



**May 18, 2020**

**Comments submitted by Angela Logomasini, Senior Fellow, Competitive Enterprise Institute**

**Docket ID Nos. Original:** EPA-HQ-OA-2018-0259-9322; **Supplemental:** EPA-HQ-OA-2018-0259-9322

**Via** <https://www.regulations.gov>

The Competitive Enterprise Institute (CEI) strongly supports EPA's effort to increase transparency in scientific research underlying agency rules and research. CEI detailed in a 2018 paper why the goals of the transparency rule are both essential to sound policymaking and are achievable.<sup>1</sup> That paper is included with these comments for reference.

Researchers around the world have recognized the need to improve transparency in scientific research by developing ways to systemize the process. In May 2014, Nature Publishing Group even launched a new publication called *Science Data*, whose mission is to find ways to promote scientific data sharing. Discussing the launch of this new publication, an editorial in *Nature Geoscience* proclaimed: "The question is no longer whether research data should be shared, but how to make effective data sharing a common and well-rewarded part of research culture."<sup>2</sup>

Unfortunately, the scientific community at large has not yet overcome resistance to data sharing among researchers, even though the community's refusal to share data undermines the scientific process. Scientists Chris Chambers of Cardiff University and Brian Nosek of the University of Virginia explain in *The Guardian* that transparency has become "undervalued" because researchers want to get published, and "the fierce competition of academia rewards those who secure large grants and publish innovative—if tentative—findings in prestigious journals, not scholars who instead focus on being transparent and careful."<sup>3</sup> As a result, researchers often avoid data sharing and transparency out of fear that their novel findings may eventually be proved wrong when someone tries to replicate their work.

This reality has fostered both unintentional bias and scientific mischief, including the propensity for researchers to work the data until it generates a positive finding. As James Mills of the National Institute of Child Health and Human Development lamented back in 1993 in the *New England Journal of Medicine*: "'If you torture your data long enough, they will tell you whatever you want to hear' has become a popular observation in our office."<sup>4</sup> The scope of this problem<sup>5</sup> is also detailed by the National Association of Scholars in its recent report, *The Irreproducibility Crisis of Modern Science*.<sup>6</sup> The authors highlight one outrageous case where Professor Brian Wansink, head of Cornell University's food behavior research lab, literally bragged in a blog post that he had schooled one of his students on how to churn data to generate positive results and get them published.<sup>7</sup>

Even when researchers do not torture data, many positive associations will occur by mere chance or unintentional biases. Stanford Professor of Medicine John Ioannidis demonstrated in a research article in 2005 that most published research findings are false positives. He explained: “Simulations show that for most study designs and settings, it is more likely for a research claim to be false than true. Moreover, for many current scientific fields, claimed research findings may often be simply accurate measures of the prevailing bias.”<sup>8</sup>

Unfortunately, regulators may rely on much of this faulty research leading to poorly designed and often far too onerous regulations that impose burdens on society for no benefits in return. In many cases, regulations prove counterproductive and do more harm to public health than good. Accordingly, it is essential that regulators demand that the studies underlying their regulations meet the highest possible standards to ensure their validity. Increasing transparency through data availability will allow others to validate research findings or demonstrate when studies are not an appropriate basis for regulatory decisions.

EPA’s Science Advisory Board (SAB) likewise noted the important role that this rule could play in terms of improving science and related agency decisions. Specifically, the SAB notes in its comments:

The SAB recognizes the importance of this rule. It can enhance public access to scientific data and analytical methods, and help ensure scientific integrity, consistency and robust analysis. Strengthening transparency by improving access to data can lead to an increase in both the quantity and the quality of evidence that informs important regulatory science and policy decisions. The scientific community is moving toward adopting the precept of sharing accurate data and information to foster credibility, high-quality outcomes and public confidence in science. The SAB supports the adoption of this precept. ...

To ensure that the rule is evidence-based EPA must provide greater clarity regarding details of the rule and how it will be implemented, as well as example analyses of how it would be deployed. The development of additional policy and/or guidance documents is strongly recommended to provide clarity on the procedures for conducting the proposed efforts.<sup>9</sup>

CEI agrees that an agency rule that promotes transparency in science is greatly needed and that the proposed rule would benefit from some improvements to ensure it meets its goals. Some revisions to the original rule made with the supplemental rule have improved several key elements overall, yet other revisions will seriously weaken the rule’s value and ability to achieve its goals. Accordingly, we offer the following input and suggestions.

## **I. Expanding the Scope of the Rule.**

CEI commends elements in Section 30.2 of the supplemental proposal that widen the scope of the rule. In particular, the addition of the definition for “influential information” and its inclusion in the definition of “pivotal regulatory science” is welcome because it indicates that the rule could also apply to research that impacts both policy decisions *and private sector decisions*. This is important because any agency science can have far-reaching marketplace impacts, both

positive and negative. It only makes sense that the agency science—including “nonregulatory activities” such as research studies or risk assessments—complies with the best available scientific standards and transparency requirements.

In addition, the expanded number of terms and clarifications in the definition section represents an overall improvement to this section.

CEI also welcomes the provision in Section 30.3 that expands the rule to apply to more than dose response studies, noting application to “data and models.” We agree that the rule should not be limited to a subset of scientific studies but instead it should strive to improve all science that the agency relies upon. However, this section limits the rule to science that is “used to justify significant regulatory decisions,” which might be construed to exempt “non-regulatory” agency research, such as risk assessments and other studies conducted inside EPA’s Office of Research and Development. This section should be improved to state that the rule applies to all science “used to justify significant regulatory decisions or research produced by agency research programs for regulatory and educational functions.”

