Laboratory Safety and Research Productivity

Alberto Galasso
University of Toronto, CEPR and NBER

Hong Luo
Harvard Business School

Brooklynn Zhu
BI Norwegian Business School*

March 24, 2023

Abstract

Are laboratory safety practices a tax on scientific productivity? We examine this question by exploiting the substantial increase in safety regulations at the University of California following the shocking accidental death of a research assistant in 2008. Difference-in-differences analyses show that relative to "dry labs" that use theoretical and computational methods, the publication rates of "wet labs" that conduct experiments using chemical and biological substances did not change significantly after the shock. At the same time, we find that wet labs that used dangerous compounds more frequently before the shock reduced their reliance on flammable materials and unfamiliar hazardous compounds afterward, even though their overall research agenda does not appear to be affected. Our findings suggest that laboratory safety may shape the production of science, but they do not support the claim that safety practices impose a significant tax on research productivity.

Keywords: economics of science, risk perception, safety regulation
JEL Codes: K13, J24, O31

1 Introduction

Workplace accidents are one of the leading causes of death and disabilities.[1] An important example of hazardous working environments is academic research laboratories. During the past decade, growing concerns have been voiced about the safety culture and protocols across universities in the United States, Europe, and China (Silver, 2021). Several proposals to introduce stricter safety regulations have been

*Weida Li, Shenye Liu, Yash Karani, Ehsen Tayyabi, and Bijan Miriabi provided excellent research assistance. We are particularly grateful to Debbie Decker, Becky Grunewald, and Imke Schroeder for providing insightful details on the responses by UCLA and UC Davis to the fatal accident. Alberto Galasso is grateful for financial support from the Social Sciences and Humanities Research Council of Canada.

[1] A recent joint study by the World Health Organization and International Labour Organization estimates that almost two million people die from work-related causes each year (WHO-ILO, 2021).

put forward, including enhanced inspections, training, hazard documentation, and the use of safety metrics in publications, promotion, and grant decisions (National Research Council, 2014).

Implementing stricter safety measures is challenging in part because of the widely held belief that they may reduce research productivity. Time spent on safety training, risk assessment, and documentation is time not spent on research. Supervision by experienced researchers and co-workers constrains resources.[2] This perception was highlighted in a 2014 National Research Council report on safe science, which states that, "*one of the most recalcitrant problems in many chemistry laboratories is the attitude, unfortunately often reinforced by principal investigators, that safety practices are time-wasting inhibitions to research productivity*" (National Research Council, 2014). Russell Phifer, chair of American Chemical Society's chemical health and safety division, commented that, "*in the labs of many competitive academic researchers, time and publication pressure favor productivity over performing and documenting safety training*" (Benderly, 2009). Indeed, survey evidence indicates that many academic scientists see safety rules as a tax on scientific productivity, and the top listed reason for barriers to improving safety in a lab is "time and hassle factors" (Van Noorden, 2013).

By contrast, proponents of stricter regulations have argued that a safe working environment is critical for sustaining research productivity (Benderly, 2009; National Research Council, 2014). In hazardous environments, adequate safety precautions allow scientists to focus on their research rather than worry about injuries. Accidents not only threaten researchers' health and safety but also result in research projects being shelved or substantially delayed. Serious injuries and death may also affect a laboratory's reputation, making it difficult to attract new researchers.

These contrasting views, which often underlie the tension between researchers and laboratory safety experts, highlight the need for empirical evidence. In this paper, we provide the first set of large-sample evidence on the relationship between lab safety and research productivity. We focus on academic chemistry laboratories. Hazardous chemicals impose health and safety risks for researchers, and they are not as stringently regulated as radioactive or biological materials (Van Noorden, 2011). Our analysis exploits a substantial and quasi-exogenous surge in academic lab safety regulations and examines its

[2] A 2010 survey by the American Chemical Society suggests that 70.5% of faculty and 52.1% of graduate students often or occasionally work alone in laboratories, which is forbidden in the industry (Van Noorden, 2013).

impact on researchers' productivity.

Specifically, on December 29, 2008, Sheharbano Sangji, a research assistant in an organic chemistry laboratory at the University of California Los Angeles (UCLA), spilled a highly flammable compound that ignited and severely burned over half of her body. Sangji died 18 days later. This accident received extensive press coverage and close attention from the academic scientific community. Investigations determined that inadequate safety training and not wearing personal protective equipment (PPE) were among the causes of this tragic incident. This led to the first-ever criminal case against a university professor due to lab safety. The UCLA accident was described as "*a pivotal development*" for academic lab safety regulations that "*had more impact on lab safety than anything else that's happened in the last 20 years*" (Basken, 2012; Trager, 2014). UCLA, as well as the entire UC system, responded by implementing major changes to its safety programs. These measures included more frequent lab inspections, more stringent protocols for dangerous chemicals, and more safety training for laboratory scientists (Gibson, et al. 2014).

From the perspective of principal investigators (PIs), who are the labs' managers, these events have led to two major changes. The first is a significant increase in the perceived risk and liability associated with laboratory accidents, and the second is a significantly more stringent regulatory environment. To guide our empirical analysis of the impact of these changes on research productivity, we develop a simple model describing the relationship between risk/liability perception, safety levels, and research output. The model builds on the multitasking literature (Holmstrom and Milgrom, 1991) and assumes that researchers direct their efforts towards either conducting research or improving the safety of their labs.

The model first clarifies that even in the absence of stricter safety regulations, an increase in risk perception alone incentivizes scientists to increase safety efforts. Ultimately, whether an increase in safety leads to an increase or a decrease in research productivity depends on the relationship between safety investments and the marginal cost of research. The debates surrounding safety regulations suggest that safety investment may influence the marginal cost of research in multiple ways and potentially in different directions. An increase in safety investment may crowd out research time, thereby increasing marginal research costs and hurting research productivity. However, if a safer work environment motivates researchers, reduces disruptions, and allows the PI to attract lab members, safety investment may decrease the marginal cost of research and promote research productivity. The overall impact would thus depend on the net effect of these various mechanisms.

Our empirical analysis relies on a sample of 592 chemistry labs affiliated with the University of California (UC) between 2004 and 2017. For each lab, we retrieve its complete record of publications from the Web of Science database. Importantly, we hire a team of chemistry PhD students to classify each lab into one of two groups. The first group (treatment group) includes "wet labs," in which scientists conduct experiments on chemical and biological substances. The second group (control group) includes "dry labs" that sspecialize in computational and theoretical research, which are not affected by changes in safety regulations that resulted from the UCLA accident.

Our first key finding is that despite the implementation of more stringent safety regulations, we do not see a significant decline in research productivity. Specifically, difference-in-differences estimations show that, relative to dry labs, UC wet labs experienced a reduction of about 3% in their yearly publications after 2008, but we *cannot* statistically distinguish this effect from zero. Event study analysis shows no pre-trend differences between these two groups, and the small and insignificant effect is robust across a variety of specifications and measurements.

Our second key finding relates to the shock's impact on the direction of research. We find no broad impact on the use of dangerous chemicals among wet labs at UCLA. However, for a small number of labs that used dangerous chemicals intensely before the shock, we observe a decrease in the use of these chemicals after the shock. Moreover, this reduction is mostly about flammable compounds and dangerous chemicals that are unfamiliar to the researchers. Zooming in on these wet labs, even though a textual analysis of their paper abstracts does not suggest an overall change in their research agenda, we find that their more hazardous research projects appear to be more similar to their pre-shock research.

The empirical evidence presented here is relevant for the academic and policy debates regarding the relationship between worker safety and productivity. Taken together, our results show that stricter safety regulations may affect scientists' behavior, increase lab safety levels, and affect the production function of science. However, the results do not support the claim that stricter safety regulations impose a significant tax on research productivity.

The rest of this paper is organized as follows. Section 2 discusses related literature. Section 3 describes the UCLA accident and what happened to chemistry lab safety regulations in the wake of this accident. Section 4 presents the theoretical model. Section 5 describes the data and econometric methods. Section 6 examines the impact of the shock on the rate of research, and Section 7 studies the impact on the direction of research. Section 8 summarizes and discusses our main findings.

2 Related literature

Our paper relates primarily to two sets of literature: the first focuses on the relationship between worker safety and firm performance and the second addresses the drivers of academic research.

2.1 Worker safety and firm performance

Studies in various disciplines, specifically economics and operations management, have investigated the relationship between workplace safety and firm performance. A central debate in these studies is whether there exists a significant trade-off between safety and productivity.[3]

On the one hand, many studies have suggested the existence of a significant trade-off; thus, an increase in safety regulation, which induces more safety, would lead to a reduction in productivity. Gray (1987) suggested that regulations such as those by the Occupational Safety and Health Administration (OSHA) were responsible for about 30% of the economic slowdown in the 1970s. Gowrisankaran et al. (2015) found that safety increases triggered by mining disasters were associated with an 11% decrease in mines' productivity. Pagell et al. (2020) found that organizations that provide a safer workplace have significantly lower odds and shorter lengths of survival. Conceptually, this trade-off may emerge from several sources. Financial and time constraints imply that safety investments may crowd out investments in productive inputs. Moreover, a greater safety requirement imposes constraints on firms, making it more difficult to take advantage of productivity-enhancing innovations. This latter argument is consistent with the findings in several studies that adopting performance-based pay and innovative production processes, such as faster work pace, short cycle time, and less slack is associated with an

[3] To this end, our paper relates to the so-called Porter Hypothesis debate on the impact of environmental regulation. According to the traditional view, environmental regulations hurt firm productivity because they force firms to allocate inputs to pollution reduction, which is unproductive (Palmer et al. 1995). Porter and van der Linde (1995) argue that more stringent but properly designed environmental regulations can trigger innovation that may partially or even more than fully offset the costs of complying with them.

increase in workplace accidents (Perrow, 1984; Adler et al. 1997; Brenner et al. 2004).

On the other hand, other studies have suggested that an increase in safety requirements does not necessarily decrease productivity and may even increase productivity. Leveraging experimental methodologies, Levine et al. (2012) found that more rigorous OSHA inspections did not have any significant impact on firm productivity. Studies arguing for the positive effects of safety on performance highlight the lost productive capacity and efficiency due to unsafe work environments. Goggins et al. (2008) showed that improving the ergonomics of a workstation leads to decreases in turnover and absenteeism, which increases productivity and quality. Building on the hierarchical motivation theory, Das et al. (2008) argued that safety is a basic human need, and workers in unsafe environments are more likely to engage in self-protection and are less motivated to pursue organizational goals. Consistent with this motivation argument, McLain (1995) showed empirically that perceptions of a less safe work environment are associated with greater work distractions. Finally, because stability is critical for efficient production, accidents would also hurt productivity by disrupting stability (Pagell et al. 2015).[4]

2.2 Drivers of the rate and direction of research

The second stream of literature to which our paper relates is the drivers of academic research. Stephan (2010) provided a comprehensive survey on this topic. Largely, this literature has investigated demand-side factors such as funding incentives (Azoulay et al. 2011) and supply-side factors such as the availability of research tools (Furman and Stern, 2011; Murray et al., 2016) and the loss of human capital such as the death of prominent collaborators (Azoulay et al. 2010; Oettl, 2012) or physical assets such as lab equipment (Baruffaldi and Gaessler, 2022).

This literature distinguishes between the rate of research—which is about productivity—and the direction of research, which is about the type of research. The rate of research is typically measured by the quantity of (which can be adjusted by quality) research outputs and is among the outcome variables investigated by most of the studies in this literature. By contrast, the direction of research is much less studied and its definition may differ by context. A set of papers focuses on the relatedness between

[4] A set of studies in operations management shows that the adoption of quality production systems may improve safety and production performance at the same time (e.g., Levine and Toffel, 2010; Lo, et al. 2014). These studies, though related, focus more on shifts to the production possibility frontier, rather than the trade-off faced on a specific frontier.

research based on paper abstracts. For example, Azoulay et al. (2019) found that the premature death of eminent life scientists allows more outsiders to enter a particular field and draw on new ideas. Myers (2020) estimated how responsive life scientists are with respect to NIH funding in specific areas. Furman and Teodoridis (2020) showed that technologies automating research tasks may induce researchers to pursue ideas that are more diverse than and distant from their original trajectories. Another set of papers focuses on the research input. Furman and Stern (2011) showed that institutional changes facilitating access to research materials can increase researchers' propensity to use knowledge associated with these materials, and Murray et al. (2016) documented how IP restrictions on research tools may lead to a reduction in the number of academic publications relying on those tools.

Whether rate or direction, this literature mainly focuses on a single aspect of the researcher's performance; that is, the research output. By contrast, dual performance objectives—safety and research outcomes—are at the center of our study. In this regard, our paper relates to other studies that focus on multiple objectives that compete for a researcher's time and resources such as the impact of academic patenting on research productivity (Azoulay, et al., 2009) and that of academic entrepreneurship on student mentoring (Roche, forthcoming). The multitasking theoretical framework that we use in this paper can also apply to these other contexts; it allows the flexibility of either a competing or a complementary relationship between the dual performance goals.

From the perspective of the safety objective alone, our paper relates to recent work that has studied the diffusion of green chemistry (Anastas and Warner, 1998). Nameroff et al. (2004) documented this phenomenon in the chemistry industry using patent data, and they relate this trend to changes in environmental regulations. Howard-Grenville et al. (2017) studied how academic chemists encouraged other chemists to practice green chemistry and the effectiveness and limitations of these efforts.

While not about academic research, another set of studies are also relevant to our work as they examine how changes in liability risk and safety regulations shape innovation in the industry. Similar to our context, these studies debate whether the impact is likely to be positive or negative. On the impact of greater liability, existing empirical evidence—Viscusi and Moore (1993) and Galasso and Luo (2017)—suggests that, on average, higher liability risk induces higher R&D spending and more patenting. Galasso and Luo (2022) did find a significant chilling effect, but they showed that it is through a specific mechanism: a sudden increase in product liability risk suppliers faced may disrupt vertical chains and negatively impact downstream innovation investments. Studies on the impact of safety regulations, such as the FDA approval process, also found mixed results. On the one hand, Peltzman (1973) showed that the 1962 drug amendments requiring proof of efficacy in addition to safety led to a significant decrease in welfare. However, a recent paper by Grennan and Town (2020) found that for coronary stents, the efficacy requirement in the US, which is more stringent than the European Union, is critical for reducing quality uncertainty and facilitating adoption.[5]

3 The UCLA accident and subsequent events

On December 29, 2008, research assistant Sheharbano (Sheri) Sangji was conducting an experiment in Professor Patrick Harran's organic chemistry lab at UCLA. When she attempted to transfer a tert-butyl-lithium solution, a highly flammable compound, from a bottle to a flask, the syringe plunger came out of the barrel and the chemical burst into flames. Sangj was not wearing a protective lab coat at the time, and her synthetic sweater caught fire. Sangj was rushed to the hospital with more than half of her body severely burned. She died from her injuries on January 16, 2009 (Kemsley, 2009).

