INSTITUTE
OF ECONOMICS

Scuola Superiore
Sant'Anna

LEM | Laboratory of Economics and Management

Institute of Economics
Scuola Superiore Sant'Anna

Piazza Martiri della Libertà, 33 - 56127 Pisa, Italy
ph. +39 050 88.33.43
institute.economics@sssup.it

LEM
WORKING PAPER SERIES

Risk, Precaution, and Regulation in Chemical Search and Innovation: The Case of the EU REACH Legislation

Gianluca Biggi [a]

[a] Institute of Economics, Scuola Superiore Sant'Anna, Pisa, Italy.

2023/06 **January 2023**
ISSN(ONLINE) 2284-0400

Risk, Precaution, and Regulation in Chemical Search and Innovation: The Case of the EU REACH Legislation

Gianluca Biggi
Institute of Economics, Scuola Superiore Sant'Anna, Pisa, Italy
gianluca.biggi@santannapisa.it

Abstract: This study investigates the impact of the introduction of the European chemical regulation (the EU REACH legislation) on chemical search and innovation by focusing on the knowledge recombination processes leading to the generation of inventions. Using a novel dataset of patents and chemical structures contained therein over the period 1978-2016, this study readapts established patent indicators to capture the complexity, novelty, and novelty in recombination of the inventive activities as a result of the chemical regulation. The separate effect of the chemical regulation reflected in +39.8% of compounds per patent, +23% of new compounds per patent, and +2% of newer recombinations of compounds per patent is supported by the Propensity Score Matching estimations. The positive and significant effect of chemical regulation on compound patenting supports prior scholarly work on the idea that regulations by altering the search space, influence the rate and intensity of technological search and innovation.

Keywords: Chemical inventions, Patent data, Regulation, Knowledge recombination.

1. Introduction

Industrialized economies have historically maintained the hope that advances in science, technology and innovation would have offered to humanity a wide range of options to improve its well-being and attain sustained economic growth. The transition towards a high technological frontier arising from the rapid advances of science, technology and innovation have opened a debate on the relations between innovative activities, related possible social and environmental threats and the role of regulations to keep up with industry developments. While a lot of scholarly and policy attention has been devoted to the undisputed role that technology and innovation have played in enhancing economic competition, growth, wealth creation, productivity, and efficiency at firm, industry, and national levels, little attention has been paid to the related social and environmental downsides (Coad et al. 2021; Biggi and Giuliani 2021). It is recognized that human and environmental conditions are deteriorating, in many cases as a result of innovation-induced industrial activities: for example, the DuPont dumping of perfluorooctanoic acid (PFOA), linked to kidney cancer, testicular cancer, ulcerative colitis into the Ohio River near its West Virginia plant for about 60 years (Judge 2016); the continuing groundwater contamination as a result of the world's worst chemical industrial disaster at the Union Carbide pesticide plant in Bhopal (Mandavilli 2018); the alleged cases of cancer linked to the exposure of the Monsanto's Roundup, the blockbuster herbicide classified in 2015 as "probably carcinogenic to humans" by the International Agency for Research on Cancer (Cressey 2015); and so on.

This study seeks to provide new evidence on the nature and the direction of inventive activities by focusing on the knowledge recombination processes in response to regulatory actions aimed at steering industrial developments away from harmful products and technologies (Coad, Biggi, and Giuliani 2021). This study readapts established indicators to capture the complexity, novelty, and novelty in recombination of the inventive activities as a result of the chemical regulation. Complexity captures the variety of knowledge components embedded in technology, novelty represents the technological distance from existing knowledge and novelty in recombination is the uniqueness of the recombination process.

The chemical industry has attracted considerable scholarly and policy attention in addressing the global challenges associated with human health and environmental degradation as a variety of chemical technologies can be linked to the emergence of numerous contemporary illnesses such as cancer, respiratory diseases, damage to natural ecosystems (see e.g., Soto and Sonnenschein 2010; Lakey et al. 2016; Ögmundarson et al. 2020). Also, the persistence of some of the most hazardous pollutants as DDT or PCBs resulting in the migration

through the food chain, bioaccumulation, and remobilization into the air creates room for the study of a global problem (Nizzetto et al. 2010; Rigét et al. 2019).

This study considers the impact of the introduction of the Substances of Very High Concern list (hereinafter SVHC list) as a part of the 2007 European REACH legislation[1] on chemical search and innovation. Besides regulating the production and use of all chemicals within the EU, the REACH legislation, continuously update a list of highly hazardous chemicals (i.e., the SVHC list) meeting the criteria for classification as a) carcinogenic, mutagenic, or toxic for reproduction, or 2) persistent, bioaccumulative and toxic or, 3) very persistent and very bioaccumulative. Once the chemical is included in the SVHC list, it cannot be produced or used within the EU market, unless authorization is granted for their specific use, or the use is exempted from authorisation. The EU REACH is considered the strictest law to date regulating chemicals with major implications for the chemical industry (Cone 2006). The case of the REACH legislation is particularly suitable for this study as it reflects recent evolutions in regulatory thinking aimed at steering responsibility in the innovation process (Owen 2009; Owen et al. 2013; Stilgoe, Owen, and Macnaghten 2013). More specifically, the REACH legislation entails preventive action in the face of risk and uncertainty, a shift in the burden of proof or responsibility onto proponents of potentially harmful technologies and promotes the exploration of a wide range of alternatives to possibly harmful technologies (Raffensperger, Schettler, and Myers 2000; Kriebel et al. 2001; Martuzzi and Bertollini 2005).

Using a novel dataset of chemical patents, this study finds that chemical regulation as the REACH legislation is correlated with an increase in the number of chemical components disclosed in patented inventions (complexity), in the number of newly invented compounds (novelty) and the number of newer recombination of chemical components (novelty in recombination). These findings highlight the role of regulatory actions as catalysts in knowledge recombination, steering distant search from regulated and harmful technologies by eliciting knowledge recombination processes.

The remainder of this paper is structured as follows. Section 2 reviews the relevant literature and formulates the research questions. Section 3 provides an overview of chemical patents and introduces the indicators based on patent chemistry for the empirical analysis conducted in section 4. Section 5 presents the results and Section 6 concludes.

