Unclassified **ENV/WKP(2017)5**

Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development **01-Mar-2017**

English - Or. English

ENVIRONMENT DIRECTORATE

ENV/WKP(2017)5
Unclassified

RETROSPECTIVE EVALUATION OF CHEMICAL REGULATIONS - ENVIRONMENTAL WORKING PAPER No. 118

By Professor Susan E. Dudley, George Washington University

Authorised for publication by Simon Upton, Director, Environment Directorate.

JEL Classification: Q5, Q51, Q52, Q58, D78, D81.

Keywords: Regulation, chemical risk, regulatory impact analysis, evaluation, environmental impact analysis, benefit-cost analysis, retrospective review.

JT03409711

Complete document available on OLIS in its original format

English - Or. English

OECD ENVIRONMENT WORKING PAPERS

OECD Working Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed are those of the author(s).

OECD Working Papers describe preliminary results or research in progress by the author(s) and are published to stimulate discussion on a broad range of issues on which the OECD works.

This series is designed to make available to a wider readership selected studies on environmental issues prepared for use within the OECD. Authorship is usually collective, but principal author(s) are named. The papers are generally available only in their original language -English or French- with a summary in the other language.

Comments on OECD Working Papers are welcomed, and may be sent to:

OECD Environment Directorate, 2, rue André Pascal, 75775 PARIS CEDEX 16, France
or by e-mail to *env.contact@oecd.org*

OECD Environment Working Papers are published on
www.oecd.org/environment/workingpapers.htm

FOREWORD

This paper was prepared by Professor Susan E. Dudley of George Washington University for the OECD Workshop on *Socioeconomic Impact Assessment of Chemicals Management* in Helsinki, 6-8 July 2016.[1]

The workshop was organised in co-operation between the Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, Pesticides and Biotechnology (Joint Meeting) and Working Party on Integrating Environmental and Economic Policies (WPIEEP), and was hosted by the European Chemicals Agency, with funding contributions from the European Commission, the European Chemicals Agency and the American Chemistry Council.

The paper underwent revision and takes into account feedback received from Delegates during and after the workshop and comments received from the Joint Meeting and WPIEEP by written procedure.

The opinions expressed and the arguments employed are those of the author.

1. Director, the George Washington University Regulatory Studies Center and Distinguished Professor of Practice, the Trachtenberg School of Public Policy and Public Administration. Aspects of this draft have benefited from conversations with and input from Nils Axel Braathen, Louis A. Cox, Tomasz Kozluk, Eeva Leinala, Brian Mannix, Sofie Miller, Kathryn Newcomer, Paul O'Brien, Marcus Peacock, Daniel Perez and Estelle Raimondo.

TABLE OF CONTENTS

Figures

Boxes

ABSTRACT

OECD countries rely on regulatory tools to manage potential risks from exposure to targeted chemicals. Ex-ante regulatory impact assessment has a long tradition in many OECD countries, with established analytical steps and oversight as well as opportunities for public engagement to hold governments accountable for conducting analysis before regulations are issued. But ex-ante analyses necessarily depend on unverifiable assumptions and models of how the world would look absent the regulation, and how responses to regulatory requirements will alter those conditions. In essence, ex-ante analyses are hypotheses of the effects of regulatory actions.

Better ex-post regulatory evaluation would allow agencies and others to test those hypotheses against actual outcomes. This would not only inform decisions related to the cost-effectiveness of existing policy, but would provide feedback that would improve future ex-ante analyses and future policies. However, ex-post analysis also poses challenges, especially when it comes to chemical risks. Once a regulation is in place, it is not always obvious what the world would have looked like without it. Measuring opportunity costs is not easy, and measuring regulatory benefits is often harder. Furthermore, once a regulation is in place, neither regulators nor regulatory entities have strong incentives for examining its actual impact.

As a result of these methodological and incentive challenges, while ex-post evaluation has a long tradition in other areas (particularly in programmes financed through the fiscal budget), it has received little attention (and even fewer resources) in the regulatory arena, despite government guidelines requiring it.

This paper attempts to address these challenges to evaluating regulatory outcomes and learning from those evaluations. Drawing on experience in OECD countries, it reviews the practices used to understand the likely impacts of regulations aimed at reducing chemical risks both before and after they are issued. It examines why efforts at retrospective review have lagged behind prospective regulatory analysis, and offers recommendations for addressing methodological and incentive challenges to better evaluation.

JEL codes: Q5, Q51, Q52, Q58, D78, D81.

Keywords: Regulation, chemical risk, regulatory impact analysis, evaluation, environmental impact analysis, benefit-cost analysis, retrospective review.

RÉSUMÉ

Pour gérer les risques potentiels liés à l'exposition à des substances chimiques particulières, les pays de l'OCDE recourent à des outils réglementaires. L'évaluation *ex ante* de l'impact de la réglementation est pratiquée depuis longtemps dans beaucoup d'entre eux, où elle est adossée à des mécanismes éprouvés d'analyse et de supervision et donne aux citoyens des occasions de participer au processus, dans le but d'assurer que les pouvoirs publics conduisent des analyses avant l'adoption d'une réglementation. Cependant, les analyses *ex ante* reposent forcément sur des hypothèses invérifiables et des modèles qui servent à déterminer la situation qui prévaudrait en l'absence de la réglementation et en quoi elle sera modifiée par les réactions aux prescriptions réglementaires. Au fond, les analyses *ex ante* livrent des hypothèses quant aux effets des mesures réglementaires.

Une meilleure évaluation *ex post* de la réglementation permettrait aux organismes responsables et à d'autres acteurs de vérifier ces hypothèses à la lumière des résultats effectifs. Non seulement cela éclairerait les décisions qui touchent au rapport coût-efficacité des politiques existantes, mais cela permettrait aussi un retour d'information qui améliorerait les analyses *ex ante* et les politiques futures. Toutefois, l'analyse *ex post* pose également certains problèmes, en particulier dans le contexte des risques chimiques. Une fois qu'une réglementation a été mise en œuvre, il n'est pas toujours évident de déterminer quelle aurait été la situation en son absence. Il n'est guère aisé de mesurer les coûts d'opportunité, et il est souvent plus difficile encore de mesurer les avantages d'une réglementation. En outre, dès lors qu'une réglementation est en place, ni ceux qui en sont à l'origine ni ceux qui y sont soumis n'ont réellement d'incitation à en examiner l'impact.

En raison de ces problèmes méthodologiques et d'incitation, l'évaluation *ex post* s'est vu consacrer peu d'attention (et encore moins de ressources) dans le domaine de la réglementation, bien qu'elle soit officiellement obligatoire et depuis longtemps appliquée ailleurs (notamment pour les programmes financés sur fonds publics).

Le présent document se penche sur ces difficultés à surmonter pour évaluer les résultats des réglementations et tirer les enseignements de ces évaluations. Sur la base de l'expérience des pays de l'OCDE, il passe en revue les pratiques appliquées pour comprendre les incidences probables des réglementations destinées à réduire les risques chimiques avant et après leur mise en œuvre. Il étudie également pourquoi les efforts d'examen rétrospectif sont à la traîne des analyses *ex ante* de la réglementation et formule des recommandations pour surmonter les problèmes méthodologiques et d'incitation qui font obstacle à une meilleure évaluation.

Codes JEL : Q5, Q51, Q52, Q58, D78, D81

Mots-clés : réglementation, risques chimiques, analyse d'impact de la réglementation, évaluation, analyse de l'impact sur l'environnement, analyse coûts-avantages, examen rétrospectif.

RETROSPECTIVE EVALUATION OF CHEMICAL REGULATIONS

1. Introduction

OECD countries rely on regulatory tools to manage potential risks from exposure to targeted chemicals. While it is standard practice to analyse and estimate how proposed regulations might affect regulated entities, consumers, citizens, etc., *before* they are issued, regulators have devoted much less analysis to evaluating the impacts of their regulations once they are in effect. This paper draws on experience in OECD countries, primarily the United States, to examine the practices used to understand the likely impacts of regulations aimed at reducing chemical risks both before and after they are issued.

Ex-ante regulatory impact assessment has a long tradition in many OECD countries, with established analytical steps (OECD, 2008) and oversight as well as opportunities for public engagement to hold governments accountable for conducting analysis before regulations are issued (Dudley and Wegrich, 2015). But *ex-ante* analyses necessarily depend on unverifiable assumptions and models of how the world would look absent the regulation, and how responses to regulatory requirements will alter those conditions. In essence, *ex-ante* analyses are hypotheses of the effects of regulatory actions.

Better *ex-post* regulatory evaluation would allow agencies and others to test those hypotheses against actual outcomes.[2] This would not only inform decisions related to the cost-effectiveness of existing policy, but would provide feedback that would improve future *ex-ante* analyses and future policies. However, *ex-post* analysis also poses challenges, especially when it comes to chemical risks. Once a regulation is in place, it is not always obvious what the world would have looked like without it. (For example, would air emissions have increased directly with economic and population growth, or would technological change and citizen preferences have driven emissions lower (Adler, 2004)?) Measuring opportunity costs (what activities or innovations were foregone to achieve regulatory goals?) is not easy, and measuring regulatory benefits is often harder. Furthermore, once a regulation is in place, neither regulators nor regulatory entities have strong incentives for examining its actual impact.[3]

As a result of these methodological and incentive challenges, while *ex-post* evaluation has a long tradition in other areas (particularly in programmes financed through the fiscal budget) (Smismans, 2015), it has received little attention (and even fewer resources) in the regulatory arena, despite government guidelines requiring it (Aldy, 2014).[4]

This paper attempts to address these challenges to evaluating regulatory outcomes, and learning from those evaluations. It begins by laying out the importance of evaluation for an effective system to manage risks in Section 2, and then offers a brief overview of the regulatory framework in many OECD countries in Section 3. Section 4 examines why efforts at retrospective review have lagged behind prospective regulatory analysis. Section 5 offers recommendations for addressing methodological challenges to better evaluation and Section 6 offers suggestions for improving incentives. Section 7 presents brief conclusions.

2. Hypothesis testing is an essential element of the scientific method, and provides the feedback necessary to inform predictions. See http://physics.ucr.edu/~wudka/Physics7/Notes_www/node6.html.