This accident received wide media coverage and close attention from the academic scientific community, especially in chemistry. News outlets, including the Los Angeles Times, Chemical & Engineering News, and Chemistry World reported the accident after Sheri Sangji's death and followed up on subsequent events. This accident and laboratory safety became a major topic of discussion at the American Chemical Society meeting in March 2009 (Benderly, 2009). Numerous editorials and blog posts in Science, Nature, Scientific American, and other publications discussed the accident's details and called for actions to improve laboratory safety and prevent future incidents.

UCLA responded to the accident immediately. Within 30 days, an interdisciplinary team of experts conducted comprehensive inspections of more than 300 laboratories, chemical storage rooms, and shops within the Department of Chemistry and Biochemistry (Gibson et al., 2014). In the following six months, a Laboratory Safety Committee was established, which conducted a thorough study of all aspects of

[5] For theoretical models in law and economics linking liability risk with innovation see Daughety and Reinganum (2013) and Dawid and Muehlheusser (2022).

lab safety at UCLA, including labs outside the chemistry department. In July 2009, the committee submitted a comprehensive list of recommendations to the UCLA Chancellor.

The data Gibson et al. (2014) provided paint a clear picture of the strengthening of safety requirements at UCLA after the accident: the number of safety class participants increased from 3,327 in 2007 to 21,789 in 2012, with the number doubling in 2009 compared to 2008 (about 13,000 versus 6,000). Similarly, the number of safety inspections at UCLA increased from about 1,100 in 2008 to about 2,000 in 2009 and to about 4,500 in 2012. Starting in 2010, scheduled inspections were complemented by unannounced inspections.

Furthermore, the California OSHA conducted an investigation. In December 2009, the investigation concluded that Sangji had not received adequate training for working with hazardous chemicals as the State of California required. The report noted that UCLA's environmental health and safety (EHS) department "*was well aware that research staff within virtually all laboratories at the University routinely did not wear lab coats and other personal protective equipment while working in the labs.... The practice was so well known by EHS that it was simply regarded 'as part of the culture.'*" (Baudendistel, 2009). "*Dr. Harran,*" the report stated, "*permitted victim Sangji to work in a manner that knowingly caused her to be exposed to a serious and foreseeable risk of serious injury or death.*" In December 2011, the Los Angeles District Attorney filed criminal charges against the Regents of the University of California and Patrick Harran for "*willful violation of safety regulations.*"

The case against the university was settled in July 2012. The terms of the settlement required UC to accept responsibility for the event, establish a scholarship in honor of Sangji, pay the OSHA litigation costs, and implement several specific laboratory safety practices throughout the entire UC system, not just at UCLA. These practices encompassed lab safety manuals, hygiene plans, training of staff and PIs, and following standard operating procedures for hundreds of chemical substances (Merlic, 2013). Harran, if convicted, could have faced up to four and a half years in prison. In June 2014, Harran reached a settlement under which he was required to: (1) teach organic chemistry to inner-city high school graduates for five years; (2) complete 800 hours of non-teaching community service; (3) speak to UCLA students about the importance of lab safety; and (4) pay a $10,000 fine to the regional burn center where Sangji was treated (Trager, 2014). In September 2018, having determined that Harran had now met the terms of the agreement, a Los Angeles County Superior Court judge dismissed the criminal charges against him (Maxmen, 2018).

Conceptually, the events presented above led to two major changes to the institutional environment from a PI's perspective . Although the magnitude of these changes evolved over time, they both started shortly after the accident, making it challenging to empirically disentangle their specific effects. This implies that the overall effect (or lack thereof) on research productivity that we investigate in this paper will reflect the joint impact of the two changes.

The first major change was a significant increase in the awareness of the risk of working in academic chemistry laboratories involving dangerous substances, as well as the PI's personal liability in case of accidents. Discussions with the EHS officers at both UCLA and UC Davis made it clear that the accident was a huge shock to the chemistry community, especially those working at UCLA and other UC campuses. They used the word "*terrified*" to describe the immediate reactions of many of the PIs, and some PIs commented explicitly at the time that "*it could have been my lab.*"[6] While the official criminal charge against Harran was not issued until late 2011, many PIs started conjecturing about Harran's potential personal consequences, including jail time, immediately after the accident.

Apart from legal liability, laboratory safety issues may also have imposed greater costs to PIs' reputations, statuses, and careers. Various proposals were made to take accident reports, laboratory investigations, and safety policy compliance into account for promotion, tenure and the allocation of grants and departmental resources. In December 2015, the American Association for the Advancement of Science (AAAS) decided to withhold recognition of Patrick Harran. In an interview with Nature, Langerman stated that "*this action is huge, and impacts every scientist who aspires to be named for national recognition or international recognition.... If I were a young chemist, and I set a career goal to win a Priestley Medal [the highest honor conferred by the American Chemical Society] this says that if my lab has a serious incident, I may never achieve my goal*" (Hayden, 2015).

The second major change was the introduction of stricter laboratory safety rules. These mandated

[6] In 2009, after reviewing the reports on the UCLA and other accidents, Dr. Neal Langerman, former chair of the Division of Chemical Health and Safety of the American Chemical Society (ACS), said that: "*I have come to the disheartening conclusion that most academic laboratories are unsafe venues for work or study*" (Langerman, 2009). AA 2013 survey published by Nature and UCLA showed that 30% of the 2,400 responding scientists reported having witnessed a lab injury that was severe enough to warrant attention from a medical professional (Van Noorden, 2013).

rules directly added to the costs of operating a lab and can be largely categorized into one-off investments and recurring costs. Examples of one-off investments include the redesign of laboratory space (e.g., removing improperly located safety showers) and updates to sprinkler systems. Recurring costs include safety training, lab inspections, and documentation of lab procedures and materials used, all of which became significantly more frequent and more thorough.[7]

The increase in safety requirements took place at UCLA immediately after the accident (Gibson et al., 2014), and the requirements were strengthened further after the July 2012 settlement between the University of California and LA County. For other UC campuses, our interviews confirmed that the Sangji accident definitively spurred a strengthening of their safety regulations, even though the timing may have lagged UCLA by several years. At UC Davis for example, a significant increase in training and inspections did not take place until after the July 2012 settlement. That said, a UC Davis EHS officer with whom we talked said that their office had substantial interactions with PIs in the Chemistry Department immediately after the accident regarding the storage and inventory of hazardous materials. It seems reasonable to conclude that such activities also took researchers' time and effort, especially for labs that used hazardous materials more intensely.

The increased focus on laboratory safety went beyond the UC system and spread to chemistry departments across universities in the United States. Immediately after the UCLA accident, Russell Phifer, chair of the American Chemical Society's (ACS) chemical health and safety division, said in an interview with Science: "*I know for a fact that many universities immediately reviewed their protocols for dealing with pyrophorics [materials inflammable when exposed to air] and many of them looked at their documentation of safety training*" (Benderly, 2009). In March 2011, the University of California also established a Center for Laboratory Safety (CLS), with the goal of supporting research in laboratory safety as well as diffusing best practices across UC campuses and other universities (National Research Council, 2014). In 2012, the ACS issued the report "Creating Safety Cultures in Academic Institutions," which described best practices and provided recommendations to university departments. This was

[7] Following the accident, the university of California substantially expanded the use and enforcement of standard operating procedures (SOPs) to cover approximately 1,000 chemicals. SOPs are written documents in which labs describe their experimental procedures involving hazardous materials and their plans to handle, store, and dispose of dangerous chemicals. According to the EHS officers with whom we talked, researchers perceive SOPs as the most burdensome requirement.

followed by an ACS presidential commission recommending the adoption of the best safety practices as a key requirement for the advancement of graduate education in chemical sciences (ACS, 2012a; 2012b).

4 Theoretical considerations

In this section, we develop a simple theoretical model that helps illustrate the potential impact of the UCLA accident and subsequent events. Section 3 suggests that these events led to two major changes from a PI's perspective: (1) a significant increase in the perceived risk (and liability) associated with laboratory safety, and (2) stricter safety regulations implemented by the university. While our empirical analysis focuses on research productivity, the model clarifies that these events influence research productivity by affecting safety practices.

To illustrate this point, we use a multitasking model in the spirit of Holmstrom and Milgrom (1991), in which a PI decides to allocate time and resources to two types of activities: $r \in (0, 1)$ is the research effort directed at conducting experiments and publishing new results, and $s \in (0, 1)$ is the safety effort aimed at reducing the risk of accidents and their consequences.

With efforts (r, s), the PI enjoys a benefit of

$$B(r, s) = r - \delta(1 - s)$$

where r represents the monetary and reputational rewards from publications, $(1 - s)$ is the risk level of the lab, and δ captures the perceived costs associated with lab accidents. For PIs, $\delta > 0$ includes both the legal liability and reputational loss due to accidents and their internalization of harms to researchers working in the lab. We also assume that $\delta < 1$ to reflect the notion that from the PI's perspective, the marginal benefit of safety is typically less than that of research output. This issue was highlighted prominently in the 2014 National Research Council report, which discussed how academic reputation and decisions about promotions, salary, and space tended to focus on research productivity.

The cost of these efforts assumes a quadratic form, which is standard in the multitasking literature (see Fryer and Holden, 2013; Benabou and Tirole, 2016; and De Philippis, 2021):

$$C(s, r) = \frac{s^2}{2} + \frac{r^2}{2} + \rho sr.$$

The parameter $\rho \in (-1, 1)$ reflects the notion that these two efforts may affect each other. Intuitively, because resources are fixed, allocating more time and budget to implementing safety protocols means less time and budget for research; that is, safety investment may increase the marginal cost of research. This is consistent with the recurring theme that an important barrier to improving laboratory safety is the perceived conflict between safety and research productivity. Harry J. Elston, editor of the *Journal of Chemical Health and Safety*, wrote that the Sangji case was "*a harbinger of things to come*" unless scientists devoted to accident prevention were willing to "*stand in the gap between worker's safety and [scientific] productivity.*"[8] This view also emerged from the UCLA-Nature survey, with one-fifth of the respondents indicating that lab safety rules had negatively impacted their research productivity (Van Noorden, 2013).

On the other hand, arguments also exist that safety investment may actually decrease the marginal cost of research; therefore, a safer work environment can promote research. Safer workspaces, by reducing the risk of accidents, enable researchers to "*focus on their tasks rather than worrying about emergencies*," because "*in addition to putting people at the risk of harm, these incidents ultimately decrease productivity, as they hinder the researchers' ability to work*" (Hersh, 2017). Labs with poor safety records may find it difficult to attract post-docs and PhD students. Moreover, laboratory safety may also affect a lab's ability to obtain research funding, given the increasing calls to include safety records as part of promotion decisions, as well as allocation of departmental resources, grants, and prizes (National Research Council, 2014).

These mechanisms are *not* mutually exclusive, and ρ reflects the net effect. If, in aggregate, safety investment increases the marginal cost of research, we have $\rho > 0$; if safety investment decreases the cost of research, $\rho < 0$; and if safety investment does not affect the cost of research, we will have $\rho = 0$.[9]

[8] Journal of Chemical Health and Safety, in a lead editorial entitled "Recipe for disaster," posted to the Internet on 31 March.

[9] Note that we can interpret $\rho = 0$ in two ways. One is that the various effects cancel each other out. The second is that safety and research efforts are truly independent of one another and do not influence one another through any of the aforementioned mechanisms.

4.1 Empirical implications

We map the two institutional changes that the UCLA accident caused to the model as follows. First, we capture the increase in risk and liability perception with an increase in the parameter δ. Specifically, we assume that $\delta = \underline{\delta}$ before the accident, and $\delta = \overline{\delta}$ after the shock, with $\Delta\delta = \overline{\delta} - \underline{\delta} > 0$. Second, we model the implementation of stricter safety protocols as an increase in minimum safety requirements. For simplicity, we normalize the minimum safety level to zero before the accident and indicate the requirement after the accident as $\underline{s} > 0$.

In Appendix A1, we examine how, together, these two changes affect the research output and safety level of a lab. Specifically, we compute the optimal safety level and research output the lab produced before the shock, which are indicated by $s^*(\underline{\delta}, 0)$ and $r^*(\underline{\delta}, 0)$. These are functions of the pre-shock risk perception and minimum safety requirement. Similarly, we compute the optimal safety level and research output after the shock, $s^*(\overline{\delta}, \underline{s})$ and $r^*(\overline{\delta}, \underline{s})$. The effect of the shock on the safety level is thus $\Delta s = s^*(\overline{\delta}, \underline{s}) - s^*(\underline{\delta}, 0)$, and that on the research level is $\Delta r = r^*(\overline{\delta}, \underline{s}) - r^*(\underline{\delta}, 0)$.

The model delivers several implications. First, the safety level will increase after the shock; $\Delta s > 0$. Consistent with prior research (Galasso and Luo, 2021), an increase in risk perceptions would lead to a voluntary increase in safety investment. Thus, even in the absence of stricter safety regulations mandated by the university, labs are likely to change their safety practices to mitigate risk. The minimum safety level mandated by the university, $\underline{s}$, may or may not be binding. If $\underline{s}$ is not binding, the post-shock safety level will be determined by the higher risk perception $\overline{\delta}$. Otherwise, the post-shock safety level will be at the required level, $\underline{s}$, which is more stringent than what PIs would find optimal.

Second, whether the shock increases or decreases research output (that is, the sign of Δr) depends on one specific parameter of the model: the (net) relationship between the two types of efforts, ρ.[10] As discussed in the previous paragraph, the shock causes scientists to increase their investment in lab safety. This translates into a reduction in research output when $\rho > 0$ because safety investment increases the marginal cost of research. However, the accident leads to an increase in research output when $\rho < 0$; that is, when safety investments lower the marginal cost of research. Finally, the accident leads to no

[10] Mathematically, this result follows directly from the first-order condition for r, which is $r = 1 - \rho s$.

changes in research output when $\rho = 0$.

Third, the model shows that the magnitude of the effect depends on some combination of ρ, $\Delta\delta$, and $\underline{s}$. This implies that, in principle, empirical estimates of a small magnitude could be driven by a low level of interaction between the two types of efforts (ρ close to zero), by small changes in perception and safety regulations ($\Delta\delta$ and $\underline{s}$ close to zero), or both. The evidence presented in Section 3 suggests that in our empirical setting, the change in risk perception and the increase in safety requirements were substantial. Thus, an empirical estimate of a small magnitude is likely to reflect a relatively small ρ.