[1] Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) is a European Union chemical legislation entered into force in 2007. Article 57 of the REACH sets the criteria for the inclusion of chemicals in the SVHC list.

2. Literature review

Inventive activities are driven by a process of knowledge recombination (Weitzman 1998; Arthur 2007; Schumpeter 2017). Knowledge recombination may derive from the exploration of new knowledge or from the exploitation of prior knowledge (March 1991). The exploration of new technological knowledge is typically costly and risky, and the balance between the two knowledge recombination modes depends on the relative costs of exploration and exploitation and on the ability to apply prior knowledge towards future inventions. The idea of induced innovation (Hicks 1963) recognizes that inventive activities are aimed at developing novelty of economic value (Nelson 2009) and that the knowledge recombination modes correspond positively to the decrease in their relative costs. Also, the characteristics of the technological search space influence the results of knowledge recombination (Fleming 2001; Fleming and Sorenson 2004). In this view, regulatory actions promoted by governments or supra-national institutions can alter the search space and hence the knowledge recombination process.

The recent evolution of regulatory thinking reflected the need to steer responsibly the inventive activities away from harmful products and technologies (Owen et al. 2013; (Owen et al. 2013, Stilgoe, Owen, and Macnaghten 2013). A poignant example of this is the requirements for submission of chemical safety assessments under the REACH legislation in Europe. These require environmental and human health and safety "data before the market" for chemical substances. This change of responsibility in providing information on the properties of chemicals is substantial compared with the current worldwide chemical regulations as it entails a shift of the burden of proof towards the industry side. This shift in the burden of responsibility in the REACH legislation also reflects a new culture of responsibility guided by precautionary principles (Stirling 2016).

Scholars indicate regulatory actions can spur the ability to search and innovate, by lowering the relative cost of exploration (Jaffe and Palmer 1997). That regulation can spur distant search as an unprecedented recombination of technological components influencing the degree of novelty of inventions (Fleming 2001). Thus, regulations are theorized as allowing firms or other inventive entities to enlarge their knowledge recombination processes. New technological knowledge occurs when new information is integrated and/or recombined with existing knowledge of the problem giving rise to breakthrough ideas and innovations (Schilling and Green 2011). In the chemical sector, regulatory actions can trigger the transformation process of existing chemical compounds.

In that context, this study argues that after the introduction of the regulation, companies and other inventing entities will take more complex and more novel knowledge recombination

paths as a result of exploratory searches. The general rationale for this is that the existing technological knowledge becomes a dangerous terrain for new investments, especially if the new technological developments draw – even if partially – on the regulated technological knowledge. Hence, in a scenario of heightened uncertainty and risk following a regulation, the changes in the nature of chemical discovery provide a unique setting to study whether firms or other inventive organizations will take an exploratory search in terms of more complex and more novel recombinations of technological knowledge (Rosenkopf and Nerkar 2001). Based on the above premises, this study formulates the following research questions.

RQ1. Do chemical inventions represent a more complex recombination of technological knowledge as a result of the regulation?
RQ2. Do chemical inventions entail more novel recombinations of technological knowledge as a result of the regulation?

This study assumes that companies and other organizations conducting research around the regulated technologies are not malevolent entities and will therefore have agreed on the threats posed by their past discoveries and will seek to find ways to address them in the future, to stop or minimize the noxious impacts on humans and the environment. The direction of their search and innovation efforts after the regulation is also assumed to be the result of calculated strategic decisions in the face of a changing regulatory environment; such decisions are expected to be taken to both address social and environmental threats and to hedge against future losses in the case of more stringent regulatory measures.

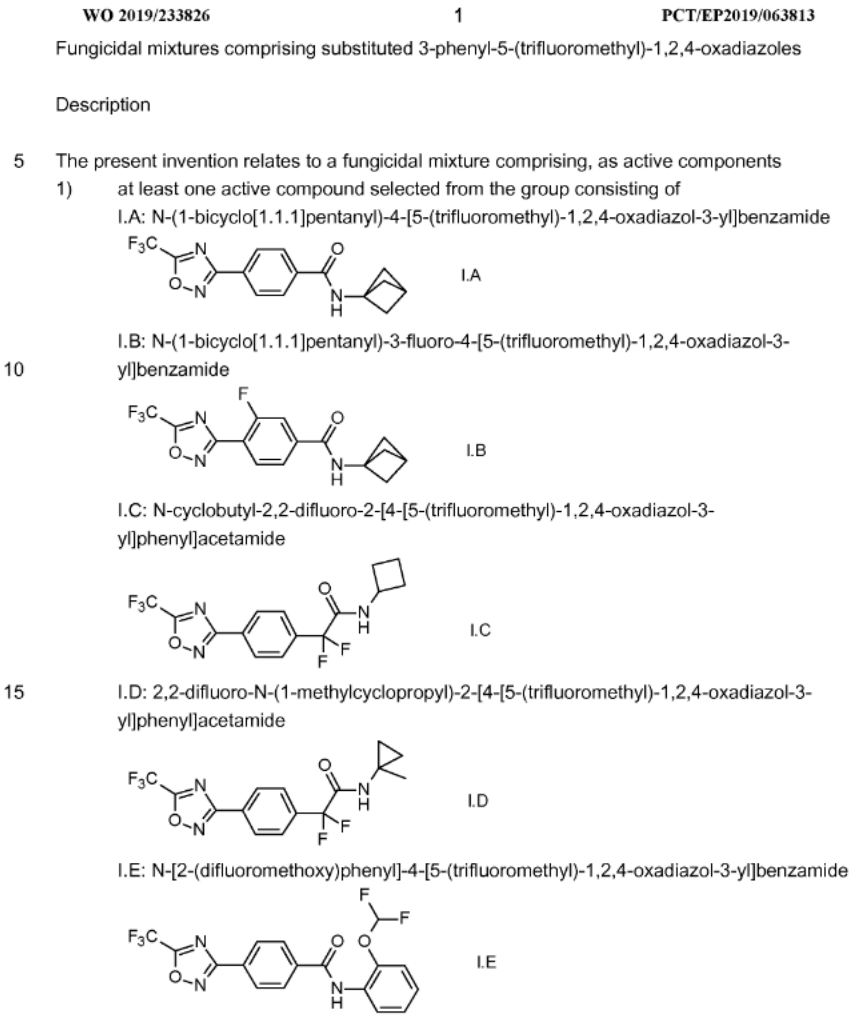
3. Chemical patenting

This study conducts an empirical analysis based on patent data.[2] In innovation studies, management and economic research, patent data are widely used and adopted measures of innovation and have been traditionally used to capture information on the knowledge base and knowledge components of inventions (Jaffe 1986; Jaffe, Trajtenberg, and Henderson 1993). Along these lines, this study relies on a novel dataset of chemical patents to capture the impact of chemical regulation on chemical search and innovation. Chemical patents have the unique feature of claiming a list of all the possible chemical components (i.e., chemical compounds)

