center for Economic Studies.
- Grosse, Scott D., [Norman J. Waitzman](#), [Patrick S. Romano](#), and [Joseph Mulinare](#). 2005. “Reevaluating the Benefits of Folic Acid Fortification in the United States: Economic Analysis, Regulation, and Public Health.” *American Journal of Public Health* 95 (11): 1917–22.
- Harrington, Winston. 2006. “Grading Estimates of the Benefits and Costs of Federal Regulation.” Discussion paper 06-39. Washington, DC: Resources for the Future.
- Harrington, Winston, and Anna Malinovskaya. 2015. “Expected vs. Actual Outcomes of Environmental Policies: The Clean Water State Revolving Fund.” Discussion paper. Washington, DC: Resources for the Future.
- Harrington, Winston, Richard Morgenstern, and Peter Nelson. 2000. “On the Accuracy of Regulatory Cost Estimates.” *Journal of Policy Analysis and Management* [19 \(2\)](#): 297–322.

- Henry, D. D., N. Z. Muller, and R. Mendelsohn. 2011. "The Social Cost of Trading: Measuring the Increased Damages from Sulfur Dioxide Trading in the United States." *Journal of Policy Analysis and Management* 30 (3): 598–612.
- Kahn, Mathew E., and Erin T. Mansur. 2013. "Do Local Energy Prices and Regulations Affect the Geographic Concentration of Employment?" *Journal of Public Economics* 101: 105–14.
- Lade, Gabriel E., C.-Y. Cynthia Lin, and Aaron Smith. 2015. "Ex Post Costs and Renewable Identification Number (RIN) Prices under the Renewable Fuel Standard." Discussion paper. Washington, DC: Resources for the Future.
- List, John A., Millimet, Daniel L., Fredriksson, Per G. & McHone, W. Warren 2003. "Effects of environmental regulations on manufacturing plant births: Evidence from a propensity score matching estimator." *Review of Economics and Statistics*, 85(4), 944–952.
- Lutter, Randall. 2012. "The Role of Retrospective Analysis and Review in Regulatory Policy." Working paper no. 12-14. Arlington, VA: Mercatus Center, George Mason University.
- . 2015. "How Effective Are Federal Food Safety Regulations? The Case of Eggs and *Salmonella* Enteritidis." Discussion paper 15-24. Washington, DC: Resources for the Future.
- Morgan, C., C. A. Pasurka, and R. J. Shadbegian. 2014. "Ex Ante and Ex Post Cost Estimates of the Cluster Rule and MACT II Rule." *Journal of Benefit-Cost Analysis* 5: 195–224.
- Nelson, Erik J., John C. Withey, Derric Pennington, and Joshua J. Lawler. 2015. "Identifying the Impacts of Critical Habitat Designation on Land Cover Change" Discussion paper 15-27. Washington, DC: Resources for the Future.
- OMB (US Office of Management and Budget). 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: Office of Information and Regulatory Affairs.
- Seong, Si Kyung, and John Mendeloff. 2004. "Assessing the Accuracy of OSHA's Projections of the Benefits of New Safety Standards." *American Journal of Industrial Medicine* [45 \(4\)](#): 313–28.

- Simpson, David. 2011. "[Do Regulators Overestimate the Costs of Regulation?](#)" Working paper series 201107. Washington, DC: National Center for Environmental Economics, US Environmental Protection Agency.
- Sunstein, Cass. 2011a. "Memorandum for the Heads of Executive Departments and Agencies: Facilitating Scientific Research by Streamlining the Paperwork Reduction Act Process." M-11-07. December 9, 2010.
<https://www.whitehouse.gov/sites/default/files/omb/memoranda/2011/m11-07.pdf>.
- . 2011b. "Memorandum for the Heads of Executive Departments and Agencies: Retrospective Analysis of Existing Significant Regulations." M-11-19. April 25, 2011.
<https://www.whitehouse.gov/sites/default/files/omb/memoranda/2011/m11-19.pdf>.
- Taylor, Margaret, C. Anna Spurlock, and Hung-Chia Yang. 2015. "Confronting Regulatory Cost and Quality Expectations: An Exploration of Technical Change in Minimum Efficiency Performance Standards." Discussion paper. Washington, DC: Resources for the Future.
- Thompson, Kimberly M., Maria Sequi-Gomez and John D. Graham. 2002. "Valuing Benefit and Cost Estimates: The Case of Airbag Regulation," *Risk Analysis*, 22:4, pp 803-811.
- Walker, W. Reed. 2013. "The Transitional Costs of Sectoral Reallocation: Evidence from the Clean Air Act and the Workforce." *Quarterly Journal of Economics*, 128(4), 1787–1836.