3. Furthermore, regulators may not have the political mandate to alter regulations once implemented.

4. In his report to the Administrative Conference of the United States (ACUS), Joseph Aldy writes that federal regulatory agencies have a mixed record on *ex post* review, despite their "long track record of prospective analysis of proposed regulations that can address these questions."

2. Importance of evaluation

Feedback is an essential element of informed action and learning in all aspects of human existence. Individually, we learn and modify our behaviour based on experience. In business, successful firms continuously learn from experience, innovating to identify and satisfy unmet consumer needs, increase quality, improve operating efficiency, and better target potential customers (Madden, 2010). In the natural sciences, hypothesis testing is an essential feature of the experimental method, in which systematic observation, measurement, and empirical testing lead to revisions to the hypothesis and greater learning. As noted above and discussed more below, programme evaluation has long been an important element of government programs, where evaluation tools are applied to understand the effectiveness of programmes designed to meet public needs (Newcomer et al., 2015; EC, 2013).

In all these areas, evaluation and feedback comprise a "systems mindset," valuable not only for understanding the effect of past actions, but for improving future decisions.

> The benefit of a systems mindset is in developing faster and better solutions to problems. Whether problems are encountered in ecology, engineering, economics, or whatever the subject, a systems mindset helps to achieve better solutions. These are solutions that result in significant improvement to the performance of the overall system, in a cost-effective manner, while minimizing unintended adverse side effects (Madden, 2016).

According to the World Bank, "[i]n a context in which policy makers and civil society are demanding results and accountability from public programmes, impact evaluation can provide robust and credible evidence on performance and, crucially, on whether a particular programme achieved its desired outcomes (Gertler et al., 2011).

Unlike other government programmes that are reassessed each time their funds are appropriated, regulations, once created, tend to exist in perpetuity. As a result, regulations would benefit particularly from a systems mindset toward retrospective review.

In the regulatory sphere, the focus of retrospective review in the United States and elsewhere has been on identifying burdensome or underperforming rules that might be revised or rescinded. This can be important, but meaningful regulatory evaluation can offer more value than simply reducing burdens. A systems approach to retrospective review would focus attention on *ex-post* regulatory evaluation of outcomes as well as costs and can also help inform future *ex-ante* analysis (by testing hypotheses and assumptions regarding causation and outcomes), and improve future regulations.

The European Commission recently observed that, "[e]valuation is a key Smart Regulation tool, helping the Commission to assess whether EU actions are actually delivering the expected results and ultimately improving conditions for European citizens and businesses and contributing to the EU's global role" (EC, 2013).

In a report for the OECD, Coglianese reinforces this observation by noting that evaluations of regulatory policies should assess both substantive programme outcomes and process outcomes. He adds that, in order to conduct meaningful measurement of regulatory policy, "governments will need both indicators to measure relevant outcomes of concern and research designs to support inferences about the extent to which a regulation or regulatory policy under evaluation has actually caused any change in the measured outcomes" (Coglianese, 2012).

Thus, understanding *causal* relationships between regulatory policies and desired outcomes is a key element of retrospective evaluation. Section 5 reviews some practices and statistical tools that can help

address these key questions of causality and close the feedback loop to inform both current and future regulatory policies.

3. Overview of the regulatory process

As part of a set of recommendations for developing "a systemic governance framework that can deliver ongoing improvements to the quality of regulations," (OECD, 2012, p. 5) the OECD's Council Recommendation on Regulatory Policy and Governance advised that its members:

> Commit at the highest political level to an explicit whole-of-government policy for regulatory quality. The policy should have clear objectives and frameworks for implementation to ensure that, if regulation is used, the economic, social and environmental benefits justify the costs, the distributional effects are considered and the net benefits are maximised (OECD, 2012, recommendation 1).

Consistent with a systems approach to regulation discussed in the last section, key elements of this governance framework are regulatory impact analysis (RIA), risk assessment, and public engagement before new regulations are issued, and evaluation of regulatory outcomes after regulations are in place (OECD, 2012). This section summarises each of these elements.

3.1 Ex-ante regulatory analysis and public comment

Ex-ante RIA has a long tradition in most OECD countries (OECD, 2008).

> RIA is a process of systematically identifying and assessing the expected effects of regulatory proposals, using a consistent analytical method, such as benefit/cost analysis. RIA is a comparative process: it is based on determining the underlying regulatory objectives sought and identifying all the policy interventions that are capable of achieving them. These "feasible alternatives" must all be assessed, using the same method, to inform decision-makers about the effectiveness and efficiency of different options and enable the most effective and efficient options to be systematically chosen (OECD, 2008, p. 2).

The OECD noted that "…RIA's most important contribution to the quality of decisions is not the precision of the calculations used, but the action of analysing – questioning, understanding real-world impacts and exploring assumptions," (OECD, 2002, p. 47).

It emphasised the importance of including *ex-ante* RIA in the policy development process, recommending that members:

> Integrate Regulatory Impact Assessment (RIA) into the early stages of the policy process for the formulation of new regulatory proposals. Clearly identify policy goals, and evaluate if regulation is necessary and how it can be most effective and efficient in achieving those goals. Consider means other than regulation and identify the trade-offs of the different approaches analysed to identify the best approach (OECD, 2012, recommendation 4).

In the United States, the Office of Information and Regulatory Affairs (OIRA), in the Office of Management and Budget (OMB), explains that the purpose of the RIA is to provide the public with a "careful and transparent analysis" of the effects of regulatory actions. It should include "an assessment and (to the extent feasible) a quantification and monetisation of benefits and costs anticipated to result from the proposed action and from alternative regulatory actions." (OMB, 2011c).

> The purpose of the RIA is to inform agency decisions in advance of regulatory actions and to ensure that regulatory choices are made after appropriate consideration of the likely consequences. To the extent permitted by law, agencies should proceed only on the basis of a reasoned determination that the benefits justify the costs (recognizing that some benefits and costs are difficult to quantify). Regulatory analysis also has an important democratic function; it promotes accountability and transparency and is a central part of open government (OMB, 2011c).

3.2 Public engagement

The OECD's Council Recommendation on Regulatory Policy and Governance emphasised the importance of transparency and public engagement, encouraging Members to:

> Adhere to principles of open government, including transparency and participation in the regulatory process to ensure that regulation serves the public interest and is informed by the legitimate needs of those interested in and affected by regulation. This includes providing meaningful opportunities (including online) for the public to contribute to the process of preparing draft regulatory proposals and to the quality of the supporting analysis.[5] Governments should ensure that regulations are comprehensible and clear and that parties can easily understand their rights and obligations (OECD, 2012, recommendation 2).

The OECD's RIA handbook also emphasises the important relationship between regulatory analysis and open government, stating, "RIA should be integrated with a public consultation process, as this provides better information to underpin the analysis and gives affected parties the opportunity to identify and correct faulty assumptions and reasoning" (OECD, 2008, p. 3). While neither *ex-ante* nor *ex-post* regulatory analysis can be complete or accurately reflect real world conditions, a transparent presentation of the analysis and assumptions allows for meaningful public scrutiny and challenge that strengthen policies and actions.

3.3 Risk assessment

The OECD's Council Recommendation on Regulatory Policy and Governance directs Members to ensure that:

> As appropriate apply risk assessment, risk management, and risk communication strategies to the design and implementation of regulations to ensure that regulation is targeted and effective. Regulators should assess how regulations will be given effect and should design responsive implementation and enforcement strategies (OECD, 2012, recommendation 9).

Regulations intended to address public health and environmental risks pose unique challenges for both *ex-ante* and *ex-post* analysis, as they depend heavily on scientific assessments of what outcomes might materialise under different exposure scenarios (risk assessment), as well as analyses of the effectiveness of different approaches to manage those risks (risk management) (Dudley, 2015a). In the context of health, safety and environmental regulation, the National Research Council (NRC) of the U.S. National Academy of Sciences described the following conceptual framework in 1983:

5. In the United States, the Administrative Procedure Act of 1946 requires agencies consider public comments filed on the public record during the comment period as they develop their final regulation (Balla and Dudley, 2015) An analysis of three regulatory actions taken by the Federal Election Commission, Nuclear Regulatory Commission, and Department of the Treasury reveals that agency responsiveness occurs in proportion to the sophistication of the information contained in comments (Cuellar, 2005).

> Regulatory actions are based on two distinct elements, risk assessment… and risk management. Risk assessment is the use of the factual base to define the health effects of exposure of individuals or populations to hazardous materials and situations. Risk management is the process of weighing policy alternatives and selecting the most appropriate regulatory action, integrating the results of risk assessment with engineering data and with social, economic, and political concerns to reach a decision (NRC, 1983).

For risk assessment, analysts will never have complete information to predict outcomes with certainty, so they rely on what the NRC called "risk assessment policy" – assumptions, judgments and rules of thumb – to guide the use of scientific information in analyses that inform policy in the face of uncertainty.

> In each step [of the risk assessment process], a number of decision points (components) occur where risk to human health can only be inferred from the available evidence. Both scientific judgments and policy choices may be involved in selecting from among possible inferential bridges, and we have used the term *risk assessment policy* to differentiate those judgments and choices from the broader social and economic policy issues that are inherent in risk management decisions (NRC, 1983).

For regulatory policies aimed at reducing risks from exposure to chemicals, regulators will often develop concentration-response relationships that allow them to estimate outcomes following a causal logic model.[6] This necessarily involves judgment about which studies to consider and which to exclude, as well as assumptions about what models best fit the selected data and how to extrapolate between animals and humans or between observed and predicted exposures. Given the sometimes large uncertainty involved in these risk assessment policy judgments, *ex-post* evaluation of predictions could be particularly important for evaluating causal relationships, and refining assumptions and models for future analysis.

A recent report from the U.S. Institute of Medicine noted the uncertainty embedded in the scientific information that underlies these health risk estimates and at each stage of the assessment process:

> Uncertainties enter the health risk assessment process at every step and can be caused by the potential confounders in observational studies, by extrapolation from animal studies to human studies, by extrapolation from high to low dose exposures, by inter-individual variability, and by modeling the relationships between concentrations, human exposures, and human health responses and evaluating the effect of interventions or risk control options on public health risk (IOM, 2013).

The uncertainties inherent in these assessments can be significant, which, in turn can have large effects on *ex-ante* estimates of regulatory impacts. For example, relying on EPA estimates, OMB reports that the benefits of regulations that reduce exposure to fine particles ($PM_{2.5}$) to be approximately USD 156 billion to USD 772 billion (OMB, 2015).[7] This constitutes between 60% and 79% of OMB's estimates of the total benefits of all U.S. regulations (OMB, 2015).