[2] See, e.g., Griliches (1990), Lanjouw et al. (1998), Arts et al. (2013) for a discussion on the possible drawbacks of empirical analyses based on patent data.

by way of standardized chemical structures that represent the sought of patent protection.[3] Figure 1 shows the example of patent number WO2019233826 claiming a range of chemical compounds composing a fungicidal mixture.

Figure 1: Example of patent number WO2019233826



WO 2019/233826 1 PCT/EP2019/063813

Fungicidal mixtures comprising substituted 3-phenyl-5-(trifluoromethyl)-1,2,4-oxadiazoles

Description

The present invention relates to a fungicidal mixture comprising, as active components

1) at least one active compound selected from the group consisting of

I.A: N-(1-bicyclo[1.1.1]pentanyl)-4-[5-(trifluoromethyl)-1,2,4-oxadiazol-3-yl]benzamide

I.A

I.B: N-(1-bicyclo[1.1.1]pentanyl)-3-fluoro-4-[5-(trifluoromethyl)-1,2,4-oxadiazol-3-yl]benzamide

I.B

I.C: N-cyclobutyl-2,2-difluoro-2-[4-[5-(trifluoromethyl)-1,2,4-oxadiazol-3-yl]phenyl]acetamide

I.C

I.D: 2,2-difluoro-N-(1-methylcyclopropyl)-2-[4-[5-(trifluoromethyl)-1,2,4-oxadiazol-3-yl]phenyl]acetamide

I.D

I.E: N-[2-(difluoromethoxy)phenyl]-4-[5-(trifluoromethyl)-1,2,4-oxadiazol-3-yl]benzamide

I.E

Typically, chemical patents try to capture a broad spectrum of chemical substances sometimes throughout multiple, often many hundreds, combinations of chemical compounds. In the chemical industry, patent applications are considered the starting point from which firms or other inventing entities begin to claim intellectual property rights by planting "flags" in chemical space (Southall and Ajay 2006). The chemical substance of interest is therefore

[3] For a review of exisiting studies exploiting the chemical content of patents see e.g., Jayaraj and Gittelman 2018; Biggi, Giuliani, and Martinelli 2020; Krieger, Li, and Papanikolaou 2021; Biggi et al. 2022.

hidden among hundreds of combinations of chemical compounds. Interestingly, chemical patents possess the unique feature of protecting the chemical substance as a result of a combination of known chemical compounds, new chemical compounds for the state of the art or new combinations of chemical compounds. Thus, chemical compounds claimed in patented technologies represent the explorative effort undertaken by firms or other organizations within the chemical space. These unique properties of chemical patents make them extremely useful for analysing the evolution of technological search and innovation and knowledge recombination processes.

3.1. Indicators based on patent chemistry

3.1.1. Complexity

Complexity captures the variety of the knowledge components embedded in technology. Most extant research measure complexity by relying on two main patent indicators: patent scope and patent originality. The patent scope is measured by looking at the distinct number of International Patent Classification (IPC) classes[4] associated with the patented invention. IPC classes link the patented invention to a specific technological field and thus, the variety of IPC classes associated with the patented invention capture its complexity (see e.g., Lerner 1994; Shane 2001). The originality index developed by Trajtenberg, Henderson, and Jaffe (1997) (see e.g., Hall, Jaffe, and Trajtenberg 2001; Hicks and Hedge 2005) is based on the patent's backward citations[5] and captures the different technological sources to which the patent belongs. This study readapts standard patent complexity measures by exploiting the unique features of chemical patents counting the distinct number of chemical compounds in patent claims[6]. Each chemical compound represents a chemical entity with a unique chemical structure. These chemical compounds compose the invention sought in the patent document and substantiate patent claims. Hence, a set of compounds in a patent indicate the firm's search activity within the chemical space and the strategic intent to protect a portion of the "chemical space". The number of compounds encoded in chemical patents is then a factor of explorative search effort, dedicated resources and strategic intent of the firm or other organizations (Drews

[4] IPC classification refers to a hierarchical system in which the patented invention is assigned to a specific technological field. For example, the IPC class G06E refers to "optical computing devices".
[5] In analogy with scientific publications, backward citations are citations made by a patent to previously issued patents.
[6] Patent claims indicate the extent (i.e., the scope) of the protection sought in a patent application (Kuhn and Thompson 2019; Bekkers, Martinelli, and Tamagni 2020; Marco, Sarnoff, and deGrazia 2019).

2000; DiMasi et al. 2010). A lot of compounds in patent claims imply large search efforts (Jaffe and Trajtenberg 1999; Henderson and Cockburn 1996). Thus, the number of patented compounds provides a quantitative representation of the intensity of technological search efforts and can be proxied as a measure of complexity.

3.1.2. Novelty

Novelty represents the uniqueness of the knowledge base of the patented invention thus capturing the extent to which the patented technology draws on previous inventions. This study measures the novelty of a chemical patent by looking at the distinct number of chemical compounds in patent claims that have never been claimed in any patent up to the year of its filing.[7] Each compound listed in patent claims can be known to the state of the art (i.e., previously claimed along with other chemical compounds) or introduced for the first time in a chemical patent. In this case, the sought of intellectual property protection is to cover a brand-new region of the chemical space in which all the combinations (i.e., the new compounds along with other compounds) result to be new for the state of the art. Hence, a set of newly invented compounds in a patent indicate the inventive activity within the chemical space and the intent to protect a new region of chemical space. Thus, the number of newly invented patented compounds provides a fine-grained representation of the intensity of the firm's search efforts and can proxy technological novelty.