Despite its simplicity, the model provides useful guidance for our empirical analysis. Specifically, it clarifies that the relationship between research and safety efforts is key to understanding these events' impact on research productivity. In the next section, we broaden our discussion in ways that are either of general policy interests or helpful for understanding additional empirical results.

4.2 Discussion

This section extends our baseline analysis in three ways: (1) it clarifies the standalone effect of the policy interventions; (2) it introduces heterogeneity in the lab hazard level; and (3) it discusses the possible effects of the shock on research direction. We use the previously described model as the basis of the discussion and extend it in simple ways when necessary.

4.2.1 Unbundling policy interventions

In our baseline analysis, we consider two changes in the economic environment—an increase in the risk perception and an increase in minimum safety standards—at the same time. This is faithful to our setting, which is not unusual as more stringent safety regulations often follow significant accidents. This said, it is instructive to clarify the standalone effect of the policy interventions. Specifically, we consider three separate policies: (1) an increase in a PI's liability in the case of an accident, (2) an increase in the minimum safety standards, and (3) a reduction in the cost of safety efforts.

Mathematically, the first two policy changes are modeled in similar ways as done in Section 4.1. Specifically, an increase in liability is modeled as a marginal increase in δ. Minimum safety standards are indicated by $\underline{s}$, and policy (2) is modeled as a marginal increase in $\underline{s}$. To study the impact of a reduction in the cost of safety efforts, we extend the model and revise the joint costs of research and safety efforts as

$$C(s,r) = (1-\varepsilon)\frac{s^2}{2} + \frac{r^2}{2} + \rho sr.$$

Thus, policy (3) can be modeled as the effect of a marginal increase in ε and evaluated at $\varepsilon = 0$. For simplicity, we do not include this third intervention in our baseline analysis, but it also occurred in our empirical setting. Gibson et al. (2014) reported that the EHS office at UCLA invested in digital tools to reduce the researchers' costs of taking training classes and documenting materials purchased and used. The EHS office also trained inspectors and revised the procedure to make inspections more efficient. Moreover, the university centralized and paid for PPE. Thus, empirically, the interpretation of our empirical results should also include these cost-reduction efforts, which took place over time.

We highlight three results from this exercise (see Appendix A2 for the details).[11] First, each of these three policy interventions, when considered alone, generates qualitatively similar results as our baseline model. Specifically, they increase the level of safety investment. As a result, the directional impact of each policy on research productivity r is uniquely determined by the sign of ρ. Thus, what we observe in our empirical setting is the joint effects of all three changes.

Second, for each policy intervention, the magnitude of the impact on research productivity r is smaller than that on safety investment s. This makes sense as these policy interventions have a direct effect on the benefit or cost of s but an indirect effect on r through the interaction between r and s.

The third set of results emerges from comparing these three policies. When $\rho > 0$ (that is, safety investment increases the marginal cost of research), increasing liability δ seems to have the greatest marginal impact on safety investment s, as well as on r. Compared to a marginal increase in minimum safety standards, for example, a marginal increase in liability has a greater impact because it allows for a feedback loop. In particular, an increase in δ induces the PI to increase s, which stimulates a reduction in r, which itself triggers a further increase in s, and so on. These additional adjustments are not present when s is fixed at a binding level. When $\rho < 0$ (that is, safety investment decreases the marginal cost of research), however, a reduction in the cost of safety may be the most effective policy to

[11] Note that the minimum safety standard needs to be binding to have an effect on safety (i.e., $\underline{s} > s^*$). Moreover, the net perceived benefits of safety are assumed to be large enough (i.e., $\delta - \rho > 0$) to avoid a corner solution in which $s^* = 0$ and $r^* = 1$. Our discussion is based on comparative statics on interior solutions.

increase safety investment and, in turn, research productivity. This is because the cost complementarity amplifies the marginal reduction in cost by triggering an increase in r. When $|\rho|$ is sufficiently large, this amplification effect can lead to changes greater than those obtained by other policy interventions.

The comparisons between different policies are only illustrative and should not be overinterpreted. Our model, built on the multitasking literature, is designed to speak mainly about how safety practices affect research productivity and not to examine the optimal policy mix. To explore these policies properly, one needs to consider a mechanism design approach in which the principal (the university) and the agent (the PI) have different objective functions and possibly different information about the riskiness of a research lab. Moreover, the model needs to consider the costs of different policy interventions, which are not modeled here. We leave these investigations to future research.

4.2.2 Heterogeneous hazard levels

The discussion so far has considered an average lab. Labs, however, differ in their hazard levels. In Appendix A3, we enrich our baseline model by decomposing the expected liability, δ, as the product between the probability of an accident, p, and the expected liability cost incurred in the case of an accident, L. In this revised model, we keep p fixed for a lab and let the events increase L.

Intuitively, labs with a greater hazard level (i.e., a higher p) may be more affected by the events we study in this paper. This is indeed the case when L is the only parameter that changes because changing L influences the marginal benefit of safety, which is roportional to the hazardous level of a lab. However, this intuition does not necessarily hold if we also consider a change in minimum safety standards. Consider an increase in minimum safety standard (i.e., an increase from 0 to $\bar{s}$) alone first. If $\bar{s}$ is uniform and binding for all labs, we see the opposite result; that is, the change in research productivity, r, actually decreases with a lab's hazardous level. Meeting the new safety standards may require more safety investments, and subsequently a greater readjustment of research investment, for low-hazard labs because their optimal pre-shock level of s is lower. With multiple interventions taking place at the same time, the impact with respect to a lab's hazard level is likely to be ex-ante ambiguous.

4.2.3 The direction of research

The prior discussion focuses primarily on the impact of these events on the rate of research. It is reasonable to expect these events to also influence the type of research on which a lab works. Changes in the type of research can be characterized in multiple ways, ranging from the input a lab uses to the content of its research output. Given the plurality of modeling approaches, we primarily take an empirical approach to this question. Nonetheless, it is instructive to provide a simple yet intuitive extension of our model to illustrate how a lab may change its research direction after the shock (see Appendix A4). Specifically, we assume that labs with underlying hazard p have the option to reduce their hazard to $p' < p$ by choosing a different research project at the cost of c. In this case, we find that higher-hazard labs have a greater incentive to redirect research toward safer research projects. In Section 7, we will build on this insight and empirically examine changes in research direction by comparing labs using dangerous chemicals with different intensities.

5 Data

Our analysis relies on a sample of chemistry labs that were affiliated with UC and were active around the time of the UCLA accident. We focus on UC labs for two reasons. First, while the accident had a nation-wide impact, UC labs were affected most directly. Second, explained below, with the smaller sample, we can manually collect information that is critical for our identification strategy.

5.1 Sample construction

To identify the sample of UC chemistry labs, we start with a set of natural publishing outlets for chemistry researchers. These include: (1) journals in the top decile of the impact factor in each of the nine chemistry subfields, as provided by the Web of Science (WoS) Journal Citation Reports database; and (2) the ten multidisciplinary scientific journals with the highest impact factors such as *Science*, *Nature*, and the *Proceedings of the National Academy of Sciences*. We download all the articles published in these journals between 1998 and 2017 from the WoS. This step gives us 698,094 articles published in 105 journals.

We then use the author and affiliation information provided by WoS to identify individual labs. In chemistry, the author listed last is typically the PI of the lab that hosted most of the research, and the corresponding author works at this lab (Venkatraman, 2010). We rely on this convention to identify the primary PI of each article and the institution with which this PI is affiliated. We refer to each unique PI-institution combination as a lab.[12] Labs that are not very active—for example, those that publish only one article or that are active only for a year or two—are likely to reflect errors in the PI names or organizations with one-off publication projects. To capture meaningful research units, we drop labs with an active life span of fewer than three years. Among the remaining labs, we further drop those with a yearly publication rate below the median (0.6 articles per year in the 105 journals). This step gives us 6,704 active chemistry labs based in the United States.

Among the labs identified from the previous step, we create the UC Sample to include labs affiliated with the UC system. Leveraging information from sources such as lab websites and news releases, we manually confirm that the labs in the UC Sample were indeed run by UC-affiliated PIs. To address potential endogeneity concerns, we drop Patrick Harran's lab from the sample. This step leads to a final sample of 592 labs affiliated with UC. UC Berkeley accounts for about 24% of the labs in the sample, UC San Diego for 15%, UCLA for 14%, and UC Davis for 11%. The remaining labs are affiliated with UC Irvine, UC Riverside, UC San Francisco, UC Santa Barbara and UC Santa Cruz.

After identifying the sample of UC labs, we retrieve from WoS all of the journal articles (not just those in the 105 journals mentioned) that these labs published between 2004 and 2017, as well as the citations these articles received up until 2020. In total, the UC labs have published 50,341 journal articles. We use these data to construct one of our key dependent variables—the level of publications—that we examine in the next section. In Section 7, we describe the dependent variables that we use to measure the type of research a lab conducts. Finally, we also collect information on the year in which

[12] Specifically, we used the following procedure. First, based on the author information provided by WoS, we identified the last name and the first initial of each author of a given article, as well as the institution (university, firm, or government agency) with which an author is affiliated. Second, we examined whether the last name and the first initial of the reprint (corresponding) author—which is provided by WoS—matched one of the last three authors listed in the article. If there was a match, we classified the re-print author as the PI. If there was not a match, we examined whether the affiliation of any of the last three authors coincides with the affiliation of the reprint author. If this is the case, we classified the author (among the last three authors) with the matched affiliation as the PI. Otherwise, we classified the re-print author as the PI.

the PIs joined and left their respective UC institutions, which we use to control for the PI's tenure and to construct a balanced panel for robustness analyses.

The journal-based method we use to identify labs has two advantages. First, relative to alternative approaches such as manual collection from archival sources, this method relies on data that are systematically available. Second, with this method, it is not costly to scale the sample to include non-UC universities, which we use to examine the external validity of our main findings in Section 6.2.

To confirm that our sample is reasonably comprehensive in capturing UC chemistry PIs, we conduct two empirical exercises. First, we retrieved the historical web page of the UCLA Chemistry and Biochemistry department website in 2008, the year of the Sangji accident, using the internet portal Wayback Machine. After excluding researchers who are not suitable for our analysis—for example, emeriti professors inactive during our sample period and cross-appointed researchers who publish mostly outside chemistry—we find only seven additional PIs that were listed on the 2008 website but are not in our UCLA sample.[13] Second, we show that by expanding the list of journals down the impact factor ranking, our ability to capture additional new PIs decreases significantly. Recall that our sample uses 105 journals, which yields 592 UC PIs. Adding ten more journals yields only 13 additional UC PIs and including another ten journals yields only three additional UC PIs. These numbers suggest a limited benefit of adding more journals.[14]

5.2 Wet and dry labs

To identify the effect of the shock, we distinguish between dry labs (the control group) and wet labs (the treatment group). We use dry labs as the control group because the shock—both the increase in risk perception of working with dangerous compounds and the more stringent safety regulations—has a limited impact on dry labs. Discussion with the EHS officers at UCLA and UC Davis confirmed this point. This is also supported by the UC Personal Protective Equipment policy guidelines, which state that the safety regulations apply to a "*location where the use or storage of hazardous materials occurs or where equipment may present a physical or chemical hazard.*"

[13] Our baseline analysis is robust to adding these seven PIs. It is useful to note that our sample also includes 18 PIs who were not listed on the 2008 website. These include those who joined UCLA after (or left before) 2008, as well as UCLA chemists that publish extensively in chemistry journals even if they are not members of this department.

[14] Our results are also robust to adding these 16 additional PIs.

We hire a team of chemistry PhD students to classify labs based on the lab webpage, as well as the PI's CV and publications. A lab is classified as a wet lab if the available information indicates that it is equipped to handle biological specimens, chemicals, drugs, and other materials used in experiments. Otherwise, the lab is coded as a dry lab. An example of a wet lab is the one run by Professor Ohyun Kwon at UCLA. Her research focuses on the transformation, catalyst, and synthesis of natural compounds. Her research group uses various research equipment including fume hoods, solvent stills, and glove boxes to manipulate hazardous materials. An example of a dry lab is the one run by Professor Anastassia Alexandrova, also at UCLA. Professor Alexandrova's research focuses on computational and theoretical design and multi-scale description of new materials. Her work relies on quantum and statistical methods, including artificial intelligence and machine learning algorithms. Of the 592 labs in the UC Sample, 512 (86.5%) are classified as wet labs, and 80 (13.5%) are classified as dry labs.

5.3 Econometric method

Our empirical strategy relies on difference-in-differences estimations in which the treatment group includes wet labs and the control group includes dry labs. The pre-treatment period is 2004-2008, and the treatment period is 2009-2017. The unit of observation is a lab-year. Specifically, we estimate

$$Y_{l,t} = \alpha + \beta WetLab_l \times AfterAccident_t + \theta X_{l,t} + \delta_t + f_l + \varepsilon_{l,t}, \tag{1}$$

where the dependent variable, $Y_{l,t}$, captures the publication level (or the type of research) of lab l in year t. The treatment variable, $WetLab_l$, is equal to one for wet labs. The dummy, $AfterAccident_t$, is equal to one for the years after 2008. The term $X_{l,t}$ captures time-varying controls at the lab level, including the lab's past productivity (measured as the total number of publications in the past three years) and for the PI's tenure at UC (in log). The terms δ_t and f_l are year and lab fixed effects. The coefficient β is a difference-in-differences estimator for the effect of the shock on the research output of wet labs relative to dry labs. The baseline results are estimated by OLS regressions, with the standard errors clustered at the lab level.

Table 1 provides summary statistics for the key empirical variables used in our empirical analysis. On average, the labs published 7.374 articles each year during our sample period, which received about

617 citations by 2020. Table A1 in the Appendix shows that wet and dry labs are not significantly different in their pre-shock publication levels, even though an average dry lab PI joined UC later than an average wet lab PI.[15] It is, therefore, important to control for the scientists' tenure in our regressions.

6 Impact on the rate of research

Table 2 presents the estimated effects of the shock on the level of publications. Column 1 shows that relative to dry labs, UC wet labs reduced their publications by 0.28 papers per year, on average, after 2008 compared to before 2008. The estimate is *not* statistically significant at the 10% level. Assuming the same difference between dry and wet labs before and after 2008, the decline for wet labs is about 3%.[16] This is small relative to the estimates of other drivers of scientists' productivity. For example, Oettl (2012) estimated a 20% performance decrease associated with the unexpected loss of a highly productive and helpful co-author, whereas Baruffaldi and Gaessler (2021) showed that the unexpected loss of lab equipment leads to a publication decline of about 15%.

