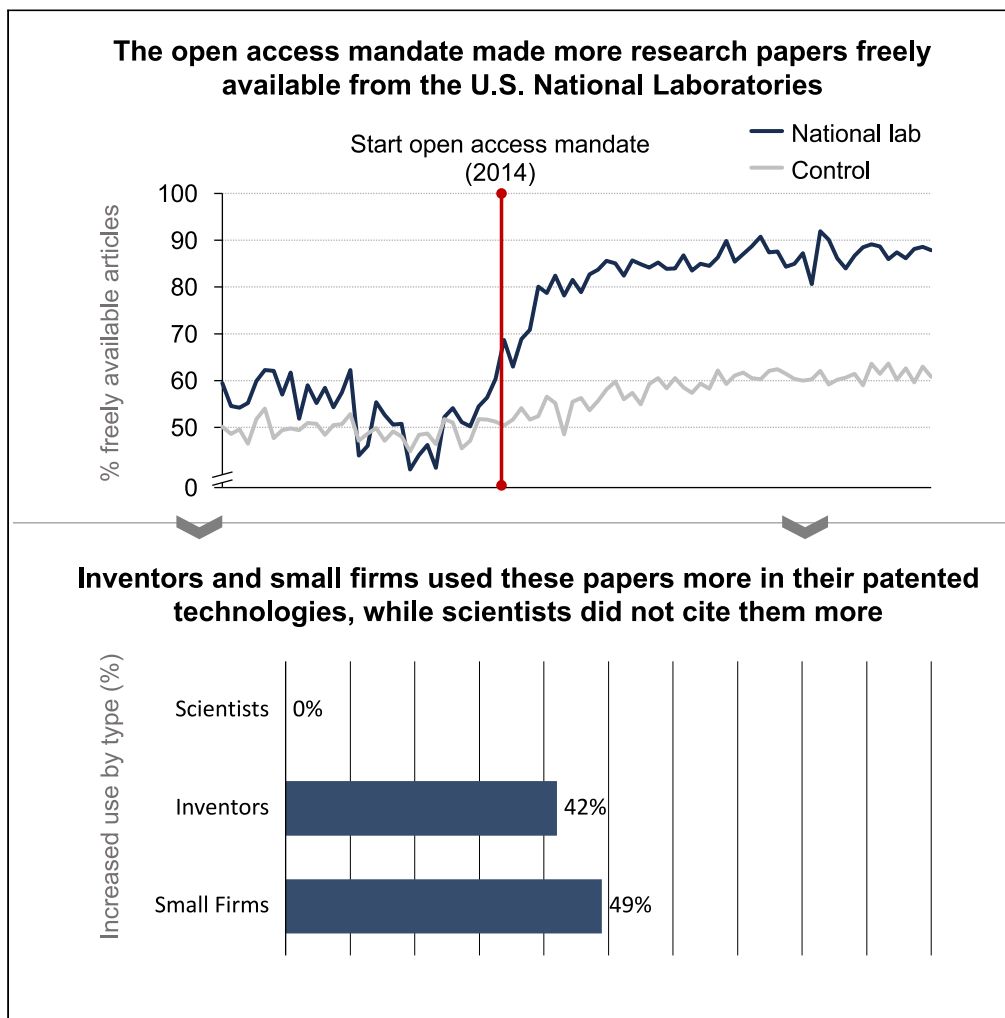


Article

# The impact of open access mandates on scientific research and technological development in the U.S.



Benedict Probst, Paul M. Lohmann, Andreas Kontoleon, Laura Díaz Anadón

bprobst@ethz.ch

**Highlights**

We examine the impact of the U.S. Department of Energy's open-access mandate

Scientific articles subject to the mandate were utilized on average 42% more in patents

Articles subject to the mandate were not cited more frequently by other academic papers

Small firms were the primary beneficiaries of the increased knowledge diffusion



## Article

# The impact of open access mandates on scientific research and technological development in the U.S.

Benedict Probst,<sup>1,3,5,\*</sup> Paul M. Lohmann,<sup>2,3</sup> Andreas Kontoleon,<sup>3</sup> and Laura Díaz Anadón<sup>3,4</sup>

**SUMMARY**

**Getting to a net-zero emissions economy requires faster development and diffusion of novel clean energy technologies. We exploit a rare natural experiment to study the impact of an open-access mandate on the diffusion of scientific research into patented technologies. From 2014 onwards, the U.S. Department of Energy (DOE) required its 17 National Laboratories (NLs) to publish all peer-reviewed scientific articles without a paywall. Using data from more than 300,000 scientific publications between 2012 and 2018, we show that scientific articles subject to the mandate were used on average 42% more in patents, despite embargo periods of up to 12 months. We also show that articles subject to the mandate were not cited more frequently by other academic articles. Our findings suggest that the mandate primarily contributed to technological development but has not led to additional academic research. Lastly, we show that small firms were the primary beneficiaries of the increased diffusion of scientific knowledge.**

**INTRODUCTION**

Achieving net-zero emissions by the second half of this century as envisaged by the Paris Agreement<sup>1</sup> requires—among other things—faster development and diffusion of novel clean energy technologies.<sup>2</sup> One of the factors that could shape the impact of scientific research on technological development is the extent to which inventors globally have knowledge about and access to those publications. As the predominant model of scientific publishing continues to rely on a subscription-based system, inventors without access to major journals paid by their universities, laboratories, or companies may have difficulties accessing scientific publications.<sup>3</sup> While the number of open access (OA) journals, publications in paywalled journals published as OA (so-called “hybrid OA”), and articles made available in repositories (so-called “green OA”) has increased substantially over the last decade, the most highly cited academic journals maintain a subscription-based model of publishing.<sup>4</sup>

For inventors, the importance of accessing scientific literature has grown over time as scientific research is playing an increasingly critical role in shaping technological progress. The share of U.S. patents reliant on federally funded science has increased steadily over the last century, with nearly one-third of patented inventions in the United States (U.S.) citing federally funded science since the 2000s. In the U.S., federal agencies play a key role in funding and conducting scientific research like the federal executive departments (e.g., defense, health, and human services), the National Science Foundation, and NASA.<sup>5</sup> In the energy domain, the Department of Energy (DOE) plays a key role in driving scientific and technological progress. For instance, the DOE directly supported around 4% of all U.S. inventions in 2017.<sup>6</sup> The DOE’s National Labs (NLs) are considered the “crown jewels” of innovation and form a critical part of the U.S. energy innovation sphere. Yet, reviews of the NL’s performance have highlighted several areas of weak organizational design, specifically on accelerating the outward diffusion of scientific advances to industry.<sup>7</sup>

