A quantitative approach of innovation

Anne Plunket
Professeur d'économie
anne.plunket@u-psud.fr
Université Paris-Sud, Université Paris Saclay

1er semestre

Master IREN - Industries de Réseau et Economie Numérique

Practical information

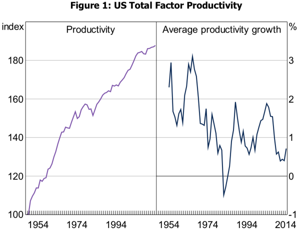
- **Course meetings** : 6 lectures and 1 meeting devoted to the exam ;
- **Language of the course** : English
- **Research-based lectures**
  - Readings - articles from scientific journals
  - Class slides and R codes - Use of R studio `https://www.rstudio.com/products/RStudio/`
- **Assignment** : Written exam based on the lectures - discuss the literature

Course Motivation

Why explain and measure technological change and innovation ?

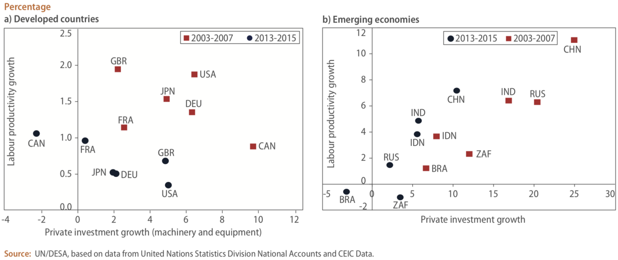
Puzzle 1. Slowdown of productivity growth, why ?

- **The slowdown in the rate of technological change** as a possible cause of the post-1973 productivity growth slowdown.



Source : Fed. Reserve Bank of St. Louis

- **What explains productivity growth ?**
  - Investment in tangible (machinery and equipement) + labor –> Residual (50 to 80%)



Puzzle 1. Slowdown of productivity growth, why ?

- **Investment in intangible capital** such as patents, know-how, human capital, etc. as a source of technological change
- **Technological change (TC)** is the rate at which new technology-based production processes and products are introduced, diffused and adopted in the economy
  - TC should enable higher levels of real output holding constant the inputs such as labor and capital $\Rightarrow$ Increase multifactor productivity
  - Faster rate of growth of technological change should lead to faster rates of productivity growth

- **The type and rate of technological change matters !**

  Incremental vs radical change - defining technological novelty/ radicalness / disruptive/breakthrough technologies
  - Measure technological change ?
  - Value technologies given their « quality » ?

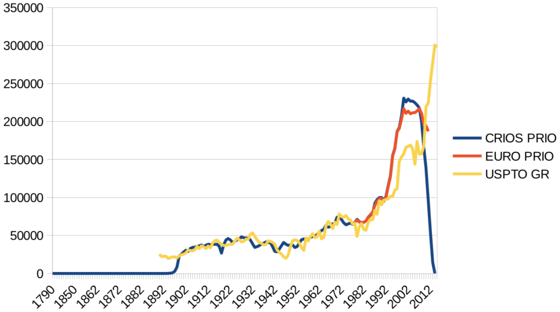
- **Chapter 1. The nature of ideas and innovation**
  - Where do ideas come from ? How do organizations search and combine knowledge ?
  - What are the types of technological change and impact ?
  - The role of science and universities
  - The generation of breakthrough technologies : (e.g. radical vs incremental innovation ; high impact ideas ; breakthrough, disruptive technologies) ?
  - **Characterizing organization's innovation using R project and patent data**

- **Chapter 2. Measuring innovation and impact**
  - Defining and measuring innovation
  - Impact on firms : productivity, market value and externalities
  - Patent statistics and measuring the value of patents

Puzzle 2. Dispersion of knowledge and actors

Explain the determinants and impact of innovation.

- **Organizational fragmentation** of knowledge creation and property ⇒ Strategic behavior - blocking patents
  - Upsurge in the number of patents



source : author calculation - US granted patents in various data bases

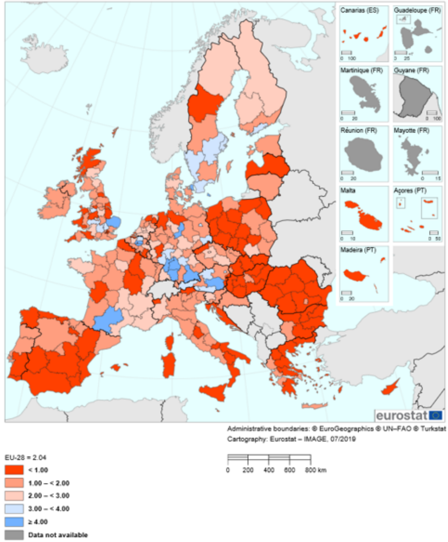
- **Knowledge sourcing and integration** - Networks - Collaborations - Open innovation
  - More difficult to master technologies ⇒ R&D Collaborations and Open innovation

Course Motivation

2. Organizational and geographical dispersion of knowledge production

- **Geographical dispersion** of knowledge $\Rightarrow$ R&D Collaborations and Open innovation
  - Increased complexity of knowledge creation
  - More difficult to be creative and produce novelty
- Rôle of geographical location
- Rôle of network embeddness - network connection - knowledge sourcing

- Research and development across Europe : R&D intensity



Note: R&D intensity is defined as gross domestic expenditure on R&D (GERD) relative to gross domestic product (GDP). 2015 data for Belgium, Germany, Ireland, Greece, Austria and Sweden (all regions), Ciudad Autónoma de Ceuta and Ciudad Autónoma de Melilla (Spain) and Basilicata (Italy); 2014 data for Umbria and Molise (Italy); 2013 data for France (all regions); data for UK are estimated.