3.1.3. Novelty in recombination

The novelty in recombination captures the uniqueness of the knowledge recombination process (Verhoeven, Bakker, and Veugelers 2016). Extant research measure novelty in recombination by looking at the newer combination of IPC classes associated with the patented inventions. A patent can be considered novel in recombination if, among all the possible combinations of its IPC classes, there is at least one combination not observed in a previously issued patent. This study readapts the standard novelty in recombination measures by looking at the newer combinations of chemical compounds filed at the EPO. The measure takes the value 1 if the patent contains at least one new combination of chemical compounds up to the year of its filing, and 0 otherwise.

[7] Since the scope of the analysis is to test the effect of the EU REACH SVCH list, this study considers the universe of chemical compounds claimed in European Patent Office (EPO) patents.

4. Data and Methods

4.1 Data

This study relies on three main databases: the chemical database of the European Molecular Biology Laboratory SureChEMBL, the EPO PATSTAT (Autumn 2016 version), and the Bureau Van Dijk ORBIS-IP. SureChEMBL is a publicly available large-scale database (Papadatos et al. 2016) which has the unique feature of providing comprehensive compound-patent associations along with the exact location of the compound in the patent document (i.e., title, description, claims, prior art, etc.) for all the chemical patents filed at the European Patent Office (EPO), the United States Patent and Trademark Office (USPTO), the Japanese Patent Office (JPO) and the World Intellectual Property Organization (WIPO). Patented chemical compounds in the SureChEMBL can be searched through their standard identifiers as the SMILES or InChIKey allowing for a precise search of all the patents related to a specific chemical compound.[8] This study focused on the universe of EPO patents claiming one or more standard InChIKey identifiers associated with the 44 chemical compounds included in the SVHC list of the EU REACH legislation published between 2008 and 2011 (see Appendix A.1. for a complete list of chemical compounds, their standard identifiers, the year of inclusion in the list and the toxicological reasons for inclusion).[9] The SVCH list aims at strictly regulating the production, use, and trade of chemical technologies deemed as highly toxic for humans and ecosystems allowing the study of the role of chemical regulation as quasi-natural experiments. Following standard practice in the literature, this study exploits the PATSTAT database to group the identified patent applications at the EPO into standard DOCDB patent families to deal with multiple equivalents of the same invention (Hall and Helmers 2013). Bureau Van Dijk ORBIS-IP has been used to retrieve indicators for knowledge base dimensions of patent families as the number of backward citations, the number of inventors and the IPC classifications. This resulted in a dataset of 3,120 patent families containing at least one chemical compound included in the SVHC list among 156,611 claimed chemical compounds over the period 1978-2016.

[8] SMILES refers to the "Simplified Molecular Input Line Entry System" which translates a chemical compound's three-dimensional structure into a universal string of symbols. Similarly, InChIKey refers to a textual identifier for chemical compounds, designed to provide a standard way to encode compound information and to facilitate the search for such information in databases and on the web.

[9] It is to note that the SVHC list is continuously updated and as of January 2022 includes a total of 223 chemical compounds https://echa.europa.eu/it/candidate-list-table [Last accessed 3 January 2023]. This study focuses on the first version of the list including 44 SVHC chemical compounds as of 2011.

Table 1: Variable description and descriptive statistics

Variable	Variable description	Obs	Mean	Std. Dev.	Min	Max
Regulated	Dummy variable equal to 1 if the patent claims a chemical compound that, as of the filing date, had been included in the SVHC list, and 0 otherwise	3,120	0.307	0.461	0	1
Complexity	Number of distinct chemical compounds claimed in the patent	3,120	69.303	26.790	1	4,705
Novelty	Number of novel chemical compounds claimed in the patent	3,120	3.339	2.295	0	1,047
Novelty in recombination	Dummy variable equal to 1 if the patent claims a novel combination of chemical compounds and 0 otherwise	3,120	0.476	0.499	0	1
Family size	Number of DOCDB patent family members	3,120	9.725	7.326	1	93
Scope	Number of IPC full-digit codes	3,120	23.915	26.063	1	609
Backward citations	Count of backward citations	3,047	4.209	7.412	0	161
Inventor team size	Number of inventors	3,120	3.120	7.091	1	62

Table 1 provides descriptive statistics of the variables employed in the empirical analysis. It is to note that around 30% of the patent families in the sample contain at least one chemical compound that, as of the filing date, had been included in the SVHC list.

4.2 Methodology

To investigate the impact of the introduction of the SVHC list on chemical search and innovation across different dimensions, such as complexity, novelty, and novelty in recombination, this study relies on a propensity score matching (PSM) approach (Rosenbaum and Rubin 1983) to take into consideration the non-exogeneity of the regulation and related problems of an ordinary least square (OLS) approach. The PSM tries to obtain an estimate of the average treatment effect on the treated (*ATT*) of the introduction of the chemical regulation list as follows:

$$ATT = E\,(Y_1 - Y_0 \mid D = 1) = E\,(Y_1 \mid D = 1) - E\,(Y_0 \mid D = 1) \qquad (1)$$

Y_1 and Y_0 are the values of the outcome variable (Y) respectively in the presence and absence of chemical regulation. D is the status of the treatment: $D = 1$, patents containing at least one regulated chemical compound (treated); $D = 0$ patents non containing regulated chemical compounds (non-treated). $(Y_0 \mid D = 1)$ is non-observable and, therefore, needs to be substituted by referring to a suitable "counter-factual" of patents that contains a regulated chemical compound, but it has not yet been regulated at the date of filing. To control for selection bias and ensure that the difference in the outcome variables of the two groups is exclusively due to the introduction of the chemical regulation, treated patents are matched with non-treated patents based on the propensity scores, $P(X)$. This represents the probability of being treated, given a set of pre-treatment characteristics, X, which are supposed to affect both the treatment and the outcome. The PSM estimate the *ATT* as follows:

$$ATT_{PSM} = E_{P(X)|D=1}\,\{E[Y_1 \mid D = 1,\, P(X)] - E[Y_0 \mid D = 0,\, P(X)]\} \qquad (2)$$

Where $P(X)$ is estimated with a standard logit model. The PSM is implemented using a set of standard procedures, assumptions, and quality tests. A set of matching procedures are employed to assess the stability and reliability of the evidence. These procedures differ in the selection and weighting of the non-treated patents to be used as matches, as well as in the capacity to trade between efficiency and bias reduction (Becker and Ichino 2002; Smith and

Todd 2005; Caliendo and Kopeinig 2008). The common support condition is imposed in all the matching procedures, to guarantee the presence of suitable counterfactual patents for each treated.