Appendix: Summaries of Case Studies

Multimedia Pollution Regulation and Environmental Performance: EPA's Cluster Rule

Wayne B. Gray and Ronald Shadbegian

In 1998, the US Environmental Protection Agency (EPA) promulgated its first integrated, multimedia (air and water) regulation, known as the Cluster Rule (CR), which aimed to reduce toxic releases from pulp and paper mills. By integrating the air and water regulations, EPA tried to reduce the overall regulatory burden on the affected plants. In this paper, Gray and Shadbegian compare EPA's ex ante expected reductions to an ex post assessment of those reductions. Using data from 1991 to 2009 for approximately 150 pulp and paper mills for both toxic and conventional pollutants, they find significant reductions in chloroform releases, nearly identical to the ex ante prediction of 99% reductions. The authors see some reductions in air toxics, smaller than the ex ante prediction and not always significant. Reductions in VOC emissions are similar in magnitude to the ex ante predictions for OLS models but smaller for fixed-effect models. No significant impact is found on PM₁₀ emissions. The authors draw conclusions for regulatory impact analyses and retrospective analyses, including the importance of carefully identifying expected compliance methods and the potential sensitivity of these analyses to the definition of the baseline.

Ex Post Costs and Renewable Identification Number (RIN) Prices under the Renewable Fuel Standard

Gabriel E. Lade, C.-Y. Cynthia Lin, and Aaron Smith

Lade et al. critically review EPA's assessment of the costs and benefits of the Renewable Fuel Standard (RFS2) as summarized in its regulatory impact analysis (RIA). They focus particularly on EPA's methods used to calculate the costs of the policy on the US fuel market, and they compare EPA's ex ante cost and benefit estimates to measures of ex post costs implied by the price of compliance credits under the policy. Overall, the authors find that the agency's assessment was inadequate. In spite of, or perhaps because of, the detailed and complex analysis underlying the RIA, EPA overlooked several fundamental factors. Lade et al. conclude by recommending a simplification of the analysis used in RIAs, as well as the use of "stress tests" in RIAs to ensure that programs like the RFS2 are designed in ways that can manage high compliance cost scenarios.

Identifying the Impacts of Critical Habitat Designation on Land Cover Change

Erik J. Nelson, John C. Withey, Derric Pennington, and Joshua J. Lawler

The US Endangered Species Act (ESA) regulates what landowners and land managers can do on lands occupied by listed species. The act does this in part through the designation of habitat areas considered critical to the recovery of listed species. Critics have argued that the designation of critical habitat (CH) has substantial economic impacts on landowners above and beyond the costs associated with listing in general. Here Nelson et al. examine the effects of CH designation on land cover change from 1992 to 2011 in areas subject to ESA regulations. They find that, on average, the rate of change in developed land (urban and residential) and agricultural land is not significantly affected by CH designation, although the results are not robust for high-value land. In addition, the authors' estimate of the effects of CH designation is not strongly correlated with the costs of CH as predicted by economic analyses published in the *Federal Register*. While CH designation, on average, does not affect the overall rates of land cover change, CH designation does appear to modify the impact of land cover change drivers. Generally, land prices had more impact (statistically) on land cover decisions within CH areas than in areas subject to ESA regulations but with no CH designation. Land cover decisions in these latter areas tended to be driven more by clustering and land availability concerns. These trends suggest that CH designation has increased landowner uncertainty and that conversion to developed and agricultural use in CH areas, on average, requires a return premium. Overall, however, this different reaction to land prices in and outside of CH areas has not been strong enough to differentiate the average rates of developed or agricultural land change in CH areas versus areas subject to ESA regulations but with no CH designation.