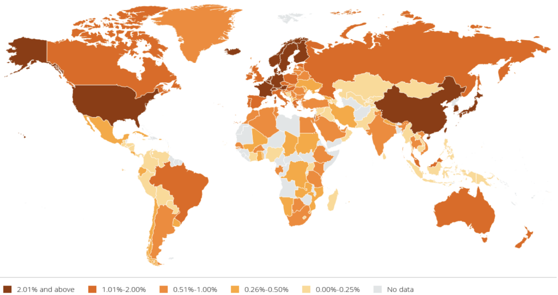
Source: Eurostat (online data code: rd_e_gerdreg)

Source : Eurostat - R&D and innovation 2020 indicators

- Research and development across the world : R&D intensity

Figure 3. A snapshot of R&D intensity

Gross domestic expenditure on R&D as a percentage of GDP, 2017 or latest year available



Source: UNESCO Institute for Statistics, June 2019.

Source : UNESCO institute for statistics - Fact Sheet No. 54

- **Chapter 3. Geography of ideas and knowledge flows**
  - Rôle of location for innovation - agglomeration economies and industrial clusters
  - Sourcing and diffusion of knowledge - Rôle of labor mobility ; patent citations ; research collaborations

- **Chapter 4. Networks and innovation**
  - Rôle of networks for knowledge diffusion and innovation
  - Network characteristics - denser vs sparse networks - Rôle of knowledge brokers
  - **Building networks** of inventors using **R Project and patent data**

Chapter 1. The nature of ideas and innovation

Section 1. What is a technological invention/innovation ?

1. What is a technological invention/innovation ?

From Ideas to Inventions to innovation

- Schumpeter (1939) defined **innovation** as the commercial application or adoption of an **invention**.
  - Schumpeter argues « the making of the invention and the carrying out of the corresponding innovation are, economically and sociologically two entirely different things » – Schumpeter (p.85)

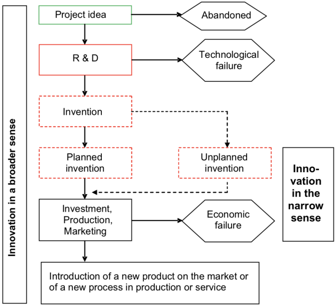


Figure 1: Correlation between invention and innovation (see Brockhoff 1997: 36; Strebel 2003: 21)

1. What is a technological invention/innovation ?

From inventions to patents and technology

- **An invention** is a unique and novel device, method, composition or process.
- Most inventions are **not technology-based** and they do not even rely on technology (e.g. marketing invention / organizational invention - innovation)
  - Some service inventions may be related to technology
  - Ex : Amazon 1-Click patent in 1999 allowed to create a strong position in the market
- **Most inventions are not patented**
  - Patents only enable to measure and evaluate part of all inventions - depends on the activity (ICT, Pharmaceuticals, Chemistry, Mechanical engineering...)
  - **Patents are technology-based inventions** which integrate distinct technological functionalities : artifacts, devices, methods, and materials - available to humans to accomplish specific tasks (Arthur, 2007).
- **New technologies lead to technological change**, when new technological functionalities are introduced into the existing repertoire of technologies

Section 2. Types of technological change

2. Types of technological change

- **Technological change** is the rate at which new technology-based production processes and products are introduced, diffused and adopted in the economy
- Technology is believed to develop along well-defined and predictable **technological trajectories**, occasionally interrupted by **discontinuities** introduced by paradigm shifts (Dosi, 1982) ⇒ **Technology life cycle** (Abernath and Utterback, 1978)
  - Example : Telephone has undergone tremendous technological change
- Schumpeter (1942) identified discontinuous technological change and related innovations as the sources of the **creative destruction** in industries.

⇒ Need to unpack the drivers and effects of radical innovations is of major interest

2. Types of technological change

- **Technological novelty - radical vs incremental** : the degree of novelty can be an important driver of radical technological innovation
  1. **Disrupt existing competences** (Tushman and Anderson, 1986)
     - Competence-destroying discontinuities - introduced by new firms and source of environmental turbulence
     - Competence-enhancing discontinuities - initiated by existing firms and associated with lower environmental turbulence
  2. **Eliminate existing players** from the market (Christensen, 2013)

$\Rightarrow$ Need to unpack the drivers and effects of radical innovations is of major interest

2. Types of technological change

- Though incremental innovations increase productivity, radical innovations are the real **engine of growth**
  - Incremental innovations run into **diminishing returns** (Akcigit and Kerr, 2010 ; Abrams, Akcigit and Popadak, 2013)
  - = benefits from subsequent product development efforts increase at a declining rate
- **Radical innovations create new technology clusters** which increase productivity directly by making another series of incremental innovations possible.
  - Example *Turbojet engine compared to former propeller engines* : introduced the concept of thrust by expelling particles to create an opposite force to accelerate an airplane
  - ⇒ In the following decades, a series of follow-on incremental improvements refined this novel approach in terms of functionning and performance allowing tremendous growth in the aviation industry and beyond
  - Example *Laser technology* enabled innovations in optics, medical technologies, music industry (CD), laser to cut metal in shipbuilding, aeronautics...