## **II. Exclusions and Exemptions in Section 30.3**

CEI is concerned that too many provisions in this rule either exclude certain science or give the EPA Administrator too much power to disregard the rule completely. While it is especially important to ensure any data released is done so in a fashion that protects the privacy of individuals, such goals should be attainable without having to completely destroy the utility of this rule, and the rule should note that such cases would be rare.

**Exclusions Related to “Conflicts” with Statutory Law.** The supplemental states that the rule would not apply when its provisions “conflict” with statutory law. Obviously, laws prevail over agency rules, so this stated exclusion is not necessary. In any case, it is not clear that any law on the books would prevent the agency from ensuring that any science it uses meets basic scientific transparency standards. And in the unlikely situation that questions arise, this issue should remain within the purview of the courts to determine whether a conflict exists. Including this provision appears to do nothing more than give the Administrator an excuse to waive the rule for rather arbitrary reasons.

**Exclusions of Research Deemed Not a “Significant Regulatory Action.”** Section 30.3 excludes anything not considered a “significant regulatory action,” such as adjudications, enforcement actions, and permit proceedings. Some provision may be necessary to ensure the agency does not needlessly delay permits and the like. But, as noted, the rule should apply to agency-produced science, including risk assessments and research that could have impacts in the marketplace.

**Exclusions of Drafts and Preliminary Analyses.** Section 30.3’s exclusions of drafts and preliminary analyses are too vague and may exclude exceptionally important research. It might make sense to exclude preliminary analyses that are not publicly available. However, anything the agency releases or discusses publicly as a possible justification for a regulatory standard or for educational purposes should meet the transparency rule. Draft risk assessments, for example,

may have profound market impacts, including impacts on litigation, and hence, should not be exempt from the rule.

**Exclusions of “Lab Samples.”** Section 30.3 explicitly states that the rule does not apply to physical objects such as lab samples. This exclusion may be intended merely to address a practical matter to make clear that researchers need not submit their lab samples to EPA or the public. However, it raises some issues about how the rule might include provisions to promote sound research lab practices. One approach might be to require that the agency place greater weight on studies that can demonstrate the application of “Good Laboratory Practices” (GLP).<sup>10</sup> In fact, poor lab controls—which can lead to such problems as inadvertent contamination of samples—can produce vastly distorted and misleading study findings. GLP-compliant studies meet higher standards in terms of study design, implementation, and peer review, which is why they should bear more weight in regulatory decisions. As a result, it is common worldwide that industry must apply GLP when conducting research for submission to regulatory bodies for the purpose of ensuring “uniformity, consistency, reliability, reproducibility, quality, and integrity of chemical (including pharmaceuticals) nonclinical safety tests.”<sup>11</sup> Accordingly, greater transparency related to the handling and management of lab samples should be included as part of the transparency rule.

### **III. Administrator’s Right to Waive the Rule as “Impractical.”**

Section 30.9 allows the EPA Administrator to issue exemptions to this rule on a “case-by-case” basis if the Administrator deems accessing the data to be impractical. This provision essentially makes the entire rule nearly meaningless. It might be used as an excuse to exempt all research conducted before implementation of the rule, which will greatly undermine the purpose of this rule. If there are to be any exemptions, the rule should delineate the specific circumstances and underscore the fact that such cases would be *rare*. And in those rare cases, the rule should place the burden on the agency to demonstrate that there were not reasonable means of releasing data without compromising privacy of the subjects or revealing confidential business information. The EPA Administrator’s opinion about practicality should not be the governing standard.

The agency’s SAB also pointed out the obvious problems with this provision as well, noting:

Case-by-case exceptions without criteria may create public concerns about inappropriate exclusion of scientifically important studies. A framework and/or guidance document could also help EPA clarify how the rule will affect current scientific review procedures. It might be useful for the EPA to consider recommendations from a scientific and advisory committee when making waiver decisions.

The SAB suggestion that the agency deploy a formal process for making such determinations is sound advice. However, *the rule itself*—rather than a future guidance document—should outline detailed standards for the rare occasion where some portion of the data might not be widely available.

As part of this framework, the rule should provide clear and detailed guidelines and a formal process the agency must follow before deciding to exclude data of any kind from new or old

studies. Such guidelines should include a provision that only the *portions of data* that would trigger violations of privacy or confidential business secrets may be excluded, and that all other data and methodologies should be made transparent and publicly available. Before excluding data, EPA should conduct and make public a systematic review of the scientific literature to determine if there is an alternative study that could be used in place of any study whose data they cannot fully access. If such studies exist and data can be made accessible, EPA should rely on the alternative study or studies in lieu of those whose data is not available. If an alternative study or studies do not exist, then the Administrator should only use the original study if the researchers demonstrate that they explored all scientific methods and tools available to make the data available, but none proved sufficient to protect individual subjects' privacy and/or confidential business secrets.

If after such efforts the Administrator deems the original study essential for agency decision making, the agency should publish its systematic review along with a detailed analysis explaining the specific legal reasons why data must be excluded. The Administrator should further detail why the study authors were unable to provide the data in a format that allows reanalysis and replication without compromising privacy or confidential business secrets. It should also include details on how much of the data fall within the exclusion, since only those specific pieces of information should gain exclusion. Details about study methods, model assumptions, and the like should always be transparent. And before the agency uses the study, this agency justification should be subject to public comment.

If after public comment the Administrator still determines that the study is essential and some data cannot be released publicly, then the agency shall set up a system for limited release. The agency would then need to set up a committee of researchers to receive all the data subject to nondisclosure agreement. These researchers should include members with a wide range of perspectives on the issue to allow independent analysis, validation, and peer review. Should the original study authors refuse to provide the data under this closed, limited-access process, the study should be excluded from EPA decision making. Refusal of such limited independent validation renders the study largely meaningless from a policymaking perspective as there is no way to determine if the findings have a relationship to scientific reality.