OMB identifies "six key assumptions underpinning the particulate matter benefits estimates, and [its] analysis of these sources of uncertainty." Of particular importance in the context of *ex-post* evaluation is the key assumption that "inhalation of fine particles is causally associated with premature death," (OMB,

6. In some cases, regulatory decisions are binary – either a chemical is determined safe and allowed, or it is not allowed in commerce. In such cases, rather than developing concentration-response relationships, regulators will base decisions on evidence of threshold effects. As discussed below, such policies are more difficult to evaluate ex post.

7. Calculated from data presented in OMB, 2015, Table 1-1. Benefits are reported in 2010 dollars.

2015, p. 15). EPA assumes a *causal* relationship based on epidemiological evidence of an *association* between PM concentrations and mortality. *Ex-post* evaluation could be used to examine statistically whether the observed correlation in fact supports a causal relationship (*cum hoc non propter hoc*).

3.4 Ex-post regulatory evaluation

The OECD's Council Recommendation on Regulatory Policy and Governance advises Members to "[c]onduct systematic programme reviews of the stock of significant regulation against clearly defined policy goals, including consideration of costs and benefits, to ensure that regulations remain up to date, cost justified, cost effective and consistent, and deliver the intended policy objectives," (OECD, 2012, recommendation 5).

In the United States, various presidential and congressional directives spanning almost 40 years[8] have called on regulatory agencies to consider the effects of regulations once they are in place. However, there, as in other OECD countries, such retrospective analysis has received much less attention and fewer resources than ex-ante analysis (Dudley, 2013a).

Most recently, President Obama has reinforced both *ex-ante* and *ex-post* regulatory procedures through a series of executive orders. EO 13563 (2011) "reaffirms the principles, structures, and definitions governing contemporary regulatory review that were established in Executive Order 12866," (EO 13563, 2011). Section 6 of that Order also called for "Retrospective Analyses of Existing Rules," encouraging agencies to "consider how best to promote retrospective analysis of rules that may be outmoded, ineffective, insufficient, or excessively burdensome, and to modify, streamline, expand, or repeal them in accordance with what has been learned."[9]

EO 13610 (2012) recognised efforts agencies had made at retrospective review but stated that "further steps should be taken, consistent with law, agency resources, and regulatory priorities, to promote public participation in retrospective review, to modernise our regulatory system, and to institutionalise regular assessment of significant regulations," (EO 13610, 2012).

Following these two Executive Orders, OIRA Administrator Cass Sunstein issued guidance to the heads of executive branch agencies and independent regulatory commissions emphasising the importance of "maintaining a consistent culture of retrospective review and analysis" in government (OMB, 2011a). The guidance instructs agencies to use the principles established for *ex-ante* analysis (EO 13563 §1 – 5) to orient their thinking during the process of retrospective analysis and specifies elements their review plans should include and timelines for sharing them with the public.

4. Why efforts at retrospective review have not been more successful

Despite directives to conduct *ex-post* regulatory evaluation, procedures for doing so have not been institutionalised to the extent that *ex-ante* regulatory impact analysis has been. The reasons for this are twofold: poor incentives and methodological challenges. These are discussed briefly below. Possible remedies for both these causes are discussed in the next two sections.

8. In 1978, President Carter directed agencies to "periodically review their existing regulations to determine whether they are achieving … policy goals" (EO 12044, 1978).

9. While, like most executive orders, EO 13563 is not binding on "independent regulatory agencies" such as the Securities and Exchange Commission or the Federal Communications Commission, EO 13579 (14 July 2011) directed independent regulatory agencies, "to the extent permitted by law, [to] comply with [the] provisions of EO 13563, particularly the retrospective review provisions."

4.1 Incentives for retrospective review are lacking

In the United States, EU, and other OECD countries, impact analysis is required before new regulations can be issued. For example, in the United States, OIRA serves a gatekeeper role, which compels regulating agencies to present analysis consistent with executive order requirements if they wish to issue new rules. On the other hand, once a regulation is issued, the consequence of not conducting *ex-post* analysis is less problematic from the agency's perspective in that the regulation will remain on the books (Miller and Dudley, 2016).

Compounding this asymmetric incentive structure is that regulated parties may be more motivated to prevent a potentially burdensome regulation from being implemented than to advocate for regulation to be removed. Companies find regulatory uncertainty challenging for long-term planning. Furthermore, once a regulation is in place, it confers a competitive advantage on some parties, especially those who have already invested in compliance (Koźluk, 2014). For example, once a company has introduced a substitute for a restricted chemical, or revamped its manufacturing processes to comply with regulation, it may stand to lose market share to new entrants if those restrictions are removed. Incumbents and other beneficiaries are thus less likely to support evaluation that may lead to changes or repeal.[10]

4.2 Retrospective review is challenging methodologically

Retrospective review of regulations is also challenging methodologically, especially when policies were not designed to facilitate *ex-post* measurement. While observers know the state of the world with the regulation, they do not know what the world would have looked like without it (the counterfactual). Regulations designed to manage risks from chemicals are complex interventions (developed through interactions between different levels of government) nested in complex adaptive systems (usually involving government mandates that impose requirements on private actors to achieve public goals, rather than direct government spending). Understanding the causal linkages between the regulation and the actual health and economic outcomes and designing evaluation to assess those impacts should take into account the sources and consequences of such complexity.

Given that agents (individual, corporations, organizations) adapt to their immediate environment, these adaptive responses may lead to consequences not anticipated by examining one variable in isolation. For example, regulations that set more stringent emission limits for new facilities than for existing plants unintentionally provide incentives for regulated entities to extend the life of older, dirtier plants. Thus, while all sources of emissions comply with environmental standards, the overall level of pollution might not achieve *ex-ante* predictions because fewer new cleaner plants are built while the operating life of existing sources is extended. (Coysh, Johstone and Kozluk, 2017).

Regulations seek to change behaviours through enforcement of standards, deterrence, modification of incentives, etc., but, as the air emissions example above illustrates, how regulated entities respond to the regulation affects its outcomes and may yield unintended results. A classic illustration of how behavioural response to regulation moderated the intended outcomes is the U.S National Highway Traffic Safety Administration's retrospective analysis of its 1983 rule requiring centre high-mounted stop lamps on new cars. The intention behind the higher brake lights was to give following drivers more time to react when drivers in front of them put on their brakes, allowing them to stop sooner and thereby avoid or reduce the consequences of crashes (OMB, 1998). Though NHTSA's *ex-ante* analysis of the effects of the rule was of a very high quality (including random trials to measure accident rates with vehicles with the higher brake lights installed) (Lutter, 2012) its *ex-post* evaluation revealed that prospective analysis had overstated the

10. US agencies suggest that "environmental advocacy groups have constituted a weak barrier to evaluation" (Hart, 2016).

effectiveness of the rule at reducing rear end collisions by a factor of 7. Apparently, once the majority of vehicles on the road were installed with the central brake light, drivers responded by driving more closely to the car in front, reducing the effectiveness of the regulation (Peltzman, 2004). Note that, though not as effective as anticipated *ex ante*, NHTSA's evaluation still showed net benefits from the requirement. The value of NHTSA's careful retrospective review is that it revealed some of the weaknesses in the initial analysis, which is useful for the development of future policies.

Understanding the causal linkages between the regulatory action and the intended outcome can be difficult. Due to the potentially non-linear and non-proportional nature of regulatory impact trajectories, careful study designs and advanced statistical methods (as discussed in the next section) may be needed to understand the interaction of various causal factors. Indeed, the overall regulatory framework can influence the effect of individual regulations, and the intersection of multiple regulations can affect both outputs and outcomes. Consequently, effects are seldom purely additive; intersection and emergence need to be taken into account.

An illustration of this is efforts to reduce tropospheric ozone, which is correlated with increased asthma symptoms and other respiratory effects. According to EPA, "ozone is not emitted directly into the air, but is created by chemical reactions between oxides of nitrogen (NO_x) and volatile organic compounds (VOC) in the presence of sunlight," (EPA, 2016a). Since VOCs are "emitted by a wide array of products numbering in the thousands," (EPA, 2016b) efforts to reduce ozone focus on reducing NO_x emissions from power plants, industrial facilities, and vehicles. However, "recent modeling indicates that NO_x emissions from urban sources in some locations have negative marginal damages due to unappreciated complexities in atmospheric chemistry" (Fraas and Lutter, 2011). In other words, reductions in NO_x emissions can, under certain conditions, actually increase ozone levels.

The causal attribution of a particular effect to chemical regulations is particularly challenging for several reasons. The linkages between the intervention, effects on concentrations and exposures, and the ultimate endpoint can be complex. (See Figure 1 below) Often, the health outcomes in question (e.g. reductions in premature mortality or cancer rates) are affected by numerous other factors, and the effects of individual regulations may be too small to measure. *Ex-ante* exposure-response models may be based on extrapolations of animal studies or associations observed from epidemiological data. Given the difficulty of distinguishing the effect of the regulatory intervention from countless other factors that may influence predicted outcomes, testing the extent to which these hypotheses are correct is challenging for an evaluator and requires established statistical techniques, as discussed in Section 5.5 below.

Even when the intended outcome is more uniquely associated with the regulatory intervention, there are often a variety of factors that may influence it (Stern et al., 2012). For example, to what extent are improvements in workplace safety the result of regulations as opposed to other factors, such as technological change, more informed workers and employers, the legal environment, etc.?

Attributing outcomes to regulatory actions is particularly hard since there are so many other potential causes of observed changes in conditions, or in behaviours of individuals, firms, communities, etc. Changes in the environment, in the units themselves, inadequate, inconsistent or incomplete implementation of the regulation, or simply faulty methods for monitoring changes over time can make it difficult to make the case for causation. In addition, there may be complex and recursive relationships among desired outcomes. Outcomes are typically measured to assess achievement of goals set by individual regulations, yet conditions such as health status are affected by many other aspects of the intended beneficiary's life, e.g. their housing, access to healthy food options, access to recreation, etc. as well as personal behaviours (seat belt use, smoking habits, etc.) A robust evaluation should attempt to attribute degrees of causation to these mediating and confounding factors.