Over the past five years, several universities and other public research institutions ranging from the University of California to the Max Planck Institute have voiced increasing discontent with the subscription-based system adopted by most academic journals. Reasons for this resistance include perceptions of unfair pricing and the assertion that the results of publicly funded science should be available to the public.<sup>8</sup> The most wide-ranging OA initiative unveiled in 2018 is dubbed “Plan-S”, which requires scientists receiving funding from European and other research organizations controlling EUR 33 billion in annual funding to make their research directly available upon publication from 2021 onward, at least in a repository.<sup>9,10</sup> Recently, the U.S. has followed suit—based on guidance by the White House Office of Science and Technology Policy<sup>11</sup>—by announcing that government-funded science needs to be freely available upon publication from 2026 onwards.<sup>12</sup>

As scientific journals provide a central diffusion pathway of scientific knowledge to firms,<sup>13–16</sup> understanding the effect of journal access on technological development is critical. Yet, the relationship between the degree of accessibility of a scientific publication and its subsequent impact on technological applications—in other words, commercial and wider economic impact—proxied by patents is, however, unclear

<sup>1</sup>Group for Sustainability and Technology, Department of Management, Technology, and Economics, ETH Zurich

<sup>2</sup>El-Erian Institute of Behavioural Economics and Policy, Judge Business School, University of Cambridge, UK

<sup>3</sup>Centre for Environment, Energy, and Natural Resource Governance, Department of Land Economy, University of Cambridge

<sup>4</sup>Harvard Kennedy School, Harvard University, United States

<sup>5</sup>Lead contact

\*Correspondence: [bprobst@ethz.ch](mailto:bprobst@ethz.ch)

<https://doi.org/10.1016/j.isci.2023.107740>



(except for Bryan and Ozcan (2021) who study the effect of OA on inventions in the biomedical sector). Understanding the commercialization of energy R&D is particularly important as the research needs to ultimately be applied and further developed by companies (in contrast to military R&D where the government is both the producer and “purchaser” of novel technologies).

We exploit a rare natural experiment to study the impact of the increased availability of scientific research on the diffusion of scientific knowledge into patented technologies. The U.S. DOE mandated that all scientific, peer-reviewed articles resulting from its 17 NLs must be made available without a paywall within a year if published on or after the 1<sup>st</sup> of October 2014.

The mandate allows us to study three interrelated questions. First, we analyze how NL researchers changed their publication strategy based on the mandate. If researchers choose to deposit accepted manuscripts in free repositories, many journals impose an embargo period, which is the time lag before it can be made publicly available in a repository. In our sample, 11 out of 16 paywalled journals had embargo periods of at least 6 months, thereby increasing access hurdles to the diffusion of research. If researchers pay for gold open access, this could divert funds away from key research activities as open access fees of 10,000 USD per article are not unusual.

Second, we analyze whether treated articles (i.e., those emanating from the NLs subject to the OA mandate) were used more in other scientific articles and patented technologies. We investigate whether articles subject to the mandate were cited more by other academic research and patents (we use the latter as a proxy for technological development). Lastly, we investigate whether the mandate primarily helped those firms with historically low access to scientific research (e.g., small firms).

Exploring whether small firms are affected by the mandate is critical for several reasons. Research scientists at universities or large, R&D-intensive firms typically have a greater awareness, access, and ability to effectively exploit scientific knowledge originating outside of their direct research vicinity. In contrast, non-academic professionals and small-and-medium-sized enterprises (SMEs) are more likely to struggle with effective knowledge exploitation. Ware and Monkman<sup>17</sup> found that half of surveyed small businesses had difficulty accessing academic research useful to their business. Houghton et al.<sup>15</sup> showed that 68% of Danish firms reported access difficulties to scientific research. Galasso and Schankerman<sup>18</sup> found that particularly small firms benefit from patent invalidation from large firms. Lastly, Doblinger et al.<sup>19</sup> showed that small firms are more resource-constrained than larger firms and therefore benefit substantially from joint development with government partners in terms of patent production. Therefore, research shows that access, internal capabilities, and external support shape the ability of firms to exploit relevant scientific knowledge.<sup>20</sup>

To explore the effect of this mandate, we first build a unique dataset of more than 300,000 scientific articles from the top 20 journals (in terms of the number of DOE-funded publications) in which DOE-funded research has appeared according to the DOE’s own PAGES database. We then check for each article whether it is available without a paywall<sup>21</sup> and whether and how often it got cited by patents. Finally, we use difference-in-difference(s) econometric modeling in line with the literature<sup>22,23</sup> to explore the impact of the exogenous change to OA availability of NL articles on scientific research and patented inventions.

### Case selection: The DOE public access plan

The public sector in the U.S. plays a distinct role in advancing innovation, particularly through different ways of funding research and development (R&D) activities.<sup>5,24</sup> In 2010, 31% of total R&D expenditures in the U.S. were funded by federal agencies, which has fallen to 19% in 2020.<sup>25</sup> Yet, of overall government-funded R&D, one-third is conducted by federal agencies in its federally funded research centers (FFRDCs). Among the FFRDCs are 16 of the 17 NLs, commonly called “the Labs”. The NLs span the geography of the U.S. and are active in areas including nuclear weapons, fundamental research, and applied energy research (Figure 1). These received around 19 USD billion in funding in 2021<sup>26</sup> and are considered the “crown jewels” of innovation and form a critical part of the energy innovation sphere of the U.S.<sup>7</sup>