Sincerely,

Angela Logomasini, Ph.D.  
Senior Fellow  
Competitive Enterprise Institute

---

<sup>1</sup> Angela Logomasini, "EPA Transparency Rule Will Bolster Science and Improve Rulemaking When Rules are Based on the Best Available Data, Regulations Are More Effective," CEI On Point 3246, July 17, 2018, <https://cei.org/sites/default/files/Angela%20Logomasini%20-%20EPA%20Transparency%20Rule%20Will%20Bolster%20Science%20and%20Improve%20Rulemaking%20%2081%29.pdf>.

<sup>2</sup> "More Bang for Your Byte," *Scientific Data*, Vol. 1 (May 27, 2014), p. 1, <https://www.nature.com/articles/sdata201410>.

---

<sup>3</sup> Chris Chambers and Brian Nosek, “The First Imperative: Science That Isn’t Transparent Isn’t Science,” *The Guardian*, June 25, 2015, <https://www.theguardian.com/science/head-quarters/2015/jun/25/the-first-imperative-science-that-isnt-transparent-isnt-science>.

<sup>4</sup> James L. Mills, “Data Torturing,” *New England Journal of Medicine*, Vol. 329 No. 16 (October 14, 1993), pp. 1196-1199, <https://www.nejm.org/doi/full/10.1056/NEJM199310143291613>.

<sup>5</sup> Monya Baker, “1,500 Scientists Lift the Lid on Reproducibility: Survey Sheds Light on the ‘Crisis’ Rocking Research,” *Nature*, Vol. 553, No. 7604 (May 25, 2016; corrected July 28, 2016): pp. 452-454, [https://www.nature.com/polopoly\\_fs/1.19970!/menu/main/topColumns/topLeftColumn/pdf/533452a.pdf](https://www.nature.com/polopoly_fs/1.19970!/menu/main/topColumns/topLeftColumn/pdf/533452a.pdf).

<sup>6</sup> David Randall and Christopher Welser, “The Irreproducibility Crisis of Modern Science: Causes, Consequences, and the Road to Reform,” National Association of Scholars, April 2018, [https://www.nas.org/images/documents/NAS\\_irreproducibilityReport.pdf](https://www.nas.org/images/documents/NAS_irreproducibilityReport.pdf).

<sup>7</sup> Ibid, p. 17.

<sup>8</sup> John P.A. Ioannidis, “Why Most Published Research Findings Are False,” *PLoS Medicine*. Vol. 2, No. (2005): p. e124, <https://doi.org/10.1371/journal.pmed.0020124>.

<sup>9</sup> Science Advisory Board (SAB) Consideration of the Scientific and Technical Basis of EPA’s Proposed Rule Titled Strengthening Transparency in Regulatory Science, April 24, 2020, p. 1, [https://yosemite.epa.gov/sab/sabproduct.nsf/LookupWebReportsLastMonthBOARD/2DB3986BB8390B308525855800630FCB/\\$File/EPA-SAB-20-005.pdf](https://yosemite.epa.gov/sab/sabproduct.nsf/LookupWebReportsLastMonthBOARD/2DB3986BB8390B308525855800630FCB/$File/EPA-SAB-20-005.pdf).

<sup>10</sup> Handbook: Good Laboratory Practices, 2nd ed. (Geneva: World Health Organization, 2009), <http://www.who.int/tdr/publications/documents/glp-handbook.pdf>.

<sup>11</sup> J. Kirk Smith, “A Comparison of the Guidance of FDA, OECD, EPA, and others on Good Laboratory Practice,” in Handbook of LC-MS Bioanalysis: Best Practices, Experimental Protocols, and Regulations, 1st Ed. Wenkui Li (Editor), Jie Zhang (Editor), Francis L. S. Tse (Editor) (Hoboken, NJ: Wiley, 2014).

July 17, 2018

No. 246

## **EPA Transparency Rule Will Bolster Science and Improve Rulemaking** **When Rules are Based on the Best Available Data, Regulations Are More Effective** *By Angela Logomasini, Ph.D.\**

Transparency is a cornerstone of the scientific process and a well-recognized goal long extolled within the scientific community. As a 2014 editorial in the British publication *Nature Geosciences* explained, “Science thrives on reproducibility.”<sup>1</sup> And reproducibility in science can only exist when researchers have access to two things: “full disclosure of the methods used to obtain and analyse data, and availability of the data that went into and came out of the analysis.”<sup>2</sup> A recently proposed Environmental Protection Agency (EPA) rule to promote transparency of scientific data underlying major regulations would help advance this goal in the realm of public policy.

Yet the EPA rule has proven highly controversial because of political and ideological influences. On one side are those who advocate stringent regulations that would err on the side of caution, and on the other are those who seek to streamline and reduce regulatory burdens. However, the rule’s ultimate goal is not to determine how much regulation we will have, but whether the regulations will be effective and necessary to achieve public health and environmental goals. By helping ensure the underlying science is valid, transparency will increase the probability that regulations will actually generate public health benefits and not unintentionally undermine public health and well-being.

Arguments against the rule focus on privacy concerns as well as doubts whether transparency of regulatory science is attainable, and appear to be based on misinformation about what the rule actually does. As the following shows, the rule is actually far more modest and flexible than depicted by its critics, and its goals are in fact achievable.