Research question :

- **Need for a theory of the process of invention** :
  - what type of firms, managers generate radical innovations ?
  - ≡ Lack of understanding of the process by which firms create new technologies
  - What is the underlying innovation process : technological trajectories, collaborations with other firms, or research units
  - How are new technologies generated ? (search process)

≡ Understand invention as a recombinant process

- Can we « measure » the degree of novelty or radicalness ?
- Two classical perspectives on the sources of technological novelty
  1. Invention is a process of recombination
  2. Invention is an inherently uncertain and therefore local search process

3. Invention as a recombinant process : defining novelty

3.1. Inventions and knowledge recombination

- **Recombination provides the ultimate source of novelty**
  - « Innovation combines components/factors **in a new way**, or it consists in carrying out **new combinations** » (Schumpeter, 1939, p.88)
  - Example
    - *Endoscopy* as a flexible tube combined with electronics and camera ;
    - *The microprocessor* as a conjunction of a computer's central processing unit with integrated circuit fabrication processes.
    - *Unsuccessful combination* such as plane-automobile combination or the nuclear powered aircraft (Bassalla, 1988)
- Recombination usually occurs between components that are **salient, proximal and available** to the inventor.
  - Potential number of combination is explosive (Weitzman, 1996)
  - Individual inventors and even communities of inventors have only limited understanding of all the potential combinations and relationships

  ⇒ Inventors, organizations and communities focus and **recombine locally from a limited set of components and combinations**

3.2. Defining and measuring technological novelty based on patents

Patent document : rich source of information

- A patent may be characterized by its technological classification - IPC codes $\Rightarrow$ trace technological trajectories also based on citations (forward / backward)

Table 5.1. **Main characteristics of IPC codes (example)**

Subdivision	Number	Symbol (code letter)	Title (code label)
Section	8	G	Physics
Subsection	20		Instruments
Class	118	G06	Computing; Calculating; Counting
Subclass	616	G06F	Electrical digital data processing
Main group	6 871	G06F-9/00	Arrangements for programme control
Subgroup	57 324	G06F-9/06	* Using stored programme
		G06F-9/46	** Multi-programming arrangements

Source: World Intellectual Property Organization (2006), *IPC Guide*, 8th edition.

Group level International Patent Classification IPC-codes to which a patent is assigned (Strasbourg 1971 agreement)

`http://www.wipo.int/classifications/en/`

3.3. Constructing the technological novelty indicator

Verhoeven D., Bakkera J., Veugelers R. (2016)

- A patent has « **Novelty in Recombination (NR)** » when it contains at least one pair of IPC codes that were **previously unconnected** over the patent history (Fleming 2001 ; Fleming et al., 2007 ; Verhoeven et al., 2016).

$$\text{Recombinant Novelty}_t = \frac{Number\ of\ new\ subclass\ pairs}{Total\ number\ of\ subclass\ pairs} = \frac{4}{28} = 0,14 \qquad (1)$$

Total combinations (n*(n-1)/2) = 8*7/2 = 28

- ▸ **Only restriction**, it requires a patent to belong to at least 2 IPC groups to be making combinations between components and/or principles, which may or may not be new.

Table 1

Illustration of construction of the indicators based on the Oncomouse patent family. Column 1 indicates the IPC groups the oncomouse patent family is assigned to. Column 2 provides the patent/scientific references of the oncomouse patent family. Column 3 provides the IPC groups/WOS Subject Categories of these references. Column 4 displays the class combinations that are considered for calculating the indicators. The last two columns indicate whether the combination is assessed as a new combination and shows the score on the dichotomous indicator variables.

The Oncomouse patent family: US4736866, EP169672, CA1341442, DE3586020, JP5048093, JP61081743, JP2058915

IPC groups (Exhaustive)	(Examples of) combinations	First occurrence?	Indicator
A01K 67	A61D 19 – C07H 21	Yes	
A61D 7	A01K 67 – C07H 21	Yes	
A61D 19	A61D 7 – C07H 21	Yes	
C07H 21	A61D 19 – C07K 14	Yes	
C07K 14	A01K 67 – A61D 7	No	NR = 1
C12N 5	A01K 67 – A61D 19	No	
C12N 15	A01 K 67 – C07H 21	No	
G01N 33	...	No	
	Total Number of Positives = 4		

Ex = A01K67 = breading of animals , A61D = veterinary instruments

https://worldwide.espacenet.com/classification?locale=en_EP#!/CPC=A01K67

Section 4. Invention as a recombinant search process

4.1 A search perspective

- How do firms search and solve problems to create new products ?
- **Search perspective** : firms are viewed as **problem solvers**
  - Organizations engage in a wide variety of searches : organizational designs, manufacturing methods, ways to implement new innovations and new product designs
  - **Search process**= organization's **problem-solving activities** that involve the creation and recombination of technological ideas (Winter, 1984).
- Individuals and firms have **bounded rationality** which limits their ability to search toward more salient areas of their own prior experience (Simon, 1978 ; Cyert and March, 1963).
- In the management literature, the search process can be defined in terms of **scope or depth**

4.2. Search scope : local vs distant

- **Search scope** : the degree of new knowledge that is explored ; It refers to the **diversity** of knowledge domains searched and to knowledge in which the actor **lacks prior experience or competence**

⇒ It is based on a continuum with two extreme search types :

1. **Local search** : address problems by using knowledge that is closely related to their prior (= preexisting) knowledge bases (Helfat, 1994 ; Stuart and Podolny, 1996)
   - *Localness* corresponds to **familiarity** with their search space
   - ***Local search or exploitation*** occurs when an invention recombines from a familiar set of technology components or refines a previously used combination (March 1991)

