Commercial Regulation and the Production of Science

Evidence from New Genomic Techniques in Agricultural Biotechnology

Elisabeth Hofmeister[1], Samantha Zyontz[2], and Rosemarie Ziedonis[3]

[1]Max Planck Institute for Innovation and Competition

[2]Boston University

[3]Boston University and NBER

1 Introduction

The rate and direction of academic science depends on the activities of scientists, who are driven by various motives and respond to external cues (Stephan, 1996). Some may only be interested in science for science's sake, but others are incentivized by the usefulness and application of their work in downstream products or services. Even if they are not directly compensated for the commercialization, these academic scientists want their innovations and discoveries to be used by corporations to benefit the public. This symbiotic relationship between academic scientists and industry can be seen, for example, in the fact that corporate patents rely on federally funded research (Fleming et al., 2019), patents closer to science are more valuable (Fleming and Sorenson, 2004; Krieger et al., 2024), and competitive advantage can be gained by being the first to commercialize advances made in public science (Arora et al., 2023). A shift away from applied commercial potential should then influence academic scientists' behavior and change the rate and/or direction of their research. In this work, we add to the understanding of scientists' motivations and study to what extent scientists place value on the applied commercial potential of their research. Specifically, we investigate whether and how a withdrawal of commercial opportunities in a market alters the rate and direction of scientific research.

To investigate the effects of constrained commercialization potential on the rate and direction of scientific research, we exploit an unexpected ruling by the European Court of Justice (ECJ)

in 2018 that curtailed the market release of gene-edited plant products but left the underlying research methods, particularly the gene-editing tool CRISPR, available to scientists. The shock created a sharp contrast with the United States, where gene-edited crops faced few commercial restrictions. We use this divergence to explore how a shock to commercial prospects affects the responses of upstream scientists. Do such regulations cause a brain drain, with scientists relocating to jurisdictions with more favorable regulatory environments? If so, which scientists are leaving? For those who remain, do they continue to work at the technological frontier, does their research begin to lag, or do they switch to alternative tools or topics?

Current literature offers countervailing predictions regarding scientists' willingness and ability to respond to a negative shock to downstream product market regulation. We consider three potential responses: relocation to more liberally regulated jurisdictions, pivots away from the affected resarch fields, and establishing cross-jurisdictional relationships to preserve access to commerical potential. For each of these responses, existing evidence is mixed and at times conflicting. One literature stream highlights scientists' preferences that are unrelated to pecuniary gains from commercialization (Stern, 2004). Several studies point to mobility constraints (Azoulay et al., 2017) as well as negative consequences from pivoting toward different research topics (Leahey, 2007; Hill et al., 2025). In contrast, a second stream challenges and adds nuance to these arguments, in particular on the role of commercialization and pecuniary attributes as part of scientists' preferences (Roach and Sauermann, 2010; Cohen et al., 2020). Whether and how scientists respond to diminished commercial prospects thus remains an empirical question.

To test the effects of the 2018 regulatory shock in the European Union on research, we employ a difference-in-differences approach at the scientist level to compare the affected European plant scientists to their unaffected US counterparts before and after the regulatory change. We rely on bibliometric data and leverage the detailed content from publications through natural language processing with large language models to construct field-specific outcome variables.

Our results suggest that scientists respond to the withdrawal of commercial opportunities in their local market. We find that at least some scientists are willing and able to relocate to other jurisdictions, while others curate cross-jurisdictional ties by gaining additional affiliations

outside their home jurisdiction. Additionally, some scientists pivot away from affected topics, in this case proxied by a shift away from using CRISPR in their research or working on plants. We find that responses vary by scientists' career stage, consistent with the notion that each response requires different opportunities and involves different constraints and costs. Our results are informed by ongoing interviews with European and US plant scientists, the first wave of which has been completed.

Our findings are of particular importance for multinational firms operating across today's globally fragmented regulatory landscape. Scientists tend to choose their direction based on personal experiences and background, including local circumstances and problems (Fry, 2023), which is particularly relevant in agricultural research where plants and pests vary geographically (Moscona and Sastry, 2025). As a result, scientific advances made in one jurisdiction may not readily translate to other geographic markets, meaning that research conducted solely in a firm's domestic market may not provide all the knowledge needed to develop products for global markets. If product regulation changes scientists' rate and direction of research, firms may encounter difficulties sourcing the external knowledge they need and may have to develop countermeasures.

2 Conceptual Framework

The effect of diminished applied commercial potential on the rate and direction of scientific research in the regulating jurisdiction is theoretically ambiguous. Scientists possess agency over the research process. Hence, any change in the focal jurisdiction's rate and direction of scientific research in response to product market regulation depends on the willingness and ability of scientists to execute such a change. Importantly, the willingness and ability of scientists to respond may vary with the precise response mechanism. This section reviews existing evidence on how scientists may respond.

Different response mechanisms - relocating, pivoting, or establishing ties with other jurisdictions - each involve different costs that may alter the willingness to execute them. Irrespective of the precise response mechanism, however, scientists must value commercialization sufficiently to even consider altering their research process when

commercialization prospects decline. Given the evidence on scientists' motives for engaging in research, such baseline willingness to respond is not self-evident. Indeed, the disinterestedness of scientists, referring to the expectation that research is pursued for the advancement of knowledge rather than personal gain, represents an ideal norm at the foundation of the scientific system (Merton, 1973). In line with this notion, prior literature provides empirical evidence of a 'taste for science', whereby scientists are willing to forego compensation in exchange for participation in open science (Stern, 2004). Characteristic traits of this 'taste for science' primarily include preferences for autonomy in research topic choice and the open dissemination of findings with the scientific community (e.g. Stern, 2004; Aghion et al., 2008; Roach and Sauermann, 2010). These attributes appear, at first glance, to be non-pecuniary and tied to the conduct of research rather than to its commercialization potential.

However, several recent studies provide nuance to this view of scientists' motives. First, 'taste for science' is a multidimensional construct composed of several motives, some of which may well align with the commercialization of research. For example, one frequently cited but rarely defined aspect of the 'taste for science' is the desire to contribute to society (Cohen et al., 2020). Since translating research into products is one way to materialize such a contribution, a desire for commercialization may actually be part of the 'taste for science' itself. Indeed, survey evidence suggests that the strength of the desire to contribute to society positively correlates with patenting activity, one proxy for commercialization, especially in the life sciences (Cohen et al., 2020).

Second, survey evidence suggests that the 'taste for science' is a spectrum, rather than a dichotomy: scientists value its characteristic attributes to varying degrees (Roach and Sauermann, 2010), suggesting that some scientists may well place value on pecuniary aspects, and hence, the additional income from commercializing their research. Whether or not the financial incentives from commercialization can affect scientists' research activities is uncertain. On the one hand, the abolition of the 'professor's privilege' in Europe - which reassigned patent ownership from scientists to universities and reduced scientists' direct commercialization income - appears to have left publication output unaffected (Hvide and Jones, 2018). On the other hand, the introduction of the Bayh-Dole Act in the US, which reassigned ownership from the government to universities and, via royalty-sharing

arrangements, enhanced scientists' prospects of income from patented research, appears to have stimulated basic research (Thursby and Thursby, 2011).

In summary, it is plausible that at least some scientists value the commercialization prospects of their research - whether because commercialization realizes a contribution to society, consistent with their presumed intrinsic motivation, or because it offers financial rewards. Yet, scientists must also have the ability to adapt their research process. To this end, there are three main actions scientists can take to maintain the commercialization prospects of their research.

First, scientists may consider relocating to a more favourably regulated jurisdiction to preserve the commercialization potential of their research. However, not all scientists have the ability to pursue this response. For example, employment opportunities are not available to every scientist: evidence shows that productivity is positively associated with relocating (Ganguli, 2015; Azoulay et al., 2017). Additionally, personal and professional circumstances may constrain the ability to move. For example, adolescent children as well as recent funding have been associated with a lower probability of relocating (Azoulay et al., 2017). Despite these constraints, empirical evidence from inventors, a group of people resembling scientists in some aspects of their work, suggests that financial prospects affect inventor mobility. Strengthened intellectual property rights, for example, can prevent independent inventors from emigrating (Ivus and MacGarvie, 2025), while an increase in the tax rate induces at least foreign inventors to leave the country (Akcigit et al., 2016).

Second, scientists may choose to remain in their original jurisdiction but pivot to a different research topic where commercialization prospects still prevail. The ability to pivot, as well as the implications of pivoting, are a subject of ongoing debate. On the one hand, the ability to pivot may be constrained by the specialization of scientists resulting from the so-called 'burden of knowledge', referring to the increasing amount of knowledge that must be understood before making a scientific contribution (Jones, 2009). Additionally, pivoting may entail penalties in terms of scholarly impact (Hill et al., 2025) as well as earnings (Leahey, 2007). In line with these arguments against scientists pivoting easily, prior work has found that inducing scientists to pivot only marginally from their usual topic requires large financial incentives, implying the

elasticity of science is low (Myers, 2020). On the other hand, prior studies have found that scientific advances often arise from outsiders to the field (Azoulay et al., 2019) and that streaks of high scholarly impact are often preceded by exploratory research (Liu et al., 2021). In contrast to the documented low elasticity of science and in favor of pivoting ability, there also appears to be an increasing trend of scientists switching between topics (Zeng et al., 2019).