5. Recommendations for improving quantification of regulatory effects

In its Report to Congress on the Benefits and Costs of Federal Regulations, OMB states that retrospective analysis can serve as an important mechanism to correct the flaws of *ex-ante* analyses. According to that report, the result of systematic retrospective review of regulations

> should be a greatly improved understanding of the accuracy of prospective analyses, as well as corrections to rules as a result of *ex-post* evaluations. A large priority is the development of methods (perhaps including not merely before-and-after accounts but also randomized trials, to the extent feasible and consistent with law) to obtain a clear sense of the effects of rules. *In addition, and importantly, rules should be written and designed, in advance, so as to facilitate retrospective analysis of their effects, including consideration of the data that will be needed for future evaluation of the rules' ex-post costs and benefits* (OMB, 2014, emphasis in original).

This section offers several recommendations for methodologies to improve evaluation of regulatory outcomes. It discusses 1) elements that regulators should include in new regulations to improve their ability to evaluate outcomes after implementation, 2) the form that regulations might take to facilitate experimentation and learning, 3) logic models, as used by programme evaluators to lay out a theory of change, 4) possible scope for evaluation, and 5) research designs that can evaluate causality.

5.1 Plan for retrospective review at the outset

As OMB states, agencies should plan for retrospective review when a regulation is first being developed: clearly identifying the problem the regulation is intended to address, laying out the causal linkages between the regulatory intervention and desired outcome, and establishing a framework for empirical testing of assumptions and hypothesized outcomes.

In their World Bank report, Gertler et al. conclude that the appropriate methods for conducting programme evaluation, or retrospective review, should be identified "at the outset of a program, through the design of prospective impact evaluations that are built into the project's implementation" (Gertler et al., 2011). This allows evaluators to better fit their evaluation methods to the programme being reviewed, and to plan for review itself through the design and implementation of the programme (or regulation).

Aldy also reinforces the importance of planning for retrospective review at the beginning of the rulemaking process, and adds that governments "should describe the methods that they intend to employ to evaluate the efficacy of and impacts caused by the regulation, using data-driven experimental or quasi-experimental designs where appropriate," (Aldy, 2014). This is consistent with the guidelines in both OMB's Report to Congress and OIRA Administrator Sunstein's memos on implementation of retrospective review, indicating that the US government recognises the importance of designing rules at the outset to facilitate measurement of their outcomes. Despite these guidelines, US agencies are not successfully incorporating such design into their proposed rules (Miller, 2015).

The European Commission's 2013 Communication, "Strengthening the foundations of Smart Regulation – improving evaluation" committed to "'evaluate first' and systematically ensure that all significant proposals for a revision are backed up by a robust evaluation of the performance of existing EU action" (EC, 2013 p. 3).

> The "evaluate first" principle locates evaluation firmly within the policy cycle. New EU intervention can only be taken after an assessment of past action has been made. While evaluation of expenditure policy has long been linked to the seven year financial programme cycle, the 2007 Communication sets the Commission on track to fit evaluation of all its action into its strategic planning and programming cycle. Most importantly, *ex-post* evaluation should

feed back into the EU system of *ex-ante* impact assessments, which has been solidly established since 2003 (Smismans, 2015).

When a rule targeting chemical risks is issued, it should include a problem statement, criteria for measuring success, a logic model that links the rule to the desired outcome, a transparent risk assessment, and a timeline for collecting data necessary for evaluation. Each of these elements is described below.

5.1.1 Identify the problem to be solved

The *ex-ante* impact analysis should clearly articulate the core problem that requires regulatory action. Generally, this should be a description of the "failures of private markets or public institutions that warrant new agency action," (EO 12866, 1993) which can include goals such as equity or distributional concerns.

The concept of "market failure" is an important one in regulation. The market exchange of goods and services between willing buyers and sellers efficiently relies on price signals to allocate scarce resources to their highest and best use. Absent a clearly identified market failure, regulation and other forms of government intervention can disrupt those market forces, and lead to misallocation of resources (OECD, 2008, p. 7).[11] Thus, targeting a fundamental problem rather than relying on anecdotes to support regulation is important. As the OECD observes, "[a] basic aspect of RIA is that it must be conducted with this "whole of society" view in mind, rather than paying undue attention to impacts on individual groups that may be lobbying for regulation (OECD, 2008, p. 6).

Generally, the fundamental argument for a government role in managing chemical risks is either externalities or asymmetric information. Environmental problems are classic cases of externalities—when one party's actions impose uncompensated benefits or costs on another party (OECD, 2008). In some cases, such as when workers are exposed to occupational risks, regulators may be concerned that, without regulation, workers do not have adequate information on the possible health effects of their exposures (OECD, 2008, p. 6).[12] An analysis cannot credibly claim net benefits from regulation until it first identifies the market failure it is designed to correct (OECD, 2008, p. 6).[13]

5.1.2 Identify criteria by which success will be measured, and plan for obtaining relevant information

In order to measure the success of any rule following implementation, it is necessary for the regulator to clearly define what constitutes a "success." Any stated metrics of success should be linked to the problems identified, and ideally should measure the actual outcomes of interest (although sometimes proxy measures may be appropriate).

For example, many risk management regulations are focused on health outcomes, and the projected benefits might include changes (over baseline estimates) in morbidity, premature mortality, cancer rates,

11. As the OECD observes: "Identifying one or more significant sources of market failure provides evidence of a potential case for regulation. However, regulation frequently fails to address the identified market failure effectively and efficiently. There is a risk that market failure may be supplanted (or compounded) by regulatory failure."

12. The OECD's RIA Handbook observes: "If market failure exists, there may be a good argument to regulate to improve the availability of information. However, it is still necessary to show that the regulation can address the market failure effectively without creating other, substantial costs."

13. A regulation may be justified on other grounds, such as equity, but even there identifying the fundamental market cause of the problem can help avoid unintended consequences. According to the OECD (2008), in these cases, "RIA helps to make transparent who benefits from regulation and who pays the costs, so that these value judgements can be made more reliably."

etc. A crucial component of effective programme evaluation is access to relevant data, so at the time regulation is developed it will be important to determine which measures to use, how to calculate changes in these measures over time, and to commit to gathering data necessary to support those measures.

In some cases, regulators will take action to prevent a product from entering commerce or to ban an existing product or component from commerce. In these cases, retrospective evaluation will be constrained due to a lack of counterfactual (i.e., what were the benefits and opportunity costs associated with the absence of that product).

5.1.3 Lay out a causal logic model

Impact assessments conducted before regulations are issued generally present a model of how the intervention is expected to achieve desired outcomes. Coglianese observes that, at its most basic level, regulation is designed to work according to three main steps:

1. *Regulation* is implemented, which leads to changes in

2. *The behaviour* of individuals or entities targeted or affected by regulation, which ultimately leads to changes in

3. *Outcomes*, such as amelioration in an underlying problem or other (hopefully positive) changes in conditions in the world (Coglianese, 2012).

As regulators commit to measuring the effects of their rules, they should be aware of mediating factors that may accomplish goals in the absence of the rule, or undermine achievement of the stated metrics. Understanding the counterfactual and determining linkages between the rule and the measured outcomes is necessary to evaluate, *ex post*, whether the policy itself resulted in the desired outcomes, rather than other factors beyond the agencies' control.

According to Aldy's analysis of US practices, "Most economically significant regulations, while subject to rigorous *ex-ante* analysis, are not designed to produce the data and enable causal inference of the impacts of the regulation in practice" (Aldy, 2014, p.9). Designing regulations to produce this information can provide important information about whether the outcomes being sought are caused by the regulation in question, rather than other factors. (See section 5.2 below.)

To facilitate *ex-post* evaluation of causal linkages, *ex-ante* impact analyses should be explicit about the assumptions and models relied on to develop estimates of health or environmental outcomes. For example, what epidemiological or animal assays formed the basis of the concentration-response relationships? What assumptions were used to fill gaps in scientific evidence? Being explicit about these risk assessment policy decisions is also important when regulators must make binary decisions on whether to allow a chemical or not based on threshold analysis, rather than estimating a concentration-response function that would predict a continuum of effects depending on dose and exposure, such as REACH authorisations or restrictions. To make risk assessment policy choices more transparent, regulatory scientists could calculate and present multiple *ex-ante* risk estimates based on a variety of scientifically plausible data sets, endpoints, models, etc., instead of presenting a single exposure-response model (Gray and Cohen, 2012). Such transparency would help evaluators disaggregate the reasons why outcomes might differ from predictions.

5.1.4 Anticipate a time frame for analysis

Many *ex-ante* analyses indicate the timeframe over which the costs and benefits of a particular rule are expected to materialise. For policies aimed at managing health and environmental risks, analyses often

use long time horizons, such as 30-years, to tally benefits and costs. However, many of the costs and benefits of these rules will become tangible in smaller time increments, such as five years after implementation for standards with upfront capital requirements or two years for standards intended to result in immediate, next-year outcomes (such as safety standards for fresh produce).

Regulators should make clear when the outcomes they value will begin to become apparent, and plan accordingly to measure those outcomes by inserting the timeframe for review when the regulations are first issued.

5.2 Design regulations to facilitate natural experimentation and learning

Experts argue persuasively that understanding causal connections between regulatory treatments and ultimate outcomes requires research designs that "emulate conditions in a laboratory experiment in order to pinpoint those effects caused by the rule or policy under study," (Coglianese, 2012, p. 7). As Greenstone observes:

> The key to a system of regulatory experimentation and evaluation is a process that accumulates credible evidence on regulations' costs and benefits. Such a system demands that regulations be structured at the outset so that they can be evaluated and that evaluations be fully funded (Greenstone, 2009, p.123).

Coglianese, in a report for the OECD, lays out a hierarchy of designs for measuring public policies:

> The best research design will be the *randomized experiment*, which could be used much more extensively than it is at present in measuring progress about many issues of public policy. When randomized experiments are not feasible, evaluations can be based on *observational studies* which use a variety of statistical methods to isolate the effects that can be causally attributed to the policy under evaluation. If quantitative observational studies are not feasible, evaluators can rely on *qualitative studies*, such as matched case studies, that seek to "control" for other influences as much as possible (Coglianese, 2012, p. 7).