**Overview of the Transparency Rule.** In April 2018, then-EPA Administrator Scott Pruitt proposed the rule “Strengthening Transparency in Regulatory Science,”<sup>3</sup> which was modeled after the concepts underlying the Honest and Open New EPA Science Treatment (HONEST) Act (H.R. 1430, S. 1794).<sup>4</sup> The text of the rule is different from that of the HONEST Act, limiting the scope and allowing for exceptions.

The rule affords the EPA administrator considerable leeway to permit regulators to use research in cases where privacy or other concerns limit public availability. In fact, under some laws, such as the newly reformed Toxic Substances Control Act, the EPA must use such research if it constitutes the “best available science” on an issue. In that case, even if data were not fully available, the agency would still be required to rely on those critical

---

\* *Angela Logomasini is a senior fellow at the Competitive Enterprise Institute.*

studies. However, in cases where data can be more transparent without privacy concerns, the EPA could not refuse to release the data on arbitrary grounds.

A few key features of the rule include the following:

- **The scope of science to which the rule applies is limited.** The rule covers “dose response data and models underlying pivotal regulatory science that are used to justify significant regulatory decisions regardless of the source of funding or identity of the party conducting the regulatory science.”<sup>5</sup> Basically, that means the rule covers the key scientific studies the EPA uses to justify major regulations—those expected to impose costs of \$100 million or more a year or that are expected to have other significant adverse impacts.<sup>6</sup> Accordingly, it does not cover every EPA regulatory activity, such as routine permitting decisions, voluntary initiatives programs, or research programs. However, any EPA research used to justify major rules would need to meet transparency standards.
- **The rule protects privacy and confidential business secrets.** It specifically directs the EPA to ensure data it relies upon is “publicly available in a manner sufficient for independent validation,” and to do so “in a fashion that is consistent with law, protects privacy, confidentiality, confidential business information, and is sensitive to national and homeland security.”
- **The rule allows the EPA to rely on research that is not publicly available in certain situations.** The rule specifically stipulates that the EPA can still use research that the agency cannot make publicly available in cases when it is “not feasible” to do so in a way that is, again, “consistent with law, protects privacy and confidentiality, and is sensitive to national and homeland security.”

**Improving transparency is an achievable goal.** Transparency in scientific research has never been particularly controversial, and ongoing efforts to improve transparency prove that greater transparency is achievable. As the previously cited *Nature Geoscience* editorial proclaimed: “The question is no longer whether research data should be shared, but how to make effective data sharing a common and well-rewarded part of research culture.”<sup>7</sup>

Similarly, scientists Chris Chambers of Cardiff University and Brian Nosek of the University of Virginia explain in the *Guardian*:

Transparency and reproducibility are the beating heart of the scientific enterprise. Transparency ensures that all aspects of scientific methods and results are available for critique, compliment, or reuse. This not only meets a social imperative, it also allows others to test new questions with existing data, makes it easier to identify and correct errors, and helps unmask academic fraud. Transparent practices such as sharing data and computer code, in turn, safeguard reproducibility: the idea that for a scientific observation to count as a discovery it must reveal something real and repeatable about the natural world.<sup>8</sup>

Yet much scientific research does not achieve the goals of transparency for a number of reasons. Chambers and Nosek explain that transparency has become “undervalued” because researchers want to get published, and “the fierce competition of academia rewards those who secure large grants and publish innovative—if tentative—findings in prestigious journals, not scholars who instead focus on being transparent and careful.”<sup>9</sup> As a result, researchers often avoid transparency out of fear that their novel findings may eventually be proven wrong when someone tries to replicate their work.

This reality has fostered both unintentional bias and scientific mischief, including the propensity for researchers to work the data until it generates a positive finding. As James Mills of the National Institute of Child Health and Human Development lamented back in 1993 in the *New England Journal of Medicine*: “‘If you torture your data long enough, they will tell you whatever you want to hear’ has become a popular observation in our office.”<sup>10</sup> The scope of this problem<sup>11</sup> is also detailed by the National Association of Scholars in its recent report, *The Irreproducibility Crisis of Modern Science*.<sup>12</sup> The authors highlight one outrageous case where Professor Brian Wansink, head of Cornell University’s food behavior research lab, literally bragged in a blog post that he had schooled one of his students on how to churn data to generate positive results and get them published.<sup>13</sup>

Even when researchers do not torture data, many positive associations will occur by mere chance or unintentional biases. Stanford Professor of Medicine John Ioannidis demonstrated in a research article in 2005 that most published research findings are false positives. He explained: “Simulations show that for most study designs and settings, it is more likely for a research claim to be false than true. Moreover, for many current scientific fields, claimed research findings may often be simply accurate measures of the prevailing bias.”<sup>14</sup>

There are a number of other factors that reduce transparency even among researchers who strive to meet high standards. In an article for *Nature*, the authors point out that researchers are sometimes reluctant to share data for three reasons:

First, high-quality data are hard to obtain. Researchers should expect their fieldwork to yield one or several primary publications. The originality of these publications might be jeopardized if the data are widely available before they are published, or if individual data sets are amalgamated in large collective synthesis publications. The reluctance towards making hard-earned data publicly available is understandable—at least as long as the measurements have not been sufficiently exploited by those who obtained them.

Second, data are context dependent. Without appropriate contextual information, for example metadata regarding locations, methods and shortcomings, they can easily be misinterpreted and misused. Opening data sets to other researchers means that the circumstances of collection—known by those who performed the measurements—need to be carefully recorded and communicated.