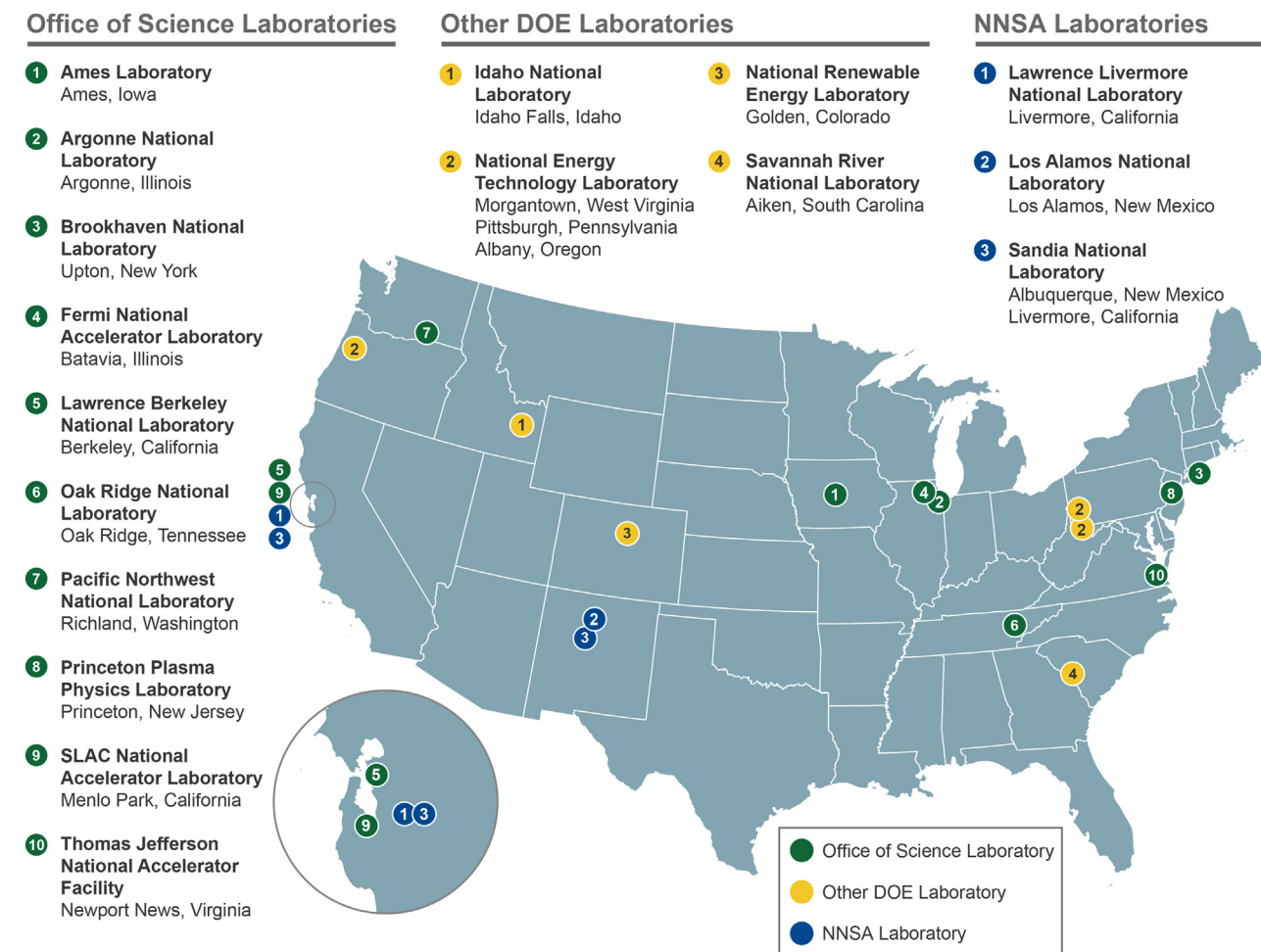
Reviews of the NL’s performance have highlighted several areas of weak organizational design, specifically on technology transfer. Most troubling is the declining engagement with the private sector, which ultimately must adopt, develop, and deploy the energy technologies invented with public R&D. Anadon et al.<sup>7</sup> document a range of declining technology transfer outcomes over the last two decades, such as NL invention disclosures and patent grants.

A potential avenue to increase the uptake of scientific advances in the private sector is to lower access costs. The DOE’s OA mandate is part of a larger plan by the U.S. federal government to increase access to federally funded science. In 2013, the White House Office of Science and Technology Policy published a memo that government-funded science should be freely accessible.<sup>28</sup> In 2014, the DOE (2014) followed with its own “Public Access Plan”.<sup>29</sup> In detailed guidance online, the DOE (2022; p.20<sup>30</sup>) explains that: “For publications emanating from DOE national laboratories and other DOE facilities, DOE requires public access to any scholarly publication published on or after October 1, 2014” and “free, public access to the full text is enabled through DOE PAGES and OSTI.GOV within 12 months after publication in a journal.”

In addition, it stated that “The OA version may be a copy of the published version of the paper if the journal permits it, a link to permanent OA version at a journal or other repository, or the author’s final peer-reviewed manuscript”.<sup>28</sup> The OA policy applied to all publications from the Labs from October 2014 onwards, whereas for DOE-funded science outside the Labs, the requirements were only added to grantees for awards after October 2014. Hence, the identification is cleaner for NLs than for all grantees since there could be a substantial lag between the award of a grant and the resulting publications. It is this exogenous mandate that we exploit to address possible selection bias by researchers, who may pay for their best research to be available without a paywall.

## RESULTS

This section provides an overview of our main results, placebo tests, and robustness checks performed in our analysis. We first discuss the changes to the publishing strategy of scientists caused by the mandate. Next, we highlight changes to the uptake of scientific knowledge



**Figure 1. Location and sub-groups of U.S. National Laboratories**  
NNSA stands for National Nuclear Security Agency. Source: DOE.<sup>27</sup>

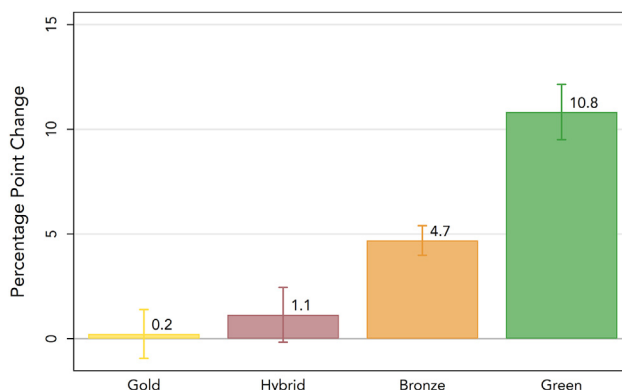
by patent holders as a proxy for technological applications of the increased diffusion of science from the NLs. Lastly, we investigate who benefited primarily from the increased openness of NL studies.

### Changes in NL researchers' publication strategy

We analyze to what extent the likelihood of different OA types changed through the OA mandate (Figure 2 see caption for a description of different OA types). If the OA mandate primarily increased green forms of OA, then this could create a substantial additional lag between the publication of the article and its availability in a free repository. We estimate a multinomial logit model of Equation 2 where the dependent variable is the categorical variable capturing the choice between bronze, green, hybrid, and gold forms of OA (with "closed" serving as the base category) to explore which OA types became more likely within the journals we analyzed (relative to our control group, which was unaffected by the mandate). The mandate primarily increased the likelihood of OA in the green category by 10.8% points, significant at the 1% level. Yet, we also find that the bronze category increased substantially (4.7% points), which are articles that are free to read on publishers' websites but without a clearly identifiable license, which does not allow re-use of the article.<sup>8</sup> In contrast, the likelihood that researchers paid for OA in the paywalled journal marginally increased by 1.1% points, significant at the 10% level. Lastly, there was no significant change in articles published in gold OA journals covered by our analysis.

### Changes in the use of National Laboratory research in scientific papers and patents

Overall, we find a strong and significant effect of the mandate on the use of NL research in patents. We differentiate between the location where the patent cites the scientific article, namely the front and body of the patent. Front citations are more likely to be added by patent lawyers, whereas body citations are more likely to be added by the inventors themselves. Body citations are typically seen as a better metric to measure knowledge diffusion as the inventor is more likely to be aware of the cited knowledge (see data section for in-depth discussion).



**Figure 2. Change in the likelihood of National Laboratory researchers publishing via different open-access routes after the mandate**

Notes: Marginal effects estimates based on multinomial logit estimates of Equation 2 where the choice alternatives are defined as gold, hybrid, bronze, and green OA, relative to the “closed” category, which is constrained to zero. Gold OA are those journals that offer immediate access to published content, hybrid OA are open-access articles in paywalled journals for which the author paid an article processing charge. Bronze OA are articles that are free to read on the publishers’ website but without an identifiable license, which limits re-use. Lastly, green OA are paywalled articles, for which an accepted version has been deposited in a free and open repository. The full regression output is presented in Table S2. Error bars indicate 95% confidence intervals. Unit of analysis is the article level (N = 352,721).