Third, scientists remaining in their home jurisdiction may establish ties with scientists in other jurisdictions to sustain commercialization prospects where regulations are less restrictive. Such reaction has been documented in response to the restrictions of US federal funding for embryonic stem cell research in the early 2000s, where scientists increasingly sought international collaborations to preserve access to funding (Furman et al., 2012). Scientists lacking the commercialization prospects in their home jurisdiction may employ a similar mechanism. However, as with all response mechanisms, its success depends on scientists' ability to pursue it. In this case, this ability also hinges on the willingness of potential partners, which itself may change after a regulatory shock.

The responses of scientists will likely differ by their tenure in academia. Technology changes can have a heterogeneous effect on workers in different stages of their careers. Even the introduction of Artificial Intelligence (AI) affects entry-level jobs more than established upper-level jobs (Hosseini and Lichtinger, 2025). Early career scientists like graduate students, postdocs, and junior professors often stick with the tools and organisms of their mentors, but they have made fewer commitments that lock them into their choices. Established career scientists who have been tenured professors and have a publication track record have made commitments that lock them into their research and labs, but they also have a much larger network and established skill set. Mid career scientists who are either about to get tenure or just recieved it have made commitments both professionally and personally but do not have as extensive a network or reputation. For these reasons, we examine not only the average effects of reduced commercial research potential across scientists, but pay particular attention to how responses vary by career stage. We also intentionally include plant scientists at different career stages in our interviews, providing deeper insight into the opportunities and constraints scientists face as their careers progress.

3 Setting

We leverage the scientific field of genome editing of plants to investigate whether and how commercialization opportunities constrained by product market regulation affect scientists' activities in the regulating jurisdiction. Genome editing in plants is a compelling setting to address our research questions for three reasons: the undisputed position of the frontier technology, CRISPR, the clear geographical regulatory divergence, and the local specificity of plant science and its commercialization. The following sections introduce the research process in plant science as well as the technological and legal developments concerning genome editing in plants in the European Union and the United States, our comparison setting.[1]

3.1 The Research Process in Plant Science

Plant science focuses entirely on the study of plants, with topics ranging from understanding ecology and genetics to creating new agricultural crops and biofuels. This stands in contrast to general biology, which covers all living organisms, including animals, and targets downstream applications in medicine and livestock improvements. We focus on scientists contributing to research on gene-edited crops, which are plants than can be cultivated. The research process to obtain a gene-edited crop involves several steps. First, scientists understand which gene should be modified to achieve the desired outcome, such as increased resistance to heat or a particular plant pest, and how it works. Typically, scientists carry out this first step in a model plant that is easily grown in a lab. Then, scientists figure out how to edit the gene, first in the model plant in the lab, then in the actual crop in a greenhouse. Lastly, scientists translate the results from the greenhouse to the field. Across these steps, scientists move from basic to applied research, from model plants to crops, and from the lab to the field. Scientists may carry out all steps as part of their research agenda or focus on particular contributions to the process.

The career trajectory of scientists in plant science is comparable to other STEM fields. Most will complete four years of undergraduate study, four to six years of graduate study through the Ph.D., between three and six years of post-doctoral work, five to seven years as a

[1]While China has become a major hub for research on gene-edited plants, we do not use Chinese scientists as a control group due to China's distinct political framework. China's strong position reflects government prioritization and funding since the 2016 Five-Year Plan, which created incentive structures for scientists and commercialization that differ fundamentally from those in the EU and US.

pre-tenure professor or research group leader, and then tenured professor. Graduate and post-doctoral scientists work under more experienced scientists in labs and learn the techniques, tools, and organisms of their mentors. Upon getting their first tenure-track position, they must setup and run their own labs. This includes making decisions on which tools and organisms they need to conduct their research. It also requires decisions on whether and how to conduct greenhouse and field trials. Established scientists often have large labs of students and have committed to certain techniques and organisms.

3.2 The Rise of New Genomic Techniques in Plant Science

Humankind has been modifying plant traits for thousands of years through traditional breeding methods, involving the selection of desired traits based on phenotypes, crossing the selected plants, and repeating these steps until the desired traits emerge. Given the lengthiness of this process, modern plant science has developed tools that can perform gene editing more efficiently. In addition to inducing mutations in the genome (mutagenesis) and introducing genetic material from sexually compatible species (cisgenesis), these tools have also allowed the insertion of new genetic material from a sexually incompatible species (transgenesis). Prior to 2012, however, standard tools and earlier scientific advances, the so-called new genomic techniques (NGTs), had not achieved the routine and especially the precision needed to broadly apply these edits to most plants.[2]

In 2012, CRISPR emerged as a breakthrough technology in the field of genome editing and quickly became the most popular type of NGT (Huang et al., 2019; Wei et al., 2022). In contrast to earlier tools, it offered an efficient and cost-effective tool with a significantly lower rate of off-target effects, contributing to the broad applicability and thus rapid diffusion of the technology (Zyontz, 2019). Importantly, CRISPR is a general-purpose technology that not only revolutionized gene-editing in plants but also in other organisms for a multitude of applications, while simultaneously spurring a new strand of research on improving the tool itself (Zyontz, 2019). CRISPR is widely viewed as the frontier technology in genome editing, with research addressing its improvement rather than discovering alternative technologies.

[2]Originally, bacteria-directed vectors were used for gene-editing. As an alternative to this tool, earlier NGTs included Transcription Activator-Like Effector Nucleases (TALENs) and Zinc Finger Nucleases (ZNFs).

In plant science, CRISPR has especially changed the efficiency and thus the ability to perform mutagenesis. Previous methods, involving exposure of the genome to ionizing radiation or certain chemicals, produced random mutations and considerable off-target effects, hindering efficient research and commercialization. Indeed, in plant science, CRISPR has been predominantly employed for targeted mutagenesis. While transgenic applications are possible and emerging, technical constraints have so far made them less common.

3.3 The Legal Landscape for Gene-Edited Plants Prior to 2018

The United States and the European Union have long differed in how they regulate the market authorization of gene-edited plants for cultivation. The United States adopted a product-based approach early on, formalized in the Coordinated Framework for Biotechnology. The U.S. Department of Agriculture (USDA) reviews plants for product characteristics, rather than the breeding process. Subsequently, plants that could have been produced through traditional breeding methods can generally avoid classification as a genetically modified organism (GMO) - a term traditionally reserved for organisms containing foreign genetic material. This classification carries more stringent and costly testing requirements. Effectively, this means that plants modified through mutagenesis or cisgenesis are not classified as GMOs; transgenic plants, however, are. This classification is largely independent of the process through which gene-editing takes place.

In contrast, the EU adopted a process-based approach, formalized in the 2001 Directive 2001/18/EC ('GMO Directive'), which is consistently applied across all member states. Under the Directive, any gene-edited plant is treated as a standard GMO, irrespective of the modification, unless actually produced by traditional breeding methods. Compared to other jurisdictions, the requirements in the Directive increase both the length and costs of the approval process, largely due to the extensive data required from field trials. The full process typically takes at a minimum six years (Garcia-Alonso et al., 2022) and costs up to $35m (Stokstad, 2018).[3] Importantly, the Directive contains a key exemption for plants obtained through mutagenesis. While this exemption was intended to cover traditional random

[3]A cost–benefit analysis of a British proposal to adopt regulations similar to those in the US highlights the potential for significant cost and time savings: the study finds that time to market can be reduced from 10 years to 12 months and estimates cost savings of up to £14.2m and £19.8m for regulatory science and dossier submisison, respectively (Parliament of the United Kingdom, 2022).

mutagenesis techniques, its precise scope was never clearly defined.

3.4 The 2018 ECJ Ruling and Its Aftermath in the EU and US

With the emergence of CRISPR, expectations in both jurisdictions about gene-edited plants using NGTs converged. In the European Union, NGTs did not exist when the Directive was enacted, hence were not explicitly addressed in its original scope. This led to considerable optimism for the deregulation of these types of gene-edited plants as part of the mutagenesis exemption. For instance, major agricultural countries, such as the Netherlands and France, were preparing to explicitly exempt NGT-edited plants indistinguishable from conventionally bred plants and not containing transgenic material from GMO regulations (Phillips and Flach, 2017; Snipes, 2017). Even after the matter moved to the European Court of Justice (ECJ) in 2015, the optimism continued and culminated when the responsible advocate general issued a statement in favor of deregulation in January 2018 (Court of Justice of the European Union, 2018a).