Column 2 focuses on the most impactful publications by counting only publications with citation counts in the top decile of our sample. The estimated effect of the shock remains small and statistically insignificant. In column 3, we again use the total number of publications as the dependent variable but restrict the sample to only labs at UCLA. The estimated effects are even smaller than those estimated for the full UC sample. Finally, column 4 re-estimates the effect of the shock in the UCLA sample using the number of articles published per lab member as the dependent variable. It is difficult to recover historical data on lab members. As a proxy, we use the number of unique researchers who published with a PI in a given year and who were also affiliated with UCLA.[17] We estimate a small, positive, and statistically insignificant coefficient, which indicates that the shock had a limited impact not only on a lab's total research output but also on the productivity per lab member.

[15] On average, dry labs publish about 6.3 articles per year before the shock and 9.2 articles per-year after the shock. If anything, this slight increase in dry lab publication rates may bias our analysis against finding a null effect.

[16] The average number of papers for dry labs after 2008 is 9.19, and the pre-2008 difference between wet and dry labs is -0.30 papers per year. Thus, the hypothetical average for wet labs would have been 8.89 publications per year after 2008.

[17] Specifically, we first identify all the unique co-authors of the articles published by a focal PI in a given year. We then keep those who are also affiliated with UCLA. Finally, we drop names that match PIs in our UCLA sample and names with a relatively long tenure at UCLA—specifically, above the 95th percentile of the tenure distribution among all the coauthors of UCLA PIs—to exclude potential faculty members in other departments at UCLA.

6.1 Pre-treatment trend and time-specific treatment effects

The previously reported results show that relative to dry labs, wet labs do not experience a significant decline in research productivity after the shock, on average. We also conduct a time-specific analysis to examine both the pre-trends and the possibility of some time-specific treatment effects. The time-specific coefficients and their 90% confidence intervals are illustrated in Figure 1. First, there are no statistically significant differences in the yearly publication level between the two groups of labs before 2008, which supports the common-trends assumption.

Second, none of the post-treatment coefficients is statistically significant, confirming an overall null effect. That said, it seems notable that there is a slight (though statistically insignificant) dip during the period 2012-2015, followed by a recovery after 2015. This is consistent with the idea that introducing stricter safety protocols has a chilling effect, but the effect is small and short-term. This may be because labs adapt and develop routines that facilitate compliance. Similarly, the EHS offices also made continual improvements to reduce compliance costs; for example, by simplifying the SoP templates, making training videos more fun, and making inspections more efficient.[18]

Taken together, the results so far show that, wet labs do not appear to experience a significant decline in research productivity after the UCLA accident, despite the significantly more stringent safety regulations. Mapped to our theoretical model, this finding implies that we cannot reject that $\rho = 0$. It is possible that the two tasks—safety and research—are truly independent of one another. It is also possible that the null finding reflects a small net combined effect; that is, the positive and negative effects of safety investment on the marginal cost of research compensate for one another. Regardless of the interpretation, the overall conclusion is the same: the increase in safety investments does not appear to affect research productivity.

[18] As discussed in Section 3, the two changes to the PIs' work environment took place around the same time after the shock and it is difficult to tease out the separate mechanisms. Unreported event studies show no evidence of significant changes in the publication levels over time even when we estimate the effect separately for UCLA and non-UCLA PIs, even though the slight (still insignificant) dip took place several years later at non-UCLA campuses. This is consistent with the fact that more stringent safety regulations were introduced later outside UCLA, but the lack of statistical significance prevents us from making meaningful conclusions.

6.2 Robustness and extensions

In the remainder of this section, we show that our main finding is robust to alternative specifications and extensions. The results are presented in Appendix Tables A2 and A3.

6.2.1 Alternative specifications

We first confirm that the shock's null result on research productivity is robust to alternative econometric models, including Poisson quasi-maximum-likelihood estimation and a weighted OLS model (where the observations are weighted by the pre-2008 publication level of the lab). The result is also robust to using the citation-weighted publication level as the dependent variable. Moreover, the result remains the same using a subsample of labs run by scientists who remained at UC during the entire sample period. Furthermore, the result is also robust to alternative ways to control for the researcher's experience, including using third-degree polynomials of tenure to account for life-cycle effects and controlling for the experience of the researcher measured as years as a PI at any institution rather than just at the current UC institution. Finally, adding additional controls of time-varying university characteristics such as the market value of the university endowment, the number of chemistry PhD students enrolled at the university, and federal science and engineering grants the institution received,[19] or replacing these time-varying controls with institution-year fixed effects also did not change the result.[20]

6.2.2 Journal-based lab classification

We also use an alternative method to distinguish between dry and wet labs. Instead of manually classifying each lab based on information available from its webpage, we ask the team of chemistry PhD students to identify the subset of journals (among the 105 described in Section 5) that specialize in theoretical and computational work. We classify a lab as a dry lab if the fraction of its pre-2008 publications in these journals is in the top decile of the sample. We again find no significant difference in the publication level between the treatment and control groups using this alternative definition of

[19] Endowment data are collected from the National Association of College and University Business Officers, the grants and PhD student data are from surveys by the National Science Foundation. The information is missing for a few institutions-years which explains the lower number of observations in this regression.

[20] We also confirm that the findings are robust to dropping from the sample the handful of labs for which the PhD students disagreed on the wet/dry classification.

dry versus wet labs in the UC sample.

6.2.3 Impact on non-UC universities

Finally, using the journal-based method to classify wet versus dry labs, we also find no differential change in the publication level between dry and wet labs for non-UC universities before versus after the shock. The coefficient, a precisely estimated zero, provides support for the external validity of our main finding. This also suggests that the null effect estimated in the baseline is not driven by the small size of the sample. Unreported results also show similar null results across the university size distribution (measured by a university's total number of labs in the US sample). This suggests that the shock also did not translate into a large tax on research productivity for smaller institutions that may not have had sufficient endowments or access to government funds to withstand potential disruptions of more stringent safety regulations.

7 Impact on the direction of research

Apart from the level of publications, the shock may also affect the type of research a lab conducts. A natural question to ask is whether the shock induced labs to redirect their research toward safer projects. In the following, we first examine whether a lab changed its tendency to work with dangerous compounds and then explore the shock's impact on a lab's overall research agenda.

7.1 Working with dangerous chemicals

We use two additional datasets to measure how much a lab works with dangerous compounds. The first is Scifinder, a proprietary chemistry database that documents all the chemical compounds associated with a journal article. Because Scifinder restricts the number of entries that can be downloaded in total and during each session, we limit the analysis in this section to only ULCA labs. The second database is the Laboratory Chemical Safety Summary (LCSS), which is publicly accessible via PubChem. LCSS uses the Globally Harmonized System (GHS) to classify the hazard levels of a compound. We create a variable "dangerous compound," which equals one if GHS associates it with a "danger" signal and equals zero if GHS either associates it with a "warning" signal or does not provide any hazardous information.

Moreover, GHS classifies a hazardous compound into nine broad classes that are not mutually exclusive: explosive, compressed gas, irritant, flammable, corrosive, health hazards, oxidizer, acute toxicity, and environment. Using these two datasets, we are able to determine whether each article published by UCLA researchers references dangerous chemicals as well as the types of hazards.[21] On average, UCLA dry labs published 2.65 articles a year referencing dangerous compounds, whereas wet labs published 3.30 articles per year referencing dangerous compounds.

Manual examination of several publications supports the idea that the references made by dry labs capture the study of the theoretical properties of these compounds through mathematical or computational models. For example, the UCLA dry lab run by Anastassia Alexandrova published an article titled, "On the mechanism and rate of spontaneous decomposition of amino acids" in the *Journal of Physical Chemistry B.* The study relied on Monte Carlo simulations to examine the properties of methylamine (CAS no 74-89-5), which is labeled by GHS as extremely flammable and as an irritant (may cause skin and eye irritation). For wet labs, many of the compounds referenced typically capture the use of these chemicals in experiments. For example, in 2014 the UCLA wet lab run by Ohyun Kwon published a paper in the *Asian Journal of Organic Chemistry* titled "Phosphine-Initiated General-Base-Catalyzed Quinolone Synthesis." The method section of this article describes an experimental procedure in which triphenylphosphine (CAS no 603-35-0) is mixed in a flask with other compounds and stirred. GHS labels this compound as an irritant and with health hazards (may cause cancer).

7.1.1 Wet versus dry labs

In Table 3, we present a series of regressions examining whether UCLA wet labs changed their propensity to publish research referencing dangerous compounds relative to dry labs after 2008. Publications from dry labs that reference dangerous chemicals are used as a control for general trends of research interest in these compounds. In column 1, the dependent variable is the number of publications involving any kind of dangerous compound. In the remaining columns, we examine the number of publications that

[21] We merged the LCSS data and the SciFinder data using the CAS registry numbers, which are unique identifiers for chemical substances. Using digital object identification (doi) information, we are able to find 80% of the WoS articles in the UCLA sample in Scifinder. The fraction of unmatched papers appears fairly constant across sample years and is not driven by specific labs. In Appendix Table A4, we replicate our analysis of the level of publications using the Scifinder instead of the WoS data, and the estimates are in line with those presented in Section 6.

involve specific types of compounds: acute toxic, explosive, and flammable. We find that on average, wet labs do not reduce research referencing dangerous chemicals relative to dry labs. The coefficients of the interaction term are all small in magnitude and statistically insignificant.[22]

7.1.2 Wet labs by their baseline hazard levels

Recall that an extension of our baseline model in Section 3.2.3 suggests that the shock may induce greater-hazard labs to redirect research toward safer projects. Thus, even if wet labs on average may not shift away from using dangerous compounds, their responses may vary by the baseline hazard level (due to the nature of their research) of individual labs. We define labs as high-hazard if their fraction of SciFinder publications referring to dangerous chemicals during the pre-shock period are among the top 20% of the sample.[23] For this analysis, we focus on labs run by PIs affiliated with UCLA for the entire sample period because we need a relatively long pre-period to define a lab's intensity of using dangerous chemicals. This leaves us with a sample of 42 wet labs, 8 of which are classified as high-hazard labs.

Table 4 reports a series of regressions comparing UCLA wet labs with high versus low hazard levels. Consistent with our baseline finding in Section 6, column 1 shows that the publication level is not statistically significant between the two types of labs. However, column 2 shows that compared to low-hazard wet labs, high-hazard wet labs publish 1.17 fewer articles (significant at the 0.05 level) referring to dangerous chemicals per year after the shock. Columns 3 through 5 examine different types of hazards. The results show that the decline found in column 2 is driven primarily by flammable chemicals. High-hazard wet labs, on average, experienced a decline of about 0.96 articles referring to flammable chemicals per year after 2008 relative to the other wet labs (significant at the 0.01 level).[24]

In Figure 2, we illustrate a graphical representation of the dynamic evolution in the differential use of flammable chemicals between UCLA high-hazard versus low-hazard wet labs. No evidence exists of pre-trend differences between the two groups. After 2008, the coefficients become negative and significant (or marginally significant) in most years until 2014. The reduction in the reliance on flammable chemicals

[22] In unreported regressions, we confirm this finding using other hazard categories reported in the LCSS data, which include corrosive or irritant substances, and compressed gas.

[23] Our analysis is robust to using the absolute number of publications referring to dangerous compounds rather than the fraction of publications to identify the labs with more intense use of hazardous chemicals.

[24] Unreported regressions show that the estimates for the other hazard types (corrosive substances, compressed gas, and irritant substances) are much smaller in magnitude and statistically insignificant.

could be driven either by the increase in the perception of risk/liability, which has led to a voluntary shift away from the use of such chemicals, or by increasingly more stringent safety rules at the university, which may have made the use of such chemicals more costly. The magnitudes are especially large in 2013 and 2014 after the settlement. The effects are smaller and statistically insignificant after 2015, suggesting that the use of flammable chemicals recovered gradually after safter practices were implemented.

We also examine whether the 2008 accident is associated with a change in the propensity to use chemicals that are unfamiliar to the lab. Specifically, we generate a dummy variable indicating that a lab used compounds in a given year that had not previously been referenced by any of the UCLA labs.[25] Assuming that labs are more likely to be familiar with the properties of compounds they or their local colleagues had already handled, working with these chemicals is likely to imply greater risk. The regressions in Table 5 show no evidence of a decline in the use of unfamiliar compounds, on average. However, separating safe compounds (column 2) from dangerous compounds (column 3) shows that while no significant change is found in the use of unfamiliar safe compounds, a strong negative effect is found for unfamiliar dangerous compounds. The magnitude of the decline is about 40% of the mean level of the dependent variable. Column 4 confirms this result using the number of papers the lab published in a given year that refer to unfamiliar dangerous compounds as the dependent variable. In this case, the estimate also indicates a significant decline in the references to dangerous chemicals not previously used at UCLA.[26]

Taken together, the results in this section thus far provide evidence for an increase in safety induced by the shock and suggest that lab safety has some effect on the production of scientific research. That said, such an effect is localized to a small subset of labs that use dangerous chemicals most intensely, rather than broadly affecting the entire institution.

[25] We classify a publication as the first use of a compound at UCLA if we cannot find a reference to that compound in previous publications (since 1998) of all the UCLA labs in our data.

[26] We also attempt to examine whether the shock changed the propensity of wet labs to discover new compounds (e.g., whether the publication is the first that has ever referenced a compound in the academic literature). This is a challenging exercise because publicly available data sources that provide comprehensive historical bibliographical data for compounds (such as PubChem) can only be matched to a small subset of the compounds in our data (roughly 20%). The available data indicate an extremely low propensity of discovering new compounds for wet labs in our sample.

7.2 Effect on research content

Does the reduced use of dangerous compounds found in some of the wet labs indicate a meaningful change in the subject of the research? To explore this question, we construct a similarity measure between a lab's research published in a given year and its core research before the shock. Specifically, we first use the Bidirectional Encoder Representations from Transformers (BERT) machine-learning language model (Devlin et al., 2018) to convert paper abstracts to 768-dimensional vectors.[27] For each lab, we calculate the average of all the article vectors before 2008. This average vector represents the "core research" the lab conducted before the shock. We then calculate the cosine similarity between the vector representing each article the lab published and its pre-shock core research vector. Finally, we aggregate these article-level similarity measures into a yearly measure (e.g., by taking the average). Appendix B provides the details on the variable construction and describes several empirical exercises that we conducted to validate the measure.