The number of citations received from the body of patents increased on average by 42% in the post-October 2014 period and is significant at the 1% level. Citations from the front of the patent increased but were not significant at the 5% level. In contrast, NL scientific literature was not cited more on the front of patents. Lastly, the mandate did not affect academic citations. We expected to see at least some positive effect on academic citations as indicated by a sizable body of literature investigating the “citation advantage” of OA publications.

A synthesis of the main regression results and the placebo test can be found in Figure 3. Overall, we find sizable and statistically significant evidence that the OA mandate was responsible for the increased uptake of peer-reviewed academic studies in patents. While the placebo tests trend in the same direction, they are not statistically different from zero, whereas our main results are statistically significant at the 1% level. We perform a range of additional robustness checks reported in the appendix (Tables S3–S6, S10, and S11).

### Beneficiaries of increased openness of NL science

Lastly, we analyze who mainly benefited from the use of the greater openness of NL research. To proxy the firm size of the filing patent, we divide the patentees based on their patent portfolio size. As there are some patentees with large portfolio sizes, we use the median to divide the firms into “large” and “small” patentees. This approach is in line with the literature<sup>18</sup> and allows us to investigate who started citing more NL research after the mandate came into force. We find strong and significant evidence that patentees with small patent portfolios cited NL research 49% more relative to a credible control group, whereas the effect on large patentees is statistically indistinguishable from zero (Figure 4).

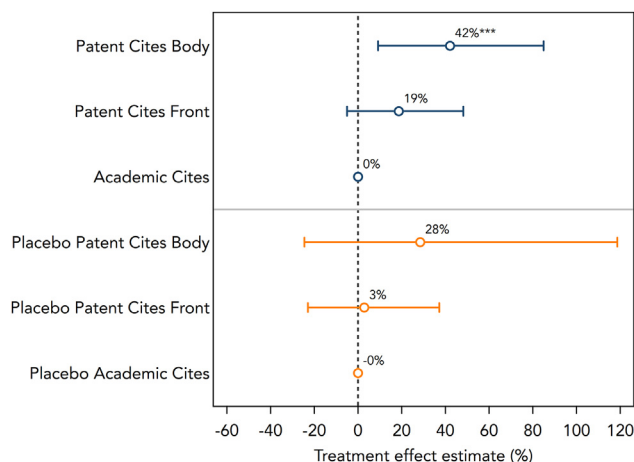
## DISCUSSION

Policymakers have enacted mandates that reflect the importance of free and public access to government-funded research and its role in technological innovation.<sup>12,31,32</sup> Here, we investigate the impact of an OA mandate applying to the “crown jewels” of energy innovation in the U.S., namely the DOE National Labs.<sup>7</sup> Our study sheds light on how an exogenous policy change to publishing rules affected the diffusion of scientific knowledge by studying the outward diffusion of scientific research from the NLS.<sup>13</sup>

While we focus on one specific case, we believe that it has the potential to illustrate mechanisms of knowledge diffusion that generalize beyond our context. For instance, National labs and publicly funded science play a key role across different energy innovation ecosystems.<sup>33</sup> Several countries have national labs, such as the European Commission Joint Research Center (JRC), the Korea Institute of Energy Research (KIER), and the Australian Commonwealth Scientific and Industrial Research Organization, among others.

Research from other fields also supports our findings. Bryan and Ozcan<sup>22</sup> study the impact of an open-access mandate in the biomedical field. They show that the National Institute of Health open access mandate led to an increase in follow-on citations from patents by 12–27%, in line with our results. Lastly, a large literature from economics and innovation research indicates that access barriers limit the diffusion of knowledge (e.g., see Keller<sup>34</sup> for an excellent overview). These literature streams highlight the fundamental principles that removing barriers to access increases the diffusion of knowledge.

We make three key contributions. First, we investigate how the mandate has affected the publication strategies of NL researchers, which is important to understand as it affects the time lag with which inventors can access scientific research emanating from the Labs. Second, we explore how the change to publication rules affected other scientific research and downstream inventions, which is critical as the NLS have



**Figure 3. Results visualization from main results and placebo tests**

Notes: Main results and placebo analysis. The top panel presents the treatment effect estimates (in % change) for the sample of academic articles published in journals that do not routinely make their articles free to read (N = 271,753). The bottom panel presents the treatment effect estimates from our placebo analysis using the sample of academic articles published in journals that routinely make their articles OA (N = 80,968). For ease of interpretation, treatment effects are displayed in % change, based on Poisson pseudo maximum likelihood estimates of Equation 2, controlling for month-by-year fixed effects and the number of academic citations. The % change is calculated as  $[e^{\beta} - 1] \times 100$ . Error bars represent 95% confidence intervals. The unit of analysis is the article level (N = 352,721).

struggled to increase the outward diffusion of scientific research.<sup>7</sup> Third, we study who benefited most from the OA mandate, shedding light on potential distributional effects. We discuss each of these contributions in turn. Lastly, we examine the limitations of our study and discuss avenues for future research.

### Changes in publication strategy

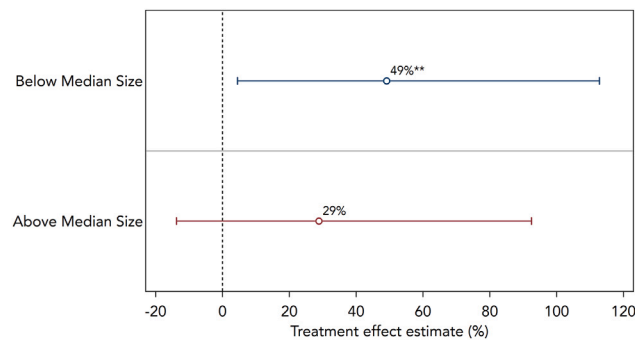
Understanding how scholars change their publishing strategies due to an OA mandate is important in several ways: First, if they opt for costly hybrid OA, this may divert research funds away from the core research to pay for publishing fees.<sup>8</sup> Second, if authors opt for free, green OA, research may have lower visibility and higher access barriers due to embargo periods.<sup>35,36</sup> Lastly, if they increasingly opt for gold OA, this might signal a more fundamental shift in the publishing strategy of scholars, indicating a move toward full models of OA.

Our results indicate that NL researchers have primarily responded to the mandate by self-archiving their articles (green OA), which has important implications for our results and the impact of the policy itself. Hybrid and gold OA allow for immediate, free, and public access to research articles, yet depositing accepted articles in repositories (green OA) commonly comes with the requirement of an embargo period of several months to over one year.<sup>35,36</sup> Publishers claim that the embargo period is critical, otherwise, libraries and research institutions might decide to cancel their journal subscriptions, leading to a substantial loss of revenue.<sup>8</sup> Researchers subject to the OA mandate must make the article free to read within one year of publication, but many journals preclude authors to deposit accepted manuscripts in repositories upon publication in a journal.