Contrary to the optimistic expectations, however, the ECJ ruled in July 2018 that all gene-edited plants fall under the GMO Directive, regardless of whether they involve foreign DNA (Court of Justice of the European Union, 2018b). This 2018 regulatory decision positioned the EU as having one of the strictest approaches globally toward plants edited with NGTs. In contrast, in the United States, the USDA in 2018 confirmed expectations that gene-edited plants that could have been produced through traditional breeding would not be subject to additional regulatory oversight, a policy formalized in 2020 in the SECURE rule. After the emergence of CRISPR, both jurisdictions held similar expectations about the regulation of gene-edited plants obtained via NGT-based mutagenesis, but the 2018 ECJ ruling introduced a sharp contrast.

4 Empirical Strategy

To investigate the effect of constrained commercialization potential on the rate and direction of scientific research, we leverage the ECJ ruling in 2018 as an exogenous shock. We employ a difference-in-differences approach at the scientist level to compare the affected European plant scientists (treatment group) to their unaffected US counterparts (control group) before and after

the regulatory change. The identifying assumption is that without the ECJ ruling, outcomes of the EU and US scientists would have followed parallel trends. Our baseline specification is as follows:

$$Outcome_{i,t} = \beta_1(Treat_i \times Post2018_t) + \gamma' X_{i,t} + \delta_i + \lambda_t + \varepsilon_{i,t} \quad (1)$$

The dataset is a panel dataset at the scientist-level. Our observation period spans the years 2013 to 2023, starting the year after CRISPR emerged and including the last year of reliable bibliometric data available. $Treat_i$ is an indicator for whether scientist i belongs to the treatment group (affected European scientists). $Post2018_t$ is an indicator equal to 1 in periods after the regulatory change takes effect (from 2019 onwards). The interaction $(Treat_i \times Post2018_t)$ is the difference-in-differences term, and its coefficient β_1 captures the impact of the policy change on affected European scientists. We include time-varying covariates $\gamma' X_{i,t}$ as well as scientist and year fixed effects, δ_i and λ_t, respectively. Standard errors are clustered at the scientist level.

The validity of our analysis depends on correctly identifying scientists whose work, regardless of where it falls along the research process, contributes to the development of gene-edited crops – and is therefore exposed to changes in their commercial regulation. To construct such a set of scientists, we leverage the fact that we can observe who adopts CRISPR, the frontier gene-editing tool, for crop research. We examine the adopters' topical focus prior to CRISPR's emergence, and identify scientists with a similar pre-CRISPR topical profile, regardless of whether they themselves adopted the technology.

The reason for prioritizing topical relevance over the adoption of CRISPR to define the analysis set is twofold. First, we avoid conditioning on CRISPR usage, an outcome itself. Second, we capture scientists who do not apply CRISPR but nonetheless contribute to the development of gene-edited crops. As confirmed by our interviews to date with plant scientists, the process of achieving a gene-edited crop is lengthy, involving many steps before and after applying the gene-editing tool. Thus, not every scientist contributing to research on gene-edited crops performs the editing steps, and some may never do so. Our approach yields a set of scientists at risk of being affected by changes to the commercial potential of research on gene-edited crops, at any stage of the research process. We then delineate treatment and

control scientists by the jurisdiction of their affiliations.

5 Data

We use bibliometric data to construct the risk set comprising scientists active in the same fields prior to CRISPR's emergence as our reference group of CRISPR adopters in crop research. We leverage the topical focus of relevant CRISPR-adopters between 2009 and 2011 to define a baseline risk set. We then refine the baseline risk set in several steps to obtain a refined set of treated (EU) and control (US) plant scientists, with characteristics similar to those of eventual CRISPR adopters. Finally, we construct outcome variables to measure scientists' responses in terms of relocation, pivoting toward different research topics, and establishing cross-jurisdictional ties.

5.1 Detecting CRISPR-Adopters in Crop Research

Scientists who adopt NGTs in crop research serve as the reference group for constructing our risk set. We identify such scientists through publications. First, we query the bibliometric database Scopus for the term 'CRISPR' in title, abstract and keywords. We chose this query term because CRISPR is the most prominent NGT, is unlikely to occur outside genome editing, and, as the central tool, is typically mentioned in titles, abstracts, or keywords. Additionally, the query term has been validated in prior work that quantifies the use of CRISPR in research and commercialization (Martin-Laffon et al., 2019; Zyontz and Pomeroy-Carter, 2021; Zyontz, 2024). This initial query yields 55,159 publications, of which we exclude 1,374 (2.49%) because of missing title and abstract or duplication.

To discern plant scientists focused on crops within the set of scientists working with CRISPR, we analyze all CRISPR publications to identify their focal organism and application area. Organism categories include archaea, bacteria, viruses, mammals, plants, and other eukaryotes; application categories consist of crops, health/medical, livestock, technical improvements, industrial applications, and environmental applications. Specifically, we employ a multi-step method leveraging the large language model Claude Haiku 3.5 to classify the organism and application category of each publication. As a first step, we use the model to filter out any social science publication, for example, those discussing ethical considerations around gene

editing. 50,935 publications proceed to the next steps. Second, we prompt the model to return a confidence score between 0 and 100 for each organism category, while all scores must add up to 100. In a third step, we prompt the model to perform the same task for application categories.[4] Appendices A.1.1 to A.1.3 contain the respective prompts given to Claude Haiku 3.5, while Appendix A.1.4 shows an illustration of our classification pipeline.

Based on the model's output, we determine the primary and secondary categories for both organism and application, taking into account the strength of the confidence scores to flag uncertainties and identify hybrid publications.[5] Figures 1 and 2 display the distribution of the primary organism and application categories classified as 'strong', meaning that the model returns a confidence score of at least 60 for the respective category. The distribution is qualitatively similar to prior work classifying CRISPR publications. For example, the most prominent categories in our sample are mammals for organisms and health/medicine for applications, consistent with the dominance of medical applications in humans reported in previous studies (Zyontz and Pomeroy-Carter, 2021; Zyontz, 2024). This consistency underscores the accuracy and reliability of our classification process.

We use the classified CRISPR publications to identify a reference group that includes only scientists committed to research on gene-edited crops. Therefore, we limit the reference group to authors of publications with plants and crops as the 'strong' primary categories, i.e. with a confidence score of 60 and higher. A focus on crops ensures that scientists do not primarily research, for example, the environmental function of certain plants, which may continue unperturbed by the regulatory shock to the *commercialization* of plants we exploit. To further ensure commitment to the field, we only consider scientists with at least two such publications between 2012 and 2024. Lastly, as we construct the risk set based on scientists' characteristics between 2009 and 2011, just before CRISPR emerged, we require at least one publication in this period. The final reference group comprises 2,300 scientists.

[4]For 503 publications, the model returns incomplete results and affected publications are currently excluded.

[5]For example, a publication on human vaccine development would receive high scores for both the 'virus' and 'mammal' organism categories.