5. Results

This section presents the results of the empirical analysis. First, Table 2 shows the logit estimation underlying the PSM estimates to test and quantify the differences between treated patents and non-treated patents by controlling for technological characteristics and other factors that might influence the three specified outcome variables (Model (1-3)). The results reported in Table 2 show that the differences between treated patents and non-treated patents persist along all the dimensions considered and controlling for geographical (*ASSIGNEE_COUNTRY DUMMIES*) and technological (*IPC_4Dig DUMMIES*) dummies. Treated patents belonging to large patent families (*FAMILY_SIZE*), that are broader in scope (*SCOPE_FULL-DIGIT)* with a large number of backward citations (*BACKWARD_CITATIONS*) and a large pool of inventors (*INVENTOR_TEAM_SIZE*) are more likely to be complex (Model (1)), Novel (Model (2)) and Novel in recombination (Model (3)). In other words, treated patents stem from a more dispersed search space, include more distinct knowledge components and are more valuable than non-treated patents.

Table 2: Estimation of the propensity score

	COMPLEXITY	NOVELTY	NOVELTY IN RECOMBINATION
	(1)	(2)	(3)
FAMILY_SIZE	0.050***	0.049***	0.023***
	(0.010)	(0.014)	(0.005)
SCOPE_FULL-DIGIT	0.020***	0.017***	0.007***
	(0.003)	(0.003)	(0.001)
BACKWARD_CITATIONS	0.039***	0.041***	0.018***
	(0.006)	(0.007)	(0.003)
INVENTOR_TEAM_SIZE	0.053***	0.048***	0.027***
	(0.007)	(0.008)	(0.005)
Constant	-1.396***	-1.226***	-0.724***
	(0.084)	(0.645)	(0.394)
Observations	3,047	2,907	2,907
IPC_4DIG DUMMIES	YES	YES	YES
ASSIGNEE_COUNTRY DUMMIES	YES	YES	YES
Pseudo R2	0.0306	0.0887	0.0865
Prob > $\chi 2$	0.0000	0.0000	0.0000

Robust standard errors in parentheses. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3 shows that the average treatment effect on treated (*ATT*) of the SVHC list is positive and significant in all the specified models. Those results are robust across the different matching procedures being used in this study. Compared to non-treated patents, treated patents are more complex (+39.8%), novel (+23%) and novel in recombination (4%) confirming the theoretical argument developed in Section 2 and its policy implications.

Table 3: Effect of the SVHC list on chemical search and innovation

	COMPLEXITY			NOVELTY			NOVELTY IN RECOMBINATION		
	5NN	Caliper	Kernel	5NN	Caliper	Kernel	5NN	Caliper	Kernel
AVERAGE TREATMENT EFFECT ON TREATED (*ATT*)	0.398***	0.398***	0.398***	0.230***	0.230***	0.230***	0.040***	0.040***	0.040***
	(15.825)	(15.825)	(15.825	(1.487)	(1.487)	(1.487)	(0.026)	(0.026)	(0.026)
N Treated on support	934	934	934	881	881	881	881	881	881
N Treated	937	937	937	880	880	880	880	880	880
N Untreated	2,110	2,110	2,110	2,026	2,026	2,026	2,026	2,026	2,026

Methods: five nearest neighbours (5NN), five nearest neigbhours with a 0.05 caliper (Caliper), Epanechnikov kernel matching (Kernel). Standard errors estimated with a 50-replication bootstrap procedure in parentheses. *** p<0.01, ** p<0.05, * p<0.1

6. Conclusions

Using a novel dataset of chemical patents, this study finds that chemical regulation as the SVHC list included in the EU REACH legislation is correlated with an increase in the number of chemical compounds disclosed in patent claims (complexity), in the number of newly invented compounds (novelty) and the number of newer recombination of compounds (novelty in recombination). These findings shed light on the role of regulation as a catalyst in knowledge recombination enabling inventing entities to explore more of the technological spaces. While most extant research points out the role of regulations on the invention of new or improved products or processes, there has been no systematic empirical evidence on the processes of search and innovation on existing and regulated technologies. This study proposed that the regulatory actions aimed at steering responsibly industry developments away from harmful products or technologies stimulate non-local and distant search - resulting in more knowledge recombination activities. These results also support prior studies on the role of regulation and technical change. By showing a significant effect on knowledge recombination in the period 1978-2016, these results empirically support theoretical models on the role of regulations on knowledge exploration and innovation. The positive and significant effect of the regulation on compound patenting supports prior work on the idea that regulations by altering the search space, alter the rate and intensity of technological innovation. Along these lines, regulation acts as a form of complementary asset that allows firms to change their innovation strategies (Teece 1986). The separate effect of the chemical regulation proxied by the SVHC list (i.e., +39.8%, +23% of new compounds per patent, and +2% of newer recombinations of compounds per patent) is supported by the PSM estimations controlling for the effect of similar technologies. From a policy perspective, this result is interesting as chemical regulation shows a stronger and separate effect on compound patenting and newly invented compound production and recombination. This is due to the fact that the regulation opens up new and unexplored territories in the chemical field. Thus, regulatory efforts aimed at steering responsibility in the innovation process and taking preventive action in the face of risk and uncertainty led to more exploration of chemical space, increased novel compounds and novel recombinations of chemical compounds per patent – in essence, creating conditions ripe for technological innovation. Understanding the economic impact of chemical regulations and how they influence innovation processes is important for public policy, R&D management and firm strategy. By linking chemical regulation with industry developments, this empirical study sheds new light on important but largely unexplored questions in the literature on technological search. These results reveal an interesting finding: regulations not only increase the rate of

search and innovation but also impact the nature and the direction of technological search and underlying search strategies.