While laboratory experiments are not possible for regulations aimed at reducing health and environmental risks, designing regulations from the outset in ways that allow variation in compliance (such as different compliance schedules in different regions) is essential if evaluators are to go beyond observing mere associations and gather data necessary to test hypotheses of the relationship between regulatory actions, hazards, and risks. Such quasi-experimental methods need not be comprehensive to be valuable at informing not only the intervention being evaluated but assumptions underlying other interventions. Many behaviours may be transferrable to different contexts, so insights from one evaluation that is conditional on certain factors may be relevant to another.

Along these lines, the U.S. OMB recommends that agencies may, for instance, implement a regulation on a trial basis or establish separate thresholds for compliance for different firms or geographic locations. Through such design an agency "could learn about the effects of its action from what emerges. Pilot projects of various sorts have informed the regulatory process, and they could be used more often for this purpose" (OMB, 2011b).

Depending on the circumstances, regulators could conduct pilot studies or "deploy different regulations where empirical evaluations of such differences will help resolve disputed issues of regulatory policy," (McGinnis, 2012). Greenstone suggests testing regulatory approaches on a small scale before applying them to a large population to provide a form of randomised control trial. By randomly assigning treatment and control groups, governments could experiment with different forms of the regulation (Greenstone, 2009, p. 118). Alternatively, allowing smaller levels of government to implement regulations

differently can provide valuable information before they are tried on a larger scale (Lutter, 2012).[14] Experimentation and competition among jurisdictions can be a powerful force for improving environmental outcomes and developing a practical knowledge of what works. For example, the United States banned the artificial sweetener, cyclamate, in 1969 based on a study that suggested it caused bladder cancer in rats. Other countries (EC, 2002) (including neighbouring Canada) did not ban the product, and subsequent research indicates that cyclamate is not a carcinogen in mice or rats (New World Encyclopedia, 2013).

5.3 Lay out the theory of change

As noted above, in terms of framing evaluation questions or identifying what to measure, developing the theory of change or logic underlying a policy or regulatory objective typically is a sound starting place (Newcomer et al., 2015). This is implicit, if not explicit, in most RIAs estimating the health or environmental outcomes that will result from particular interventions. Gertler et al. emphasise the importance of being cognizant of causal pathways, rather than simply assuming that government programmes result in outcomes:

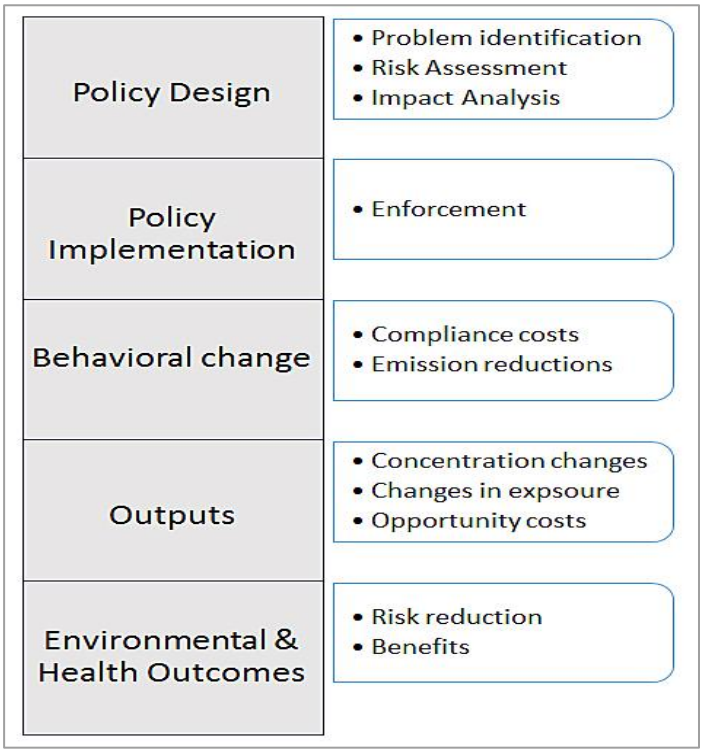
> Finally, we strongly encourage policy makers and program managers to consider impact evaluations in a logical framework that clearly sets out the causal pathways by which a program works to produce outputs and influence final outcomes, and to combine impact evaluations with monitoring and complementary evaluation approaches to gain a full picture of performance (Gertler et al., 2011).

As Coglianese observes, "the world consists of numerous and complex causal relationships that contribute to social and economic problems; regulations take aim at one or more steps on the causal pathways leading to those problems," (Coglianese, 2012, p. 9). Modelling programmatic elements, that is, the causal mechanisms that are expected to produce desired results, and the contextual factors that may mediate or moderate the risk management processes is not only helpful for designing evaluation studies, but in determining what to measure on a routine basis, and then in analysing and interpreting what the performance data mean.

Figure 1 presents a simple diagram illustrating a typical process for designing new regulation aimed at achieving environmental and health outcomes. Regulatory agencies in the United States and other OECD countries often rely, at least implicitly, on a framework like this to estimate likely benefits of actions to reduce exposure to targeted chemicals. This *ex-ante* model can provide a framework for *ex-post* evaluation, which will also need to take into account any mediating factors that may not have been anticipated that can affect outcomes, or determinations of whether outcomes can plausibly be attributed to the regulatory action.

14. Lutter (2012) suggests, "One idea might be simply to authorise special programs intended to collect information about regulatory effectiveness. A stronger form of the same idea would authorise an agency to issue regulations only after it completed a pilot study of their likely effect."

Figure 1. Illustrative simple causal model



Context matters in understanding when and how regulations work (Parker and Kirkpatrick, 2012). Since regulatory impacts are the product of both the rule and its implementation, they are bound to differ according to the context surrounding implementation. For example, *ex-ante* analyses may assume complete implementation and 100% compliance; however, with any intergovernmental programme, variations in implementation at different levels of government (e.g. federal, state and local governments in the United States, member states in the EU, etc.) may affect how a regulation actually operates.

Further, unlike on-budget programmes, regulations command or prohibit private behaviour in order to achieve social goals. Regulations generally attempt to change outcomes (e.g., reduce anthropogenic causes of climate change) not through direct spending or programmatic action by government entities, but by imposing obligations on private parties (e.g., on vehicle manufacturers to manufacture more fuel-efficient vehicles) in order to change consumer purchasing decisions. Not only will the regulated party's choice of compliance (by developing lighter-weight vehicles, alternative fuel vehicles, etc.) affect outcomes, but so will consumer response to the altered choices available. For example, both theory and empirical evidence suggest that people who own more fuel-efficient vehicles drive them more miles because the cost per mile of driving is reduced (Linn, 2013). The additional driving offsets some of the environmental benefits otherwise gained from a more fuel-efficient vehicle, but arguably offers consumers other benefits. Such a "rebound effect" should be considered as a mediating factor in the causal model (VTPI, 2014).

External factors beyond the regulators' control may also affect regulatory outcomes. Evaluators will need to include in their statistical analyses relevant factors that might correlate with the outcomes of concern. Continuing the fuel economy illustration, gasoline prices also influence both the type of vehicle consumers buy (demand for fuel efficient vehicles is positively correlated with gasoline prices) as well as their driving habits.

Coglianese (2012) points out that other regulations may also affect implementation and enforcement, and that "targeted organisations and individuals experience a variety of non-regulatory factors, such as economic and community or social pressures, that will affect their behaviour as well." Engaging pertinent stakeholders in conversations on how regulations are intended to operate and to identify factors outside of the regulator's control that may hinder achievement of desired outcomes can be extremely helpful in modelling how a regulation is likely to work in practice.

5.4 Determine the proper scope of evaluation

The proper scope of an evaluation will depend on its purpose and the availability of quality data (Coglianese, 2012). A few considerations are offered here.

5.4.1 Should evaluation be regulation-specific or cover a broader range of policies?

One relevant question will be whether the evaluation should focus on an individual regulation or a group of rules that target at the same health outcomes. Lutter (2012) encourages the review of programmes:

> The focus on retrospective analysis and review of regulations, as opposed to regulatory programs more broadly, may be too narrow. In particular, this focus may have led to retrospective reviews that do not appear to use information developed from broader retrospective analyses of regulatory programs (Lutter, 2012).

Ex-ante regulatory analysis is usually conducted on rule-by-rule basis, but each regulation is often part of a bundle of actions that seek to affect a single outcome of interest. It is thus often challenging to disentangle a single regulatory intervention from others. For example, an RIA might estimate the health benefits of removing one hazard from the environment (e.g. a pesticide) without considering interactions with other policies aimed at the same outcome or the health effects of substitutes for the banned chemical.

5.4.2 Should costs as well as benefits be evaluated?

So far, this paper has discussed evaluating health and environmental outcomes and how they compare to predictions. The extent to which the *costs* of implementing the risk management policy are more or less than projected may also be relevant, particularly if the evaluation aims to examine whether the policy yielded net benefits. As Coglianese notes:

> few if any regulatory problems call for solutions to be made at any cost. Thus, in addition to a regulation's impact on its ultimate outcome of concern, it could ultimately lead to other outcomes as well. Depending on the specific regulatory problem, these … might include the costs of regulation, impacts on technological innovation, equity, and so forth (Coglianese, 2012, p. 12).

Measuring costs presents its own challenges, however, since some important opportunity costs may be unseen. For example, what are the costs associated with security measures that infringe upon airline travellers' privacy? What are the costs of pre-market approval requirements that prevent a promising, but yet unknown, product from reaching consumers? Simply measuring administrative costs or compliance costs does not capture the full social costs of regulations.

A related question is whether the evaluation should examine the *valuation* of desired health or environmental outcomes (the monetised benefits). Some of these problems have been considered in the context of *ex-ante* analysis, but they still present challenges for evaluators.

Box 1. EPA criteria for ex-post evaluation

In its 2011 "Plan for Improving Retrospective Review", the U.S. EPA laid out the following "Criteria for Regulatory Review," which offer useful considerations for determining scope (EPA, 2011a).