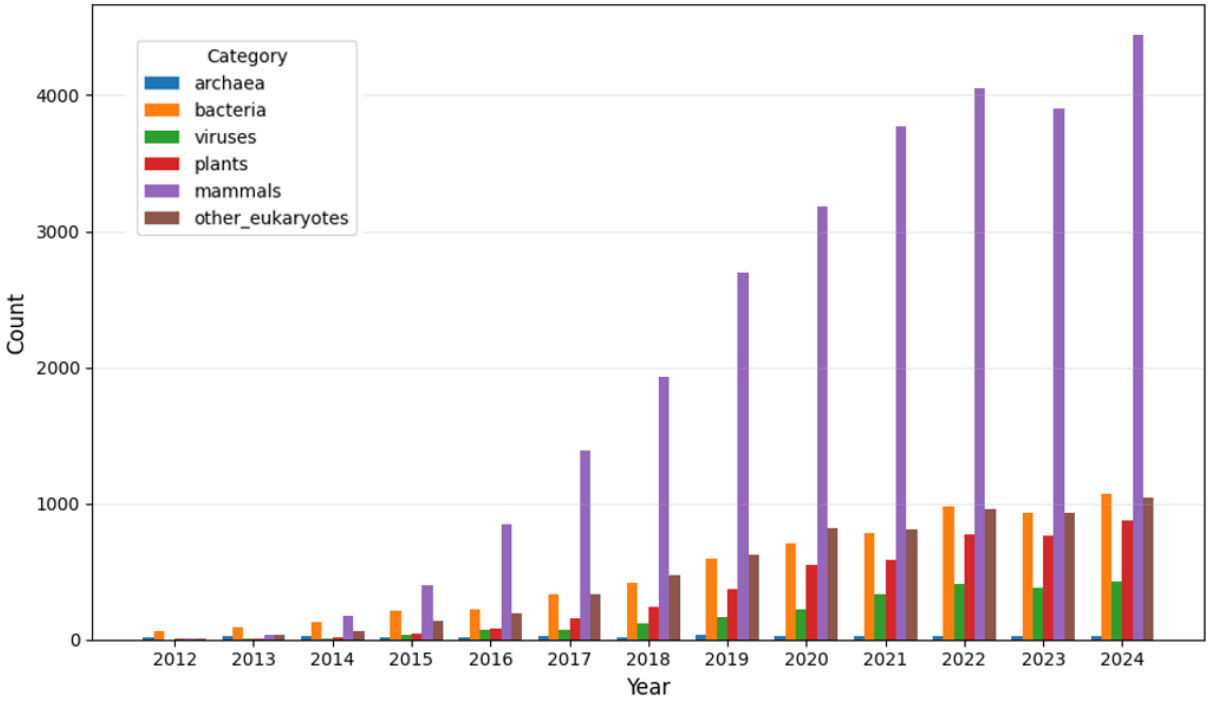


Figure 1: CRISPR Publications by Primary Organism over Time, 2012-2024

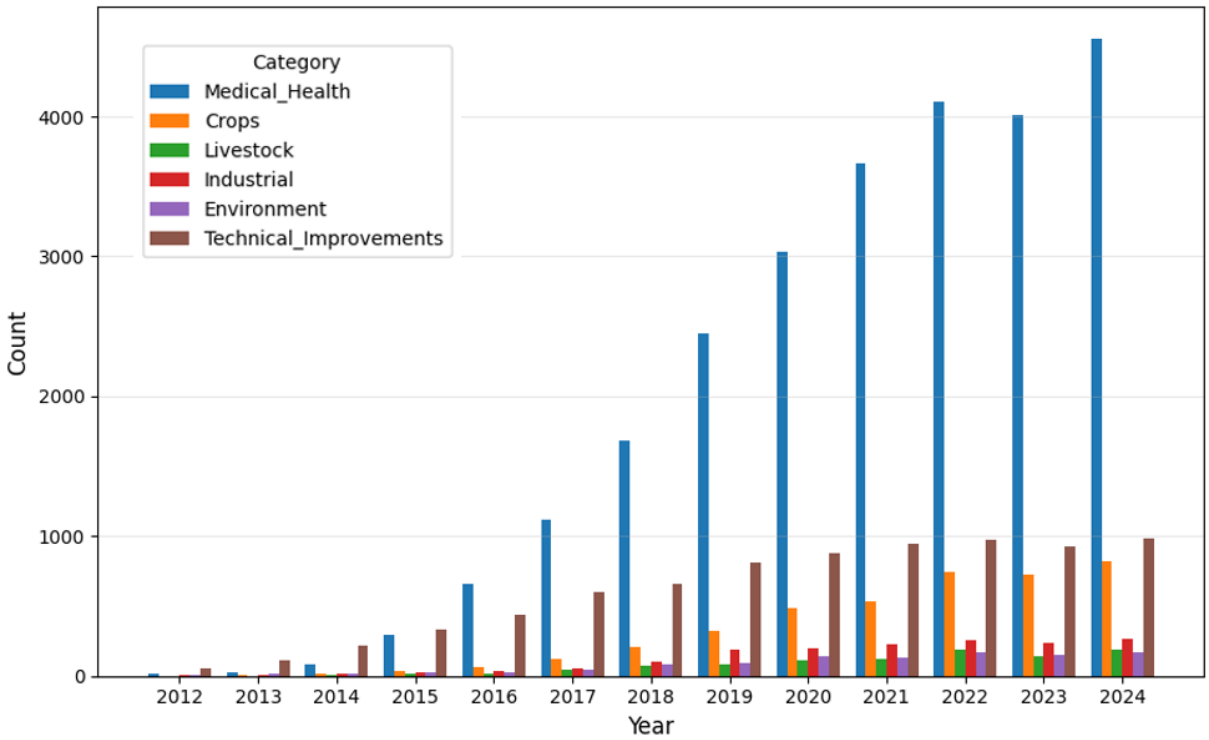


Figure 2: CRISPR Publications by Primary Application over Time, 2012-2024

5.2 Constructing the Risk Set

First, we construct the baseline risk set that includes all scientists working on the same topics as those in which the reference group was predominantly engaged between 2009 and 2011. To identify scientists working on the same research topics as our reference group between 2009 and 2011, we classify reference scientists' publications in the period using Scopus SciVal topics. Scopus assigns one SciVal topic to every publication using direct citation analysis (Waltman and Van Eck, 2012), yielding a fine-grained grouping of topically related publications. We consider relevant the topics that reference scientists are predominantly engaged in, specifically those that appear more than ten times among the publications of our reference group between 2009 and 2011. In total, 243 topics meet this requirement, accounting for roughly 65% of author-publication pairs. Reassuringly, the topics treated as relevant overwhelmingly reflect crop research.[6] We then retrieve all scientists with at least one publication in these topics between 2009 and 2011. We consider the resulting 135,048 scientists the baseline risk set.

Second, we refine the baseline risk set in three steps to identify treatment and control scientists who plausibly contribute to the development of gene-edited crops – and are thus exposed to changes in their commercial regulation. First, we restrict the baseline set to scientists who geographically fit our treatment and control groups. As treated scientists, we consider those working in the European Union, as control scientists those working in the United States. We treat the United Kingdom part of the European Union. Specifically, we require scientists to be affiliated exclusively with an EU or US institution in the last observed publication between 2009 and 2011. Second, we ensure scientists are sufficiently invested in the 243 topics we deem relevant to the reference group. The median share of publications within these topics in the reference group is 75%. Hence, we require eligible scientists to have published at least 75% of their papers between 2009 and 2011 in the relevant topics, excluding those whose research activities are not concentrated on crop research. Third, we only include scientists with their last observed publication in 2023 or later. This restriction mitigates the confounding of our results by retirement and ensures we observe active scientists only. The refined risk set includes 6,294 scientists.

[6] Topics include, for example, "Drought Response Mechanisms in Crop Plants" and "Arabinoxylans and Xylanase Inhibitors in Wheat".

Third, we apply coarsened exact matching (CEM) to ensure similar pre-treatment covariate distributions between treated and control scientists, thus improving the plausibility of the parallel trends assumption. The matching criteria account for differences in productivity, output quality, scientists' resources and quality of training, as well as their topic diversity. Matching is performed over the 2009–2011 period. As a measure of productivity, we use the total number of publications in the time period. To match on output quality, we use the SCImago Journal Rank Indicators for the respective years and take the average across publications. Next, we retrieve the institutional rank of scientists' affiliations. We use the subject-specific QS rankings for Biological Sciences as well as Agricultural Sciences aand Forestry in 2013, and the general QS ranking over all subjects from 2011.[7] For matching, we choose a scientist's last affiliation's ranking between 2009 and 2011, taking the best rank if multiple affiliations or rankings exist. Addtionally, we match on the organization type (company, research institute, university including college and medical school, other institution), retrieved from Scopus. Next, we match on the number of relevant SciVal topics a scientist publishes in between 2009 and 2011, signalling how topically diverse a scientists' research is. Lastly, we match on scientists' career stage, proxied by their first publishing year. We exclude scientists with a first publication before 1980, as these records likely conflate multiple individuals given the implausibly long career spans. Of 6,294 scientists in the refined risk set, 5,391 (85.7%) scientists can be matched. When analyses are conducted on subsets of the full sample, we re-apply CEM within each subset.

5.3 Measuring Scientists' Responses

We construct outcome variables to investigate all three conceivable responses by scientists: relocating to a more favourably regulated jurisdiction, pivoting away from research contributing to the development on gene-edited crops, and establishing ties with other jurisdictions to maintain commercialization potential. We also consider effects on the rate of research, which may plausibly be affected by the different types of responses. All outcome variables are constructed at the scientist level. First, to identify relocation decisions, we trace changes in scientists' affiliations as reported on publications. We define relocation as a binary

[7] 2013 is the first year where subject-specific QS rankings are available.

variable equal to one when all of a scientist's previous affiliations disappear from one year to the next and are replaced by one or more new affiliations. The new affiliations must be in a different jurisdiction, thus we exclude domestic moves within the same jurisdiction. When a scientist has no publication in a given year and their affiliation is unknown, we assign the relocation event to the earliest year after the last recorded affiliation.

Our second set of outcome variables captures pivoting away from research contributing to the development of gene-edited crops. To place these variables into context and spot potential trade-offs, we initially examine the rate of research, measured by the total number of publications produced in a given year. We then capture signs of pivoting along two dimensions. First, we examine CRISPR usage as an outcome, as it captures the most explicit form of contribution to gene-edited crop research. We construct both a binary variable, capturing whether the scientist publishes any publication using CRISPR, and a count measure of the number of CRISPR-related publications in a given year. We use our earlier set of CRISPR publications retrieved via a keyword search on Scopus to construct these variables. Second, we capture broader shifts away from plant research altogether. Using Claude Haiku again, we classify each publication as plant-related or not, with details reported in Appendix A.2. We then calculate a scientist's share of plant-related publications in each year. In years where a scientist does not publish, the plant share is set to zero, effectively treating the absence of publications as a lack of contribution to plant science. To address the prevalence of zeros in count variables, we use the inverse hyperbolic sine (IHS) of the count variable in the main linear specification.

