INSTITUTE
FOR LAW & AI

LawAI's thoughts on proposed updates to U.S. federal benefit-cost analysis

Christoph Winter, Suzanne Van Ardsale, and Kevin Frazier | Institute for Law & AI | June 2023

This analysis is based on a comment submitted in response to the Request for Comment on proposed Circular A-4, "Regulatory Analysis".

We support the many important and substantial reforms to the regulation review process in the proposed Circular A-4. The reforms, if adopted, would reduce the odds of regulations imposing undue costs on vulnerable, underrepresented, and disadvantaged communities both now and well into the future. In this piece, we outline a few additional changes that would further reduce those odds: expanding the scope of analysis to include catastrophic and existential risks, including those far in the future; including future generations in distributional analysis; providing more guidance regarding model uncertainty and regulations that involve irreversible outcomes; lowering the discount rate to zero for irreversible effects; and in a narrow set of cases or, minimally, lowering the discount rate in proportion to the temporal scope of a regulation.

1. Circular A-4 contains many improvements, including consideration of global impacts, expanding the temporal scope of analysis, and recommendations on developing an analytical baseline.

Circular A-4 contains many improvements on the current approach to benefit-cost analysis (BCA). In particular, the proposed reforms would allow for a more comprehensive understanding of the myriad risks posed by any regulation. The guidance for analysis to include global impacts[1] will more accurately account for the effects of a regulation on increasingly interconnected and interdependent economic, political, and environmental systems. Many global externalities, such as pandemics and climate change, require international regulatory cooperation; in these cases, efficient allocation of global resources, which benefits the United States and its citizens and residents, requires all countries to consider global costs and benefits.[2]

The instruction to tailor the time scope of analysis to "encompass all the important benefits and costs likely to result from regulation" will likewise bolster the quality of a risk assessment[3]—though, as mentioned below, a slight modification to this instruction could aid regulators in identifying and mitigating existential risks posed by regulations.

[1] Circular A-4 at 9.
[2] We echo the comment of Robert Stavins in endorsing this approach as not only strategically and ethically reasonable, but also economically efficient.
[3] *See* Circular A-4 at 11.

The recommendations on developing an analytic baseline have the potential to increase the accuracy and comprehensiveness of BCA by ensuring that analysts integrate current and likely technological developments and the resulting harms of those developments into their baseline.[4]

A number of other proposals would also qualify as improvements on the status quo. A litany of commentors have discussed those proposals, so the remainder of this piece is reserved for suggested amendments and recommendations for topics worthy of additional consideration.

2. The footnote considering catastrophic risks is a welcome addition that could be further strengthened with a minimum time frame of analysis and clear inclusion of catastrophic and existential threats in "important" and "likely" benefits and costs.

The proposed language will lead to a more thorough review of the benefits and costs of a regulation by expanding the time horizon over which those effects are assessed.[5] We particularly welcome the footnote encouraging analysts to consider whether a regulation that involves a catastrophic risk may impose costs on future generations.[6]

We recommend two suggestions to further strengthen the purpose of this footnote in encouraging the consideration of catastrophic and existential risks and the long-run effects of related regulation. First, we recommend mandating consideration of long-run effects of a regulation.[7] Given the economic significance of a regulation that triggers review under Executive Orders 12866 and 13563, as supplemented and reaffirmed by Executive Order 14094, the inevitable long-term impacts deserve consideration—especially because regulations of such size and scope could affect catastrophic and existential risks that imperil future generations. Thus, the Office should consider establishing a minimum time frame of analysis to ensure that long-run benefits and costs are adequately considered, even if they are sometimes found to be negligible or highly uncertain.