Our estimates of the NL mandate's impact on downstream inventions should therefore be seen as the likely lower bound as impacts would likely even be higher if articles were free to read upon publication. We also observe imperfect compliance with the mandate, suggesting that our estimated impacts could even be larger in the case of full compliance (around 10% of NL authors do not comply with the mandate in our study period).

Lastly, we also document an increase in Bronze OA, which are articles that have no clearly identifiable license but are free to read on the publishers' website. While these articles can be accessed without cost, the lack of a clear license means that they cannot be re-used.<sup>8</sup> In addition, for Bronze OA the rights remain with the publishers, making it unclear whether OA is granted temporarily or permanently, which could lead to unstable diffusion through this OA pathway.<sup>4,37</sup>

Our results hold important implications for policymakers. First, the exact policy design significantly influences the diffusion of scientific knowledge. Given that scholars strive to minimize the cost of complying with the mandate, the imposed embargo periods further delay public access to scientific publications. Therefore, the U.S. government's decision to require immediate public access to federally funded publications from 2026 onwards is a critical step toward ensuring timely access.<sup>12</sup> Second, compliance enforcement is another key component in increasing the diffusion of scientific knowledge. We found that approximately 10% of researchers do not comply with the mandate. Thus, policymakers should prioritize clear communication of the requirements and consider implementing penalties for non-compliance. Lastly, the rise in articles published under a bronze license—which are free to read but lack a clearly identifiable license—signals a need for policymakers to collaborate more closely with publishers to ensure clear licensing of scientific articles.



**Figure 4. Percentage increase in body-only patent citations by assignee size**

Notes: Treatment effect estimates (in % change) for the sample of academic articles published in journals that do not routinely make their articles free-to-read ( $N = 271,753$ ), split into two sub-groups corresponding to above median and below median assignee size. For ease of interpretation, treatment effects are displayed in % change, based on Poisson pseudo maximum likelihood estimates of Equation 2, controlling for month-by-year fixed effects and the number of academic citations. The % change is calculated as  $[e^{\beta} - 1] \times 100$ . Error bars represent 95% confidence intervals. The full regression output can be found in Table S7.

### Outward diffusion of NL scientific knowledge into academic research and patented inventions

Scientific knowledge entails both tacit and explicit knowledge and various diffusion channels, which makes it difficult to investigate. In our study, we leverage a range of recent advances in data availability by Marx and Fuegi<sup>38,39</sup> which standardize and report the connections between patents and scientific publications on a large scale. Combining bibliometric, patent, and open-access data on a large scale, we demonstrate the causal effect of the policy on academic research and downstream inventions.

Our results indicate that the OA mandate is most useful in spurring additional technical inventions but not additional scientific research (as proxied by additional citations from research articles). Our results suggest that at least on average researchers do not struggle to access scientific research, though there might be inequities that our results do not fully reflect (e.g., researchers from low-income countries and other potentially disadvantaged groups). The application of NL research in patented inventions is particularly important for the NLS as they have historically struggled with increasing the outflow of knowledge to more applied domains.<sup>7</sup> While there is a range of important diffusion channels apart from scientific publications,<sup>40</sup> increasing the diffusion of scientific knowledge—via scientific papers—through open-access mandates appears to be an effective strategy to increase outward diffusion.

### Beneficiaries of the increased scientific knowledge diffusion

Lastly, we find that particularly small firms (those with below-median patent portfolios) benefited from the OA mandate. After the mandate, small firms cite NL research around 50% more (compared to articles that were not subject to the mandate). This change indicates that particularly smaller firms are adversely affected by the journals' paywalls. Several survey-based studies indicate access barriers, but we are the first study to demonstrate the effect of an OA mandate on downstream patented inventions by smaller firms. We, therefore, make an important extension to the work of Galasso and Schankerman (2015) who show that patent invalidation particularly helps small firms. Yet, we document that this also holds for the diffusion of scientific knowledge, which tends to be further removed from the actual application than the patent-to-patent citations that Galasso and Schankerman (2015) draw on for their analysis.

SMEs play an integral role in driving economic development in Europe and the U.S. but appear to be particularly affected by difficult access to scientific publications.<sup>15</sup> In the U.S. SMEs contribute 44% of U.S. economic activity, though their share has declined over time.<sup>41</sup> Similarly, Europe's economy is underpinned by 25 million SMEs employing around 100 million people and contributing more than half of Europe's GDP.<sup>42</sup> While an OA mandate may not seem to be an SME promotion policy, it is important to stress that these appear to be the key beneficiaries, though more work needs to be done to explore the exact mechanisms behind the increased uptake and the use of patents in this subgroup.<sup>20,43</sup> While we do not show this empirically, high-growth start-ups are likely not the main beneficiaries, as they are often spun out from universities,<sup>44</sup> which should give them easier access to academic publications. In contrast, SMEs are likely further removed from the academy.<sup>15</sup>

### Future research

Our work can be extended in several ways. First, it would be critical to investigate more closely how OA mandates impact other forms of knowledge diffusion.<sup>13</sup> While the OA mandate primarily affected the ways peer-reviewed scientific publications are stored and used, the mandate may have led to the additional transfer of tacit knowledge due to subsequent academic consulting, co-supervision of students, or other science-industry spillovers.<sup>45</sup> Understanding these complementary knowledge transfers is critical<sup>40</sup> but difficult to study. Similarly, while paywalls typically only cover the peer-reviewed article, the DOE also encourages researchers to deposit their data for further use. Analyzing to what extent data in repositories is re-used and sparks novel research and technological applications, may be another fruitful avenue to pursue.<sup>20</sup>

While we show that the mandate primarily affected patented inventions by small firms, studying whether academics from low- and middle-income countries benefited from the mandate, leading to spillovers outside of the U.S. would be an important addition to the literature.<sup>34</sup> Understanding where and how the beneficiaries are located, how they use the knowledge, and what economic consequences it has, is important to understand. Estimating the increased return to R&D due to greater knowledge spillovers would be another interesting research avenue.<sup>8</sup> Lastly, investigating whether the patents were eventually commercialized would be a further step toward understanding the full effect of the policy. Recent advances in virtual patent marking papers could be a promising research avenue.<sup>46</sup>

### Limitations of study

Our work draws on one diffusion channel of scientific research, namely, scientific publications. Yet, there are many ways that scientific knowledge can diffuse, including co-supervision of students between industry and academia, licensing, and consulting.<sup>13,14,16</sup> Hence, our work covers only one potential pathway for diffusion, yet one that is commonly seen among the most important diffusion channels.<sup>15</sup> In addition, as is commonly challenging in innovation research, we can only follow the “paper trail”, meaning that patents offer a standardized, well-documented data source for studying the diffusion of scientific knowledge into more practical applications. Yet, while patents are widely used in the innovation literature, they come with certain limitations.<sup>47</sup> Not all inventions are patented, and some firms may choose to protect their inventions by secrecy and lead times, rather than by patents.<sup>13</sup> While our results indicate that smaller firms benefited most strongly from the NL mandate, more work in this regard is needed, particularly in understanding how small firms find, understand, and use knowledge emanating from research facilities.<sup>20</sup>

### STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- METHOD DETAILS
  - Data
- QUANTIFICATION AND STATISTICAL ANALYSIS

### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2023.107740>.