References

Arthur, W Brian. 2007. "The structure of invention." *Research policy* 36 (2):274-287.

Becker, Sascha O, and Andrea Ichino. 2002. "Estimation of average treatment effects based on propensity scores." *The stata journal* 2 (4):358-377.

Bekkers, Rudi, Arianna Martinelli, and Federico Tamagni. 2020. "The impact of including standards-related documentation in patent prior art: Evidence from an EPO policy change." *Research Policy* 49 (7):104007. doi: https://doi.org/10.1016/j.respol.2020.104007.

Biggi, Gianluca, and Elisa Giuliani. 2021. "The noxious consequences of innovation: what do we know?" *Industry and Innovation* 28 (1):19-41. doi: 10.1080/13662716.2020.1726729.

Biggi, Gianluca, Elisa Giuliani, and Arianna Martinelli. 2020. "What is Boiling in the Pot? On the Patent Toxicity of Globally Banned Chemicals." *Academy of Management Proceedings* 2020 (1):11567. doi: 10.5465/AMBPP.2020.11567abstract.

Biggi, Gianluca, Elisa Giuliani, Arianna Martinelli, and Emilio Benfenati. 2022. "Patent Toxicity." *Research Policy* 51 (1):104329. doi: https://doi.org/10.1016/j.respol.2021.104329.

Caliendo, Marco, and Sabine Kopeinig. 2008. "Some practical guidance for the implementation of propensity score matching." *Journal of economic surveys* 22 (1):31-72.

Coad, Alex, Gianluca Biggi, and Elisa Giuliani. 2021. "Asbestos, leaded petrol, and other aberrations: comparing countries' regulatory responses to disapproved products and technologies." *Industry and Innovation* 28 (2):201-233. doi: 10.1080/13662716.2020.1830041.

Coad, Alex, Paul Nightingale, Jack Stilgoe, and Antonio Vezzani. 2021. "Editorial: the dark side of innovation." *Industry and Innovation* 28 (1):102-112. doi: 10.1080/13662716.2020.1818555.

Cone, Marla. 2006. "Europe passes tough chemical law." *Los Angeles Times*.

Cressey, Daniel. 2015. "Widely used herbicide linked to cancer." *Nature* 24:1-3.

DiMasi, J. A., L. Feldman, A. Seckler, and A. Wilson. 2010. "Trends in risks associated with new drug development: success rates for investigational drugs." *Clin Pharmacol Ther* 87 (3):272-7. doi: 10.1038/clpt.2009.295.

Drews, J. 2000. "Drug discovery: a historical perspective." *Science* 287 (5460):1960-4. doi: 10.1126/science.287.5460.1960.

Fleming, Lee. 2001. "Recombinant uncertainty in technological search." *Management science* 47 (1):117-132.

Fleming, Lee, and Olav Sorenson. 2004. "Science as a map in technological search." *Strategic Management Journal* 25 (8-9):909-928. doi: https://doi.org/10.1002/smj.384.

Hall, Bronwyn H, and Christian Helmers. 2013. "Innovation and diffusion of clean/green technology: Can patent commons help?" *Journal of Environmental Economics and Management* 66 (1):33-51.

Hall, Bronwyn H, Adam B Jaffe, and Manuel Trajtenberg. 2001. "The NBER patent citation data file: Lessons, insights and methodological tools."

Henderson, Rebecca, and Iain Cockburn. 1996. "Scale, Scope, and Spillovers: The Determinants of Research Productivity in Drug Discovery." *RAND Journal of Economics* 27 (1):32-59.

Hicks, Diana, and Deepak Hegde. 2005. "Highly innovative small firms in the markets for technology." *Research policy* 34 (5):703-716.

Hicks, John. 1963. *The theory of wages*: Springer.
Jaffe, Adam. 1986. "Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value." *American Economic Review* 76 (5):984-1001.
Jaffe, Adam B., and Karen Palmer. 1997. "Environmental Regulation and Innovation: A Panel Data Study." *The Review of Economics and Statistics* 79 (4):610-619.
Jaffe, Adam, and Manuel Trajtenberg. 1999. "International Knowledge Flows: Evidence From Patent Citations." *Economics of Innovation and New Technology* 8 (1-2):105-136.
Jaffe, Adam, Manuel Trajtenberg, and Rebecca Henderson. 1993. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *The Quarterly Journal of Economics* 108 (3):577-598.
Jayaraj, Sebastian, and Michelle Gittelman. 2018. "Impact of Scientific Maps on Technological Search and Novelty." *Academy of Management Proceedings* 2018 (1):15414. doi: 10.5465/ambpp.2018.234.
Judge, J Matthew, Phil Brown, Julia Green Brody, and Serena Ryan. 2016. "The exposure experience: Ohio River Valley residents respond to local perfluorooctanoic acid (PFOA) contamination." *Journal of Health and Social Behavior* 57 (3):333-350.
Kriebel, David, Joel Tickner, Paul Epstein, John Lemons, Richard Levins, Edward L Loechler, Margaret Quinn, Ruthann Rudel, Ted Schettler, and Michael Stoto. 2001. "The precautionary principle in environmental science." *Environmental health perspectives* 109 (9):871-876.
Krieger, Joshua, Danielle Li, and Dimitris Papanikolaou. 2021. "Missing Novelty in Drug Development*." *The Review of Financial Studies*. doi: 10.1093/rfs/hhab024.
Kuhn, Jeffrey M, and Neil C Thompson. 2019. "How to measure and draw causal inferences with patent scope." *International Journal of the Economics of Business* 26 (1):5-38.
Lakey, Pascale SJ, Thomas Berkemeier, Haijie Tong, Andrea M Arangio, Kurt Lucas, Ulrich Pöschl, and Manabu Shiraiwa. 2016. "Chemical exposure-response relationship between air pollutants and reactive oxygen species in the human respiratory tract." *Scientific reports* 6 (1):1-6.
Lerner, Joshua. 1994. "The importance of patent scope: an empirical analysis." *The RAND Journal of Economics*:319-333.
Mandavilli, Apoorva. 2018. "The World's Worst Industrial Disaster Is Still Unfolding." *The Atlantic* 10.
March, James G. 1991. "Exploration and exploitation in organizational learning." *Organization science* 2 (1):71-87.
Marco, Alan C., Joshua D. Sarnoff, and Charles A. W. deGrazia. 2019. "Patent claims and patent scope." *Research Policy* 48 (9):103790. doi: https://doi.org/10.1016/j.respol.2019.04.014.
Martuzzi, Marco, and Roberto Bertollini. 2005. "The precautionary principle, science and human health protection." *Human and Ecological Risk Assessment: An International Journal* 11 (1):63-68.
Nelson, Richard R. 2009. *An evolutionary theory of economic change*: harvard university press.
Nizzetto, Luca, Matthew Macleod, Katrine Borgå, Ana Cabrerizo, Jordi Dachs, Antonio Di Guardo, Davide Ghirardello, Kaj M. Hansen, Andrew Jarvis, Anders Lindroth, Bernard Ludwig, Donald Monteith, Judith A. Perlinger, Martin Scheringer, Luitgard Schwendenmann, Kirk T. Semple, Lukas Y. Wick, Gan Zhang, and Kevin C. Jones. 2010. "Past, Present, and Future Controls on Levels of Persistent Organic Pollutants