- Benefits justify costs
  - Now that the regulation has been in effect for some time, do the benefits of the regulation still justify its costs?
- Least burden
  - Does the regulation impose requirements on entities that are also subject to requirements under another EPA regulation? If so, what is the cumulative burden and cost of the requirements imposed on the regulated entities?
  - Does the regulation impose paperwork activities (reporting, record-keeping, or third party notifications) that could benefit from online reporting or electronic recordkeeping?
  - If this regulation has a large impact on small businesses, could it feasibly be changed to reduce the impact while maintaining environmental protection?
  - Do feasible alternatives to this regulation exist that could reduce this regulation's burden on state, local, and/or tribal governments without compromising environmental protection?
- Net benefits
  - Is it feasible to alter the regulation in such a way as to achieve greater cost effectiveness while still achieving the intended environmental results?
- Performance objectives
  - Does the regulation have complicated or time-consuming requirements, and are there feasible alternative compliance tools that could relieve burden while maintaining environmental protection?
  - Could this regulation be feasibly modified to better partner with other federal agencies, state, local, and/or tribal governments?
- Alternatives to direct regulation
  - Could this regulation feasibly be modified so as to invite public/private partnerships while ensuring that environmental objectives are still met?
  - Does a feasible non-regulatory alternative exist to replace some or all of this regulation's requirements while ensuring that environmental objectives are still met?
- Quantified benefits and costs/qualitative values
  - Since being finalized, has this regulation lessened or exacerbated existing impacts or created new impacts on vulnerable populations such as low-income or minority populations, children, or the elderly?
  - Are there feasible changes that could be made to this regulation to better protect vulnerable populations?
- Open exchange of information
  - Could this regulation feasibly be modified to make data that is collected more accessible?
  - Did the regulatory review consider the perspectives of all stakeholders?
- Co-ordination, simplification, and harmonization across agencies
  - If this regulation requires coordination with other EPA regulations, could it be better harmonized than it is now?

- If this regulation requires coordination with the regulations of other federal or state agencies, could it be better harmonized with those regulations than it is now?

- Innovation
  - Are there feasible changes that could be made to the regulation to promote economic or job growth without compromising environmental protection?
  - Could a feasible alteration be made to the regulation to spur new markets, technologies, or jobs?
  - Have new or less costly methods, technologies, and/or innovative techniques emerged since this regulation was finalized that would allow regulated entities to achieve the intended environmental results more effectively and/or efficiently?
- Flexibility
  - Could this regulation include greater flexibilities for the regulated community to encourage innovative thinking and identify the least costly methods for compliance?
- Scientific and technological objectivity
  - Has the science of risk assessment advanced such that updated assessments of the regulation's impacts on affected populations such as environmental justice communities, children or the elderly could be improved?
  - Has the underlying scientific data changed since this regulation was finalized such that the change supports revision to the regulation?

5.4.3 European Commission evaluation focus

In its communication "Strengthening the foundations of Smart Regulation – improving evaluation" the European Commission defined evaluation as "a critical evidence-based judgement of whether EU action(s) has met the needs it aimed to satisfy and actually achieved its expected effects," (EC, 2013, p.7). According to Smismans,

> All evaluations are supposed to look at effectiveness (do the verified effects correspond to the original objectives?), efficiency (were the costs justified?), relevance (do the original objectives still correspond to the needs of the EU?), coherence (internally and with other interventions with similar objectives), and EU added value (compared to what could be achieved by Member States) (Smismans, 2015).

5.5 Apply research tools to measure causality

As noted above, when regulations are designed at the outset to allow for experimentation (for example, using small scale pilot studies or differential implementation for different populations) analysts have access to data and experiences that can help evaluate the steps in the causal chain linking actions to outcomes. Even when regulations have not been designed with such experimentation in mind, researchers can take advantage of observed variations, and with the benefit of established statistical tools, examine the extent to which different factors (both within and beyond the regulators control) are likely to cause different outcomes (Chay and Greenstone, 2003).

According to Greenstone, reliable estimates of regulatory costs and benefits must begin with the identification of a causal hypothesis, the key feature of which is that

> …it contains a manipulable treatment that can be applied to a subject and an outcome that may or may not respond to the treatment. For example, we may hypothesise that a regulation aiming to reduce air pollution in a city will reduce mortality rates among residents. For a causal hypothesis

> to have any practical relevance, we must be able to subject it to a meaningful test. Such a test requires that all other determinants of the outcome be held constant so that the effect of the treatment can be isolated (Greenstone, 2009, p 115).

When such randomised controlled experiments are not available, evaluators must rely on observational data. Coglianese notes that evaluations can be "*attributional*, that is, support inferences about the causal relationship between the treatment and the indicators, or *non-attributional*, that is, not supportive of any causal claim but assessing the level of the indicators against other benchmarks," (Coglianese, 2012, p. 15).

Non-attributional evaluations can often be easier to conduct, but they "cannot explain *why* problems are getting better or worse. They do not show whether the *treatment* actually worked," (Coglianese, 2012, p. 15). For example, Section 812 of the Clean Air Act Amendments of 1990 requires EPA periodically to assess the benefits and costs of the Act (EPA, 2011b), but EPA's assessment under this provision has relied on the same modelling it used for *ex-ante* analysis, so it has not provided information necessary to validate estimates or underlying risk assessment assumptions and procedures (Dudley, 2015a).

A more useful, though more challenging, evaluation would measure population changes with respect to the predicted outcome following the regulatory intervention. For example, actual reductions in cancer rates would be compared to predicted reductions to determine if actual experience corroborates or challenges the hypothesised benefits. Statistical tools can help determine the extent to which associations are causally linked; they can test "how changes in inputs (such as exposure) propagate through a network of validated causal mechanisms to cause resulting changes in outputs (such as health effects)," (Cox, 2015). To establish causal relationships, Cox argues that causal analytics algorithms "must do more than quantify past statistical associations between or among variables (such as policy inputs and outputs): they must also identify the pathways by which changes in some variables can produce changes in others," (Cox, 2015).

Standard and well-supported statistical methods are available for "detecting, analysing, quantifying, and visualising associations and other relations (such as information relations among multiple variables)" without depending on hypothetical modelling assumptions. (Cox, 2016). Cox has developed a publicly available "Causal Analytics Toolkit" (CAT) to make these sophisticated statistical tests more accessible to non-statisticians (Cox, 2016). The following principles can guide attributional *ex-post* evaluation:

> *Information principle for identifying potential causes:* For X to be a cause of Y, X must be *informative about* Y. That is, it must be possible to predict Y better if X is known than if it is not, and it must not be possible to find some other set of observations that would make Y statistically independent of X after conditioning on the observed values of other variables. There are a number of statistical tests for determining whether this information condition holds.
>
> *Temporal principles for identifying potential causes:* Not only must causes precede their effects, but past changes in causes should help to predict and explain future changes in their effects.
>
> *Ensemble principle for model uncertainty:* To avoid making causal conclusions contingent on modelling assumptions having unknown or uncertain validity, [evaluators can rely] on "model-free" (*non-parametric*) statistical methods and on the use of *model ensembles* (Scikit Learn, 2010), which are collections of hundreds of non-parametric models that provide plausible descriptions of the data. It has been discovered that averaging the predictions from many such models yields much more accurate and reliable predictions than using any single model.

> *Conditional expectation principle for quantifying causal relations:* The direct causal relation between one variable and another (e.g. between exposure and effect) can be quantified by studying how the effect changes as only that one cause is varied, holding all other variables fixed at their observed values. (Cox, 2016)

Statistical analysis requires data. Although randomised control trials are ideal, observational studies may be able to provide the data necessary for statistical techniques such as these. These are called quasi-experiments. According to Coglianese:

> Variation in observational studies can arise in one of two ways: either over time or across jurisdictions. When regulations vary over time within a single jurisdiction, researchers can compare outcomes longitudinally, that is, before and after the adoption of the regulation. When the variation exists across jurisdictions, researchers can compare outcomes cross-sectionally, that is, comparing outcomes in jurisdictions with the regulation being evaluated with those in jurisdictions without that regulation.

Domenici, Greenstone and Sunstein explain that quasi-experimental techniques may be particularly useful in the context of health effects from fine particulates by improving the understanding of the relationship between the two. Although in quasi-experimental evaluations, treatment groups and control groups are not chosen randomly and treatment status is determined by factors outside of the researcher's control, "it is possible to draw causal inferences from the differences in outcomes (by 'outcomes,' we refer to both air pollution levels and human health) between the treatment and control groups in a quasi- or natural experiment, provided certain assumptions are met" (Dominici et al., 2014).

6. Recommendations for improving incentives for robust evaluation

Efforts in the United States and other OECD countries to institute *ex-post* regulatory evaluation have met with limited success, largely because they did not change underlying incentives (Dudley, 2013b). This section suggests some institutional changes that might improve incentives for better retrospective review.

6.1 Make attributional retrospective evaluation a prerequisite for issuing new regulations

To incentivise more robust evaluation along the lines identified above, agencies could be required to test the validity of risk-reduction predictions before commencing new regulation (rather than relying on the same or revised hypothetical models that do not benefit from rigorous evidence of causal relationships). The European Commission's 2013 Communication "Strengthening the foundations of Smart Regulation – improving evaluation" articulates an "evaluate first" principle that, according to Smismans, requires that "[n]ew EU intervention can only be taken after an assessment of past action has been made."[15]

In the United States, EPA is required to review the National Ambient Air Quality Standards (NAAQS) every five years. Before it issued a new standard, it could be required to apply quasi-experimental techniques to gather and analyse epidemiology data and health outcome trends in different regions of the country and compare them against predictions (Cox, 2015, and Domenici et al., 2014). Regulations such as these lend themselves to quasi-experimental techniques to examine regulatory benefits and costs, because different areas of the US must respond differently depending on their attainment status.

15. "The 'evaluate first' principle locates evaluation firmly within the policy cycle. New EU intervention can only be taken after an assessment of past action has been made. While evaluation of expenditure policy has long been linked to the seven year financial programme cycle, the 2007 Communication sets the Commission on track to fit evaluation of all its action into its strategic planning and programming cycle. Most importantly, ex post evaluation should feed back into the EU system of ex ante impact assessments, which has been solidly established since 2003."

For example, Greenstone (2002) found that between 1972 and 1987, "nonattainment counties (relative to attainment ones) lost approximately 590 000 jobs, USD 37 billion in capital stock, and USD 75 billion (1987 dollars) of output in polluting industries," (Greenstone, 2002).

Similarly, oversight bodies could not allow new regulations to be issued without robust framework for later evaluation and a commitment to gather necessary data.