### ACKNOWLEDGMENTS

We thank Matt Marx for support with the Reliance on Science dataset and Jonas Mark for research assistance.

### AUTHOR CONTRIBUTIONS

B.P., P.L., A.K., and L.D.A. were responsible for developing the research idea. B.P. collected the data. P.L. and B.P. did the empirical analysis. B.P., P.L., A.K., and L.D.A. were responsible for interpreting the outputs from the empirical analysis. B.P. wrote the manuscript with support from P.L., and B.P., P.L., A.K., and L.D.A. edited the final draft.

### DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 26, 2023

Revised: August 7, 2023

Accepted: August 24, 2023

Published: August 26, 2023

### REFERENCES

1. Viñuales, J.E. (2015). The Paris Climate Agreement: an initial examination. SSRN Journal 9, 1–16. <https://doi.org/10.2139/ssrn.2704670>.
2. Bosetti, V., Carraro, C., Duval, R., Sgobbi, A., and Tavoni, M. (2009). The Role of R&D and Technology Diffusion in Climate Change Mitigation: New Perspectives Using the WITCH Model. *Energy* 664, 1–41. <https://doi.org/10.1787/227114657270>.
3. Tennant, J.P., Waldner, F., Jacques, D.C., Masuzzo, P., Collister, L.B., and Hartgerink, C.H.J. (2016). The academic, economic and societal impacts of Open Access: An evidence-based review. *F1000Res.* 5, 632–655. <https://doi.org/10.12688/f1000research.8460.1>.
4. Piwowar, H., Priem, J., Larivière, V., Alperin, J.P., Matthias, L., Norlander, B., Farley, A., West, J., and Haustein, S. (2018). The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ* 6, e4375. <https://doi.org/10.7717/peerj.4375>.