2. **Distant search or exploration** : conscious effort to **move away** from current organizational routines and knowledge bases (March, 1991).
   - Distant search occurs when a firm tries **completely new components or combinations**
   - *Example : early automobile industry* : many successful and unsuccessful search including power sources, pneumatic tires, brakes for each passenger in the 1930s (Bassala, 1988)

4.2. Advantages and limits of search Scope

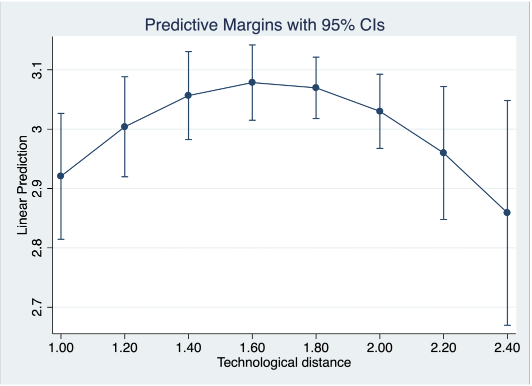
- **High search scope** affects products innovation **positively** through two mechanisms
  - Introduces **new variations** which provide sufficient amount of choice to solve problems (March, 1991)
    - Evolutionary theory : « selection effect of variation »
  - **Enhances recombinatory search** (Fleming and Sorenson, 2001 ; Nelson and Winter, 1982) : higher probability of finding new combinations

- **High search scope** may also have **negative** consequences
  - **Increased cost** of technological and organization integration of new knowledge
    - Technologically : common interfaces need to be established among knowledge elements
    - Organizationally : new knowledge requires changes in networks of relations and communication relationships both within and outside the organization (Henderson and Clark, 1990 ; Fleming, 2001) : the wider the scope of the knowledge to be integrated and the more **complex** are the problems of creating and managing **integration**.
  - Eventually, **the cost of integration** will exceed the benefits of acquiring new knowledge

4.2. Advantages and limits of search Scope

- **High search scope** may also have **negative** consequences
  - Decrease in the firm's **reliability**, that is, its ability to respond to new information correctly (e.g. martin and Mitchell, 1998) : innovation projects in which the proportion of new knowledge is high are less likely to succeed than projects that search closely related knowledge (Cyert and March, 1963)
- **Search scope is curvilinearly** (= take an inverted U-shape) related to the number of new products introduced by a firm.

- Graphical example of a curvilinear relationship : inverted u-shape



4.3. Search depth : specialization

- **Search depth** : the degree to which search **revisits a firm's prior knowledge**, that is existing knowledge is **reused** or exploited
  - Search depth : refers to the extensiveness of search within a given area - the deeper the search effort, the more **cumulative experience** an actor accrues in a domain of knowledge and greater competence (Katila, 2000)
  - Organizations attain a deep and narrow knowledge reservoir accrued through years of **specialization**
  - Individuals often require at least a decade of intense study in a particular domain of knowledge prior to making a significant contribution in that domain (Simonton, 1999)

4.3. Advantages and limits of search depth

Search depth affects **positively** product innovation through **three kinds of experience effects**

1. Make search more **reliable**
   - Using the same knowledge repeatedly reduces the likelihood of errors and false starts and facilitates the development of routines, making search more reliable (Levinthal and March, 1991)
2. Make search more **predictable** : familiarity in knowledge and requirements better understood
   - Product development can be decomposed into solvable subproblems, activities can be sequenced in efficient order
3. Make search more **creative** : ability to identify valuable knowledge elements and combinations :
   - Develop connections among them and to combine them in many different and significant ways not apparent to less experienced users of those concepts

4.3. Advantages and limits of search depth

- **Excessive depth** has **negative** consequences : the relationship between depth and innovation is nonlinear
  - **Limits to improvement** along a technological trajectory (Dosi, 1988)
    Further developments based on the same knowledge become increasingly expensive and the solutions excessively complicated : cost of depth > benefits of depth
  - ⇒ **Diminishing returns** to building on the same knowledge : benefits from subsequent product development efforts increase at a declining rate
  - **Rigidity** : reuse of existing knowledge can make an **organisation rigid** : solutions and problem-solving strategies may become ineffective but because of past effectiveness the organizations get **locked-in** (Argyris and Schoen, 1978)
- **Search depth is curvilinearly** (inverted U-shaped) related to the number of new products introduced by the firm

4.4. Interaction between depth and scope search

- **Scope and depth are mutually beneficial** and have interactive effects through two mechanisms
  - **Absorptive capacity** : firms can use their accumulated knowledge to recognize and assimilate new (external) knowledge (Cohen and Levinthal, 1989) and thus facilitate further development of new knowledge (Winter, 1994) ;
  - **Uniqueness of recombinations** : by combining firm-specific accumulated understanding of certain knowledge elements (depth) with new solutions (scope), firms are more likely to create new, unique combinations that can be commercialized (Winter, 1984)
    - increases in scope may be costly : the probability of finding valuable knowledge elements is small and even if a firm succeeds in doing so, it is possible that the same product idea has already been discovered

4.5. Empirical analysis of the search process

Empirical analysis of the search process

Katila, R. and Ahuja, G., (2002), Something old, something new : a longitudinal study of search behavior and new product introduction, Academy of Management Journal, 45(6), 1183–1194