A Retrospective Study of EPA's Air Toxics Program under the Revised Section 112 Requirements of the Clean Air Act

Arthur G. Fraas and Alexander Egorenkov

Under the 1990 Clean Air Act Amendments, EPA was required to establish standards limiting air toxics emissions from industrial plants. This paper examines the effects of five of the largest-cost rules issued by EPA in the initial round of air toxics rulemaking over 1995 to 2000. Fraas and Egorenkov's estimates suggest that plants in the printing and publishing and pulp and paper industries realized important reductions in their air toxics emissions in the period between publication of the final rule and the effective date for compliance with the rule—although the reduction in air toxics emissions by pulp and paper mills fell short of EPA's ex ante projections. However, the authors' estimates suggest that plants in the other three industries—petroleum

refining, pharmaceutical, and wood furniture—achieved little or no additional reduction in air toxics emissions over the compliance period in response to EPA’s air toxics rules. Finally, the paper explores steps that EPA should take in setting up future retrospective analyses.

How Effective Are Federal Food Safety Regulations? The Case of Eggs and *Salmonella* Enteritidis

Randall Lutter

In 2009, the Food and Drug Administration (FDA) estimated that its egg rule would reduce illness from *Salmonella* Enteritidis (SE) by about 79,000 cases annually (37%), with a range from about 30,000 to 191,000 cases avoided. Lutter assesses the effectiveness of this rule, which requires farmers who sell eggs to adopt SE control measures, by comparing illness from SE with illness from other *Salmonella* serotypes, using a differences-in-differences approach. The data reject the hypothesis that the rule reduced illnesses by FDA’s best 2009 estimate but do not reject a hypothesis of no effect. The percentage of young broilers that test positive for SE has a modest effect on the incidence of human cases of salmonellosis caused by SE. Recent literature offers two other reasons to adjust FDA’s prospective 2009 calculations. One adjustment would follow the Centers for Disease Control’s use of a lower multiplier to infer the total number of (unobserved) cases of illness from those confirmed by positive lab tests. A second adjustment would lower the average cost of *Salmonella* cases by recognizing lower risk of severe sequelae. These adjustments and the new retrospective assessment of the effectiveness of the rule together suggest that the benefits of FDA’s egg rule may be a small fraction of the prospective estimate of benefits and less than the prospective estimate of costs. Lutter concludes with some policy recommendations to make food safety regulations more effective.

The Net Benefits of the Acid Rain Program: What Can We Learn from the Grand Policy Experiment?

H. Ron Chan, B. Andrew Chupp, Maureen L. Cropper, and Nicholas Z. Muller

This study quantifies the cost savings from the Acid Rain Program (ARP) compared with a command-and-control alternative and also examines the impact of trading under the ARP on health damages. To quantify cost savings, Chan et al. compare compliance costs for non-NSPS (New Source Performance Standards) coal-fired electricity generating units (EGUs) under the ARP with compliance costs under a uniform performance standard that achieves the same aggregate emissions. For 2002, the third year of Phase II of the program, the authors estimate annual cost savings of approximately \$250 million (1995\$). To examine the health effects of

trading, they compute the health damages associated with observed sulfur dioxide (SO₂) emissions from all units regulated under the ARP in 2002—approximately 10.2 million tons—and compare them with damages from a no-trade counterfactual in which each unit emits SO₂ at a rate equal to its allocation of permits for the year 2002, plus any drawdown of its allowance bank. Damages under the no-trade scenario are \$2.4 billion (2000\$) lower than under the ARP. This reflects the transfer of allowances from EGUs west of the Mississippi River to units in the eastern United States with higher exposed populations.