The third factor is related to the previous one: significant effort is often necessary to prepare the data for reuse before they can be made available to the scientific community. In addition to contextualizing data, formats may have to be adjusted and a point of contact may be necessary in case there are questions.<sup>15</sup>

However, these challenges are not an excuse for continued lack of transparency. Rather, the *Nature* article authors make recommendations to change the incentives. For example, journals should require data transparency for articles they publish, and government agencies should also require it for any study they fund. In addition, they note, the scientific community needs to develop systems to reward researchers for releasing data. One way to reward those who release data involves creating systems to credit the researchers who release data. When others use the data and cite it, the researchers who generated it gain stature and credibility within their fields. One example is the creation of *Scientific Data*, an entire publication—launched by Nature Research in 2014—focused on publishing data sets to reward those who share that data.<sup>16</sup>

The Charlottesville-based Center for Open Science (COS), led by Nosek, is also taking a lead by creating open science guidelines and systems to reward researchers who follow them. “Our mission is to increase openness, integrity, and reproducibility of research,” COS explains on its website. “We envision a future scholarly community in which the process, content, and outcomes of research are openly accessible by default.”<sup>17</sup>

Some COS initiatives include:

- **Registered Reports.** This program promotes a process that journals can use to essentially peer review a study design before data is collected. COS explains that registered reports “emphasize the importance of the research question and the quality of the methodology by conducting peer review prior to data collection. High-quality protocols are then provisionally accepted for publication if the authors follow through with the registered methodology.”<sup>18</sup> The Open Science website says that 108 journals apply this approach.
- **Transparency and Openness Promotion (TOP) Guidelines.** TOP guidelines “require open data, research materials, and analytic code, while encouraging disclosure of additional steps that can increase the credibility of inferential research.”<sup>19</sup>
- **Preregistration of Research Plans.** This program encourages researchers to preregister their research plan with COS before beginning to collect data. This approach reduces risk that researchers might “work” data to get positive results. Instead, COS notes, “you’re simply committing to your plan in advance.”
- **Badges.** Under this program, journals can award COS-designed badges to researchers who follow certain transparency protocols. When a researcher follows specified transparency guidelines, a badge appears along with his or her article. These badges add credibility to a researcher’s published works, providing a strong incentive for researchers to follow standards for open science.<sup>20</sup>

Yet, challenges remain, including concerns about the privacy of human research subjects. The TOP guidelines, for example, state:

In rare cases, despite authors' best efforts, some or all data or materials cannot be shared for legal or ethical reasons. In such cases, authors must inform the editors at the time of submission. This will be taken into account during the review process. Authors are encouraged to anticipate data and material sharing at the beginning of their projects to provide for these circumstances. It is understood that in some cases access will be provided under restrictions to protect confidential or proprietary information.<sup>21</sup>

The TOP guidelines highlight that exceptions to data release may be necessary, but only in "rare cases." Such exceptions are not the rule or the goal. In fact, in the interest of making research more valuable, transparency guidelines will facilitate ways to release data as much as possible. Planning ahead—before collecting data and finalizing study design—should help researchers collect and code data in ways that protect privacy and allow legal release of data sets once they are complete. As COS notes on its website: "Requiring more transparent research practices, with reasonable exceptions for difficult situations, is possible."<sup>22</sup>

**Facts and Misperceptions about the Transparency Rule.** Robert Hahn, visiting professor at Oxford University and senior fellow at the Brookings Institution, addressed criticisms of the rule in a *Washington Post* op-ed appropriately titled, "Many mocked this Scott Pruitt proposal. They should have read it first." First, he detailed why the rule is needed:

Taking steps to increase access to data, with strong privacy protections, is how society will continue to make scientific and economic progress and ensure that evidence in rule-making is sound. The EPA's proposed rule follows principles laid out in 2017 by the bipartisan Commission on Evidence-Based Policymaking—humility, transparency, privacy, capacity and rigor—and moves us toward providing greater access to scientific data while protecting individual privacy.<sup>23</sup>

Hahn had served on the U.S. Commission on Evidence-Based Policymaking, which was formed pursuant to the Evidence-Based Policymaking Commission Act of 2016.<sup>24</sup> Sponsored by Rep. Paul Ryan (R-Wisc.) and signed by President Barack Obama on March 30, 2016, it set up a bipartisan commission of 15 experts "to study and develop a strategy for strengthening government's evidence-building and policymaking efforts,"<sup>25</sup> and produce a report within 18 months. Published in September 2017, the commission report addresses privacy issues, finding that transparency of data can be achieved without compromising privacy concerns:

Traditionally, increasing access to confidential data presumed significantly increasing privacy risk. The Commission rejects that idea. The Commission believes there are steps that can be taken to improve data security and privacy protections beyond what exists today, while increasing the production of evidence. Modern technology and statistical methods, combined with transparency and a strong legal

framework, create the opportunity to use data for evidence building in ways that were not possible in the past.<sup>26</sup>

The commission report provides detailed recommendations on how such data can be collected and managed to ensure privacy and transparency. Such efforts can inform regulators as they move forward with implementing the EPA's transparency rule, which complements similar initiatives already underway in the private sector. Private researchers who want their research to inform regulators should follow such transparency guidelines as well.