Second, the final draft should clarify what constitutes an "important" benefit and cost as well as when those effects will be considered "likely".[8] We recommend that those concepts clearly encompass potential catastrophic or existential threats, even those that have very low likelihood.[9] An expansive definition of both

[4] *See id.* at 12.
[5] *Id.* at 11.
[6] *Id.* at 11 n.19.
[7] *Id.*
[8] The findings of Milburn and others document human biases with respect to forecasts. In many contexts, people underestimate the risks of high-impact, low-probability events, such as many catastrophic and existential risks. *See, e.g.*, Max H. Bazerman & Michael D. Watkins, *Predictable Surprises: The Disasters You Should Have Seen Coming, and How to Prevent Them* 84-87 (2004); Nassim Nicholas Taleb, *The Black Swan: The Impact of the Highly Improbable* (2007); *cf.* Cass Sunstein, *Irreversible and Catastrophic*, 91 Cornell L. Rev. 841, 870-71, 895 n.252 (2007) (citing the availability heuristic). This effect may be compounded for events further in the future, with one study finding that people tend to underestimate the likelihood of future negative events, and the further in the future they are, the less likely they seem. Michael A. Milburn, *Sources of Bias in the Prediction of Future Events*, 21 Org. Beh. & Human Perf. 17 (1978), https://doi.org/10.1016/0030-5073(78)90035-1.
Rather than rely solely on subjective triggers for long-term analysis, the Office could establish a minimum time frame of analysis to ensure at least some consideration of future generations. The Office could also address these biases in part by providing expansive definitions. Just as the proposed Circular acknowledges the need for regulation to address behavioral biases, *see* Circular A-4 at 15, 18-19, it is appropriate for agency guidance to address biases present during analysis.
[9] Such language would further the purpose of the Global Catastrophic Risk Management Act, and the identification and review of catastrophic and existential risks would be made easier by the requirement under the Act that the

qualifiers would allow the BCA to provide stakeholders with a more complete picture of the regulation's short- and long-term impact.

3. Distributional analysis should become the default of regulatory review and include future generations as a group under consideration.

The potential for disparate effects of regulations on vulnerable, underrepresented, and disadvantaged groups merits analysis in all cases. Along with several commentors, we recommend that distributional analysis become the default of any regulatory review. When possible, we further recommend that such analysis include future generations among the demographic categories.[10] Future generations have no formal representation and will bear the costs imposed by any regulation for longer than other groups.[11]

The Office should also consider making this analysis mandatory, with no exceptions. Such a mandate would reduce the odds of any group unexpectedly bearing a disproportionate and unjust share of the costs of a regulation. The information generated by this analysis would also give groups a more meaningfully informed opportunity to engage in the review of regulations.

4. Treatment of uncertainty is crucial for evaluating long-term impacts and should include more guidance regarding models, model uncertainty, and regulations that involve irreversible outcomes.

Circular A-4 directs agencies to seek out and respond to several different types of uncertainty from the outset of their analysis.[12] This direction will allow for a more complete understanding of the impacts of a regulation both in the short- and long- term. Greater direction would accentuate those benefits.

The current model uncertainty guidance, largely confined to a footnote, nudges agencies to "consider multiple models to establish robustness and reduce model uncertainty."[13] The brevity of this instruction conflicts with

---

Secretary of Homeland Security submit a report containing a detailed assessment of such risks. *See* Public Law No. 117-263, § 7304, codified at 6 U.S.C. § 823.

[10] In a distributional analysis, the benefits and costs could be disaggregated broadly across current and future generations, or more specific units of analysis or subgroups could be selected according to the regulation under consideration and to observe relevant distributional effects. *Cf.* Circular A-4 at 62-63. This type of analysis would be supported by the other requirements of the Circular A-4, in particular the instruction to assess costs and benefits over the appropriate time horizon, and would take it one step further by evaluating whether a disproportionate amount of the costs would be borne by future generations in general or in a specific time frame.