4.5. Empirical analysis of the search process

- **Aim of the paper** : How firms search and the effect on innovation performance
  - Scope and Depth search on the number of new products
- Data : 124 Industrial robotics companies in Europe (19), Japan (78) and North America (27) ; average firm size 39,000 employees
- Robotics technology is a combination of multiple rapidly changing technological disciplines such as electronics, new materials, and optics
- Data source : new product introduction announcement in trade magazines and catalogues and patent data (USPTO)

4.5. Empirical analysis of the search process

- Dependent variable : **# of new products**
- Independent variables : need to asses **intrafirm search activities** over time
  - Patent data : description of a technical problem and a solution to that problem (Katila, 2002 ; Rosenkopf and Nerkar, 2001 ; Stuart and Podolny, 1996)
  - **Search depth** = **accumulation of search experience** with the same knowledge elements = measured as the average number of times a firm repeatedly used the citations in the patents it applied for

$$\text{Depth}_{it-1} = \frac{\sum_{y=t-6}^{t-2} \text{repetition count}_{iy}}{\text{total citations}_{it-1}}$$

  - **Search scope** : corresponds to the notion of **exploration** of new knowledge. It is the proportion of previously unused citations (new citations$_{it-1}$) in a firm's focal year's list of citations = measured through the share of citations in a focal year's citations, that could not be found in the previous five year's list of patents and citations by that firm.

$$\text{Scope}_{it-1} = \frac{\#\ \text{new citations}_{it-1}}{\text{total citations}_{it-1}}$$

4.5. Empirical analysis of the search process

- Control variables : neutralize variables that may affect the innovation process
  - *Firm performance* (= increase in performance increases exploration),
  - *R&D expenditures* (= amount of search activities),
  - *Degree of product diversification* (= diversified firms have more opportunities for internal use of new knowledge),
  - *Firm size* (= positive effect of firm size on product innovation)
- Statistical method
  - Poisson regression : # of new products is a **count variable**
  - GEE - generalized estimating equations to control for firm heterogeneity - i.e. accounts for autocorrelations due to repeated yearly measurements of the same firms

4.5. Empirical analysis of the search process

TABLE 1
Descriptive Statistics and Correlations[a]

Variable	Mean	s.d.	Min.	Max.	1	2	3	4	5	6	7	8	9
1. Number of new products	0.92	2.36	0.00	24.00									
2. Search depth	0.22	0.24	0.00	1.74	.04								
3. Search scope	0.74	0.30	0.00	1.00	.08**	−.003							
4. Collaboration frequency	0.16	0.50	0.00	6.00	.33***	.02	.07*						
5. Firm performance	0.02	0.05	−0.73	0.45	.01	.04	.03	.04					
6. R&D expenditure	0.32	0.78	0.01	6.78	.07*	.22***	.07*	.07*	.03				
7. Diversification	0.64	0.37	0.00	1.62	−.02	.20***	.15***	.03	−.02	.11***			
8. European firm	0.15	0.36	0.00	1.00	−.06*	−.06*	.06*	−.01	.03	.17***	.20***		
9. American firm	0.22	0.41	0.00	1.00	.15***	−.02	−.06*	.01	.04	.05	−.19***	−.22***	
10. Firm size	39.66	100.77	0.03	876.80	−.03	.15***	.08**	.06*	.05	.86***	.14***	.29***	.17***

[a] $n = 1.185$.
* $p < .05$
** $p < .01$
*** $p < .001$

4.5. Empirical analysis of the search process

TABLE 2
Results of Regression Analysis for Number of New Products[a, b]

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
Intercept	−0.58	−0.57	−0.57	−0.57	−0.57	−0.63	−0.62	−0.63	−0.53	−0.51
	(0.39)	(0.40)	(0.39)	(0.40)	(0.41)	0.42	0.42	0.42	0.42	0.42
Search depth		0.55†		0.59*	0.95*	3.07***	2.70	3.11***	3.45***	3.24***
		(0.34)		(0.34)	(0.44)	(1.04)	(3.24)	(1.05)	(0.96)	(0.90)
Search scope			0.14	0.22†	1.34**	0.76*	1.32	0.76*	0.85*	0.82*
			(0.15)	(0.16)	(0.49)	(0.44)	(4.25)	(0.45)	(0.42)	(0.39)
Search depth × search scope					5.69**	3.10†	3.33†	3.08†	3.48*	3.43*
					(2.22)	(1.97)	(2.13)	(2.01)	(1.83)	(1.70)
Search depth squared						−3.05*	−2.71	−3.12*	−3.32**	−2.99**
						(1.35)	(2.92)	(1.40)	(1.22)	(1.09)
Search scope squared							−0.51			
							(4.05)			
Collaboration frequency	0.30**	0.30**	0.30***	0.30**	0.29**	0.25*	0.25*	0.26*	0.23*	0.21
	(0.11)	(0.11)	(0.11)	(0.11)	(0.10)	(0.11)	(0.11)	(0.11)	(0.12)	(0.13)
Firm performance	−1.42	−1.46	−1.39	−1.43	−1.44	−1.46	−1.47	−1.46	−1.12	−1.14
	(1.07)	(1.08)	(1.08)	(1.11)	(1.14)	(1.21)	(1.21)	(1.22)	(1.26)	(1.20)
R&D expenditure	0.40**	0.40*	0.40**	0.40*	0.40*	0.30*	0.30*	0.30*		
	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)		
Diversification	−0.50	−0.54	−0.51	−0.55	−0.56	−0.58	−0.58	−0.58	−0.60	−0.59
	(0.36)	(0.38)	(0.36)	(0.38)	(0.38)	(0.38)	(0.38)	(0.37)	(0.39)	(0.39)
European firm	−1.14*	−1.08*	−1.14*	−1.07*	−1.05*	−1.05*	−1.04*	−1.03*	−0.97†	−1.01†
	(0.50)	(0.50)	(0.50)	(0.50)	(0.51)	(0.51)	(0.51)	(0.50)	(0.57)	(0.58)
American firm	−2.05***	−2.04***	−2.05***	−2.03***	−1.97***	−1.97***	−1.97***	−1.95***	−1.82***	−1.86***
	(0.43)	(0.43)	(0.43)	(0.43)	(0.40)	(0.40)	(0.41)	(0.39)	(0.39)	(0.40)
Number of new products (lagged)								0.01		−0.03
								(0.03)		(0.06)
Firm size									0.0001	0.0001
									(0.002)	(0.001)
Deviance	2,694.1	2,685.2	2,679.4	2,667.2	2,646.8	2,618.8	2,617.5			
Difference in log likelihood vis-à-vis the base model		8.9**	14.7***	26.9***	47.3***	75.3***	76.6***			
df	19	20	20	21	22	23	24	24	23	24