To investigate our third response, establishing cross-jurisdictional ties, we again draw on affiliation data. We identify gains of affiliations, a binary variable taking the value of one in years where new affiliations are added while at least one affiliation from the previous year is retained. As before, we exclude gaining affiliations within the same jurisdiction and attribute the outcome to the earliest year after the last recorded affiliation in case of years without publication. We interpret this measure as scientists establishing ties with institutions in other jurisdictions, without physically relocating.

We use the entire sample of scientists to estimate the effect of the ECJ ruling in 2018 on

scientists' relocation choices. For all other outcomes, we exclude scientists that relocate to another jurisdiction during our observation period, focusing on scientists that remain in their original jurisdiction. Table 1 reports descriptive statistics on the outcome variables prior to treatment for scientists included in the risk set. Statistics are computed on the CEM-matched sample using the matching weights. While EU and US scientists exhibit similar pre-treatment relocation rates and overall productivity, we observe differences in outcomes on CRISPR usage and plant focus as well as their propensity to gain affiliations. However, identification in our difference-in-differences approach relies on parallel trends in outcomes rather than equal levels, which we discuss in further detail in Section 6.1.1. We also note that several outcomes, such as relocating, gaining affiliations, or using CRISPR, are rare events, which may introduce noise and limit statistical power. While this implies that some estimates may be imprecise, it does not diminish the importance of these outcomes. Even rare changes can carry lasting consequences: relocation, for example, entails a loss of resources, knowledge, and the training of the next generation of scientists, while refraining from using a frontier tool may impact the future competitiveness of a jurisdiction.

6 Results

We report results from our difference-in-differences approach, examining scientists' responses to curtailed commercialization opportunities in terms of relocation to different jurisdictions, pivoting away from the affected research topic, and establishing ties with other jurisdictions to preserve the commercial potential of their research. All specifications are estimated on matched samples and incorporate the weights generated from coarsened exact matching. Throughout the analysis, we control for past productivity and actvity, as well as increasing seniority. We first present the effects for the full sample of scientists before investigating heterogeneity by career stage.

6.1 Main Results

6.1.1 Identifying Assumptions

The causal interpretation of our results relies on the parallel trends assumption, namely that European scientists' outcome trends would have evolved similarly to those of US scientists, the control group, absent the treatment. Through coarsened exact matching, we ensure scientists

Table 1: Descriptive Statistics for US (Control) and EU (Treatment) Plant Scientists in the Matched Sample

N in Full Matched Sample = 5,391	**Control Group** (US plant scientists) *N = 1,841*	**Treatment Group** (EU plant scientists) *N = 3,550*	**Difference** (Treat. – Control)
First publication year	2002.0 (8.25)	2002.0 (8.06)	−0.007
Relocation rate per year 2013–2018	0.010 (0.100)	0.008 (0.091)	−0.002
Number of publications per year 2013–2018	1.569 (1.952)	1.482 (1.798)	−0.086
Any CRISPR publication per year 2015–2018	0.018 (0.135)	0.010 (0.102)	−0.008**
Number of CRISPR publications per year 2015–2018	0.020 (0.154)	0.011 (0.113)	−0.009***
Plant share per year 2013–2018	0.484 (0.490)	0.516 (0.487)	0.032*
Affiliation gain rate per year 2013–2018	0.027 (0.162)	0.018 (0.132)	−0.009***

Notes: Values are means with standard deviations in parentheses. Differences represent the weighted difference-in-means between control and treatment groups, respecting the weights generated through coarsened exact matching. Standard errors are clustered at the scientist level. The mean first publication year and relocation rate are calculated on the entire sample of scientists, all further mean outcomes for scientists that do not relocate during the observation period. For CRISPR-related outcomes, we exclude the years 2013 and 2014 as CRISPR researchers focused on improving the tool, rather than applying it to plants or other organisms. When a scientist has no publication in a given year and their affiliation is unknown, we conservatively assign relocation and affiliation gain to the earliest year after the last recorded affiliation. Plant share equals zero in years where a scientist does not publish, treating non-publication as part of the measure tracking contributions to plant science.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

in the treatment and control group enter the observation period with similar observable characteristics that typically shape career trajectories. As outlined previously, CRISPR poses a breakthrough technology for the entire field, independent of geography, and regulatory expectations were similar in both jurisdictions until the ECJ ruling in 2018. Accordingly, we see no compelling reason to expect differential trends between European and US scientists in absence of the ECJ ruling. While we cannot definitively prove the parallel trends assumption, we find support for it in the pre-treatment data. In particular, we conduct the respective event study for each outcome variable.[8] In support of parallel trends prior to the 2018 regulatory shock, we find that pre-treatment coefficients in the event study are jointly statistically insignificant for all outcomes.

6.1.2 Treatment Effects

Our main results on the full sample of scientists are reported in Table 2. For the decision to relocate, results in Column (1) indicate a statistically significant increase in the probability that, relative to US scientists, EU scientists relocate to a different jurisdiction after 2018. Though the increase is small in absolute terms - translating to a 0.5 percentage point increase in relocation probability - it represents a 62.5% increase relative to the pre-treatment mean (0.8% in a given year) of EU scientists. Additionally, the effect is likely a lower-bound estimate given our post-treatment period covers the COVID19 pandemic, which limited mobility substantially. In any case, scientists emigrating may have spillovers beyond the scientist themselves to their former institution where research groups, students, and co-authors lose colleagues and mentors, contributing to long-term consequences from emigrating scientists.

Results on scientists pivoting away from research on gene-edited plants are reported in Columns (2) to (5). Overall, these are indicative of a shift away from research on gene-edited plants, though not away from plants as such. To put these outcomes into context, Column (2) shows results on scientists' rate of research, their overall publication output, regardless of topic or tool. We do not find any statistically significant effect of the regulatory shock on the overall productivity of treated scientists. Next, Columns (3) and (4) report results on the extensive and intensive margins of CRISPR usage in scientists' research. Both coefficients are

[8]The full event study results are available upon request.

statistically significant and negative, indicating that EU scientists become less likely to use CRISPR and publish fewer CRISPR-related papers relative to their US counterparts following the 2018 regulation. For the IHS-transformed count variable, we calculate retransformed marginal effects on the original scale following Norton (2022). The retransformed effect is –0.023 publications per author-year (p = 0.001), a substantial reduction given that the pre-treatment mean among EU scientists was 0.011 - reflecting a technology still in the early stages of adoption.[9] The significant negative effect on the binary indicator in Column (3) suggests that the result is primarily driven by the extensive margin - whether scientists publish any CRISPR-related work at all - which is consistent with the low baseline rate of CRISPR publications in the sample. In contrast, as shown in Column (5), we find no evidence of a signifcant shift away from plant research as such. These estimates point to a dampening of research with CRISPR, that may substantially impact long-term research trajectories as CRISPR is only emerging during our observation period.

Lastly, Column (6) presents results on whether affected scientists establish cross-jurisdictional ties through affiliations in response to the 2018 regulatory shock. We find a statistically signficant increase in the probability for EU scientists to add affiliations in other jurisdictions after 2018, relative to their US counterparts. While we note that the effect is only weakly significant at the 10% level, this finding suggests that some scientists expanded their institutional footprint outside the European Union. Again, a 0.6 percentage point increase may be practically meaningful given the treated scientists' 1.8% average probability to gain a cross-jurisdictional affiliation in the pre-period, representing a 33% relative increase. How exactly scientists may use affiliations in other jurisdictions to their advantage warrants further attention in the future development of this paper.

6.2 Results by Career Stage

6.2.1 Approach

The different response types - relocating, pivoting topic, establishing cross-jurisdictional ties - hinge on different kinds of opportunities, constraints, and costs. As these factors plausibly change

[9] At low values of the outcome, the IHS transformation approximates the identity function, and the retransformed marginal effect is nearly identical to the IHS-scale coefficient (Norton, 2022). As a further robustness check, we plan to estimate a Poisson pseudo-maximum-likelihood specification.