[11] The Legal Priorities Project has conducted substantial research regarding the importance of protecting future generations, whose interests are vastly underrepresented in current legal systems. *See, e.g.*, Eric Martínez & Christoph Winter, *Protecting Future Generations: A Global Survey of Legal Academics*, Legal Priorities Project Working Paper 1-2021, https://www.legalpriorities.org/research/protecting-future-generations.html ("Although our laws and policies may have historically unique consequences for future generations, their interests are rarely represented in current legal systems."); Tyler M. John & William MacAskill, *Longtermist Institutional Reform*, in Natalie Cargill & Tyler M. John (Eds.), The Long View: Essays on Policy, Philanthropy, and the Long-term Future (2023), available as Legal Priorities Project Working Paper 4-2021, https://www.legalpriorities.org/research/longtermist-institutional-reform.html (discussing the need to align government incentives with the interests of future generations and suggestions for reform); Tyler M. John, *Empowering Future People by Empowering the Young?*, in Greg Bognar & Alex Gosseries (Eds.), Ageing without Ageism? (forthcoming 2023), also available as Legal Priorities Project Working Paper 5-2021, https://www.legalpriorities.org/research/empowering-future-people.html (suggesting methods for apportioning greater political power to the young to counteract short-termist policy).

[12] Circular A-4 at 66.

[13] *Id.* at 67 n.118.

the complexity of this process. Absent more guidance, agencies may be poorly equipped to assess and treat uncertainty, which will frustrate the provision of "useful information to decision makers and the public about the effects and the uncertainties of alternative regulatory actions."[14] A more participatory, equitable, and robust regulation review process hinges on that information.

We encourage the agency to provide further examples and guidance on how to prepare models and address model uncertainty, in particular regarding catastrophic and existential risks, as well as significant benefits and costs in the far future.[15] A more robust approach to responding to uncertainty would include explicit instructions on how to identify, evaluate, and report uncertainty regarding the future. Several commentors highlighted that estimates of costs and benefits become more uncertain over time. We echo and amplify concerns that regulations with forecasted effects on future generations will require more rigorous treatment of uncertainty.

We similarly recommend that more guidance be offered with respect to regulations that involve irreversible outcomes, such as exhaustion of resources or extinction of a species.[16] The Circular notes that such regulations may benefit from a "real options" analysis; however, this simple guidance is inadequate for the significance of the topic. The Circular acknowledges that "[t]he costs of shifting the timing of regulatory effects further into the future may be especially high when regulating to protect against irreversible harms." We agree that preserving option value for future generations is of immense value. How to value those options should receive more attention in subsequent drafts. Likewise, guidance on how to identify irreversible outcomes and conduct real options analysis merits more attention in forthcoming iterations.

---

[14] *Id.* at 67.

[15] Agencies tasked with regulating around events that pose catastrophic or existential risks may benefit from adopting a "storyline" approach to modeling uncertainty. *See, e.g.*, Theodore Shepherd et al., *Storylines: An Alternative Approach to Representing Uncertainty in Physical Aspects of Climate Change*, 151 Climatic Change 555 (2018), https://doi.org/10.1007/s10584-018-2317-9; *see generally* David Althaus & Lukas Gloor, *Reducing Risks of Astronomical Suffering: A Neglected Priority*, Center on Long-Term Risk (updated Aug. 2019), https://longtermrisk.org/reducing-risks-of-astronomical-suffering-a-neglected-priority/#IIIV_Unknown_unknowns_and_model_uncertainty (discussing the connection between model uncertainty and mitigating existential risk); *cf.* Toby Ord et al., *Probing the Improbable*, 2 J. Risk Res. 191 (2010), https://doi.org/10.1080/13669870903126267 (discussing the need for "robust estimates that can handle theory, model, and calculation errors", particularly for "low-probability, high-stake events").