[a] The table gives parameter estimates; the standard error is below each parameter estimate in parentheses.
[b] There were 124 firms and 1,185 firm-year observations. Year dummies were included but are not shown.
† $p < .10$
* $p < .05$
** $p < .01$
*** $p < .001$

4.5. Empirical analysis of the search process

- **Results**
  - Search depth has a **curvilinear relationship** with new product innovation, whereas
  - Search scope has a **positive and linear** relationship
  - The interaction between depth and scope is positive : they leverage each other
  - Search is a two-dimensional construct : adding depth and the interaction variables (models 4-6) significantly improves the model fit
    - For a hypothetical firm at the mean of the depth (0.22) and scope (0.74) variables, a 10% increase in depth (an increase of 0.022) leads to a 10% increase in new product introduction
    - For the same firm, a 10% increase in search scope (an increase of 0.074) leads to an 11% increase in new product introduction.

4.5. Empirical analysis of the search process

- **Limits**
  - Sample may drive some results : the linear impact of scope due to the fact that few companies « oversearched » along this dimension because of the cost
    in line with the proposition that firms search locally and the tendency to reduce uncertainty
  - Use of archival data
  - The # of new products does not say anything on the disruptive nature of the innovation/invention
    - Need to study the quality of innovations

Section 5. Does science favor technological search ?

5. Does science favor technological search ?

- **Scientific research** stimulates technological innovation, increases the rate of technological advance and accelerates economic growth (Stephan, 1996)
  - Cumulative research output, in the form of published papers, accelerate growth (Adams, 1990) ;
- Strong empirical support for a link between science research, technological innovation and economic growth

Adams, J. (1990), Fundamental stocks of knowledge and productivity growth, Journal of Political Economy, 98 : 673 - 702

5. Does science favor technological search ?

- Despite the value of scientific research to innovation, the degree to which technological search relies on the **use of science varies** tremendously across fields
  - Universities disproportionately patent in some high tech sectors : biotechnology and life sciences (drugs, medicine), chemicals, nanotechnologies
- Why do firms draw more heavily on scientific research in some areas ?

  Scientific knowledge may lead to different type of search
  - Provides a stylized representation of the area being searched = **Map**
  - Scientific endeavor attempts to generate and test theories - why phenomena occur and provide a means of **predicting the results of untried experiments** and the usefulness of previously uncombined configurations of technological components.
  - The Map likely **modifies the search process**

Example : Generation of chip technology : researchers predicted that single-wall carbon nano-tubes would either conduct or semi-conduct, depending on the diameter of the tube and the angles of the carbon bonds $\Rightarrow$ resulted in the fabrication of conducting and semiconducting nano-tubes (Fleming and Sorenson, 2004)

5. Does science favor technological search ?

- Theory does not predict everything $\Rightarrow$ researchers must still engage in **trial-and-error search** to produce results but it accelerates results
- Theory **eliminates fruitless investigations** and thus reduce the size of the combinatorial search pace (Nelson, 1982)
  - Even if science has an inaccurate or incomplete understanding of a problem, it might still alter the search process in a useful manner.

Empirical analysis : science and technological search

Fleming, L. and Sorenson, O. (2004), Science as a map in technological search, Strategic Management Journal, (2004)

Empirical analysis : science and technological search

- Aim : estimate the returns to the application of science using patent data
- Data : US patents granted in May and June of 1990 (n= 17,265)
- Dependent variable : the future citation count = measure of the value of the invention
- Independent variables :
  - **Science** : dummy (0,1) indicating if a patent referenced scientific publications identifying the usage of science through **non-patent references** within patent documents (Narin, Hamilton and Olivastro, 1997)
  - **Coupling** : indicates the degree to which an invention's components have been previously combined
    Measure is computed based on combinations of subclasses over the 200 previous years

Empirical analysis : science and technological search

- Control variables
  - Technology control - controls the fact that patents in some technologies receive on average more patents
  - # of prior art citations - citations to other patents
  - # of repeated trials : # of previous patents that combined exactly the same set of subclasses
  - % of patents that cite a non-patent reference in the same subclass Science stock :
- Estimation procedure
  - Negative binomial (Poisson but allowing for overdispersion, where variance exceeds the mean) : # citations is a count data