Table 2: Effects in Full Sample

	(1) Relocation	(2) All Pubs. (IHS)	(3) Any CRISPR Pub.	(4) CRISPR Pubs. (IHS)	(5) Plant Share	(6) Cross-Jur. Ties
EU Scientist × Post-2018	0.005**	−0.019	−0.020***	−0.023***	0.004	0.006*
	(0.002)	(0.021)	(0.007)	(0.007)	(0.011)	(0.003)
Scientist FE and Year FE	✓	✓	✓	✓	✓	✓
Seniority	✓	✓	✓	✓	✓	✓
Publication Count (t-1)	✓	✓	✓	✓	✓	✓
Publication Stock (t-2)	✓	✓	✓	✓	✓	✓
Active (t-1, t-2)	✓	✓	✓	✓	✓	✓
Observations	59,301	55,429	45,351	45,351	55,429	55,429
Scientists	5,391	5,039	5,039	5,039	5,039	5,039
Treated Scientists	3,550	3,346	3,346	3,346	3,346	3,346

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the scientist level and shown in parentheses. The sample includes EU and US plant scientists with affiliations exclusively in their respective jurisdiction during their last observed year with a publication (2009–2011). Column (1) uses the full sample of scientists. Columns (2)–(6) restrict the sample to scientists who do not relocate during the observation period. For CRISPR-related outcomes, we exclude the years 2013 and 2014 as CRISPR researchers focused on improving the tool, rather than applying it to plants or other organisms. All specifications include scientist and year fixed effects. Time-varying controls include seniority, the inverse hyperbolic sine of prior-year publication counts and publication stock (t-2), as well as an indicator for any publication activity within the past two years. Seniority is captured by a time-varying categorical variable based on years since first publication, with cutoffs at 5, 11, and 19 years.

over a scientist's career, we investigate how responses vary by career stage in our setting. We split the sample by career stage, allowing fixed effects, error variances, and control variable coefficients to vary flexibly across subsamples. We use the year of a scientist's first publication to deduce their career stage and chose cutoffs to generate three subsamples with different positions at the time of the regulatory shock in the European Union. Figure 3 depicts the career development of the median scientist by academic age in each subsample.[10] Early career scientists, with a first publication between 2008 and 2011, are on the verge of changing from a postdoctoral position to Group Leader or Assistant Professor at the end of 2018. Mid career scientists, having published their first paper between 2003 and 2007, are in the midst of their first Group Leader or Assistant Professor position. Established career scientists, who published their first paper before 2003, have secured a permanent position as Full Professor in 2018. Again, we find support for our assumption of parallel trends absent the treatment within each subsample in the respective event study.

[10] We assume the following career path: five-year PhD program with the first publication in the second year, five years of postdoctoral research, six years as a Group Leader or Assistant Professor, followed by a permanent postion as a Full Professor.

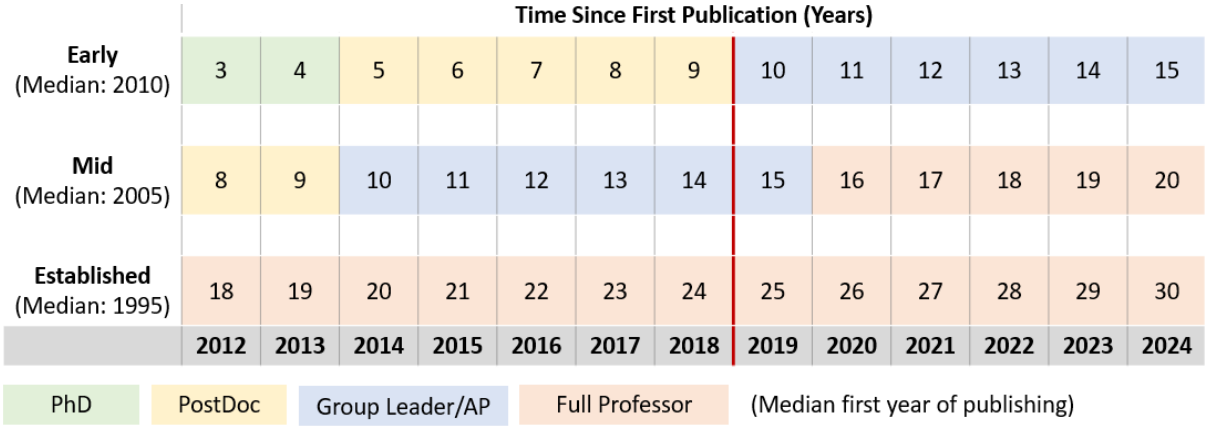


Figure 3: Career Development over the Observation Period by Career Stage

6.2.2 Treatment Effects

Table 3 presents results from the difference-in-differences specification estimated separately for early career, mid career and established career scientists. For early career scientists, results in Columns (3) and (4) show a statistically significant negative effect on the intensive and extensive margins of CRISPR usage: treated EU scientists are less likely to use CRISPR and publish fewer papers using CRISPR in response to reduced commercial potential, relative to early career scientists in the United States. This indicates a shift away from using the frontier gene-editing tool at the time. In contrast, we do not find statistically significant effects for any other reponse, including the share of plant publications. Taken together, early career scientists appear to pivot away from the frontier gene-editing tool, CRISPR, but not from research on plants as such.

For mid career scientists, we find a statistically significant decline in the rate of research by treated EU scientists relative to US scientists at the same career stage, as shown in Column (2). This decline is meaningful: the retransformed marginal effect amounts to a reduction of 0.209 publications per author-year ($p = 0.006$), or 16.0% relative to the pre-period mean of 1.307 publications among EU mid career scientists. Additionally, the share of publications on plants declines significantly for EU scientists after the regulatory shock, as indicated by results in Column (5). One possible explanation consistent with these results is that mid career scientists attempt to pivot away from plants to other organisms. Changing the focal organism represents a major pivot and requires acquiring new knowledge, which may cause the observed rate of research to decline. Nonetheless, our setting does not allow us to distinguish this

explanation from an equally possible one: other factors associated with the regulatory shock negatively impact scientists' productivity, which falls entirely on their plant research. We probe these two explanations in ongoing interviews with European plant scientists.

For established scientists, results in Columns (1) and (6) indicate that these scientists rely on their network to respond to the regulatory shock. We find a statistically significant increase for EU scientists to establish cross-jurisdictional ties through gaining affiliations, and a weakly significant increase in relocation, relative to their US counterparts. Additionally, Column (5) shows a significantly positive effect on established EU scientists' share of plant publications, suggesting that scientists increase their focus on plant research.

Collectively, these results suggest that scientists' responses vary by career stage. The observed treatment effects are consistent with the career stage moderating the different opportunities, constraints, and costs associated with each response. Established scientists relocating, for instance, aligns with findings in prior literature that especially productive scientists and those past having adolescent children are mobile (Ganguli, 2015; Azoulay et al., 2017). Similarly, established scientists may have plausibly accumulated the connections and reputation to gain cross-jurisdictional affiliations. Results suggest early and mid-career scientists both pivot, but along different dimensions – a pattern that could reflect differences in which aspects of their research agendas remain flexible enough to adjust at each career stage. In particular, expertise in a research tool may become more constraining than expertise in a focal organism over time, explaining why early-career scientists appear to shift away from the tool while mid-career scientists shift away from plants as their focal organism. Again, we are investigating these explanations in ongoing interviews with plant scientists.

7 Conclusion

Does commercial regulation shape the production of science? To study this question, we investigate a 2018 regulatory shock to commercializing gene-edited plants in the European Union, which left European plant scientists with little commercial potential in their domestic market. This stood in stark contrast to their US counterparts. Our results indicate that some scientists respond by moving outside the affected jurisdction, others attempt to establish

international ties through gaining affiliations outside their home jurisidction. We also observe some scientists pivoting away from the affected research topic. Additionally, our results indicate that responses differ by career stage, which we explore further in ongoing interviews with EU and US plant scientists at different stages in their careers. Beyond adding to the understanding of scientists' motivations, our work informs policy debates on regulating gene-edited plants and emerging frameworks for general-purpose technologies, highlighting a critical issue: if commercial regulation affects the production of science, this may alter the supply of knowledge and inventions to firms. Our results will thus prove informative for firms that have to decide the extent to which they invest in research and development versus relying on academic science to provide useful knowledge and inventions. For policymakers, our results underscore that regulations targeting commercial applications can have upstream consequences for scientific production itself - with potential implications for the long-run competitiveness of a jurisdiction.

8 Acknowledgements

Hofmeister acknowledges support from a fellowship by the Joachim Herz Foundation. The authors declare no other financial relationships or conflicts of interest.