Unfortunately, some researchers and transparency advocates seem to misunderstand how the rule would work. For example, Ioannidis says: "A new standard currently proposed for the Environmental Protection Agency aims to ban the use of scientific studies for regulatory purposes unless all their raw data are widely available in public and can be reproduced. If the proposed rule is approved, science will be practically eliminated from all decision-making processes. Regulation would then depend uniquely on opinion and whim."<sup>27</sup> And even Nosek appears critical of the transparency rule, recently tweeting a *Washington Post* article critical of the rule, with the comment: "WaPo: 'Thanks a lot, Scott Pruitt. You've made transparency a bad word.'"<sup>28</sup>

Hahn articulately addresses such criticisms in his *Washington Post* op-ed:

Here's what the rule would actually do. First, it would require the EPA to identify studies that are used in making regulatory decisions. Second, it would encourage studies to be made publicly available "to the extent practicable." Third, it would define "publicly available" by listing examples of information that could be used for validation, such as underlying data, models, computer code, and protocols. Fourth, the proposal recognizes not all data can be openly accessible in the public domain and that restricted access to some data may be necessary. Fifth, it would direct the EPA to work with third parties, including universities and private firms, to make information available to the extent reasonable. Sixth, it would encourage the use of efforts to de-identify data sets to create public-use data files that would simultaneously help protect privacy and promote transparency. Seventh, the proposal outlines an exemption process when compliance is "impracticable." Finally, it would direct the EPA to clearly state and document assumptions made in regulatory analyses.

Here's what the EPA's rule *wouldn't* do: nullify existing environmental regulations, disregard existing research, violate confidentiality protections, jeopardize privacy or undermine the peer-review process.<sup>29</sup>

Advocates for transparency in science should applaud the EPA's proposed rule. While much of the work can and should be done in the private sphere, government policy can provide better incentives to researchers on whose work they rely. Ironically, in the same article where he claims the transparency rule would "ban science," Ioannidis makes some good points that actually support the need for the proposed transparency rule. He notes:

In the USA and elsewhere, governments are major funders of research and their regulatory mandates provide powerful incentives for best science. Making widely applicable, reproducible research practices and sharing the default option for research (with sparse exceptions, when appropriately justified) will strengthen scientific investigation and maximize its benefits to society at large. Governments can bolster their legacy through such initiatives and scientists would be broadly supportive of such a transformative vision to promote a standard of openness in science.<sup>30</sup>

**Efforts to Undermine Transparency.** Many critics of EPA’s transparency rule claim it is “anti-science” and represents an ideological attack on regulation.<sup>31</sup> It is true that those who prefer less regulation hope that the rule would eliminate unnecessary regulations that are based on poor-quality science. And it is also true that many oppose the rule because they fear it will weaken regulation. But irrespective of these ideological views, increasing transparency in science, whether used for government regulation or not, is an inherently pro-science goal.

Consider a key argument against the transparency rule, as reported in a May 1, 2018 *Science* magazine news story, with the headline, “Critics allege EPA’s new transparency rule has hidden pro-industry agenda.” According to these critics, the rule is designed to undermine clean air regulations related to airborne particles smaller than 2.5 micrometers in diameter, known as PM2.5. One key study with unavailable data is the Harvard University Six Cities Study, which the EPA uses to justify very stringent air quality regulations.<sup>32</sup> This 1993 study is a statistical analysis that reported an association between the life spans of people in six cities and the levels of PM2.5 found in the air. It claims that people who live in cities with higher PM2.5 have a life expectancy that is two to three years shorter than those in cities with lower PM2.5 levels.<sup>33</sup>

Harvard researchers maintain that the study participants—whose medical information is part of that data—never agreed for their information to be released. And some critics say it is too difficult, if not impossible, to release the data and keep the subjects’ identities completely anonymous.<sup>34</sup> Whether that is a legitimate concern remains to be seen, but it is not a good argument against the transparency rule.

Privacy concerns *might* be a legitimate challenge for releasing some or all of the Six Cities data. If that is the case, the rule, as noted, provides exemptions for rare cases where data cannot be made anonymous and privacy must be maintained. Accordingly, regulators can still use the Six Cities data, if legitimate privacy concerns prevent full release.

In cases where the data can be made anonymous, it should be released regardless of whether it supports weakening or strengthening regulations. After all, if a study’s findings are valid, releasing the data will only strengthen claims about the benefits of these regulations. If the findings are not valid, then we know that regulatory costs may not be justified and that society actually suffers net negative effects because of those costs. Indeed, regulation can translate into higher prices for food, transportation, consumer products, and even

medicines. The debate over the rule is not about whether it benefits industry or not, but about how it impacts public health and well-being overall.

According to the most recent Office of Management and Budget (OMB) report on the cost of federal regulations, EPA rules are the most expensive in the federal government. OMB explains “the large estimated benefits of EPA rules” emanate from the EPA air quality program and are “mostly attributable to” reductions in PM<sub>2.5</sub>.<sup>35</sup> In the interest of public health and well-being, policy makers have an obligation to do their best to ensure that the science actually demonstrates those benefits; data access and replication are the gold standards to ensure it does.

Activists have also cited costs imposed on industries it seeks to regulate. For example, the Union of Concerned Scientists recently released EPA internal emails, obtained via a Freedom of Information Act request, that some say indicate the rule will harm businesses.<sup>36</sup> Yet the only supposed evidence they show is that EPA Office of Chemical Safety and Pollution Prevention Deputy Administrator Nancy Beck had indicated via email that she was concerned about costs to industry and potential challenges with the pesticide and chemical approval processes. It appears that Beck simply recommended revisions to a draft version to address such potential concerns. Certainly, it is reasonable to avoid unnecessary barriers to innovation, but transparency should not be sacrificed because it is inconvenient.

Another criticism comes from a staff writer for *The Atlantic*, who warns that the costs associated with compliance are simply too high, pointing to a Congressional Budget Office report that said implementing the HONEST Act would cost \$250 million a year, as the EPA would have to devote staff to removing confidential information from data.<sup>37</sup> Even if true, this is a relatively modest amount for the agency, given that EPA regulations cost around \$394 billion a year, according to estimates developed by Clyde Wayne Crews of the Competitive Enterprise Institute.<sup>38</sup> Moreover, the EPA has plenty of funds within its \$8 billion budget to pay for transparency. Heritage Foundation analyst Diane Katz suggests a number of programs that could easily be cut to save the agency billions of dollars a year.<sup>39</sup> In the end, transparency might reduce costs for businesses and consumers, who ultimately pay the high costs of misguided regulations.