6.2 Institutionalise independent review

Retrospective review should not be left exclusively to entities responsible for issuing and implementing regulations, which have little incentive to find fault with their regulations, but should be subject to third-party evaluation.[16] Legislation has been introduced in the US Congress to create an independent commission with responsibility for reviewing the accumulated stock of regulations and making recommendations to repeal rules or sets of rules.[17] This model has potential to address some of the accumulated regulatory burden (Mandel and Carew, 2013), and improve regulatory evaluation. An independent third-party review would offer an objectivity that past efforts (which depend on regulatory agencies themselves to identify outmoded regulations) lacked, and would likely identify reform opportunities agencies would miss (Dudley, 2015b).

Perhaps most importantly, institutionalising a third-party review could motivate better data collection and more rigorous efforts to evaluate whether risk management regulation is actually achieving its desired effect.

6.3 Change the default rules regarding how long regulations remain in effect

Unlike many programs supported through fiscal budgets, regulations, once implemented, tend to stay in place. Several countries, including the Netherlands (Government of the Netherlands, 2016), the United Kingdom (HM Government, 2011), and Canada (Government of Canada, 2015) have initiated programs that require new regulatory costs to be offset by removal of existing regulatory burdens (Dudley, 2015c). Some US legislators are reviewing these programmes and exploring the application of budgeting tools to motivate better evaluation (as well as to constrain regulatory activity) (US Senate, 2015). Such approaches do not explicitly balance the benefits of regulations against the costs (Pierce, 2016), and governments may want to consider whether they are appropriate for all types of regulations (including those addressing chemical risks). However, they could motivate government and non-governmental agents to develop approaches to quantify the effectiveness and costs of regulations so they can trade off less-cost effective rules and retain those that are achieving their goals. Along these lines, sunset provisions could provide incentives for evaluation of regulations' effects.

Sunset or offset provisions may not be appropriate or practical in all cases, and could increase uncertainty regarding regulations that require large up front investments.

16. As Greenstone observed, "the process of self-evaluation is challenging for all organisations, as it requires complete objectivity. Indeed, history is unkind to organizations that fail to get outside reviews of their work." Statement of Michael Greenstone, Milton Friedman Professor of Economics, University of Chicago, Director, Energy Policy Institute at Chicago, before the United States Senate Subcommittee on Regulatory Affairs and Federal Management Roundtable on "Examining Practical Solutions to Improve the Federal Regulatory Process", 4 June 2015.

17. For example, Senate bills S. 708 and S. 1683.

6.4 Resources

Of course, more rigorous analysis demands resources. In the United States, regulatory agencies as a whole have received consistent increases in budget resources, growing much faster than funding increases for other government programmes. After adjusting for inflation, federal spending on US regulatory activities increased 18-fold over the 55-year period between 1960 and 2015 to well over USD 60 billion (Dudley and Warren, 2016). As a matter of comparison, all federal spending increased approximately 5-fold over the same time period.[18] And while growth has slowed, as of 2015 spending on regulatory programmes has more than doubled, in real terms, since 1995. In such an environment it may simply be a case of regulators setting different priorities in order to fund retrospective review.

A recent evaluation of the US EPA's hazardous waste regulations found the resource question to be "paradoxical." (Hart, 2016) During periods of restricted resources, the agency had a greater incentive to conduct evaluation to determine which programmes should be prioritised to achieve the greatest environmental benefits. When resources were less constrained, the agency put higher priority on developing new regulations to meet legislative and judicial requirements. Entities responsible for appropriating funds thus might choose to reallocate resources from *ex-ante* analysis to allow agencies to gather the information and evaluation tools necessary to validate *ex-ante* predications (Peacock, 2016). Shifting resources from *ex-ante* analysis to *ex-post* review would not only help with evaluation, but, by improving understanding of causal linkages, would improve *ex-ante* hypotheses of regulatory effects (EC, 2013).[19]

7. Conclusions

A systems approach to understanding the efficacy of chemical risk management regulations depends on evaluation and feedback, and would contribute both to learning and accountability. More consistent and robust evaluation of regulatory outcomes would be valuable not only for understanding the effect of past actions, but for improving future decisions. Depending on the circumstances, such evaluation could be performed for individual regulations or groups of related actions.

In the United States and other countries, while risk regulations are subject to rigorous *ex-ante* analysis, once they are in place they are often not evaluated with the same care. This may be due to methodological challenges, and lack of incentives.

Conducting reliable retrospective evaluation of regulations' impacts is not a trivial exercise, but tools are available to examine the causal relationship between risk management interventions and desired outcomes. While no evaluation will be perfect, transparent presentation of assumptions and analysis will allow testing and facilitate constructive public engagement and learning.

Agencies should plan for retrospective review at the outset, establishing a framework for empirical testing of assumptions and hypothesised outcomes. They should design regulations to allow for experimental learning – by initiating small scale trials or by randomising treatments in different areas. Regulations should be based on a theory of change or logic model that lays out the process by which the action is expected to lead to outcomes, so that evaluators can measure moderating or confounding factors that may have affected outcomes. Advanced statistical tools should be applied to establish whether observations are casually related.

18. Calculated from US Office of Management and Budget, Budget of the United States Government for Fiscal Year 2017: Historical Tables, Table 1.3.

19. According to the European Commission, "good impact assessments should draw on the lessons learnt from evaluations, which should identify problems, deficiencies, challenges and successes."

Institutional changes could improve incentives that support a culture of regulatory evaluation. Just as many jurisdictions will not approve new regulations without appropriate analytical justification, they could make issuance contingent upon a framework and plan for conducting retrospective review at a specified time in the future. Retrospective review should not be left exclusively to regulatory agencies, who have little incentive to find fault with their own regulations, but should be subject to third-party evaluation. Some OECD countries are applying mechanisms such as sunset provisions or offsets, which, while not appropriate in all situations, could provide incentives for evaluation of regulations' effects. Finally, shifting resources from *ex-ante* analysis to *ex-post* review would not only help with evaluation, but would provide valuable information that could improve *ex-ante* hypotheses of regulatory effects.

BIBLIOGRAPHY

Adler, J. (2004), "The Fable of Federal Regulation: No, States Didn't Ignore Environmental Problems", *PERC Report*, 22(4), Property and Environment Research Center, http://www.perc.org/articles/fable-federal-regulation.

Adventures in DM (2013), "Partial Dependency Plots and GBM", http://adventuresindm.blogspot.fr/2013/01/partial-dependency-plots-and-gbm.html.

Aldy, J. (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", A report for the Administrative Conference of the United States, https://www.acus.gov/report/retrospective-review-report.

Balla, S.J. and S.E. Dudley (2015), "Stakeholder Participation and Regulatory Policymaking in the United States", in *Regulatory Policy in Perspective: A Reader's Companion to the OECD Regulatory Policy Outlook 2015*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264241800-en.

Chay, K. and M. Greenstone (2003), "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession". *Quarterly Journal of Economics*, 2003, 118(3).

Coglianese, C. (2012), "Evaluating the Impact of Regulation and Regulatory Policy", in Measuring Regulatory Performance, OECD, Paris, https://www.oecd.org/gov/regulatory-policy/1_coglianese%20web.pdf.

Cox, L.A. (2016), "A Causal Analytics Toolkit (CAT) for Assessing Potential Causal Relations in Data", https://regulatorystudies.columbian.gwu.edu/causal-analytics-toolkit-cat-assessing-potential-causal-relations-data.

Cox, L.A. (2015), "Public Interest comment on The Environmental Protection Agency's Proposed Rule: National Ambient Air Quality Standards for Ozone", https://regulatorystudies.columbian.gwu.edu/sites/regulatorystudies.columbian.gwu.edu/files/downloads/TCox-NAAQS-ozone-2015.pdf.

Cox, L.A. (2012), "Reassessing the human health benefits from cleaner air", *Risk Analysis*, 32(5):816-29.

Coysh, D. et al. (2017), *Effects of Vintage Differentiated Environmental Regulations - Evidence from Survival Analysis of Coal-Fired Power Plants*, OECD Economics Department Working paper, forthcoming.

Cuellar, M-F. (2005), "Rethinking Regulatory Democracy", *Administrative Law Review*, 57.

Dominici, F., M. Greenstone and C. Sunstein (2014), "Particulate Matter Matters", *Science*, 344, American Association for the Advancement of Science.

Dudley, S. (2015a), "Regulatory Science and Policy: A Case Study of the National Ambient Air Quality Standards," GW Regulatory Studies Center working paper, https://regulatorystudies.columbian.gwu.edu/regulatory-science-and-policy-case-study-national-ambient-air-quality-standards.

Dudley, S. (2015b), Testimony before the Homeland Security and Governmental Affairs Committee: A Review of Regulatory Reform Proposals, https://regulatorystudies.columbian.gwu.edu/sites/regulatorystudies.columbian.gwu.edu/files/downloads/Dudley_HSGAC_Statement_RegReformBills_09-16-2015.pdf.

Dudley, S. (2015c), "Can Fiscal Budget Concepts Improve Regulation?" *NYU J. Legislation & Policy*, 19(259), http://www.nyujlpp.org/wp-content/uploads/2016/06/Dudley-Can-Fiscal-Budget-Concepts-Improve-Regulation-19nyujlpp259.pdf.

Dudley, S. (2013a), "A Retrospective Review of Retrospective Review", https://regulatorystudies.columbian.gwu.edu/sites/regulatorystudies.columbian.gwu.edu/files/downloads/20130507-a-retrospective-review-of-retrospective-review.pdf.

Dudley, S. (2013b), Testimony before the Joint Economic Committee: Reducing Unnecessary and Costly Red Tape through Smarter Regulations, http://regulatorystudies.columbian.gwu.edu/sites/regulatorystudies.columbian.gwu.edu/files/downloads/2013_06_26_Dudley_JEC_statement.pdf.

Dudley, S. and K. Wegrich (2015), "Achieving Regulatory Policy Objectives: An Overview and Comparison of U.S. and EU Procedures," GW Regulatory Studies Center working paper, https://regulatorystudies.columbian.gwu.edu/sites/regulatorystudies.columbian.gwu.edu/files/downloads/Dudley-Wegrich_US-EU_RegOverview.pdf

Dudley, S. and M. Warren (2016), "Regulators' Budget from Eisenhower to Obama: An Analysis of the U.S. Budget for Fiscal Years 1960 through 2017", https://wc.wustl.edu/files/wc/imce/2017_regulators_budget_05-17-2016.pdf.