[16] Several other situations may involve irreversibility, such as those involving research and publication of information. Several strands of biological research, for example, may pose an information hazard; historical examples of perceived or actual information hazards include, but are not limited to mousepox, gain-of-function research on H5N1, and botulinum toxin H. Research and publication on such consequential topics requires new and better tools to thoughtfully weigh the trade-offs between secrecy and openness, and regulation involving such research and publication warrants equally thoughtful consideration. *See* Gregory Lewis et al., *Information Hazards in Biotechnology*, 39 Risk Anal. 975 (2018), https://doi.org/10.1111/risa.13235 (discussing biological information hazards, examples, and ways to mitigate these hazards). In the context of software, the release of source code or, more consequentially, of model weights and other components of transformative artificial intelligence may be irreversible and lead to misuse. *See, e.g.*, Josephine Wolff, *How to Improve Cybersecurity for Artificial Intelligence*, Brookings Institute (Jun. 9, 2020), https://www.brookings.edu/research/how-to-improve-cybersecurity-for-artificial-intelligence/. Of course, irreversibility is material to regulations that relate to catastrophic and existential risk. For more information about irreversible harms posed by emerging technologies, see Anthony Barrett & Seth Baum, *A Model of Pathways to Artificial Superintelligence Catastrophe for Risk and Decision Analysis*, 29 J. Exp. & Theo. AI 397 (2017), https://doi.org/10.1080/0952813X.2016.1186228; Arthur Duforest, *To What Extent Can Cyberattacks Constitute A Global Catastrophic Risk?*, Institute for Applied Geopolitical Studies (2022), https://www.institut-ega.org/l/to-what-extent-can-cyberattacks-constitute-a-global-catastrophic-risk/; Matthijs Maas et al., *Military Artificial Intelligence as Contributor to Global Catastrophic Risk in Era of Global Risk*, in SJ Beard et al. (Eds), The Era of Global Risk (forthcoming 2023), http://dx.doi.org/10.2139/ssrn.4115010.

We recommend similar caution for regulations involving harms that are persistent and challenging to reverse, but not irreversible.

5. A lower discount rate and declining discount rate are necessary to account for the impact of regulations with significant and long-term effects on future generations.

The discount rate in a BCA is one signal of how much a society values the future. We join a chorus of commentors in applauding both the overall lowering of the discount rate as well as the idea of a declining discount rate schedule.

The diversity of perspectives in those comments, however, indicate that this topic merits further consideration. In particular, we would welcome further discussion on the merits of a zero discount rate. Though sometimes characterized as a blunt tool to attempt to assist future generations,[17] zero discount rates may become necessary when evaluating regulations that involve irreversible harm.[18] In cases involving irreversibility, a fundamental assumption about discounting breaks down—specifically, that the discounted resource has more value in the present because it can be invested and, as a result, generate more resources in subsequent periods.[19] If the regulation involves the elimination of certain resources, such as nonrenewable resources, rather than their preservation or investment, then the value of the resources remain constant across time periods.[20] Several commentors indicated that they share our concern about such harms, suggesting that they would welcome this narrow use case for zero discount rates.[21]

We likewise support the general concept of declining discount rates and further conversations regarding the declining discount rate (DDR) schedule,[22] given the importance of such schedules in accounting for the impact of regulations with significant and long-term effects on future generations.[23] US adoption of a DDR schedule

---

[17] Cass Sunstein & Arden Rowell, *On Discounting Regulatory Benefits: Risk, Money, and Intergenerational Equity*, 74 U. Chi. L. Rev. 171, 198-99 (2007) (pointing out that a zero discount rate can in some cases lead to the "postponement of protective programs" as well as a diminished capacity among future generations to benefit from long-term prosperity).

[18] William Baumol, *On the Social Rate of Discount*, 58 Am. Econ. Rev. 788, 801 (1968) (reserving low discount rates for "irreversibilities"). Regulations that "involve" irreversible harm include any regulations that change the probability of that harm materializing, whether it might cause, alleviate, or avert such harm.

[19] *See* John J. Donohue III, *Why We Should Discount the Views of Those Who Discount Discounting*, 108 Yale L.J. 1901, 1905 (1998).

[20] *See* Lisa Heinzerling, *Regulatory Costs of Mythic Proportions*, 107 Yale L.J. 1981, 2051-54 (1998) (citing Baumol in recommending a zero discount rate in cases involving irreversible harms).

[21] Note that the adoption of a temporal discount rate of zero would further the principles advanced by a geographical discount rate of zero, as encouraged by Circular A-4, which has embraced a more international, equitable approach. The application of a lower or zero temporal discount rate in specific circumstances would extend that approach to questions of intergenerational equity.

[22] Office of Management and Budget, Exec. Office of the President, Preamble: Proposed OMB Circular No. A-4, "Regulatory Analysis" (April 6, 2023), 29-31.