Armed with these yearly similarity measures, we examine how a lab's research may have changed after the shock relative to before. Table 6 presents a series of regressions that compare high-hazard to low-hazard wet labs at UCLA. The dependent variable of column 1 is the average similarity score of all the articles published in a given year by a lab, as described in the previous paragraph. Columns 2 through 5 focus on research that refers to dangerous compounds. The dependent variable is still a yearly similarity measure, but the measure is constructed using only articles that refer to dangerous compounds: column 2 uses articles that refer to at least one dangerous compound; columns 3 and 4 use articles for which the fraction of dangerous compounds referenced exceeds 50% and 75%; and column 5 uses articles for which all the compounds referenced are dangerous.[28]

Table 6 yields two results. First, the small and largely insignificant estimates of columns 1 and 2

[27] The simplest way to construct a similarity measure is to use the actual vocabulary terms used in a textual document, whereby each unique word in the entire corpus constitutes a dimension. A main shortcoming of these vocabulary-based matrices is that they are very sparse. More importantly, such methods do not account for relationships between words. Terms related to one another (e.g., "cars" and "automobiles") are not treated as more similar than words that are not (e.g., "cars" and "dogs"). BERT addresses both problems by reducing the dimensionality of the matrices and considering the relationships between words.

[28] Notice that the sample size shrinks as we move from columns 3 to 5, because we need to drop lab-year observations that do not include any article above the specific threshold. For consistency, in columns 2-5 of Table 6, we construct the similarity measure using a benchmark 'core research' vector constructed only using all the articles a lab published before the shock that refer to at least one dangerous compound.

suggest that compared to low-hazard wet labs, the shock does not have a substantial impact on the research content of high-hazard wet labs, even when we look at only articles that make some references to dangerous compounds. Second, at the same time, columns 3 through 5 show that for articles that refer more substantially to dangerous compounds, compared to low-hazard wet labs, those published by high-hazard wet labs after 2008 become significantly more similar to their pre-shock core research. Table A5 in the Appendix shows robustness to these two results using either the minimum or the median similarity score of the articles published in a year.

Taken together, the results presented in this section show that the shock does not appear to have substantially altered the research direction of wet labs in our sample. However, it did induce a subset of wet labs that more heavily use dangerous chemicals to become relatively more conservative in their most hazardous research. These interpretations resonate with our discussions with the UC EHS officers. All of them conjectured that the shock was unlikely to have altered a lab's research agenda and that, for core experiments, the lab will find ways to run them "*no matter what.*" On the margin, however, the EHS officers thought that the researchers might have modified their process by changing their start materials (the materials that are used as input in a chemical experiment, which are substitutable) or reduced research involving dangerous compounds if it is less core to the lab.[29]

8 Discussion and conclusion

In this paper, we study the relationship between lab safety and research productivity by examining the impact of a substantial increase in safety regulations at the University of California following the shocking death of a UCLA research assistant. We have two key findings. First, compared to dry labs that use theoretical and computational methods, the publication rates of wet labs that conduct experiments on chemical and biological substances did not change significantly after the shock. This finding, thus, does not support the common view among academic researchers that "*safety is a tax on research productivity.*" Second, while the shock had no broad impact on the type of research a lab

[29] Interestingly, Gibson et al (2014) also document that the more stringent paperwork requirement after the accident has induced more disposal of hazardous chemicals. Our conversations with the EHS officers confirmed that this is the case and suggest that these disposed of materials are likely to be extra or old materials that the labs do not frequently use in their current research.

conducted, it led a small set of labs that worked intensely with dangerous materials before the shock to be more conservative in their research with dangerous materials. This manifested in a reduction in the use of flammable chemicals and dangerous chemicals that were unfamiliar to the researchers, as well as a reluctance to deviate from prior research in their new projects that involved dangerous chemicals.

It is important to recognize that these findings may have alternative explanations. Specifically, safety investment may actually have a negative impact on research productivity, but this effect might have been masked by a relative increase in research funding for wet labs versus dry labs around the same time. We do not have data on the funding each lab received, but two pieces of evidence mitigate this concern. First, we find that wet labs do not experience a significant change in the size of their labs after the shock relative to dry labs. Second, surveys the National Science Foundation conducted showed that after 2009, research funds (both federal and private) for computer sciences actually increased more sharply than they did for chemistry. To the extent that computational fields in chemistry mirror the general trend in computational sciences, this would make it easier, rather than harder, for us to find a relative decline in wet labs' productivity.

As for the second finding on the reduced use of dangerous chemicals, a possible explanation is the contemporaneous diffusion of the green chemistry movement, which aims to minimize the use or generation of hazardous chemicals.[30] To address this concern, we first show that our key results are not driven by articles related to green chemistry.[31] Furthermore, leveraging the data on different hazard types, we confirm that our shock does not lead to a significant reduction in the use of environmentally hazardous compounds, which are central to the green chemistry movement. Instead, the reduction is driven mostly by flammable compounds that are specific to the UCLA accident and the subsequent revision of safety standards.

Our paper contributes to the literature and the policy debates on the trade-offs between safety

[30] The green chemistry movement gained prominence in the late 90s, after the publication of a framework called "the 12 Principles of Green Chemistry" by Anastas and Warner (1998). While the movement has focused on environmental hazards, Principle 12 of this framework does state that "*substances and the form of a substance used in a chemical process should be chosen to minimize the potential for chemical accidents, including releases, explosions, and fires.*"

[31] To identify green chemistry articles, we followed one of the approaches developed by Nelson et al. (2014) and searched for the phrase 'green chemistry' in the title and abstract of the articles. We found that only a small set of articles published by our UC PIs involved green chemistry (less than 50 articles). The baseline regressions (Table 2) and those showing a reduction in the use of dangerous compounds (Table 4) are robust to controlling for the number of green chemistry papers a lab published each year.

and productivity in several ways. First, it adds much-needed empirical evidence from an important setting. Academic research fuels research and development and, ultimately, economic growth. While the relationship between lab safety and research productivity has been much debated in the recent decade, to the best of our knowledge, ours is the first study that provides large-sample evidence on this topic. Second, this paper also adds new data—hazardous information at the compound level—to examine changes in research direction. The safety dimension has been overlooked by the literature, even though it is greatly relevant to researchers' well-being and the type of research generated.

This study is not without limitations. First, rather than a random policy change, the increase in safety regulations in our setting follows a high-profile accident. This makes it difficult to empirically differentiate the effect of the mandated safety regulations from that of an increase in the perception of risk and liability. This is a common problem for many studies examining changes in industry regulation (Gowrisankaran et al., 2015; Nameroff, et al. 2004), because accidents, for better or worse, are typically the impetuses for the strengthening of safety practices that we observe in practice (Barnett and King, 2008). In this paper, we make some progress in our understanding of this problem by clarifying theoretically the potential standalone impacts of these various changes. Second, we also cannot precisely conclude whether our null finding reflects a truly independent relationship between safety investment and research productivity or that various positive and negative effects of safety investment on the marginal cost of research compensate for one another, leading to a small net effect. Given the contentious debates about the relationships between safety and productivity in the literature and in practice, we favor the latter interpretation. But we cannot rule out the former. We recognize all of these limitations, which also point to interesting paths for future research.

References

[1] ACS (2012a), Creating safety cultures in academic institutions: A report of the safety culture task force of the ACS committee on chemical safety, report available at www.acs.org/content/acs/en/education/students/graduate/creating-safety-cultures-in-academic-institutions.html

[2] ACS (2012b), Advancing graduate education in the chemical sciences, report available at www.acs.org/content/acs/en/about/governance/acs-presidential-commission-on-graduation-education-in-the-chemical-sciences.html

[3] Adler, Paul, Barbara Goldoftas, and David Levine (1997) "Ergonomics, employee involvement, and the Toyota Production System: A case study of NUMMI's 1993 model introduction," *ILR Review* 50: 416-437

[4] Anastas, Paul, and John Warner (1998) *Green chemistry: Theory and practice*, Oxford University Press

[5] Azoulay, Pierre, Christian Fons-Rosen, and Joshua Graff-Zivin (2019) "Does science advance one funeral at a time?," *American Economic Review* 109: 2889-2920

[6] Azoulay, Pierre, Joshua Graff Zivin, and Gustavo Manso (2011) "Incentives and creativity: Evidence from the academic life sciences," *The RAND Journal of Economics* 42: 527-554

[7] Azoulay, Pierre, Joshua Graff-Zivin, and Jialan Wang (2010) "Superstar extinction," *Quarterly Journal of Economics* 125: 549-589

[8] Azoulay, Pierre, Waverly Ding, and Toby Stuart (2009) "The impact of academic patenting on the rate, quality and direction of (public) research output," *The Journal of Industrial Economics* 57: 637-676

[9] Barnett, M. L. and A. A. King (2008) "Good fences make good neighbors: A longitudinal analysis of an industry self-regulatory institution," *Academy of Management Journal* 51, 1150-1170

[10] Baruffaldi, Stefano, and Fabian Gaessler (2021) “The returns to physical capital in knowledge production: Evidence from lab disasters,” Max Planck Institute for Innovation & Competition Research Paper 21-19

[11] Basken, Paul (2012) “UCLA criminal charges focus universities’ attention on laboratory safety,” *Chronicles of Higher Education*, April 15

[12] Baudendistel, Brian (2009) “Investigation report University of California, Los Angeles,” Case No. S1110-003-09, (Department of Industrial Relations, Division of Occupational Safety and Health, Los Angeles)

[13] Bénabou, Roland, and Jean Tirole (2016) “Bonus culture: Competitive pay,screening, and multitasking,” *Journal of Political Economy* 124: 305-370

[14] Benderly, Beryl (2009) “The burning question of laboratory safety,” *Science* online article available at www.science.org/content/article/burning-question-laboratory-safety

[15] Brenner, Mark, David Fairris, and John Ruser (2004) “Flexible work practices and occupational safety and health: exploring the relationship between cumulative trauma disorders and workplace transformation,” *Industrial Relations: A Journal of Economy and Society* 43: 242-266

[16] Das, Ajay, Mark Pagell, Michael Behm, and Anthony Veltri (2008) “Toward a theory of the linkages between safety and quality,” *Journal of Operations Management* 26: 521-535

[17] De Philippis, Marta (2021) “Multi-task agents and incentives: The case of teaching and research for university professors,” *Economic Journal* 131: 1643-1681

[18] Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova (2018) “Bert: Pre-training of deep bidirectional transformers for language understanding,” arXiv preprint arXiv:1810.04805

[19] Daughety, Andrew and Jennifer Reinganum (2013) "Economic analysis of products liability: theory," In *Research handbook on the economics of torts*, Edward Elgar Publishing

[20] Dawid, Herbert, and Gerd Muehlheusser (2022) "Smart products: Liability, investments in product safety, and the timing of market introduction," *Journal of Economic Dynamics and Control* 134: 1-20

[21] Feng Lu, Susan (2012) "Multitasking, information disclosure, and product quality: Evidence from nursing homes," *Journal of Economics & Management Strategy* 21: 673-705

[22] Fryer, Roland, and Richard Holden (2013) "Multitasking, dynamic complementarities, and incentives: A cautionary tale," Harvard University working paper

[23] Furman, Jeffrey, and Florenta Teodoridis (2020) "Automation, research technology, and researchers' trajectories: Evidence from computer science and electrical engineering," *Organization Science* 31: 330-354

[24] Furman, Jeffrey, and Scott Stern (2011) "Climbing atop the shoulders of giants: The impact of institutions on cumulative research," *American Economic Review* 101: 1933-1963

[25] Galasso, Alberto, and Hong Luo (2017) "Tort reform and innovation," *Journal of Law and Economics* 60: 385-412

[26] Galasso, Alberto, and Hong Luo (2021) "Risk-mitigating technologies: The case of radiation diagnostic devices," *Management Science* 67: 3022-3040

[27] Galasso, Alberto, and Hong Luo (2022) "When does product liability risk chill innovation? Evidence from medical implants," *American Economic Journal: Economic Policy* 14: 366–401

[28] Gibson, James, Imke Schröder, and Nancy Wayne (2014) "A research university's rapid response to a fatal chemistry accident: Safety changes and outcomes," *Journal of Chemical Health and Safety* 21: 18-26

[29] Gray, Wayne (1987) "The cost of regulation: OSHA, EPA and the productivity slowdown," *American Economic Review* 77: 998–1006

[30] Greenstone, Michael, Richard Hornbeck, and Enrico Moretti (2010) "Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings," *Journal of Political Economy* 118: 536-598

[31] Grennan, Matthew, and Robert Town (2020) "Regulating innovation with uncertain quality: Information, risk, and access in medical devices," *American Economic Review* 110: 120-161

[32] Goggins, Richard, Peregrin Spielholz, and Greg Nothstein (2008) "Estimating the effectiveness of ergonomics interventions through case studies: Implications for predictive cost-benefit analysis," *Journal of Safety Research* 39: 339-344

[33] Gowrisankaran, Gautam, Charles He, Eric Lutz, and Jefferey Burgess (2015) "Productivity, safety, and regulation in underground coal mining: Evidence from disasters and fatalities," National Bureau of Economic Research working paper 21129

[34] Hayden, Erika (215) "Scientific society withholds honour over lab death," *Nature News*, December

[35] Hersh, Erica (2017) "Five ways effective laboratory design impacts health, safety, and productivity," electronic article available at www.hsph.harvard.edu/ecpe/five-ways-effective-laboratory-design-impacts-health-safety-and-productivity/

[36] Holmstrom, Bengt, and Paul Milgrom (1991) "Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design," *Journal of Law Economics and Organizations*: 24-52

[37] Howard-Grenville, Jennifer, Andrew Nelson, Andrew Earle, Julie Haack, and Douglas Young (2017) "If chemists don't do it, who is going to? Peer-driven occupational change and the emergence of green chemistry," *Administrative Science Quarterly* 62: 524-560

[38] Kemsley, Jyllian (2009) "Learning from UCLA," *Chemical & Engineering News* 87: 29-31

[39] Langerman, Neal (2009) "Laboratory safety?" *Journal of Chemical Health and Safety*, 16: 49-50

[40] Levine, David, and Michael Toffel (2010) "Quality management and job quality: How the ISO 9001 standard for quality management systems affects employees and employers," *Management Science* 56: 978-996

[41] Levine, David, Michael Toffel, and Matthew Johnson (2012) "Randomized government safety inspections reduce worker injuries with no detectable job loss," *Science* 336: 907-911

[42] Lo, Chris, Mark Pagell, Di Fan, Frank Wiengarten, and Andy Yeung (2014) "OHSAS 18001 certification and operating performance: The role of complexity and coupling," *Journal of Operations Management* 32: 268-280

[43] Maxmen, Amy (2018) "Judge dismisses criminal charges against US university chemist over lab death," *Nature* online article available at www.nature.com/articles/d41586-018-06726-1

[44] Mclain, David (2010) "Responses to health and safety risk in the work environment," *Academy of Management Journal* 38: 1726-1743

[45] Merlic, Craig (2013) "Enhancing laboratory safety - Academic laboratory safety in a post 12/29 World," unpublished presentation deck - April 11

[46] Mullen, Kathleen, Richard Frank, and Meredith Rosenthal (2010) "Can you get what you pay for? Pay for performance and the quality of healthcare providers," *The Rand Journal of Economics* 41: 64-91

[47] Murray, Fiona, Philippe Aghion, Mathias Dewatripont, Julian Kolev, and Scott Stern (2016) "Of mice and academics: Examining the effect of openness on innovation," *American Economic Journal: Economic Policy* 8: 212-252

[48] Myers, Kyle (2020) "The elasticity of science," *American Economic Journal: Applied Economics* 12: 103-134

[49] Nameroff, Tamara, Raymond Garant, and Michael Albert (2004) "Adoption of green chemistry: An analysis based on US patents," *Research Policy* 33: 959-974

[50] National Research Council (2014) *Safe science: Promoting a culture of safety in academic chemical research*, National Academies Press.