5. Moretti, E., Steinwender, C., and Reenen, J. (2019). The Intellectual Spoils of War? Defense R&D, Productivity and International Spillovers. National Bureau of Economic Research, 1–76. <https://doi.org/10.3386/W26483>.
6. Fleming, L., Greene, H., Li, G., Marx, M., and Yao, D. (2019). Government-funded research increasingly fuels innovation. *Science* 364, 1139–1141. [https://doi.org/10.1126/SCIENCE.AAW2373/SUPPL\\_FILE/AAW2373\\_FLEMING\\_SM.PDF](https://doi.org/10.1126/SCIENCE.AAW2373/SUPPL_FILE/AAW2373_FLEMING_SM.PDF).
7. Anadón, L.D., Chan, G., Bin-Nun, A.Y., and Narayanamurti, V. (2016). The pressing energy innovation challenge of the US National Laboratories. *Nat. Energy* 1, 16117. <https://doi.org/10.1038/nenergy.2016.117>.
8. UKRI (2020). Economic Implications and Benefits Assessment of an Updated UKRI Open Access Policy for Peer Reviewed Research Articles.
9. Else, H. (2018). Radical open-access plan could spell end to journal subscriptions. *Nature* 561, 17–18. <https://doi.org/10.1038/d41586-018-06178-7>.
10. Plan, S. (2021). Accelerating Open Access cOAlition S.
11. The White House (2022). OSTP Issues Guidance to Make Federally Funded Research Freely Available without Delay (Press Release).
12. Tollefson, J., and Van Noorden, R. (2022). US government reveals big changes to open-access policy. *Nature* 609, 234–235. <https://doi.org/10.1038/d41586-022-02351-1>.
13. Agrawal, A., and Henderson, R. (2002). Putting Patents in Context: Exploring Knowledge Transfer from MIT. *Manage. Sci.* 48, 44–60. <https://doi.org/10.1287/mnsc.48.1.44.14279>.
14. Agrawal, A. (2000). Importing Scientific Inventions: Direct Interaction, Geography, and Economic Performance.
15. Houghton, J., Swan, A., and Brown, S. (2011). Access to Research and Technical Information in Denmark.
16. Cohen, W., Florida, R., Randazzese, L., and Walsh, J. (1998). Industry and the academy: Uneasy partners in the cause of technological advance. In *Challenges to the Research University*, R. Noll, ed. (Brookings Institution).
17. Ware, M., and Monkman, M. (2009). Access by UK Small and Medium-Sized Enterprises to Professional and Academic Information.
18. Galasso, A., and Schankerman, M. (2015). Patents and cumulative innovation: Causal evidence from the courts. *Q. J. Econ.* 130, 317–369. <https://doi.org/10.1093/qje/qju029>.
19. Doblinger, C., Surana, K., and Anadon, L.D. (2019). Governments as partners: The role of alliances in U.S. cleantech startup innovation. *Res. Policy* 48, 1458–1475. <https://doi.org/10.1016/j.respol.2019.02.006>.
20. Huber, F., Wainwright, T., and Rentocchini, F. (2020). Open data for open innovation: managing absorptive capacity in SMEs. *R&D Management* 50, 31–46. <https://doi.org/10.1111/RADM.12347>.
21. Unpaywall (2019). Unpaywall. <https://unpaywall.org/>.
22. Bryan, K.A., and Ozcan, Y. (2021). The impact of open access mandates on invention. *Rev. Econ. Stat.* 103, 954–967. [https://doi.org/10.1162/rest\\_a\\_00926](https://doi.org/10.1162/rest_a_00926).
23. Furman, J.L., and Stern, S. (2011). Climbing atop the shoulders of giants: The impact of institutions on cumulative research. *Am. Econ. Rev.* 101, 1933–1963. <https://doi.org/10.1257/aer.101.5.1933>.
24. Anadón, L.D. (2012). Missions-oriented RD&D institutions in energy between 2000 and 2010: A comparative analysis of China, the United Kingdom, and the United States. *Res. Policy* 41, 1742–1756. <https://doi.org/10.1016/j.respol.2012.02.015>.
25. National Center for Science and Engineering Statistics (2022). U.S. R&D expenditures, by performing sector and source of funds: 2010–20. <https://nces.nsf.gov/pubs/nsf22330>.
26. DOE (2021) (Department of Energy FY 2022 Congressional Budget Request).
27. DOE (2022). DOE National Laboratories. <https://www.energy.gov/doe-national-laboratories>.
28. Lawrence Berkeley National Laboratory (2022). DOE Open Access Policy: What is it?. <https://it.lbl.gov/service/publications/what-is-the-does-policy/>.
29. DOE (2014). Public Access Plan. [https://www.energy.gov/sites/prod/files/2014/08/f18/DOE\\_Public\\_Access%20Plan\\_FINAL.pdf](https://www.energy.gov/sites/prod/files/2014/08/f18/DOE_Public_Access%20Plan_FINAL.pdf).
30. DOE (2022). FAQs | DOE PAGES. <https://www.osti.gov/pages/faqs>.
31. Siegel, D., Bogers, M.L., Jennings, P.D., and Xue, L. (2023). Technology transfer from national/federal labs and public research institutes: Managerial and policy implications. *Res. Policy* 52, 104646. <https://doi.org/10.1016/j.respol.2022.104646>.
32. Choi, H., Yoon, H., Siegel, D., Waldman, D.A., and Mitchell, M.S. (2022). Assessing differences between university and federal laboratory postdoctoral scientists in technology transfer. *Res. Policy* 51, 104456. <https://doi.org/10.1016/j.respol.2021.104456>.
33. Meckling, J., Galeazzi, C., Shears, E., Xu, T., and Anadon, L.D. (2022). Energy innovation funding and institutions in major economies. *Nat. Energy* 7, 876–885. <https://doi.org/10.1038/s41560-022-01117-3>.
34. Keller, W. (2004). International Technology Diffusion. *J. Econ. Lit.* 42, 752–782. <https://doi.org/10.1257/0022051042177685>.
35. Elsevier (2019). Journal Embargo Finder FAQ (Elsevier Embargo Finder). <https://www.elsevier.com/about/open-science/open-access/journal-embargo-finder>.
36. Ottaviani, J. (2016). The post-embargo open access citation advantage: It exists (probably), its modest (usually), and the rich get richer (of course). *PLoS One* 11, e01596144. <https://doi.org/10.1371/journal.pone.0159614>.
37. Brock, J. (2018). 'Bronze' open access supersedes green and gold | News | Nature Index. <https://www.nature.com/nature-index/news-blog/bronze-open-access-supersedes-green-and-gold>.
38. Marx, M., and Fuegi, A. (2022). Reliance on science by inventors: Hybrid extraction of in-text patent-to-article citations. *J. Econ. Manag. Strategy* 31, 369–392. <https://doi.org/10.1111/jems.12455>.
39. Marx, M., and Fuegi, A. (2020). Reliance on science: Worldwide front-page patent citations to scientific articles. *Strateg. Manag. J.* 41, 1572–1594. <https://doi.org/10.1002/smj.3145>.
40. Clayton, P., Lanahan, L., and Nelson, A. (2022). Dissecting diffusion: Tracing the plurality of factors that shape knowledge diffusion. *Res. Policy* 51, 104389. <https://doi.org/10.1016/j.respol.2021.104389>.
41. SBA (2019). Small Businesses Generate 44 Percent of U.S. Economic Activity – SBA's Office of Advocacy. <https://advocacy.sba.gov/2019/01/30/small-businesses-generate-44-percent-of-u-s-economic-activity/>.
42. European Commission (2020). An SME Strategy for a Sustainable and Digital Europe. <https://doi.org/10.1016/j.respol.2021.104389>.
43. de Rassenfosse, G. (2012). How SMEs exploit their intellectual property assets: Evidence from survey data. *Small Bus. Econ.* 39, 437–452. <https://doi.org/10.1007/S11187-010-9313-4/TABLES/6>.
44. Bryan, K., and Guzman, J. (2021). Entrepreneurial Migration. SSRN J. 26483. <https://doi.org/10.2139/ssrn.3914010>.
45. Azoulay, P., Graff Zivin, J.S., and Sampat, B.N. (2012). The Diffusion of Scientific Knowledge across Time and Space. In *The Rate and Direction of Inventive Activity Revisited* (University of Chicago Press), pp. 107–155. <https://doi.org/10.7208/chicago/9780226473062.003.0006>.
46. de Rassenfosse, G., Mages, A., Shultz, C., Wadsworth, P., Waldron, R., Barnett, J.J., Brant, E., Claeys, D., Karshstedt, J., Kesan, A., et al. (2018). Notice failure revisited: Evidence on the use of virtual patent marking. Nber working paper:1–21, working paper number: 24288. <https://doi.org/10.3386/W24288>. <https://www.nber.org/papers/w24288>.
47. Probst, B., Touboul, S., Glachant, M., and Dechezleprêtre, A. (2021). Global trends in the invention and diffusion of climate change mitigation technologies. *Nat. Energy* 6, 1077–1086. <https://doi.org/10.1038/s41560-021-00931-5>.
48. Bryan, K.A., Ozcan, Y., and Sampat, B. (2020). In-text patent citations: A user's guide. *Res. Policy* 49, 103946. <https://doi.org/10.1016/J.RESPOL.2020.103946>.
49. Baker, A.C., Larcker, D.F., and Wang, C.C. (2022). How much should we trust staggered difference-in-differences estimates? *J. Financ. Econ.* 144, 370–395. <https://doi.org/10.1016/J.JFINECO.2022.01.004>.
50. Rambachan, A., and Roth, J. (2023). A More Credible Approach to Parallel Trends. *Rev. Econ. Stud.* 664, 1745303. <https://doi.org/10.1093/restud/rdad018>.
51. Rios-Avila, F., Sant'Anna, P., and Naqvi, S.A.A. (2022). DRDID: Stata module for the estimation of doubly robust difference-in-difference models. [https://friosavila.github.io/playingwithstata/main\\_drdid.html](https://friosavila.github.io/playingwithstata/main_drdid.html).

## STAR★METHODS

## KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
Stata 17	Stata	<a href="https://www.stata.com/stata-news/news36-2/">https://www.stata.com/stata-news/news36-2/</a>
DRDID	Rios Avilla et al. 2022	<a href="https://friosavila.github.io/playingwithstata/main_drdid.html">https://friosavila.github.io/playingwithstata/main_drdid.html</a>

## RESOURCE AVAILABILITY

## Lead contact

Further information requests should be directed to the lead contact Benedict Probst ([bprobst@ethz.ch](mailto:bprobst@ethz.ch)).

## Materials availability

This study did not use or generate reagents.

## Data and code availability

- The data are available upon request from the lead author.
- The code is available upon request from the lead author.
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

## METHOD DETAILS

## Data

The data compilation process is shown in [Figure S3](#) and contains five steps. We first compiled a list of the top 20 journals that DOE-funded research has appeared in the past according to the PAGES database (a full list can be found in [Table S8](#)). In a second step, we downloaded all articles from the Dimensions.AI database that appeared in these journals between 2012 and 2018, which gives us at least three years before and after the mandate (352,721 articles). For each of the scientific articles, we check whether at least one author is associated with one of the Labs and therefore was subject to the OA mandate. Third, we used the 'reliance on science' database to check whether articles are cited by a patent or not.<sup>38,39</sup> The database differentiates between patent citations to academic articles that stem from the body or the front of the patent. Citations from the body of the patent are considered to be a better proxy for knowledge diffusion as these citations are commonly added by inventors rather than patent lawyers. Fourth, we used the Unpaywall (2019)<sup>21</sup> API to check for each article (via its digital object identifier) whether the article is available without a paywall. Unpaywall is a comprehensive database for the OA status and is widely used in the literature (Else, 2018b; Unpaywall, 2019). Fifth, we scraped Google Patents to determine the size of the patent assignees' portfolio (of patents that cite scientific research). Our final dataset contains 352,721 scientific articles published between 2012 and 2018. The descriptive statistics of the dataset can be found in [Table S9](#).