A Retrospective Study of EPA's Rules Setting Best Available Technology Limits for Toxic Discharges to Water under the Clean Water Act

Arthur G. Fraas and Alexander Egorenkov

Under the Clean Water Act, EPA is required to establish standards limiting water toxics discharges from industrial plants. In this paper, Fraas and Egorenkov examine the effect of EPA's 1998 Pulp and Paper Cluster Rule on the discharge of toxics to the waters of the United States. Their estimates suggest that mills covered by the water limits in the rule achieved little or no additional reduction in water toxics discharges in response to publication of the final rule compared with mills not subject to the water discharge limits of the rule. The authors also offer some recommendations for improving both ex ante and future retrospective analyses.

Expected vs. Actual Outcomes of Environmental Policies: The Clean Water State Revolving Fund

Winston Harrington and Anna Malinovskaya

This paper examines the performance of the Clean Water State Revolving Fund (CWSRF), a federally funded program to provide loans to local publicly owned treatment works (POTWs) in four states: Iowa, Indiana, Maryland, and Texas. Harrington and Malinovskaya find that between 2007 and 2014, the typical plant receiving a loan in these states substantially improved the quality of effluent discharges of biochemical oxygen demand (BOD) and organic nitrogen (N) in all four states compared with a sample of plants in the same states that did not receive loans. They also find, however, that plants receiving loans tended to have better performance than plants that did not receive loans in 2007, before the funds were distributed. While loans were effective in improving water quality, it appears that the plants receiving loans were not those most in need of improvement, but those already among the best. Thus, state authorities responsible for choosing which plants received loans favored those with a record of prior success rather than a record of current need. The authors recommend further review of

these findings by both federal and state authorities, with an eye on options to reform both the priority-setting process and the data collection procedures.

Confronting Regulatory Cost and Quality Expectations: An Exploration of Technical Change in Minimum Efficiency Performance Standards

Margaret Taylor, C. Anna Spurlock, and Hung-Chia Yang

Since at least the 1970s, it has been recognized that appliances and other energy-using products make significant contributions to the energy consumption of the residential building sector (currently estimated at 22% of US energy use). The basic concept underlying the minimum efficiency performance standards (MEPS) is that they remove from the market certain models of covered appliances and other energy-using products that do not meet specified energy efficiency thresholds. The covered products studied in this project are room air conditioners, refrigerator-freezers, dishwashers, clothes washers, and clothes dryers. These five products, which heavily saturate US households, are the full set of large domestic appliances that were subject to MEPS regulatory analyses conducted by DOE prior to 2012. The analysis draws on several sources of ex post data, including (i) extensive 2003–11 US point-of-sale data on appliance models, which the authors helped match to model energy use data (for all but clothes dryers) to construct a monthly panel of model-specific prices, quality characteristics, and market shares; (ii) datasets the authors constructed from independent third-party appliance testing and surveys that speak to the quality and reliability of all five products; and (iii) a dataset the authors constructed and coded that consists of the features identified in the product manuals of 1,109 clothes washer models sold in the United States in 2003–11 (these models represent 95% of the identifiable models in the point-of-sale data, which account for 29% of US units sold over that period).

Overall, the authors find that the ex ante regulatory cost projections significantly overestimated the majority of model price observations across all efficiency levels. By product, significantly overestimated observations were > 95% of room AC observations, 54–66% of refrigerator observations, 42–72% of dishwasher observations, and 50–81% of clothes washer observations (clothes dryers had to be modeled in a different way, but the results were similar). In addition, the authors show that within-model (i.e., with fixed effects) average prices declined for the refrigerators and clothes washers that US consumers bought in 2003–11.

The monthly sales-weighted average energy use of these products was better than the standard for almost all the data points of the study period, and in the one case for which the authors have data (clothes washers), they indicate that regulatory analyses underestimated how

enthusiastically US consumers would buy highly efficient products. The results show that these better-than-expected price and efficiency outcomes did not occur to the general detriment of the availability of products with high-quality performance attributes other than energy use, despite stakeholder concerns and the revealed concerns of the designers of the MEPS rulemaking process in 1996. Instead, in most cases, the statistically significant changes that occurred in third-party quality variables assessed across MEPS events were quality improvements. Similarly, the rate of significant repairs after five years of product ownership declined across the study period, according to third-party surveys, which points to increased product quality over the period of time that products have been regulated by MEPS.