[51] Oettl, Alexander (2012) “Reconceptualizing stars: Scientist helpfulness and peer performance,” *Management Science* 58: 1122-1140

[52] Pagell, Mark, Mary Parkinson, Anthony Veltri, John Gray, Frank Wiengarten, Michalis Louis, and Brian Fynes (2020) “The tension between worker safety and organization survival,” *Management Science* 66: 4863-4878

[53] Palmer, Karen, Wallace Oates, and Paul Portney (1995) “Tightening environmental standards: The benefit-cost or the no-cost paradigm?,” *Journal of Economic Perspectives* 9: 119-132

[54] Peltzman, Sam (1973) “An evaluation of consumer protection legislation: The 1962 drug amendments,” *Journal of Political Economy* 81: 1049-1091

[55] Perrow, Charles (1984) *Normal Accidents: Living with High Risk Technologies*, Princeton University Press

[56] Porter, Michael, and Claas van der Linde (1995) “Toward a new conception of the environment-competitiveness relationship,” *Journal of Economic Perspectives* 9: 97-118

[57] Roche, Maria (forthcoming) “Academic entrepreneurship: Entrepreneurial advisors and their advisees’ outcomes,” *Organization Science*

[58] Silver, Andrew (2021) “Fatal lab explosion in China highlights wider safety fears,” *Nature* electronic publication available at www.nature.com/articles/d41586-021-03589-x

[59] Stephan, Paula (2010) “The economics of science,” in *Handbook of the Economics of Innovation*, Vol 1, North-Holland

[60] Trager, Rebecca (2014) “UCLA chemist avoids prison time for lethal lab accident,” *Chemistry World*, 26 June

[61] Van Noorden, Richard (2011) “A death in the lab,” *Nature* 472: 270-271

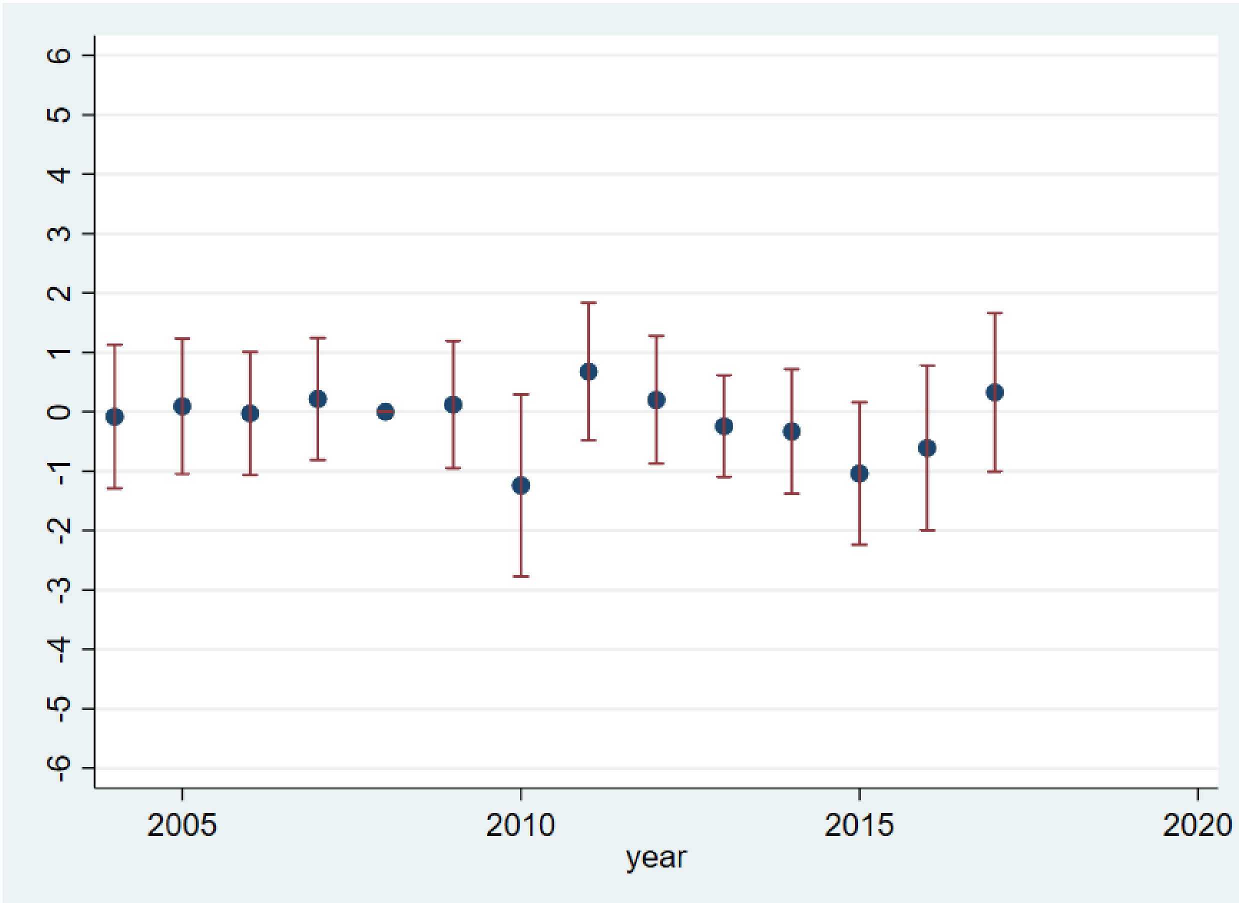
[62] Van Noorden, Richard (2013) “Safety survey reveals lab risks,” *Nature* 493: 9-10

[63] Venkatraman, Vijaysree (2010) "Conventions of scientific authorship," *Science* 12: 243-244

[64] Viscusi, Kip, and Michael Moore (1993) "Product liability, research and development, and innovation," *Journal of Political Economy* 101: 161-184

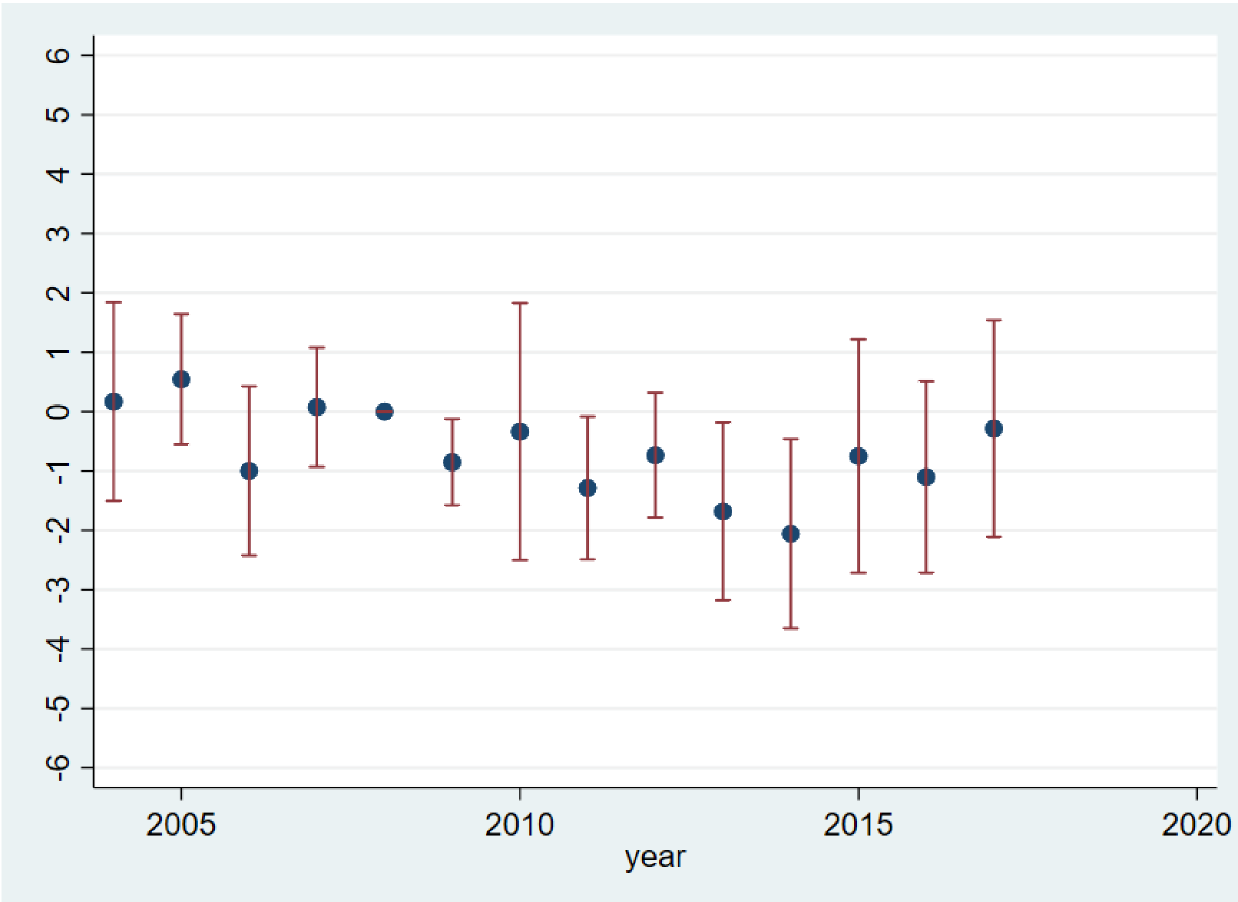
[65] World Health Organization and International Labour Organization (2021) WHO/ILO Joint Estimates of theWork-related Burden of Disease and Injury, 2000–2016, technical report available at www.who.int/publications/i/item/9789240034945

Figure 1: Annual treatment effects – total publications



Note: The figure plots the coefficients (and the 90% confidence intervals) of the interaction terms between year dummies and the wet lab dummy, which equals one if the lab conducts experiments using chemical and biological substances.

Figure 2: Annual treatment effects – publications referring to flammable chemicals



Note: The figure plots the coefficients (and the 90% confidence intervals) of the interaction terms between year dummies and a dummy capturing wet labs with high use of dangerous chemicals (i.e., if the fraction of SciFinder publications referring to dangerous chemicals during the pre-shock period are among the top 20% of the sample). The control group includes the remaining wet labs.

Table 1 - Summary statistics

Panel A: UC sample	obs.	mean	sd	min	max
Articles	6827	7.374	6.821	0	79
Wet Lab	6827	0.868	0.339	0	1
Year	6827	2010.825	3.933	2004	2017
Panel B: UCLA sample	**obs.**	**mean**	**sd**	**min**	**max**
Articles	976	7.740	7.287	0	57
Wet Lab	976	0.826	0.380	0	1
Year	976	2010.826	3.920	2004	2017

NOTES: Unit of observation is a lab-year. Panels A and B report summary statistics for UC and UCLA samples, respectively.Articles = the number of articles the lab published in year t. Wet Lab = 1 if the lab conducts experiments using chemical and biological substances.

Table 2: Introduction of stricter lab safety regulations is not associated with changes in wet labs' publication levels relative to dry labs

	(1)	(2)	(3)	(4)
Dep. Var.	Articles	Highly cited articles	Articles	Articles/lab members
Wet Lab × After Accident	-0.279	0.003	-0.002	0.196
	(0.488)	(0.122)	(1.048)	(0.402)
Sample	UC labs	UC labs	UCLA labs	UCLA labs
Year effects	YES	YES	YES	YES
Lab effects	YES	YES	YES	YES
Observations	6827	6827	976	976

NOTES: OLS regressions. Articles = the number of articles the lab published in year t. Highly cited articles = the number of articles the lab published in year t in the top decile of citations. Wet Lab = 1 if the lab conducts experiments using chemical and biological substances. After Accident = 1 if after year 2008. All regressions control for the total lab publications in the previous three years and the logarithm of the lab's tenure. Lab members = the number of unique local non-PI and non-faculty researchers listed as coauthors in papers the lab published in year t. Robust standard errors clustered at the lab level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Introduction of stricter lab safety regulations is not associated with changes in references to dangerous chemicals in UCLA wet lab publications relative to UCLA dry labs

	(1)	(2)	(3)	(4)
Dep. Var.	Dangerous substances	Acute toxic substances	Explosive substances	Flammable substances
Wet Lab × After Accident	0.303	-0.112	-0.085	-0.118
	(0.918)	(0.734)	(0.102)	(0.887)
Sample	UCLA labs	UCLA labs	UCLA labs	UCLA labs
Year effects	YES	YES	YES	YES
Lab effects	YES	YES	YES	YES
Observations	976	976	976	976

NOTES: OLS regressions. Acute toxicity substances = the number of articles referring to acute toxicity substances the labs published in year t. Explosive substances = the number of articles referring to explosive substances the lab published in year t. Flammable substances = the number of articles referring to flammable substances the lab published in year t. After Accident = 1 if after year 2008. All regressions control for the total lab publications in the previous three years and the logarithm of the lab's tenure. Robust standard errors clustered at the lab level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: The introduction of stricter lab safety regulations is associated with a reduction in the use of flammable substances in UCLA wet labs with high use of dangerous substances relative to other UCLA wet labs

	(1)	(2)	(3)	(4)	(5)
Dep. Var.	SciFinder articles	Dangerous substances	Acute toxic substances	Explosive substances	Flammable substances
High Hazard × After Accident	-0.307	-1.171***	-0.303	-0.030	-0.967**
	(0.782)	(0.492)	(0.390)	(0.045)	(0.477)
Sample	UCLA wet labs active between 2004 and 2017	UCLA wet labs active between 2004 and 2017	UCLA wet labs active between 2004 and 2017	UCLA wet labs active between 2004 and 2017	UCLA wet labs active between 2004 and 2017
Year effects	YES	YES	YES	YES	YES
Lab effects	YES	YES	YES	YES	YES
Observations	588	588	588	588	588