Table 3: Effects by Career Stage

	(1) Relocation	(2) All Pubs. (IHS)	(3) Any CRISPR Pub.	(4) CRISPR Pubs. (IHS)	(5) Plant Share	(6) Cross-Jur. Ties
Subsample A: Early Career						
EU Scientist × Post-2018	0.004	−0.036	−0.030***	−0.033***	−0.011	−0.001
	(0.003)	(0.039)	(0.009)	(0.011)	(0.020)	(0.006)
Observations	20,823	18,546	15,174	15,174	18,546	18,546
Scientists	1,893	1,686	1,686	1,686	1,686	1,686
Treated Scientists	1,250	1,112	1,112	1,112	1,112	1,112
Subsample B: Mid Career						
EU Scientist × Post-2018	0.003	−0.123***	−0.014	−0.022	−0.048**	0.009
	(0.005)	(0.044)	(0.012)	(0.014)	(0.022)	(0.007)
Observations	13,519	12,551	10,269	10,269	12,551	12,551
Scientists	1,229	1,141	1,141	1,141	1,141	1,141
Treated Scientists	744	702	702	702	702	702
Subsample C: Established Career						
EU Scientist × Post-2018	0.007*	0.046	−0.014	−0.017	0.039**	0.009**
	(0.004)	(0.032)	(0.012)	(0.012)	(0.017)	(0.004)
Observations	24,959	24,332	19,908	19,908	24,332	24,332
Scientists	2,269	2,212	2,212	2,212	2,212	2,212
Treated Scientists	1,556	1,532	1,532	1,532	1,532	1,532
Scientist FE and Year FE	✓	✓	✓	✓	✓	✓
Publication Count (t-1)	✓	✓	✓	✓	✓	✓
Publication Stock (t-2)	✓	✓	✓	✓	✓	✓
Active (t-1, t-2)	✓	✓	✓	✓	✓	✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the scientist level and shown in parentheses. The sample includes EU and US plant scientists with affiliations exclusively in their respective jurisdiction during their last observed year with a publication (2009–2011). Column (1) uses the full sample of scientists in each career stage. Columns (2)–(6) restrict the sample to scientists who do not relocate during the observation period in each career stage. For CRISPR-related outcomes, we exclude the years 2013 and 2014 as CRISPR researchers focused on improving the tool, rather than applying it to plants or other organisms. All specifications include scientist and year fixed effects. Time-varying controls include seniority, the inverse hyperbolic sine of prior-year publication counts and publication stock (t-2), as well as an indicator for any publication activity within the past two years. Seniority is captured by a time-varying categorical variable based on years since first publication, with cutoffs at 5, 11, and 19 years.

References

Aghion, P., Dewatripont, M., and Stein, J. C. (2008). Academic Freedom, Private-Sector Focus, and the Process of Innovation. *The RAND Journal of Economics*, 39(3):617–635.

Akcigit, U., Baslandze, S., and Stantcheva, S. (2016). Taxation and the International Mobility of Inventors. *American Economic Review*, 106(10):2930–2981.

Arora, A., Belenzon, S., and Dionisi, B. (2023). First-Mover Advantage and the Private Value of Public Science. *Research Policy*, 52(9):104867.

Azoulay, P., Fons-Rosen, C., and Zivin, J. S. G. (2019). Does Science Advance One Funeral at a Time? *American Economic Review*, 109(8):2889–2920.

Azoulay, P., Ganguli, I., and Zivin, J. G. (2017). The Mobility of Elite Life Scientists: Professional and Personal Determinants. *Research Policy*, 46(3):573–590.

Cohen, W. M., Sauermann, H., and Stephan, P. (2020). Not in the Job Description: The Commercial Activities of Academic Scientists and Engineers. *Management Science*, 66(9):4108–4117.

Court of Justice of the European Union (2018a). Judgment in Case C-528/16: Confédération paysanne and Others v Premier ministre and Ministre de l'Agriculture, de l'Agroalimentaire et de la Forêt.

Court of Justice of the European Union (2018b). Organisms Obtained by Mutagenesis Are GMOs and Are, in Principle, Subject to the Obligations Laid Down by the GMO Directive.

Fleming, L., Greene, H., Li, G., Marx, M., and Yao, D. (2019). Government-Funded Research Iincreasingly Fuels Innovation. *Science*, 364(6446):1139–1141.

Fleming, L. and Sorenson, O. (2004). Science as a Map in Technological Search. *Strategic Management Journal*, 25(8-9):909–928.

Fry, C. V. (2023). Local Researchers, Local Problems. *Working Paper*.

Furman, J. L., Murray, F., and Stern, S. (2012). Growing Stem Cells: The Impact of Federal Funding Policy on the US Scientific Frontier. *Journal of Policy Analysis and Management*, 31(3):661–705.

Ganguli, I. (2015). Who Leaves and Who Stays? Evidence on Immigrant Selection from the Collapse of Soviet Science. In *Global Mobility of Research Scientists*, pages 133–154. Elsevier.

Garcia-Alonso, M., Novillo, C., Kostolaniova, P., Martinez Parrilla, M., Alcalde, E., and Podevin, N. (2022). The EU's GM Crop Conundrum: Did the EU Policy Strategy to Convert EFSA GMO Guidance into Legislation Deliver on Its Promises? *EMBO Reports*, 23(5):e54529.

Hill, R., Yin, Y., Stein, C., Wang, X., Wang, D., and Jones, B. F. (2025). The Pivot Penalty in Research. *Nature*, pages 1–8.

Hosseini, S. M. and Lichtinger, G. (2025). Generative AI as Seniority-Biased Technological Change: Evidence from U.S. Resume and Job Posting Data. *Working Paper*.

Huang, Y., Porter, A., Zhang, Y., and Barrangou, R. (2019). Collaborative Networks in Gene Editing. *Nature Biotechnology*, 37:1107–1109.

Hvide, H. K. and Jones, B. F. (2018). University Innovation and the Professor's Privilege. *American Economic Review*, 108(7):1860–1898.

Ivus, O. and MacGarvie, M. (2025). Reclaiming Brain Power: Intellectual Property Rights and Inventor Migration. *Working Paper*.

Jones, B. F. (2009). The Burden of Knowledge and the "Death of the Renaissance Man": Is Innovation Getting Harder? *The Review of Economic Studies*, 76(1):283–317.

Krieger, J. L., Schnitzer, M., and Watzinger, M. (2024). Standing on the Shoulders of Science. *Strategic Management Journal*, 45(9):1670–1695.

Leahey, E. (2007). Not by Productivity Alone: How Visibility and Specialization Contribute to Academic Earnings. *American Sociological Review*, 72(4):533–561.

Liu, L., Dehmamy, N., Chown, J., Giles, C. L., and Wang, D. (2021). Understanding the Onset of Hot Streaks across Artistic, Cultural, and Scientific Careers. *Nature Communications*, 12(1):5392.

Martin-Laffon, J., Kuntz, M., and Ricroch, A. E. (2019). Worldwide CRISPR Patent Landscape Shows Strong Geographical Biases. *Nature Biotechnology*, 37(6):613–620.

Merton, R. K. (1973). The Normative Structure of Science. In Merton, R. K., editor, *The Sociology of Science: Theoretical and Empirical Investigations*, pages 267–278. University of Chicago Press, Chicago. Originally published in 1942.

Moscona, J. and Sastry, K. A. (2025). Inappropriate Technology: Evidence from Global Agriculture. *National Bureau of Economic Research Working Paper*.

Myers, K. (2020). The Elasticity of Science. *American Economic Journal: Applied Economics*, 12(4):103–134.

Norton, E. C. (2022). The Inverse Hyperbolic Sine Transformation and Retransformed Marginal Effects. *The Stata Journal*, 22(3):702–712.

Parliament of the United Kingdom (2022). Impact Assessment - Genetic Technology (Precision Breeding) Bill. *Publications of the Parliament of the United Kingdom*.

Phillips, S. and Flach, B. (2017). Dutch Proposal to Legislate NBTs. *United States Department of Agriculture, Foreign Agricultural Service*.

Roach, M. and Sauermann, H. (2010). A Taste for Science? PhD Scientists' Academic Orientation and Self-Selection into Research Careers in Industry. *Research Policy*, 39(3):422–434.

Snipes, K. (2017). Agricultural Biotechnology Annual: France. *United States Department of Agriculture, Foreign Agricultural Service*.

Stephan, P. E. (1996). The Economics of Science. *Journal of Economic Literature*, 34(3):1199–1235.

Stern, S. (2004). Do Scientists Pay to Be Scientists? *Management Science*, 50(6):835–853.

Stokstad, E. (2018). European Court Ruling Raises Hurdles for CRISPR Crops. *Science*.

Thursby, J. G. and Thursby, M. C. (2011). Has the Bayh-Dole Act Compromised Basic Research? *Research Policy*, 40(8):1077–1083.

Waltman, L. and Van Eck, N. J. (2012). A New Methodology for Constructing a Publication-Level Classification System of Science. *Journal of the American Society for Information Science and Technology*, 63(12):2378–2392.

Wei, X., Pu, A., Liu, Q., Hou, Q., Zhang, Y., An, X., Long, Y., Jiang, Y., Dong, Z., and Wan, X. (2022). The Bibliometric Landscape of Gene Editing Innovation and Regulation in the Worldwide. *Cells*, 11(17):2682.

Zeng, A., Shen, Z., Zhou, J., Fan, Y., Di, Z., Wang, Y., Stanley, H. E., and Havlin, S. (2019). Increasing Trend of Scientists to Switch between Topics. *Nature Communications*, 10(1):3439.