Inventors can cite scientific literature in two parts of the patent, namely the 'front' of the patent and the 'body' of the patent. The latter is also called in-text citations in the literature. Several studies<sup>22,48</sup> indicate that citations from the body of the patent may better reflect true knowledge flows between scientific research and patented technologies as these citations are more commonly added by inventors rather than patent examiners. The inventor may be unaware of citations to the academic literature added by patent examiners and that citation therefore would not constitute the true state of the inventors' knowledge when filing the patent. Bryan and Ozcan (2021) therefore use body citations to track the diffusion between peer-reviewed scientific articles and patented inventions but do not test whether citations from the front of the patents also become more likely because of OA mandates. To test for differences in the effect of these different types of citations, we include both in our empirical analysis.

## QUANTIFICATION AND STATISTICAL ANALYSIS

To estimate the causal effect of the NL mandate on patent citations we use a difference-in-differences (DID) model for two main reasons. First, to address selection bias, which is a key concern in causal inference. In our case, researchers could pay to make their best articles published open access. This self-selection can confound the estimates of the mandate's effect. A DID-model controls for time-invariant confounders that simultaneously influence the selection into open access and the outcome. Second, to tackle time-varying confounders. Before and after the implementation of the policy, there might be concurrent changes in the wider political, technical, and funding landscape. By comparing changes in the treatment group (subject to the mandate) to a credible control group (not subject to the mandate, but otherwise affected by the same broader changes), the DID model can isolate the causal effect of the policy by addressing these time-varying confounders.

As a baseline specification, we estimate a canonical DID model with a binary treatment indicator and a binary post-treatment indicator. The unit of analysis is at the individual article level, denoted by the subscript  $i$ :

$$Y_i = \rho NL + \gamma PostOct14 + \delta^{DD} (NL \times PostOct14) + \varepsilon_i \quad (\text{Equation 1})$$

where  $Y_i$  is the dependent variable capturing the number of article-level patent citations (emanating from the body or front of the patent), or the number of academic citations.  $NL$  is the treatment indicator identifying articles that were published by authors affiliated with the National Labs.  $PostOct14$  is the post-treatment indicator identifying articles that were published on or after 1<sup>st</sup> October 2014, the date from which NL articles had to be published OA. The DID estimator ( $\delta^{DD}$ ) is obtained from the interaction of both main effects ( $NL \times PostOct14$ ) and provides an unbiased estimate of the effect of the NL OA mandate on the dependent variables (Note that the DID estimator captures the effect of the NL OA mandate on patent citations, not the effect of OA per-se. The estimate thus reflects the incremental effect of increasing the share of OA articles published by NLS by 30 percentage points (from 60% to 90%) compared to the control group that saw little change in the share of OA articles.).

To probe the robustness of our baseline specification, we estimate a generalized DID model, which additionally captures general trends in patent citations over time using a set of temporal fixed effects. Moreover, we include relevant article-level control variables:

$$Y_i = \rho NL + \delta^{DD} (NL \times PostOct14) + \lambda_t + X_i + \varepsilon_i \quad (\text{Equation 2})$$

where  $Y_i$  is the dependent variable capturing the article-level citations (body, front and academic) (The article-level citations refers to whether the citation to the scientific article emanates from the body of the patent, the body of the patent, or a scientific article (see [data](#) section for detailed description).). For our analysis of OA type (Hypothesis 1),  $Y_i$  is a categorical outcome variable that captures four OA publishing alternatives (hybrid, green, bronze and gold) relative to 'closed' articles (the base category).  $\lambda_t$  are month-by-year fixed effects which account for temporal trends common to both NL and non-NL articles, and  $X_i$  represents a vector of article-level control variables including the number of authors and the number of academic citations (inverse hyperbolic sine transformed, which is a common alternative to log-transformation in the presence of 0-values), a measure of article quality, as well as the journal in which the article was published (using a set of journal dummies). The journal fixed effects capture all journal-specific unobserved characteristics (e.g., Journal quality, readership and reach).

We estimate [Equations 1 and 2](#) using Poisson Pseudo Maximum Likelihood under the assumption that the OA mandate caused a multiplicative increase in total patent citations (i.e., following a 0-inflated Poisson distribution). Importantly, the generalized DID model provides unbiased estimates of the treatment effect in settings where there is a single treatment period (i.e., no staggered implementation of the treatment) (See Baker et al. (2022)<sup>49</sup> for details.). DID models are sensitive to the parallel trends assumption. In support of parallel trends, we find no significant pre-mandate differences in the number of body patent citations between treatment and control groups, aggregated over six-month periods (see [Figure S1](#)). Moreover, we use the approach developed in Rambachan and Roth (2023)<sup>50</sup> to show that our results are robust to violations in the parallel trends assumption of up to 5% (see [Figure S2](#)). Finally, we use the doubly robust difference-in-differences estimator (DRDID)<sup>51</sup> to test the robustness of our main results. The doubly robust estimator combines both a regression adjustment approach and an augmented inverse-probability weighting approach, which produces consistent and robust estimates even if either the outcome regression model or the propensity score model are misspecified.<sup>51</sup>

A key hypothesis underpinning our study is that the OA mandate effectively increased the OA likelihood of studies emanating from NLS but not from studies coming from non-NLS. As [Figure S4](#) shows, the OA likelihood of studies between NL and non-NL research groups was very similar before the mandate, oscillating between 50% and 60%. After the introduction of the mandate, the OA likelihood increased to around 90% for NLS, whereas non-NLS saw only slight increases relative to the baseline of around 60%. Overall, the mandate increased the OA likelihood relative to the baseline by 30%-points. As our study investigates only this marginal change in the OA availability (and not the difference between fully closed vs. fully open), the true empirical effect of OA articles could be up to three times as large (if one assumes linear scaling of effects).

All models are estimated using Stata 17. Full regression outputs can be found in the Appendix, which accompany the visualizations presented in the manuscript. We estimate the DRDID using the user-written Stata package 'drdid' (Rios-Avila et al., 2022). Results are presented in [Table S11](#). Please note that the outcome regression model is estimated by ordinary least squares (OLS) and coefficient estimates thus differ in magnitude from our main analysis outcomes obtained from Poisson Pseudo Maximum Likelihood estimation. In the figures, error bars indicate the 95% confidence intervals. Significance levels indicated by \*\*\* (1%), \*\* (5%) and \* (10%).