NOTES: OLS regressions. SciFinder articles = the number of articles in the SciFinder database the lab published in year t. Dangerous substances = the number of articles referring to dangerous substances the lab published in year t. Acute toxicity substances = the number of articles referring to acute toxicity substances the lab published in year t. Explosive substances = the number of articles referring to explosive substances the lab published in year t. Flammable substances = the number of articles referring to flammable substances the lab published in year t. High Hazard = 1 if the lab is in the top quintile in terms of publications referencing dangerous substances. After Accident = 1 if after year 2008. All regressions control for the total publications by the lab in the previous three years and the logarithm of the lab's tenure. Robust standard errors clustered at the lab level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Introduction of stricter lab safety regulations is associated with a reduction in the use of relatively unfamiliar dangerous compounds

	(1)	(2)	(3)	(4)
Dep. Var.	Use of compounds new to UCLA	Use of safe compounds new to UCLA	Use of dangerous compounds new to UCLA	Articles referring to new dangerous compounds
High Hazard × After Accident	0.012	0.054	-0.174**	-0.633**
	(0.053)	(0.065)	(0.078)	(0.247)
Sample	UCLA wet labs active between 2004 and 2017	UCLA wet labs active between 2004 and 2017	UCLA wet labs active between 2004 and 2017	UCLA wet labs active between 2004 and 2017
Year effects	YES	YES	YES	YES
Lab effects	YES	YES	YES	YES
Observations	588	588	588	588

NOTES: OLS regressions. Use of compounds new to UCLA = 1 if at least one of the compounds referenced in the labs' publications in year t was never used before at UCLA. Use of safe compounds new to UCLA = 1 if at least one of the safe compounds referenced in the labs' publications in year t was never used before at UCLA. Use of dangerous compounds new to UCLA = 1 if at least one of the dangerous compounds referenced in the labs' publications in year t was never used before at UCLA. Articles referring to new dangerous compounds = the number of articles with dangerous substances first used at UCLA the labs' published in year t. High Hazard = 1 if lab in top quintile in terms of articles published using dangerous substances. After Accident = 1 if after year 2008. All regressions control for the number of compounds referenced by the lab in year t, total lab publications in the past three years, and the logarithm of the lab's tenure. Robust standard errors clustered at the lab level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Introduction of stricter lab safety regulations is associated with an increase in textual similarity for articles referring to many dangerous compounds

	(1)	(2)	(3)	(4)	(5)
Dep. Var.	Average semantic similarity score	Average semantic similarity score	Average semantic similarity score	Average semantic similarity score	Average semantic similarity score
High Hazard × After Accident	-0.018	-0.021*	0.042**	0.090***	0.115***
	(0.011)	(0.012)	(0.020)	(0.032)	(0.037)
Publications sample	All	Articles referring to dangerous compounds	Articles with fraction of dangerous compounds referenced > 50%	Articles with fraction of dangerous compounds referenced > 75%	Articles with fraction of dangerous compounds referenced = 100%
Lab Sample	UCLA wet labs active between 2004 and 2017	UCLA wet labs active between 2004 and 2017	UCLA wet labs active between 2004 and 2017	UCLA wet labs active between 2004 and 2017	UCLA wet labs active between 2004 and 2017
Year effects	YES	YES	YES	YES	YES
Lab effects	YES	YES	YES	YES	YES
Observations	517	517	144	75	65

NOTES: OLS regressions. Average semantic similarity score = average similarity score of all the articles published in a focal year by a lab relative to the same lab's pre-shock core research. High Hazard = 1 if lab in top quintile in terms of articles published using dangerous substances. After Accident = 1 if after year 2008. All regressions control for the total publications by the lab in the previous three years and the logarithm of the lab's tenure. Robust standard errors clustered at the lab level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Online Appendix for: Laboratory Safety and Research Productivity

March 24, 2023

Appendix A: Proofs and Model Extensions

A.1 Empirical implications of the model

The maximization problem for the scientist before the shock is

$$\max_{s,r} r + \underline{\delta} s - \underline{\delta} - \frac{s^2}{2} - \frac{r^2}{2} - \rho s r$$
$$\text{s.t.} \quad s \geq 0 \text{ and } \ r \geq 0.$$

The first-order conditions are:

$$r = 1 - \rho s$$
$$s = \underline{\delta} - \rho r.$$

Considering the constraint $s \geq 0$, the optimal solution for safety investment is

$$s^*(\underline{\delta}, 0) = \begin{cases} \dfrac{\underline{\delta} - \rho}{1 - \rho^2} & \text{if } \underline{\delta} - \rho > 0 \\ 0 & \text{if } \underline{\delta} - \rho \leq 0 \end{cases} ,$$

and the optimal solution for research output is:

$$r^*(\underline{\delta}, 0) = \begin{cases} \dfrac{1 - \underline{\delta}\rho}{1 - \rho^2} & \text{if } \underline{\delta} - \rho > 0 \\ 1 & \text{if } \underline{\delta} - \rho \leq 0 \end{cases} .$$

After the shock, this problem becomes

$$\max_{s,r} r + \overline{\delta} s - \overline{\delta} - \frac{s^2}{2} - \frac{r^2}{2} - \rho s r$$
$$\text{s.t.} \quad s \geq \underline{s} \text{ and } \ r \geq 0.$$

Similar to before the shock, the optimal solutions depend on whether the minimum safety standard is binding; that is, the optimal solution for safety investment is

$$s^*(\overline{\delta}, \underline{s}) = \begin{cases} \dfrac{\overline{\delta} - \rho}{1 - \rho^2} & \text{if } \dfrac{\overline{\delta} - \rho}{1 - \rho^2} > \underline{s} \\ \underline{s} & \text{if} \dfrac{\overline{\delta} - \rho}{1 - \rho^2} < \underline{s} \end{cases} ,$$

and the optimal research output becomes

$$r^*(\overline{\delta},\underline{s}) = \begin{cases} \dfrac{1-\overline{\delta}\rho}{1-\rho^2} & \text{if } \dfrac{\overline{\delta}-\rho}{1-\rho^2} > \underline{s} \\ 1-\rho\underline{s} & \text{if} \dfrac{\overline{\delta}-\rho}{1-\rho^2} < \underline{s} \end{cases}.$$

It is straightforward to see that $\Delta s = s(\overline{\delta},\underline{s}) - s(\underline{\delta},0)$ is positive in each of the following four scenarios:

$$\Delta s = \begin{cases} \dfrac{\Delta\delta}{1-\rho^2} & \text{if } \underline{\delta}-\rho > 0 \;\;\&\;\; \dfrac{\overline{\delta}-\rho}{1-\rho^2} > \underline{s} \\ \dfrac{\overline{\delta}-\rho}{1-\rho^2} & \text{if } \underline{\delta}-\rho \leq 0 \;\;\&\;\; \dfrac{\overline{\delta}-\rho}{1-\rho^2} > \underline{s} \\ \underline{s} & \text{if } \underline{\delta}-\rho \leq 0 \;\;\&\;\; \dfrac{\overline{\delta}-\rho}{1-\rho^2} < \underline{s} \\ \underline{s} - \dfrac{\underline{\delta}-\rho}{1-\rho^2} & \text{if } \underline{\delta}-\rho > 0 \;\;\&\;\; \dfrac{\overline{\delta}-\rho}{1-\rho^2} < \underline{s} \end{cases}.$$

The change in research output, $\Delta r = r(\overline{\delta},\underline{s}) - r(\underline{\delta},0)$, is:

$$\Delta r = \begin{cases} -\rho\dfrac{\Delta\delta}{1-\rho^2} & \text{if } \underline{\delta}-\rho > 0 \;\;\&\;\; \dfrac{\overline{\delta}-\rho}{1-\rho^2} > \underline{s} \\ \dfrac{1-\overline{\delta}\rho}{1-\rho^2} - 1 & \text{if } \underline{\delta}-\rho \leq 0 \;\;\&\;\; \dfrac{\overline{\delta}-\rho}{1-\rho^2} > \underline{s} \\ -\rho\underline{s} & \text{if } \underline{\delta}-\rho \leq 0 \;\;\&\;\; \dfrac{\overline{\delta}-\rho}{1-\rho^2} < \underline{s} \\ 1-\rho\underline{s} - \dfrac{1-\underline{\delta}\rho}{1-\rho^2} & \text{if } \underline{\delta}-\rho > 0 \;\;\&\;\; \dfrac{\overline{\delta}-\rho}{1-\rho^2} < \underline{s} \end{cases}.$$

Inspections of the first and the third cases immediately imply that $\Delta r < 0$ if $\rho > 0$ and $\Delta r > 0$ if $\rho < 0$. For the second case, notice that $\Delta r = \dfrac{1-\overline{\delta}\rho}{1-\rho^2} - 1 = \dfrac{(\rho-\overline{\delta})\rho}{1-\rho^2}$ and that this happens when $\dfrac{\overline{\delta}-\rho}{1-\rho^2} > \underline{s}$. The latter condition implies that $\rho - \overline{\delta} < 0$. Thus, we again have $\Delta r < 0$ if and only if $\rho > 0$. For the last case notice that this occurs when $\underline{s} \geq \dfrac{\overline{\delta}-\rho}{1-\rho^2}$. This implies that if $\rho > 0$

$$\Delta r = 1-\rho\underline{s} - \frac{1-\underline{\delta}\rho}{1-\rho^2} \leq 1-\rho(\frac{\overline{\delta}-\rho}{1-\rho^2}) - \frac{1-\underline{\delta}\rho}{1-\rho^2} = -\rho\frac{\Delta\delta}{1-\rho^2} < 0.$$

Conversely, when $\rho < 0$, we have that

$$\Delta r = 1-\rho\underline{s} - \frac{1-\underline{\delta}\rho}{1-\rho^2} \geq 1-\rho(\frac{\overline{\delta}-\rho}{1-\rho^2}) - \frac{1-\underline{\delta}\rho}{1-\rho^2} = -\rho\frac{\Delta\delta}{1-\rho^2} > 0.$$

A.2 Unbundling the policy interventions

Building on the results from Appendix A1, the PI maximization problem gives the following first-order conditions:

$$r = 1 - \rho s$$
$$s = \delta - \rho r.$$

With interior solutions, the optimal effort levels are:

$$r^* = \frac{1 - \delta\rho}{1 - \rho^2}$$
$$s^* = \frac{\delta - \rho}{1 - \rho^2}.$$

Notice that it is always the case that $r^* > 0$ as $1 \geq \delta\rho$.

We have $s^* > 0$, as long as $\rho < \delta$. If $\delta - \rho < 0$, we have that $s^* = 0$ and $r^* = 1$. In this case, a marginal increase in δ has no impact on research and safety. In the analysis below we focus on the case of an interior solution; that is, when $\rho < \delta$.

A.2.1 An increase in liability

We interpret the parameter δ, which is the marginal benefit of safety, as the expected liability. If $s^* > 0$ the effect of increasing δ is:

$$\frac{dr^*}{d\delta} = \frac{d}{d\delta}\left(\frac{1 - \delta\rho}{1 - \rho^2}\right) = -\frac{\rho}{1 - \rho^2} \gtrless 0$$
$$\frac{ds^*}{d\delta} = \frac{d}{d\delta}\left(\frac{\delta - \rho}{1 - \rho^2}\right) = \frac{1}{1 - \rho^2} > 0.$$

The impact on r^* is positive when $\rho < 0$ and negative when $\rho > 0$. Moreover, given our assumption that $|\rho| < 1$, the magnitude of the change in r is always lower than the magnitude of change in s. This makes sense as the change in δ has a direct effect on the marginal benefit of s but an indirect effect on r only through its relationship with s.

A.2.2 An increase in minimum safety standards

Notice that when safety mandates are not binding, $\underline{s} < s^* = \dfrac{\delta - \rho}{1 - \rho^2}$, a marginal increase in $\underline{s}$ won't have an effect. Thus, we focus on the binding case, in which

$$r^* = 1 - \rho \underline{s}$$

$$s^* = \underline{s}$$

In this case,

$$\frac{dr^*}{d\underline{s}} = -\rho \gtrless 0$$
$$\frac{ds^*}{d\underline{s}} = 1 > 0.$$

As with the previous policy, the directional impact on r^* depends on the sign of ρ and that the change in r is smaller in magnitude than that in s. Moreover, comparing $\dfrac{ds^*}{d\underline{s}}$ in this section and $\dfrac{ds^*}{d\delta}$ in the previous section shows that the magnitude of the change in s (as well as in r) due to a marginal change in minimum safety standards is smaller than that of an increase in liability. This is because the increase in δ is amplified by feedback effects through a change in r. The case of mandated $\underline{s}$ does not have this feedback effect.

A.2.3 A reduction in the cost of safety

In the baseline model, the marginal costs for s and r are symmetric. Here, we add a ε term to the marginal cost of s as follows:

$$U(r, s) = r - \delta(1 - s) - \frac{r^2}{2} - (1 - \varepsilon)\frac{s^2}{2} - \rho r s.$$

We examine what happens if ε marginally increases from zero to $\varepsilon > 0$. The optimal investment levels are then

$$r^* = \frac{1 - \varepsilon - \delta\rho}{1 - \varepsilon - \rho^2}$$
$$s^* = \frac{\delta - \rho}{1 - \varepsilon - \rho^2},$$

which lead to the following comparative statics (evaluated at $\varepsilon = 0$):

$$\frac{dr^*}{d\varepsilon} = -\rho \frac{\delta - \rho}{(\rho^2 - 1)^2} \gtrless 0$$
$$\frac{ds^*}{d\varepsilon} = \frac{\delta - \rho}{(\rho^2 - 1)^2} > 0.$$

Similar to the previous two policies, the directional impact on r^* depends on the sign of ρ and that the change in r is smaller in magnitude than that in s.

To compare this policy to the other policies, we consider two different cases, depending on the sign of ρ. Notice that the previous section shows that regardless of the sign of ρ, the magnitude of changes due to a marginal change in minimum safety standards is smaller than that resulting from a marginal change in liability. Thus, in the following, we focus on comparing only liability change to a change in the cost of safety.

When $\rho > 0$, we have:

$$\frac{\delta - \rho}{(\rho^2 - 1)^2} < \frac{1}{1 - \rho^2},$$

because this can be re-written as $\delta - \rho < 1 - \rho^2$, which is satisfied because $\delta < 1$ and $\rho > \rho^2$. Comparing $\frac{ds^*}{d\delta}$ in this section and $\frac{ds^*}{d\varepsilon}$ in Section A2.1 shows that the change in safety due to a marginal decrease in the cost of safety is smaller than that due to a marginal increase in liability.