in the Global Environment." *Environmental Science & Technology* 44 (17):6526-6531. doi: 10.1021/es100178f.
Ögmundarson, Ólafur, Markus J Herrgård, Jochen Forster, Michael Z Hauschild, and Peter Fantke. 2020. "Addressing environmental sustainability of biochemicals." *Nature Sustainability* 3 (3):167-174.
Owen, Richard, David Baxter, Trevor Maynard, and Michael Depledge. 2009. Beyond regulation: risk pricing and responsible innovation. ACS Publications.
Owen, Richard, Jack Stilgoe, Phil Macnaghten, Mike Gorman, Erik Fisher, and Dave Guston. 2013. "A framework for responsible innovation." *Responsible innovation: managing the responsible emergence of science and innovation in society* 31:27-50.
Papadatos, George, Mark Davies, Nathan Dedman, Jon Chambers, Anna Gaulton, James Siddle, Richard Koks, Sean A Irvine, Joe Pettersson, and Nicko Goncharoff. 2016. "SureChEMBL: a large-scale, chemically annotated patent document database." *Nucleic acids research* 44 (D1):D1220-D1228.
Raffensperger, Carolyn, Ted Schettler, and Nancy Myers. 2000. "Precaution: Belief, regulatory system, and overarching principle." *International Journal of Occupational and Environmental Health* 6 (4):266-269.
Rigét, Frank, Anders Bignert, Birgit Braune, Maria Dam, Rune Dietz, Marlene Evans, Norman Green, Helga Gunnlaugsdóttir, Katrin S. Hoydal, John Kucklick, Robert Letcher, Derek Muir, Stacy Schuur, Christian Sonne, Gary Stern, Gregg Tomy, Katrin Vorkamp, and Simon Wilson. 2019. "Temporal trends of persistent organic pollutants in Arctic marine and freshwater biota." *Science of The Total Environment* 649:99-110. doi: https://doi.org/10.1016/j.scitotenv.2018.08.268.
Rosenbaum, Paul R, and Donald B Rubin. 1983. "The central role of the propensity score in observational studies for causal effects." *Biometrika* 70 (1):41-55.
Rosenkopf, Lori, and Atul Nerkar. 2001. "Beyond local search: boundary-spanning, exploration, and impact in the optical disk industry." *Strategic management journal* 22 (4):287-306.
Schilling, Melissa A, and Elad Green. 2011. "Recombinant search and breakthrough idea generation: An analysis of high impact papers in the social sciences." *Research Policy* 40 (10):1321-1331.
Schumpeter, Joseph A. 2017. *The theory of economic development: An inquiry into profits, capita I, credit, interest, and the business cycle*: Routledge.
Shane, Scott. 2001. "Technological opportunities and new firm creation." *Management science* 47 (2):205-220.
Smith, Jeffrey A, and Petra E Todd. 2005. "Does matching overcome LaLonde's critique of nonexperimental estimators?" *Journal of econometrics* 125 (1-2):305-353.
Soto, Ana M, and Carlos Sonnenschein. 2010. "Environmental causes of cancer: endocrine disruptors as carcinogens." *Nature Reviews Endocrinology* 6 (7):363-370.
Southall, Noel T., and Ajay. 2006. "Kinase Patent Space Visualization Using Chemical Replacements." *Journal of Medicinal Chemistry* 49 (6):2103-2109. doi: 10.1021/jm051201m.
Stilgoe, Jack, Richard Owen, and Phil Macnaghten. 2013. "Developing a framework for responsible innovation." *Research Policy* 42 (9):1568-1580.
Stirling, Andy. 2016. "Precaution in the Governance of Technology." *Forthcoming Oxford Handbook on the Law and Regulation of Technology, SWPS* 14.
Teece, David J. 1986. "Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy." *Research Policy* 15 (6):285-305. doi: https://doi.org/10.1016/0048-7333(86)90027-2.

Trajtenberg, Manuel, Rebecca Henderson, and Adam Jaffe. 1997. "University versus corporate patents: A window on the basicness of invention." *Economics of Innovation and new technology* 5 (1):19-50.
Verhoeven, Dennis, Jurriën Bakker, and Reinhilde Veugelers. 2016. "Measuring technological novelty with patent-based indicators." *Research policy* 45 (3):707-723.
Weitzman, Martin L. 1998. "Recombinant growth." *The Quarterly Journal of Economics* 113 (2):331-360.