Empirical analysis : science and technological search

Table 2. Descriptive statistics of the variables used in the models

	Full sample $N = 16{,}822$		Cite science $N = 2{,}919$		Do not cite science $N = 13{,}903$	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Cites	2.75	3.54	3.43	4.25	2.60	3.34
Mean technology control	1.19	0.41	1.27	0.45	1.17	0.39
Number of prior art cites	7.63	6.99	7.90	8.66	7.27	6.04
Single subclass dummy	0.08	0.27	0.08	0.28	0.08	0.27
Number of subclasses	4.31	3.31	4.97	4.86	4.02	2.79
Number of classes	1.78	0.95	1.90	1.06	1.74	0.91
Number of trials	2.94	15.23	2.55	18.01	2.98	14.07
Recent technology	3.96	0.65	4.17	0.49	3.92	0.67
Science stock	0.21	0.22	0.45	0.24	0.16	0.17
Reference to science	0.18	0.38	1.00	0.00	0.00	0.00
Coupling	0.63	0.35	0.50	0.28	0.66	0.36

Empirical analysis : science and technological search

Table 4. Negative binomial estimates of citation counts (5-year window, standard errors in parentheses)[a]

	Model 1 baseline	Model 2	Model 3 10-year variable	Model 4
Mean technology control	**0.246**	**0.245**	**0.212**	**0.246**
	(0.036)	**(0.036)**	**(0.036)**	**(0.036)**
Number of prior art cites	**0.015**	**0.015**	**0.015**	**0.015**
	(0.001)	**(0.001)**	**(0.001)**	**(0.001)**
Single subclass dummy	**−0.186**	**−0.187**	**−0.201**	**−0.185**
	(0.036)	**(0.036)**	**(0.036)**	**(0.038)**
Number of subclasses	**0.027**	**0.027**	**0.029**	**0.027**
	(0.003)	**(0.003)**	**(0.002)**	**(0.003)**
Number of classes	**0.059**	**0.059**	**0.062**	**0.059**
	(0.009)	**(0.009)**	**(0.009)**	**(0.009)**
Number of trials	−0.000	−0.000	−0.001	−0.000
	(0.001)	(0.001)	(0.001)	(0.003)
Recent technology	**0.316**	**0.316**	**0.290**	**0.316**
	(0.016)	**(0.016)**	**(0.016)**	**(0.016)**
Science stock in subclass	0.083	0.085	0.073	0.083
	(0.052)	(0.051)	(0.051)	(0.052)
Cite to scientific publication	**0.098**	0.105	0.079	**0.098**
	(0.022)	(0.058)	(0.052)	**(0.022)**
Coupling	**0.462**	**0.485**	**0.632**	**0.471**
	(0.071)	**(0.076)**	**(0.079)**	**(0.078)**
Coupling^2	**−0.116**	**−0.136**	**−0.155**	**−0.121**
	(0.028)	**(0.030)**	**(0.036)**	**(0.032)**
Coupling × cite to scientific publication		−0.097	−0.023	
		(0.127)	(0.125)	
Coupling^2 × cite to scientific publication		**0.118**	**0.099**	
		(0.049)	**(0.051)**	
Number of trials × coupling				−0.000
				(0.003)
Number of trials × coupling^2				0.000
				(0.001)
Constant	**−1.891**	**−1.893**	**−1.783**	**−1.894**
	(0.084)	**(0.085)**	**(0.078)**	**(0.085)**
Fixed effects	363 Classes	363 Classes	363 Classes	363 Classes
Log-likelihood	−33,619.4	−33,614.5	−33,571.3	−33,619.4
N	16,822	16,822	16,822	16,822

[a] Bold type indicates that the coefficient estimate differs signification from zero with 95% confidence.

Empirical analysis : science and technological search

- Results :
  - Science has a positive effect on citations counts
  - Coupling has an inverted u-shape - positive but diminishing impact = detrimental effect of high coupling
  - Interaction with the number of trials : unlike science, experience does not mitigate the detrimental effects of high coupling
  - Interaction between coupling and scientific citations has a positive effect
  - In sum : strong evidence that the returns of science depend on the difficulty of the inventive problem (= highly coupled components) being addressed
- Limitations
  - Non-patent references is an indicator but it does not precisely indicate the degree to which the patent search has relied on science or basic research.

6. Recombinant search and the generation of breakthroughs

Section 6. Recombinant search and the generation of breakthroughs

6. Recombinant search and the generation of breakthroughs

- Does recombinant search favor the generation of breakthrough ideas ?
  - Some ideas or technologies have dramatically more impact than others
    - Einstein's theory of relativity
    - Watson and Crick's discovery of the double helix structure in DNA
- *High impact ideas* are cited at a substantially higher rate in future work

$=$ yield subsequent technological developments (new inventions) $\Rightarrow$ scientific and social progress

$\neq$ from *incremental work*

  - Overturning existing paradigms
  - Launching new areas of scientific inquiry
  - Radical inventions are the starting point of a new technology and a new technology cycle or paradigm
    - Radical inventions need further improvements, that is, incremental inventions/innovations that are decisive to transforming a radical invention into a marketable product or process

6. Recombinant search and the generation of breakthroughs

- Technological and scientific breakthroughs
  - Highly cited inventions
  - = Indicator of *technological value of inventions*
- Are these high impact ideas or breakthrough technologies the result of
  - successul connections forged between seemingly disparate bodies of knowledge = seemingly random connections occurring through a free associative process
  - $\Rightarrow$ Evolutionary process (Darwinian process ?) : individuals generate many unusual combination between different bodies of knowledge and go through a screening process of selective retention, keeping only the best variations
- Where do these high impact ideas come from ?