[23] The use of progressively lower discount rates to account for future interest rate uncertainty is in line with the approach recommended in the literature. *See* Maureen Cropper et al., *Declining Discount Rates*, 104 Am. Econ. Rev. 538 (2014), https://doi.org/10.1257/aer.104.5.538; Kenneth J. Arrow et al., *Should Governments Use a Declining Discount Rate in Project Analysis?*, 8 Rev. Env't. Econ. Pol'y 145 (2014), https://doi.org/10.1093/reep/reu008; *see also* Richard G. Newell & William A. Pizer, *Discounting the Distant Future: How Much Do Uncertain Rates Increase Valuations?*, 46 J. Env't Econ. & Mgmt. 52 (2003).

would bring us into alignment with two peers—namely, the UK and France.[24] The former, which is based on the Ramsey formula rather than a fixed DDR schedule proposed, deserves particular attention given that it estimates time preference ρ as the sum of "pure time preference (δ , delta) and catastrophic risk (L)",[25] defined in the previous Green Book as the "likelihood that there will be some event so devastating that all returns from policies, programmes or projects are eliminated".[26] This approach to a declining discount schedule demonstrates the sort of risk aversion, considering catastrophic and existential risk, that is necessary in light of regulations that present significant uncertainty.

6. Regulations that relate to irreversible outcomes, catastrophic risk, or existential risk warrant review as being significant under Section 3(f)(1).

In establishing thresholds for which regulations will undergo regulatory analysis, Section 3(f)(1) of Executive Order 12866 includes a number of sufficient criteria in addition to the increased monetary threshold. We note that regulations that might increase or reduce catastrophic or existential risk should be reviewed as having the potential to "adversely affect in a material way the economy, a sector of the economy, productivity, competition, jobs, the environment, public health or safety, or State, local, territorial, or tribal governments or communities."[27] Even "minor" regulations can have unintended consequences with major ramifications on our institutions, systems, and norms—those that might influence such grave risks are of particular import. For similar reasons, the Office should also review any regulation that has a reasonable chance of causing irreversible harm to future generations.[28]

7. Conclusion

Circular A-4 contains important and substantial reforms to the regulation review process. The reforms, if adopted, would reduce the odds of regulations imposing undue costs on vulnerable, underrepresented, and disadvantaged communities both now and well into the future. A few additional changes would further reduce those odds—specifically, expanding the scope of analysis to include catastrophic and existential risks, including those far in the future; including future generations in distributional analysis; providing more guidance regarding model uncertainty and regulations that involve irreversible outcomes; lowering the discount rate to zero for irreversible effects; and in a narrow set of cases or, minimally, lowering the discount rate in proportion to the temporal scope of a regulation.

---

[24] Maureen Cropper et al., *Declining Discount Rates*, 104 Am. Econ. Rev. 538, 545 (2014), https://doi.org/10.1257/aer.104.5.538

[25] Green Book: Central Government Guidance on Appraisal and Evaluation, HM Treasury, 116-19 & Table 5 (2022), https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1063330/Green_Book_2022.pdf (setting forth a DDR schedule and further noting: "Where the possible effects of an intervention being examined as part of an appraisal are long term and involve very substantial or irreversible wealth transfers between generations further sensitivity analysis is appropriate. This could include irreversible changes to the natural environment. This involves applying both the standard Green Book discount rate and a reduced discount rate.").

[26] Green Book: Appraisal and Evaluation in Central Government, HM Treasury, 97 (2003), https://webarchive.nationalarchives.gov.uk/ukgwa/20080305121602/http://www.hm-treasury.gov.uk/media/3/F/green_book_260907.pdf. The most recent edition of the Green Book also includes systemic risk among the "unpredictable risks not normally included in appraisal", including "technological disruption, natural disasters … and other unforeseeable occurrences." Green Book (2022) at 112,117.

[27] Executive Order 12866, as supplemented and reaffirmed by Executive Order 14094.

[28] Emil Moldovan offered a comment that would support this approach, pointing out that regulations with forecasted effects over longer time horizons will have larger bounds of uncertainty with respect to their cumulative impact. Moldovan's recommendations that OIRA consider "optionality" and "cost of risk" as part of its review also deserve attention from the Office.