Apparently, many of these criticisms are based on misunderstandings about how the rule would work and what it contains. Much of the hype is from those who fear that transparency will weaken regulations, but whether it makes regulations more extensive or scales them back is not the point. The goal is to ensure that the best science informs regulators so that they can design rules to best protect public health. Promoting the fundamentals of good science—transparency and reproducibility—will help achieve that goal.

**Conclusion.** The EPA rule on transparency should be viewed as yet another tool in a larger effort to promote transparency and reproducibility within the scientific community, most of which is happening outside of government. Ultimately, the value and credibility of scientific research rests on the community’s ability to rein in questionable practices that

undermine the scientific discovery process. Much work remains to be done, because the problem spans a wide range of scientific disciplines.

The EPA's transparency rule will help bolster private pro-transparency efforts and should be the beginning of a government-wide effort. Transparency guidelines could also be applied to government research and government-funded research within all agencies, from the EPA to the National Institutes of Health, the Food and Drug Administration, and the Consumer Product Safety Commission.

Arguments against the rule that claim that it constitutes an assault on science, would violate privacy, or would prevent useful research from informing regulators are without merit. The rule simply supports fundamental principles of scientific research. It contains provisions to protect privacy and provide exceptions that allow regulators to consider quality research that is not fully transparent because of privacy concerns. It will help ensure that regulators are informed by the best science so they can design rules that protect public health while not imposing counter-productive burdens on society.

## Notes

---

<sup>1</sup> "Towards Transparency," *Nature Geoscience*, Vol. 7 (October 30, 2014), p. 777, <https://www.nature.com/articles/ngeo2294>.

<sup>2</sup> Ibid.

<sup>3</sup> Ibid.

<sup>4</sup> 115<sup>th</sup> Congress, 1<sup>st</sup> Session, H. R. 1430, Report No. 115-59, <https://www.congress.gov/115/bills/hr1430/BILLS-115hr1430rh.pdf>.

<sup>5</sup> *Federal Register*, Vol. 83, No. 83 (April 30, 2018), pp. 18768-18774, <https://www.gpo.gov/fdsys/pkg/FR-2018-04-30/pdf/2018-09078.pdf>.

<sup>6</sup> "Significant Regulatory Actions" are defined by the Office of Management and Budget's Executive Order 12866, *Federal Register*, Vol. 58, No. 190 (October 4, 1993), <https://www.archives.gov/files/federal-register/executive-orders/pdf/12866.pdf>; see section 3(f) for more details about other adverse impacts that make a regulatory action "significant."

<sup>7</sup> "More Bang for Your Byte," *Scientific Data*, Vol. 1 (May 27, 2014), p. 1, <https://www.nature.com/articles/sdata201410>.

<sup>8</sup> Chris Chambers and Brian Nosek, "The first imperative: Science that isn't transparent isn't science," *The Guardian*, June 25, 2015, <https://www.theguardian.com/science/head-quarters/2015/jun/25/the-first-imperative-science-that-isnt-transparent-isnt-science>.

<sup>9</sup> Ibid.

<sup>10</sup> James L. Mills, "Data Torturing," *New England Journal of Medicine*, Vol. 329 No. 16 (October 14, 1993), pp. 1196-1199, <https://www.nejm.org/doi/full/10.1056/NEJM199310143291613>.

<sup>11</sup> Monya Baker, "1,500 Scientists Lift the Lid on Reproducibility: Survey Sheds Light on the 'Crisis' Rocking Research," *Nature*, Vol. 553, No. 7604 (May 25, 2016; corrected July 28, 2016): pp. 452-454, [https://www.nature.com/polopoly\\_fs/1.19970!/menu/main/topColumns/topLeftColumn/pdf/533452a.pdf](https://www.nature.com/polopoly_fs/1.19970!/menu/main/topColumns/topLeftColumn/pdf/533452a.pdf)

<sup>12</sup> David Randall and Christopher Welser, "The Irreproducibility Crisis of Modern Science: Causes, Consequences, and the Road to Reform," National Association of Scholars, April 2018, [https://www.nas.org/images/documents/NAS\\_irreproducibilityReport.pdf](https://www.nas.org/images/documents/NAS_irreproducibilityReport.pdf).

<sup>13</sup> Ibid, p. 17.

<sup>14</sup> John P.A. Ioannidis, "Why Most Published Research Findings Are False," *PLoS Medicine*. Vol. 2, No. (2005): p. e124, <https://doi.org/10.1371/journal.pmed.0020124>.

<sup>15</sup> Jens Kattge, Sandra Díaz and Christian Wirth, "Of carrots and sticks," *Nature Geoscience*, Vol. 7 (2014), pp. 778-779, <https://www.nature.com/articles/ngeo2280>.

<sup>16</sup> "More Bang for Your Byte."

<sup>17</sup> Center for Open Science, Mission Statement, accessed June 18, 2018, <https://cos.io/about/mission>.