By contrast, when $\rho < 0$, it is possible that $\delta - \rho > 1 - \rho^2$. Thus, if $|\rho|$ is sufficiently large, acting on the cost side may be more effective than acting on the benefit side.

A.3 Heterogeneous hazard levels

To capture different hazard levels of different labs, we assume that $\delta = pL$, where p is the expected risk of an accident and L is the liability cost in the case of an accident. As in the baseline model, we normalize the minimum required safety level to zero before the accident and indicate the requirement after the accident as $\underline{s} > 0$. We capture the increase in perceived liability with an increase in the parameter L. Specifically, we assume that $L = \underline{L}$ before the accident, and $L = \overline{L}$ after the shock. Without loss of generality, we assume a continuum of labs, each characterized by a hazard level p, distributed over the interval $[\underline{p}, \overline{p}]$.

Let's first consider the case in which $\underline{s}$ is not binding for any p. As such, we have

$$\Delta r = -\rho \frac{p(\overline{L} - \underline{L})}{1-\rho^2}.$$

Thus, a greater L implies a greater safety investment by the labs with a higher p. This also translates to greater changes in research investment for these labs.

When $\underline{s}$ is binding for all labs, however, we have:

$$\Delta r = 1 - \rho \underline{s} - \frac{1 - p\underline{L}\rho}{1-\rho^2}$$

which implies that the magnitude of the change in research actually decreases in p. This is intuitive as the binding safety standard pushes all labs to invest at $1 - \rho\underline{s}$. When $\rho > 0$, the lower p is, the higher is the pre-shock research investment. This implies that the drop will be larger for less risky labs, because they were doing more research and less safety investment before the shock. The intuition is similar for the case in which $\rho < 0$. In principle, $\underline{s}$ may be binding for some labs but not for others, and this can generate a non-monotonicity in p for the effect of the shock.

A.4 Change in research direction

One simple but intuitive way to examine potential changes in research type is to give a lab with underlying hazard p the option to reduce its hazard level to $p' < p$ after the shock by choosing a different research project. We assume that redirecting research toward a new project costs c.

Indicate the maximized utility of a PI with hazard level p and endogenously chosen r and s as $U(p, L, \underline{s})$, where L indicates the liability level and $\underline{s}$ the level of mandated safety. We make two assumptions. First, $U(\overline{p}, \underline{L}, 0) - U(p', \underline{L}, 0) < c$. This assumption means that liability risk and minimum safety requirement before the shock are sufficiently low such that even the riskiest lab does not find it worthwhile to switch. This assumption simplifies the analysis as it allows us to take the initial distribution of p as exogenous. We will discuss the implications of relaxing this assumption later. The second assumption is that $U(p', \overline{L}, 0) - U(\overline{p}, \overline{L}, 0) > c$. This assumption implies that the riskiest lab would prefer to switch to a safer project with the post-shock level of liability risk, even in the absence of mandated safety. This assumption guarantees that the shock is large enough to have some effect.

In the following, we analyze two separate cases, depending on whether the post-shock mandated safety level, $\underline{s}$, is binding.

When $\underline{s}$ is binding, Appendix A1 shows that $r^* = 1 - \rho\underline{s}$ and $s^* = \underline{s}$. This gives the following utility to the PI:

$$U(p, \overline{L}, \underline{s}) = 1 - \rho\underline{s} - p\overline{L}(1 - \underline{s}) - \frac{\underline{s}^2}{2} - \frac{(1 - \rho\underline{s})^2}{2} - \rho r\underline{s}.$$

To make redirecting research fro p to p' profitable, we need to have $U(p', \overline{L}, \underline{s}) - U(p, \overline{L}, \underline{s}) = (p - p')\overline{L}(1 - \underline{s}) > c$. This implies that re-directing the research project is profitable only if p is above a threshold:

$$p > \frac{c}{\overline{L}(1 - \underline{s})} + p'.$$

When $\underline{s}$ is not binding, the utility of a PI with risk level p is

$$U(p, \overline{L}, \underline{s}) = U(p, \overline{L}, 0) = \frac{(1 - p\overline{L})^2}{2(1 - \rho^2)} - \frac{p\overline{L}\rho}{(1 + \rho)}.$$

Notice that $U(p, \overline{L}, 0)$ decreases with p. Thus, $U(p', \overline{L}, 0) - U(p, \overline{L}, 0)$ is an increasing function of p. Combined with the assumption that $U(p', \overline{L}, 0) - U(\overline{p}, \overline{L}, 0) > c$ for the highest possible p, there exists a threshold in p above which it is worthwhile for the PI to switch to the new project.

The analysis above provides a simple illustration of how the shock may induce high-hazard labs to redirect research toward safer research projects. This was conducted under the assumption of no-redirection before the shock, which is equivalent to assuming that a particular PI's p before the shock was exogenously given. To properly investigate a PI's decision of research direction and the impact of a shock of liability and safety regulation, we need to provide a substantive micro-foundation of why a researcher chooses to work on projects of a certain risk level in the first place. There may be multiple factors that influence a researcher's choice, including the ability to manage risk, the intensity of competition in a certain research area, a taste for risk-taking, and how much the PI cares about the potential harm to the lab's researchers. The impact of a liability and safety regulation shock is potentially different, depending on which motivation is most salient. These topics, while interesting, are outside the scope of this paper.

Appendix B: Textual Similarity Measure

The method we use is called the Bidirectional Encoder Representations from Transformers (BERT) language models published in 2018 by Jacob Devlin and his colleagues at Google (Devlin et al., 2018). To the best of our knowledge, BERT models are among the state-of-the-art natural language processing methods used in management and economics research.

The simplest way to construct a similarity measure is to use the actual vocabulary terms used in a textual document, whereby each unique word in the entire corpus constitutes a dimension. A main shortcoming of these vocabulary-based matrices is that they are very sparse. More importantly, such methods do not account for relationships between words. Terms related to one another (e.g., "cars" and "automobiles") are not treated as more similar than words that are not (e.g., "cars" and "dogs"). Recent word-embedding techniques such as word2vec and BERT models address both problems. They reduce the dimensionality of the matrices and consider the relationships between words. For example, BERT models convert texts into 768-dimensional vectors. Compared to word2vec, which is a group of related earlier models, BERT models also factor in the semantic context of each word. For example, BERT models will yield different vectors for the word "bank" as a financial institution from the bank of a river, whereas word2vec models will produce the same vector.

Specifically, we used a pre-trained BERT model to construct our similarity measure in three steps:

1. We first convert each paper's abstract into a 768-dimension vector.
2. For each lab, we compute a benchmark vector as the average of the vectors of all the papers this lab published before 2008. Intuitively, this average vector characterizes the core research the lab produced before the accident (Whalen, et al., 2020).
3. Then, for each of the lab's publications, we construct the cosine similarity measure between the focal article vector and the benchmark vector. The higher the similarity score, the more similar the focal paper is compared to the same lab's pre-shock core research. Note that this measure is defined for all articles both before and after the shock.

Validation of the measure

We use both a case study and a systematic approach to get a sense of how well the measure works.

Take the publications of Professor Michael Jung of UCLA for example. Jung specializes in synthetic organic chemistry. Manually checking his publications before 2008, we noticed that 'C1-C-11 FRAGMENT' is the most frequent keyword that WoS associates to his articles. (Keywords are a separate field from abstracts that we use to construct the above measure.) C1-C-11 is a compound that could be synthesized for anti-tumor purposes and is listed as a keyword for four articles Jung published before the shock.

Among Jung's publications after 2008, we find that articles with the highest similarity measures with respect to his pre-shock average vector also list 'C1-C-11 FRAGMENT' as a keyword. An example is the article "Selectivity in Non-Aldol Aldol Rearrangements of Cyclic Epoxides" published in *Organic Letters* in 2011 with a similarity measure of 0.86. Note that the title of an article is also a different field from the keywords. By contrast, the articles with the lowest similarity measures not only do not refer to "C1-C-11 FRAGMENT," but they also have no keywords in common with any of the pre-2008 publications. An example is "Broad-spectrum antiviral JL122 blocks infection and inhibits transmission of aquatic rhabdoviruses" published in *Virology* in 2018 with a similarity score of 0.55.

We also provide a more systematic validation of the measure using regression analysis. We construct a dummy that equals one if the article refers to a compound that is not previously referenced by the lab. Among all the articles published by UCLA wet labs in our sample, the raw correlation between the dummy variable and our similarity measure is -0.004 (the standard error is 0.002). This indicates that articles exploring new compounds have a lower similarity score with respect to the lab's past research than articles that do not. The correlation remains robust (and significant at the 5% level) when adding additional controls to the regression. The negative relationship is also robust to the inclusion of year and PI fixed effects in addition to the control variables.

Table A1 - Wet and dry lab comparison

Panel A: UC sample	Wet Lab	Dry Lab	P-value
Articles per year before 2008	6.172	6.307	0.843
Highly cited articles per year before 2008	0.861	1.023	0.311
Year joined UC	1997.543	2000.863	0.017
Panel B: UCLA sample	**Wet Lab**	**Dry Lab**	**P-value**
Articles per year before 2008	5.647	5.775	0.928
Highly cited articles per year before 2008	0.919	0.628	0.502
Year joined UC	1995.758	1998.429	0.452

NOTES: Unit of observation is a lab. Panels A and B report summary statistics for UC and UCLA samples, respectively. Articles per year before 2008 = the average number of articles the lab published each year before 2008. Highly cited articles per year before 2008 = the average number of articles the lab published in the top decile of citations per year before 2008. Year joined UC = the year the lab is established within the UC system. Wet Lab = 1 if the lab conducts experiments using chemical and biological substances.

Table A2: Laboratory safety and publication levels -- Robustness to alternative econometric models and tenure controls

	(1)	(2)	(3)	(4)	(5)
Dep. Var.	Articles	Citation weighted articles	Articles	Articles	Articles
Model	Poisson	Poisson	OLS	Weighted-OLS	OLS
Wet Lab × After Accident	-0.053	-0.176	-0.382	-0.464	-0.182
	(0.061)	(0.207)	(0.619)	(0.685)	(0.500)
Sample	UC labs	UC labs	UC labs active between 2004 and 2017	UC labs	UC labs
Year effects	YES	YES	YES	YES	YES
Lab effects	YES	YES	YES	YES	YES
Tenure control	UC	UC	UC	UC	FULL
Observations	6818	6818	4564	6234	6827

NOTES: Articles = the number of articles the lab published in year t. Citation weighted papers = the number of articles weighted by citations received as of 2020 the lab published in year t. Wet Lab = 1 if the lab conducts experiments using chemical and biological substances. After Accident = 1 if after year 2008. All regressions control for the total publications by the lab in the previous three years. Columns 1-4 control for the logarithm of the lab's tenure and column 5 controls for the lab's full tenure since the PI's first publication as a PI. Robust standard errors clustered at the lab level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: Laboratory safety and publication levels -- Robustness to alternative controls and samples

	(1)	(2)	(3)	(4)
Dep. Var.	Articles	Articles	Articles	Articles
Wet Lab × After Accident	-0.208	-0.244		
	(0.558)	(0.486)		
Endowment	0.000**			
	(0.000)			
Chem PhDs	0.008			
	(0.008)			
Science & Engineering Grants	-0.000			
	(0.000)			
Journal Based Wet Lab × After Accident			0.131	-0.043
			(0.540)	(0.087)
Sample	UC labs with available data	UC labs	UC labs	Non-UC US academic labs
Year effects	YES	YES	YES	YES
Lab effects	YES	YES	YES	YES
Institution-Year Effects	NO	YES	NO	NO
Observations	5553	6827	6827	38763

NOTES: OLS regressions. Articles = the number of articles the lab published in year t. Wet Lab = 1 if the lab conducts experiments using chemical and biological substances. After Accident = 1 if after year 2008. Endowment = total market value of the endowment held by the institution in year t (in 1,000s USD). Chem PhDs = the number of individuals receiving a research doctorate in the field of chemistry in the institution in year t. Science & Engineering Grants = the amount of federal science and engineering (S&E) funding received by the institution in year t (in 1,000s USD). Journal Based Wet Lab defines wet versus dry labs using classification of journals and publications between 2004 and 2008. All regressions control for the total publications by the lab in the previous three years and the logarithm of the lab's tenure. In Column 4, the DV is constructed using the subset of WoS journals specialized in chemistry and ten multidisciplinary science journals with the highest impact factors. Robust standard errors clustered at the lab level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: Laboratory safety and publication levels -- Robustness using SciFinder data

	(1)	(2)	(3)
Dep. Var.	Articles	SciFinder articles	SciFinder articles
Model	OLS	OLS	Poisson
Wet Lab × After Accident	-0.002	-0.039	-0.035
	(1.048)	(1.172)	(0.087)
Sample	UCLA labs	UCLA labs	UCLA labs
Year effects	YES	YES	YES
Lab effects	YES	YES	YES
Observations	976	976	975

NOTES: Articles = the number of articles the lab published in year t. SciFinder articles = the number of articles recorded by the SciFinder database the lab published in year t. Wet Lab = 1 if the lab conducts experiments using chemical and biological substances. After Accident = 1 if after year 2008. All regressions control for the total publications by the lab in the past three years and the logarithm of the number of years that the lab has existed as of year t. Robust standard errors clustered at the lab level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: Lab research direction after the shock -- Robustness checks

	(1)	(2)	(3)	(4)
Dep. Var.	Minimum semantic similarity score	Median semantic similarity score	Minimum semantic similarity score	Median semantic similarity score
High Hazard × After Accident	0.050**	0.043**	0.083**	0.090***
	(0.022)	(0.020)	(0.033)	(0.032)
Publications sample	Articles with fraction of dangerous compounds referenced > 50%	Articles with fraction of dangerous compounds referenced > 50%	Articles with fraction of dangerous compounds referenced > 75%	Articles with fraction of dangerous compounds referenced > 75%
Lab Sample	UCLA wet labs active between 2004 and 2017	UCLA wet labs active between 2004 and 2017	UCLA wet labs active between 2004 and 2017	UCLA wet labs active between 2004 and 2017
Year effects	YES	YES	YES	YES
Lab effects	YES	YES	YES	YES
Observations	144	144	75	75

NOTES: OLS regressions. Minimum semantic similarity score = minimum similarity score of all the articles published in a focal year by a lab relative to the same lab's pre-shock core research. High Use = 1 if lab in top quintile in terms of articles published using dangerous substances. Median semantic similarity score = median similarity score of all the articles published in a focal year by a lab relative to the same lab's pre-shock core research. High Hazard = 1 if lab in top quintile in terms of articles published using dangerous substances. After Accident = 1 if after year 2008. All regressions control for the total publications by the lab in the previous three years and the logarithm of the lab's tenure. Robust standard errors clustered at the lab level in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01