European Commission (2013), Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, "Strengthening the foundations of Smart Regulation – improving evaluation," 686 final.

European Commission (2000), Opinion of Health & Consumer Protection Directorate-General, "Revised Opinion on Cyclamic Acid and Its Sodium and Calcium Salts", *Scientific Committee on Food*, https://ec.europa.eu/food/sites/food/files/safety/docs/sci-com_scf_out53_en.pdf

Fraas, A. and R. Lutter (2011), "Do Some NOx Emissions Have Negative Environmental Damages? Evidence and Implications for Policy", *Environmental Science & Technology*, 45(17), http://pubs.acs.org/doi/abs/10.1021/es202622z.

Gertler, P.J. et al. (2011), "Impact Evaluation in Practice", *The World Bank*, http://siteresources.worldbank.org/EXTHDOFFICE/Resources/5485726-1295455628620/Impact_Evaluation_in_Practice.pdf.

Government of Canada (2015), "Backgrounder- Legislating the One-for-One Rule", Treasury Bd. Of Can. Secretariat, https://www.canada.ca/en/treasury-board-secretariat/services/federal-regulatory-management/backgrounder-legislating-one-for-one-rule.html

Government of the Netherlands (2016), "Regulatory Burden on Businesses", *Reducing the Regulatory Burden*, https://www.government.nl/topics/reducing-the-regulatory-burden/contents/regulatory-burden-on-businesses.

Gray, G. and J. Cohen (2012), "Policy: Rethink chemical risk assessments", *Nature*, 489, pp. 27–28, http://www.nature.com/nature/journal/v489/n7414/full/489027a.html

Greenstone, M. (2015), Statement before the United States Senate Subcommittee on Regulatory Affairs and Federal Management Roundtable on "Examining Practical Solutions to Improve the Federal Regulatory Process", 4 June 2015.

Greenstone, M. (2009), "Toward a Culture of Persistent Regulatory Experimentation and Evaluation", in D. Moss and J. Cisternino, (eds.), *New Perspectives on Regulation*, The Tobin Project, Cambridge, MA.

Greenstone, M. (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), pp. 1175-1219. P1213.

Hart, N. (2016), "Evaluation at EPA: Determinants of the U.S. Environmental Protection Agency's Capacity to Supply Program Evaluation", the George Washington University Doctoral Dissertation, http://pqdtopen.proquest.com/pqdtopen/doc/1812333751.html?FMT=ABS.

HM Government (2011), "One-in, One-out: Statement of New Regulation", Dept. for Bus., Innovation & Skill, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/31617/11-p96a-one-in-one-out-new-regulation.pdf.

Koźluk, T. (2014), "The Indicators of the Economic Burdens of Environmental Policy Design: Results from the OECD Questionnaire", *OECD Economics Department Working Papers*, No. 1178, OECD Publishing, Paris. DOI: http://dx.doi.org/10.1787/5jxrjnbnbm8v-en

Krstić, G. (2013), "A reanalysis of fine particulate matter air pollution versus life expectancy in the United States", J Air Waste Manag Assoc. 2013 Feb;63(2):133-5.

Linn, J. (2013), "The Rebound Effect for Passenger Vehicles", Resources for the Future, http://www.rff.org/files/sharepoint/WorkImages/Download/RFF-DP-13-19.pdf.

Lutter, R. (2012), "The Role of Retrospective Analysis and Review in Regulatory Policy, *Working paper*, No. 12-14, The Mercatus Center at George Mason University.

Madden, B.J. (2016), *Value Creation Thinking*, LearningWhatWorks, Naperville, IL, http://www.valuecreationthinking.com.

Madden, B.J. (2010), *Wealth Creation: A Systems Mindset for Building and Investing in Businesses for the Long Term*, Wiley Finance.

Mandel, M. and D. Carew (2013), "Regulator Improvement Commission: A Politically-Viable Approach to U.S. Regulatory Reform", Policy Memo, Progressive Policy Institute, http://www.progressivepolicy.org/wp-content/uploads/2013/05/05.2013-Mandel-Carew_Regulatory-Improvement-Commission_A-Politically-Viable-Approach-to-US-Regulatory-Reform.pdf

McGinnis, J.O. (2012), *Accelerating Democracy: Transforming Governance through Technology*, Princeton University Press.

Miller, S.E. and S.E. Dudley (2016), "Regulatory Accretion: Causes and Possible Remedies", *Administrative Law Review Accord*, 3 March 2016, http://www.administrativelawreview.org/wp-content/uploads/2016/03/MillerDudley_PublishedVersion-1.pdf.

Miller, S.E. (2015) "Learning from Experience: Retrospective Review of Regulations in 2014", GW Regulatory Studies Center Working Paper, p. 20, https://regulatorystudies.columbian.gwu.edu/learning-experience-retrospective-review-regulations-2014.

National Research Council (NRC) (1983), *Risk Assessment in the Federal Government: Managing the Process*, National Academies Press, Washington D.C.

New World Encyclopedia (2013), "Cyclamate," *New World Encyclopedia*, http://www.newworldencyclopedia.org/entry/Help:Writers_Manual.

Newcomer, K.E. et al. (2015), *Handbook of Practical Program Evaluation*, 4th Edition, Wiley.

OECD (2012), *Recommendation of the Council on Regulatory Policy and Governance*, OECD Publishing, Paris. DOI: http://dx.doi.org/10.1787/9789264209022-en

OECD (2008), *Introductory Handbook for Undertaking Regulatory Impact Analysis* (RIA), OECD Publishing, Paris, https://www.oecd.org/gov/regulatory-policy/44789472.pdf.

OECD (2002), *Regulatory Policies in OECD Countries: From Interventionism to Regulatory Governance*, OECD Publishing, Paris.
DOI: http://dx.doi.org/10.1787/9789264177437-en

Parker, D. and C. Kirkpatrick (2012), "The Economic Impact of Regulatory Policy: a Literature Review of Quantitative Evidence", in Measuring Regulatory Performance, OECD, http://www.oecd.org/gov/regulatory-policy/3_Kirkpatrick%20Parker%20web.pdf

Peacock, M. (2016), "How Declining Budgets at U.S. Regulatory Agencies Could Improve Performance", The George Washington University Regulatory Studies Center Working Paper, https://regulatorystudies.columbian.gwu.edu/how-declining-budgets-us-regulatory-agencies-could-improve-performance.

Peltzman, S. (2004), "Regulation and the Natural Progress of Opulence", *AEI-Brookings Joint Center Distinguished Lecture*, https://www.aei.org/publication/regulation-and-the-natural-progress-of-opulence

Pierce, R. (2016), "The Regulatory Budget Debate," *NYU J. Legislation & Policy*, 19(249), http://www.nyujlpp.org/wp-content/uploads/2016/06/Pierce-Regulatory-Budget-Debate-19nyujlpp249.pdf.

President of the United States, Executive Order 13610 (2012), "Identifying and Reducing Regulatory Burdens," Federal Register, 77(93), https://www.gpo.gov/fdsys/pkg/FR-2012-05-14/pdf/2012-11798.pdf

President of the United States, Executive Order 13563 (2011), *Federal Register*, 76(14), https://www.gpo.gov/fdsys/pkg/FR-2011-01-21/pdf/2011-1385.pdf

President of the United States, Executive Order 12866 (1993), Federal Register, 58(190), https://www.regulations.gov/docs/EO_12866.pdf

President of the United States, Executive Order 12044 (1978), http://www.presidency.ucsb.edu/ws/?pid=30539

Scikit Learn (2010), "Ensemble Methods", http://scikit-learn.org/stable/modules/ensemble.html.

Smismans, S. (2015), "Policy Evaluation in the EU: The Challenges of Linking *Ex Ante* and *Ex Post* Appraisal", *European J. Risk Regulation*, 1, pp.6-26.

Stern, N. et al. (2012), "Broadening the range of designs and methods for impact evaluations", *Department for International Development*, Working Paper 38, pp. 39.

U.S. Senate Committee on Homeland Security and Governmental Affairs and the Committee on the Budget (2015), "Accounting for the True Cost of Regulation: Exploring the Possibility of a Regulatory Budget", https://www.hsgac.senate.gov/hearings/measuring-the-true-cost-of-regulations-lessons-from-great-britain-and-canada-on-implementing-regulatory-reforms.

U.S. EPA (Environmental Protection Agency) (2016a), *Ozone Pollution*, U.S. EPA, Washington D.C., https://www.epa.gov/ozone-pollution.

U.S. EPA (2016b), *Indoor Air Quality (IAQ)*, U.S. EPA, Washington D.C., https://www.epa.gov/indoor-air-quality-iaq.

U.S. EPA (2011a), *EPA's Final Plan for Periodic Retrospective Reviews*, U.S. EPA, Washington D.C., https://www.epa.gov/sites/production/files/2015-09/documents/eparetroreviewplan-aug2011_0.pdf

U.S. EPA (2011b), The Benefits and Costs of the Clean Air Act, 1990-2012, http://www.epa.gov/air/sect812/feb11/fullreport.pdf.

U.S. Institute of Medicine (IOM), Board on Population Health and Public Health Practice (2013), *Environmental Decisions in the Face of Uncertainty*, https://www.nap.edu/catalog/12568/environmental-decisions-in-the-face-of-uncertainty

U.S. Office of Management and Budget (OMB) (2015), "2015 Report to Congress on the Benefits and Costs of Federal Regulations and Agency Compliance with the Unfunded Mandate Reform Act", U.S. Office of Management and Budget, Washington D.C.

U.S. Office of Management and Budget (OMB) (2011a), *Memorandum for the Heads of Executive Departments and Agencies: Retrospective Analysis of Existing Significant Regulations* by Cass Sunstein, U.S. Office of Management and Budget, Washington D.C.

U.S. Office of Management and Budget (OMB) (2011b), "Report to Congress on the Costs and Benefits of Federal Regulation", U.S. Office of Management and Budget, Washington D.C.

U.S. Office of Management and Budget (OMB) (2011c), *Regulatory Impact Analysis: A Primer*, U.S. Office of Management and Budget, Washington D.C.

Victoria Transport Policy Institute (VTPI) (2014), "Rebound Effects: Implications for Transport Planning," *TDM Encyclopedia*, http://www.vtpi.org/tdm/tdm64.htm.