- 
- <sup>18</sup> Center for Open Science, Registered Reports, accessed June 18, 2018, <https://cos.io/rr>.
- <sup>19</sup> Center for Open Science, TOP Guidelines, accessed June 18, 2018, <https://cos.io/our-services/top-guidelines>.
- <sup>20</sup> Center for Open Science, Open Science Badges, accessed June 18, 2018, <https://cos.io/our-services/open-science-badges>.
- <sup>21</sup> Center for Open Science, “Guidelines for Transparency and Openness Promotion (TOP) in Journal Policies and Practices: ‘The TOP Guidelines’,” Version 1.0.0, published in the Open Science Framework, 2014, [https://osf.io/ud578/?\\_ga=2.170903520.114412219.1528480552-1647426821.1528480552](https://osf.io/ud578/?_ga=2.170903520.114412219.1528480552-1647426821.1528480552). The guidelines were developed at a November 2014 workshop held in Charlottesville, Virginia, November 3-4, 2014. For details and contributors see: <https://docs.google.com/document/d/1JvswEe5X0aCY02zyCoyiFbdKlij24xUnuO2Qa6HJIo/edit>.
- <sup>22</sup> Center for Open Science, “The TOP Guidelines were created by journals, funders, and societies to align scientific ideals with practices,” accessed July 6, 2018, <https://cos.io/our-services/top-guidelines>.
- <sup>23</sup> Robert Hahn, “Many mocked this Scott Pruitt proposal. They should have read it first,” *Washington Post*, May 10, 2018, [https://www.washingtonpost.com/opinions/many-mocked-this-scott-pruitt-proposal-they-should-have-read-it-first/2018/05/10/31baba9a-53c2-11e8-abd8-265bd07a9859\\_story.html?noredirect=on&utm\\_term=.d2637e2cc2a1](https://www.washingtonpost.com/opinions/many-mocked-this-scott-pruitt-proposal-they-should-have-read-it-first/2018/05/10/31baba9a-53c2-11e8-abd8-265bd07a9859_story.html?noredirect=on&utm_term=.d2637e2cc2a1).
- <sup>24</sup> Public Law No: 114-140, <https://www.congress.gov/114/plaws/publ140/PLAW-114publ140.pdf>.
- <sup>25</sup> The Promise of Evidence-Based Policymaking: Report of the Commission on Evidence-Based Policymaking, September 2017, p. 1, <https://www.cep.gov/content/dam/cep/report/cep-final-report.pdf>, p. 1.
- <sup>26</sup> Ibid.
- <sup>27</sup> John P. A. Ioannidis, “All Science Should Inform Policy and Regulation,” *PLoS Medicine*, Vol. 15, No. 5 (May 3, 2018): e1002576, <http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1002576>.
- <sup>28</sup> Brian Nosek, Twitter, April 6, 2018, 5:12 pm, <https://twitter.com/BrianNosek/status/982410695663869954>.
- <sup>29</sup> Hahn.
- <sup>30</sup> Ioannidis, “All Science Should Inform Policy and Regulation.”
- <sup>31</sup> Sid Shapiro, “EPA ‘Transparency’ Rule Confuses Science and Regulatory Science,” *The Hill*, May 15, 2018, <http://thehill.com/opinion/energy-environment/387831-epa-transparency-rule-epa-confuses-science-and-regulatory-science>.
- <sup>32</sup> Douglas W. Dockery, C. Arden Pope, Xiping Xu, John D. Spengler, James H. Ware, Martha E. Fay, Benjamin G. Ferris, Jr., and Frank E. Speizer, “An Association between Air Pollution and Mortality in Six U.S. Cities,” *New England Journal of Medicine*, Vol. 329, No. 4 (December 9, 1993): pp. 1753-1759, <https://www.nejm.org/doi/full/10.1056/NEJM199312093292401>.
- <sup>33</sup> Warren Cornwall, “Critics Allege EPA’s New Transparency Rule Has Hidden Pro-Industry Agenda,” *Science*, May 1, 2018, <http://www.sciencemag.org/news/2018/05/critics-allege-epa-s-new-transparency-rule-has-hidden-pro-industry-agenda>.
- <sup>34</sup> Amelia Urry, “The EPA is Jeopardizing Scientific Research and Privacy in the Name of ‘Transparency,’” *Popular Science*, April 27, 2018, <https://www.popsci.com/epa-transparency-public-health-data>.
- <sup>35</sup> Office Of Management and Budget, Office of Information and Regulatory Affairs, *2017 Draft Report to Congress on the Benefits and Costs of Federal Regulations and Agency Compliance with the Unfunded Mandates Reform Act*, [https://www.whitehouse.gov/wp-content/uploads/2017/12/draft\\_2017\\_cost\\_benefit\\_report.pdf](https://www.whitehouse.gov/wp-content/uploads/2017/12/draft_2017_cost_benefit_report.pdf).
- <sup>36</sup> Scott Waldman and Niina Heikkinen, “Trump’s EPA Wants to Stamp Out ‘Secret Science,’ Internal Emails Show it is Harder Than Expected,” *E&E News*, April 20, 2018, <http://www.sciencemag.org/news/2018/04/trump-s-epa-wants-stamp-out-secret-science-internal-emails-show-it-harder-expected>.
- <sup>37</sup> Ed Yong, The Transparency Bills That Would Gut the EPA, *The Atlantic*, March 15, 2017, <https://www.theatlantic.com/science/archive/2017/03/how-to-gut-the-epa-in-the-name-of-honesty/519462>.
- <sup>38</sup> Wayne Crews, *10,000 Commandments: An Annual Snapshot of the Federal Regulatory State, 25th Anniversary Edition*, Competitive Enterprise Institute, 2018, <https://cei.org/10kc2018>.
- <sup>39</sup> Diane Katz, “How Trump Can Clean up the EPA’s Budget,” *Daily Signal*, March 01, 2017, <https://www.dailysignal.com/2017/03/01/how-trump-can-clean-up-the-epas-budget>.