APPENDIX

A.1. List of chemical compounds included in the EU REACH SVHC list

Chemical compound name	InChIKey	Year of inclusion	Reason for inclusion
Cobalt(II) chloride (cobalt dichloride)	GVPFVAHMJGGAJG-UHFFFAOYSA-L	2008	Carcinogen
Arsenic pentoxide (diarsenic pentaoxide)	RKELNIPLHQEBJO-UHFFFAOYSA-N	2008	Carcinogen
Arsenic trioxide (diarsenic trioxide)	CFXQEHVMCRXUSD-UHFFFAOYSA-N	2008	Carcinogen
4,4'-Diaminodiphenylmethane (MDA)	SECXISVLQFMRJM-UHFFFAOYSA-N	2008	Carcinogen
Lead hydrogen arsenate	OAKJQQAXSVQMHS-UHFFFAOYSA-N	2008	Carcinogen Toxic for reproduction
Sodium dichromate	ILVKYQKHSCWQAW-UHFFFAOYSA-N	2008	Carcinogen Mutagen Toxic for reproduction
Triethyl arsenate	NVKTUNLPFJHLCG-UHFFFAOYSA-N	2008	Carcinogen
Benzyl butyl phthalate (BBP)	SVONRAPFKPVNKG-UHFFFAOYSA-N	2008	Toxic for reproduction
Bis(2-ethylhexyl) phthalate (DEHP)	KRVSOGSZCMJSLX-UHFFFAOYSA-L	2008	Toxic for reproduction
Dibutyl phthalate (DBP)	WGLPBDUCMAPZCE-UHFFFAOYSA-N	2008	Toxic for reproduction
Anthracene	ZNQVEEAIQZEUHB-UHFFFAOYSA-N	2008	PBT
Hexabromocyclododecane (HBCDD)	XNWFRZJHXBZDAG-UHFFFAOYSA-N	2008	PBT
Tributyltin oxide (Bis(tributyltin) oxide, TBTO)	QAHREYKOYSIQPH-UHFFFAOYSA-L	2008	PBT
Musk xylene (5-tert-butyl-2,4,6-trinitro-m-xylene)	ZOTKGJBKKKVBJZ-UHFFFAOYSA-L	2008	PBT
Chromic acid, Oligomers of chromic acid and dichromic acid, Dichromic acid	UFMZWBIQTDUYBN-UHFFFAOYSA-N	2010	Carcinogen
Chromium trioxide	KTVIXTQDYHMGHF-UHFFFAOYSA-L	2010	Carcinogen Mutagen
2-Ethoxyethanol	PXLIDIMHPNPGMH-UHFFFAOYSA-N	2010	Toxic for reproduction
2-Methoxyethanol	XMXNVYPJWBTAHN-UHFFFAOYSA-N	2010	Toxic for reproduction
Cobalt(II) diacetate	JOSWYUNQBRPBDN-UHFFFAOYSA-P	2010	Carcinogen Toxic for reproduction
Cobalt(II) carbonate	KMUONIBRACKNSN-UHFFFAOYSA-N	2010	Carcinogen Toxic for reproduction
Cobalt(II) dinitrate	UQGFMSUEHSUPRD-UHFFFAOYSA-N	2010	Carcinogen Toxic for reproduction
Cobalt(II) sulfate	CDMADVZSLOHIFP-UHFFFAOYSA-N	2010	Carcinogen Toxic for reproduction

A.1. (Continued) List of chemical compounds included in the EU REACH SVHC list

Chemical compound name	InChIKey	Year of inclusion	Reason for inclusion
Sodium chromate	KGBXLFKZBHKPEV-UHFFFAOYSA-N	2010	Carcinogen Mutagen Toxic for reproduction
Potassium chromate	XSTXAVWGXDQKEL-UHFFFAOYSA-N	2010	Carcinogen Mutagen
Ammonium dichromate	HRPVXLWXLXDGHG-UHFFFAOYSA-N	2010	Carcinogen Mutagen Toxic for reproduction
Potassium dichromate	HQUQLFOMPYWACS-UHFFFAOYSA-N	2010	Carcinogen Mutagen Toxic for reproduction
Tetraboron disodium heptaoxide, hydrate	RMBFBMJGBANMMK-UHFFFAOYSA-N	2010	Toxic for reproduction
Disodium tetraborate, anhydrous	MGWAVDBGNNKXQV-UHFFFAOYSA-N	2010	Toxic for reproduction
Boric acid	MOUPNEIJQCETIW-UHFFFAOYSA-N	2010	Toxic for reproduction
Trichloroethylene	AUNAPVYQLLNFOI-UHFFFAOYSA-L	2010	Carcinogen
Acrylamide	KUAYCHYPUWBYFY-UHFFFAOYSA-L	2010	Carcinogen Mutagen
Tris(2-chloroethyl)phosphate	COHDHYZHOPQOFD-UHFFFAOYSA-N	2010	Toxic for reproduction
2,4-Dinitrotoluene	QTLQKAJBUDWPIB-UHFFFAOYSA-N	2010	Carcinogen
Diisobutyl phthalate (DIBP)	YBRVSVVVWCFQMG-UHFFFAOYSA-N	2010	Toxic for reproduction
Lead chromate	UWRBYRMOUPAKLM-UHFFFAOYSA-L	2010	Carcinogen Toxic for reproduction
Lead chromate molybdate sulfate red (C.I. Pigment Red 104)	JYDRNIYTFCBIFC-UHFFFAOYSA-N	2010	Carcinogen Toxic for reproduction
Lead sulfochromate yellow (C.I. Pigment Yellow 34)	MGRFDZWQSJNJQP-UHFFFAOYSA-N	2010	Carcinogen Toxic for reproduction
1,2-benzenedicarboxylic acid, di-C6-8-branched alkyl esters, C7-rich	IRIAEXORFWYRCZ-UHFFFAOYSA-N	2011	Toxic for reproduction
1,2,3-Trichloropropane	BJQHLKABXJIVAM-UHFFFAOYSA-N	2011	Carcinogen Toxic for reproduction
1-methyl-2-pyrrolidone	DOIRQSBPFJWKBE-UHFFFAOYSA-N	2011	Toxic for reproduction
Hydrazine	MWPLVEDNUUSJAV-UHFFFAOYSA-N	2011	Carcinogen
1,2-benzenedicarboxylic acid, di-C7-11-branched and linear alkyl esters	DEIGXXQKDWULML-PQTSNVLCSA-N	2011	Toxic for reproduction
Strontium chromate	APQHKWPGGHMYKJ-UHFFFAOYSA-N	2011	Carcinogen
2-ethoxyethyl acetate	XMWRWTSZNLOZFN-UHFFFAOYSA-N	2011	Toxic for reproduction