Zyontz, S. (2019). Making the Cut: the Rate and Direction of CRISPR Innovation. *MIT Doctoral Dissertation*.

Zyontz, S. (2024). Mapping CRISPR Research, Innovation, and Diffusion Activities in Agriculture – Cross-national. *Commission of Experts for Research and Innovation (EFI)*, (7-2024).

Zyontz, S. and Pomeroy-Carter, C. (2021). Mapping of the Research, Innovation and Diffusion Activity of CRISPR Across Countries. *Commission of Experts for Research and Innovation (EFI)*, (12-2021).

A Appendix

A.1 Classifying CRISPR Publications

A.1.1 Prompt to Exclude Social Science Publications

You are a scientific text classification system. Determine whether the following publication, based on the title and abstract, fulfills at least one of the following definitions:

Definition 1: a review, summary, overview, or commentary that reviews, synthesizes, summarizes, or analyzes the current status of research on a specific scientific topic. The mention of potential applications, implications, or future directions is standard and expected in such publications and does not disqualify a publication from matching this definition. In contrast, the publication must not present new findings (particularly new experimental findings) or new scientific methods, including new databases and tools.

Definition 2: a study focused on analyzing historical, philosophical, societal, policy, legal, or ethical perspectives and implications, or the market potential and industry environment of a technology or scientific field.

Do not view any definition as satisfied based on implied, perceived, or speculative conclusions; a definition is only fulfilled when the text explicitly and directly discusses the above-mentioned aspects as its primary and sole focus.

Return a match score between 0 and 100 to what extent the publication matches at least one of the definitions. Use the full scale including 0 (no match) and 100 (perfect match) when warranted.

Format your response as a JSON array, with one object per abstract. Each object must include the following fields: {"abstract_id": i, "match": number}

Categorize strictly based on the explicit, verbatim content of the title and abstract, using zero inference or speculation about potential applications, implications, or effects not directly stated in the text. The score assigned must correspond only to the content that is unambiguously and literally mentioned in the document, with zero tolerance for reading between the lines or extrapolating beyond the precise language used. View controversial societal topics from a neutral and fact-based perspective. Return ONLY the JSON array without additional text or markdown. Do not return an explanation.

A.1.2 Prompt to Determine the Focal Organism

You are a scientific text classification system. Process each paper, with its title and abstract, independently. For each category, return a confidence score indicating the likelihood that the paper refers to the following organisms: archaea, bacteria, viruses, plants, mammals, other eukaryotes. Do not create additional categories beyond those specified. The confidence score represents the relative importance of each category compared to all others. The total scores across all categories must sum to exactly 100. Partial scoring across categories is acceptable, but multiple categories can receive zero points. Only assign points to categories that genuinely match the text — do not force-fit categories.

Guidelines for determining organism relevance:

- Always classify based on the biological subject which is being modified with CRISPR-Cas
- Classify modified cell lines based on their biological origin (e.g., human cancer cells = mammals)
- For diagnostic methods, classify based on the organism being detected (e.g., studies of bacterial detection in human samples = bacteria)
- If the modified organism is not explicitly stated as one of the possible categories, classify as mammals if the abstract discusses diseases (especially cancer), therapies (especially gene therapy), or cell types predominantly found in humans (e.g., T-cells, human stem cells)
- If the study focuses on the CRISPR-Cas system itself (e.g., its mechanism or modifications) classify as bacteria only if the study does not apply the system to any specific organism or organism group. If the system is being applied to an organism, classify based on the organism being modified (e.g., mammals for human studies, plants for crop studies)
- 'Other eukaryotes' is not a fallback category, but a distinct classification used explicitly for eukaryotic organisms that do not fit into the mammals or plants categories, as per the rules above. This includes, but is not limited to, fungi, yeasts, protozoans, parasites, insects, and other invertebrates

Format your response as a JSON array, with one object per paper. Each object must include the following fields: {"paper_id": 0, "archaea": number, "bacteria": number, "viruses": number, "plants": number, "mammals": number, "other_eukaryotes": number}

Return only the JSON array without additional text or markdown. Before returning, check that the total score is 100. Do not return an explanation.

A.1.3 Prompt to Determine the Focal Application

You are a scientific text classification system. Process each paper, with its title and abstract, independently. For each category, return a confidence score (0–100) indicating the likelihood that the paper refers to the following application areas: Medical/Health, Crops, Livestock, Industrial, Environment, Technical Improvement. Do not create additional categories beyond those specified. The confidence score represents the relative importance of each category compared to all others. The total scores across all categories must sum to exactly 100. Partial scoring across categories is acceptable, but multiple categories can receive zero points. Only assign points to categories that genuinely match the text — do not force-fit categories.

Guidelines for determining application area relevance:

- *Medical/Health* refers to studies focused specifically on improving human health outcomes, including both the understanding of the mechanisms behind human development and disease, as well as the creation of new medical therapies, treatments, drugs, screening methods for the detection of diseases, human biological materials, and foreign compounds in them.

- *Crops* relates to crop plants, including their cultivation, genetic modification, plant breeding, stress responses, nutrient content, yield, and other plant biology research aimed at understanding or enhancing crop characteristics for agricultural purposes. This category excludes research on animals.

- *Livestock* relates to animal husbandry, including cattle, pigs, poultry, fish, and other animals raised for food, commercial purposes, or farming.

- *Industrial* includes studies that focus on biofuels, chemicals, industrial enzymes, industrial microorganisms, and commercial foods that are not plant or animal based (e.g., yogurt production).

- *Environment* includes studies of naturally occurring ecological interactions, wildlife behaviors, pollutant distributions, wild insect populations, non-commercial animal interactions, and conservation research focused on understanding unmanipulated natural

systems. Research must focus on ecosystem-level or organism-environment interactions, not molecular, genetic, or cellular mechanisms, even when studying naturally occurring organisms or sampling from natural environments.

- *Technical Improvement* applies specifically to studies primarily focused on discovering or developing new or significantly improving existing CRISPR-Cas-related methodologies, technological platforms, and analytical approaches. This includes discovering and developing new tools, model organisms, techniques, or significantly modifying existing scientific methods, the CRISPR-Cas delivery mechanisms and screening. Important: simply using and applying gene-editing and CRISPR-Cas as a tool alone does not justify points in this category. The focus must be on improving the tool itself, not just using the tool.

Categorize strictly based on the explicit, verbatim content of the title and abstract, using zero inference or speculation about potential applications, implications, or effects not directly stated in the text. Any points assigned must correspond only to categories that are unambiguously and literally mentioned in the document, with zero tolerance for reading between the lines or extrapolating beyond the precise language used.

Format your response as a JSON array, with one object per paper. Each object must include the following fields: {"paper_id": i, "Medical_Health": number, "Crops": number, "Livestock": number, "Industrial": number, "Environment": number, "Technical_Improvements": number}

Return only the JSON array without additional text or markdown. Do not return an explanation. Before returning, check that the total score is 100.

A.1.4 Overview of the Classification Pipeline

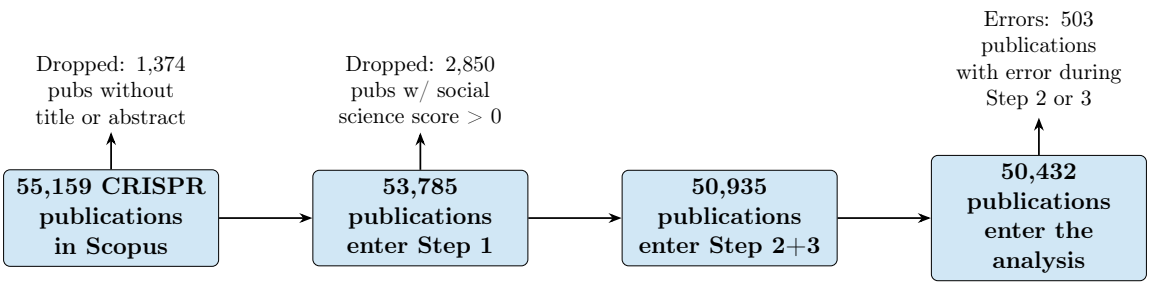


Figure 4: Overview of the Three-Step Classification Pipeline and Publication Counts per Stage.

A.2 Prompt to Identify Plant Publications

You are a scientific classification system. Determine, from the title and abstract, whether this scientific publication is about plants.

A publication is about plants if any of the following criteria apply:

- Focal organism is a plant
- Focus is on any aspect of plant science (e.g., biology, physiology, genetics, ecology, cultivation, but not limited to these)
- Focus is on organisms that directly interact with plants in cultivation contexts (e.g., pests, pathogens, pollinators, but not limited to these)

Return a binary score: 1 if the publication is about plants, 0 otherwise. Format your response as a JSON array, with one object per abstract. Each object must include the following fields: {"abstract_id": i, "plants": binary score}

Return ONLY the JSON array without additional text or markdown. Do not return an explanation.