Technology familiarity, recombinant novelty, and breakthrough invention

Arts S. and R. Veugelers (2014), Technology familiarity, recombinant novelty, and breakthrough invention, Industrial and Corporate Change, 24(6), p. 1215-1246

Technology familiarity, recombinant novelty, and breakthrough invention

- **Question** : Relationship between the Recombinant search process and the quality/success of inventions (= breakthrough inventions)
  - Using familiar components : reusing technologies based on local search ⇒ learning traps
  - Technology brokering : create new inventions by combining previously uncombined technologies ⇒ Recombinant novelty
- **Aim** : Study the effect of combining formerly uncombined but familiar technology components on the likelihood of creating more useful inventions
- **Methodology** :
  - Econometrics - use of patent data over 26 years of the US patent record in biotechnology

Technology familiarity, recombinant novelty, and breakthrough invention

- **Dependent variables** : breakthrough patents = count of forward citations $\Rightarrow$ Top citations
- **Independent variables** :
  - **New combinations** [0,1] : measure of recombinant novelty = # of new combinations over total number of pairs
  - **Component familiarity** : reuse of familiar components - components used in prior inventions = Average number of times a subclass has been previously used
- Controls : team size, average experience, experience diversity, network size, patent characteristics
- **Model**

$$\begin{aligned}\text{Breakthrough}_{it} &= \beta_0 + \beta_1 \text{New combinations}_{it} + \beta_2 \text{Component familiarity}_{it} \\ &\quad + \beta_3 \text{New combinations}_{it} \times \text{Component familiarity}_{it} + \text{Controls}_{it} + \epsilon_{it}\end{aligned} \tag{2}$$

Empirical analysis of the link between search and breakthrough

Table 1 Descriptive statistics ($n=84{,}119$)

Variable	Description	Mean	Stdev.	Min.	Max.
Breakthrough	Binary: 3 standard deviation outlier in distribution of forward citations	0.015	0.122	0.00	1.00
Breakthrough 2 stdev	Binary: 2 standard deviation outlier in distribution of forward citations	0.029	0.169	0.00	1.00
Breakthrough 4 stdev	Binary: 4 standard deviation outlier in distribution of forward citations	0.008	0.092	0.00	1.00
Forward citations	Number of forward citations received within 5 years	2.86	5.86	0.00	201.00
Failure	Patent received no forward citations within 5 years	0.36	0.48	0.00	1.00
New combinations	The focal patent's number of pair-wise subclass combinations which appear for the first time in history divided by the focal patent's total number of pair-wise subclass combinations	0.19	0.30	0.00	1.00
Component familiarity	Recent and frequent usage of the focal patent's subclasses by all prior US patents (see Fleming, 2001)	4.71	1.96	0.00	8.65
Team size	Number of inventors	1.27	0.42	0.69	3.50
Average experience	The average number of prior patents by the focal patent's inventors	1.33	1.02	0.00	6.00
Experience diversity	The number of technology classes at least one of the focal patent's inventors has patented in before	1.57	0.95	0.00	5.04
Network size	The number of inventors at least one of the focal patent's inventors has collaborated with before	1.71	1.34	0.00	6.22
Patent references	The number of backward patent citations	1.37	0.99	0.00	6.60
Non-patent references	The number of citations to non-patent literature	2.26	1.41	0.00	6.98
Number of classes	The number of technology classes	0.68	0.48	0.00	2.77
Number of subclasses	The number of technology subclasses	1.62	0.67	0.00	4.50
Single subclass	Binary: single technology subclass	0.04	0.20	0.00	1.00
Newest subclass	The minimum number of previous uses among the focal patent's subclasses	3.11	1.72	0.00	8.65
New subclass	Binary: at least one subclass appears for the first time in history	0.07	0.26	0.00	1.00

All explanatory count variables are logged after adding 1 for those variables with zero values

Empirical analysis of the link between search and breakthrough

Table 3 Probit models of technology breakthrough

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Variables	Breakthrough	Breakthrough	Breakthrough	Breakthrough 2stdev	Breakthrough 4stdev	Breakthrough	Breakthrough	Breakthrough	Breakthrough
Sample	1976–2001	1976–2001	1976–2001	1976–2001	1976–2001	1990–2001	1976–2001	New combinations > 0	Exclusive failures
New combinations		0.4819*** [0.079]	0.6121*** [0.064]	0.5413*** [0.052]	0.6948*** [0.081]	0.5673*** [0.078]	0.6096*** [0.068]	0.3768*** [0.098]	0.5898*** [0.068]
Component familiarity		0.0697*** [0.018]	0.0288 [0.027]	0.0176 [0.018]	0.0299 [0.032]	0.0202 [0.031]	0.0280 [0.028]		0.0299 [0.029]
New combinations *Component familiarity			0.0968*** [0.033]	0.1014*** [0.023]	0.1012** [0.041]	0.0969*** [0.035]	0.0927** [0.041]		0.0880*** [0.034]
Team size	−0.0369 [0.045]	−0.0290 [0.045]	−0.0280 [0.045]	−0.0504 [0.036]	−0.0230 [0.054]	−0.0527 [0.053]	−0.0281 [0.045]	−0.0039 [0.056]	−0.0194 [0.047]
Average experience	−0.1094*** [0.041]	−0.1022** [0.040]	−0.0952** [0.040]	−0.0928*** [0.030]	−0.1218** [0.054]	−0.0839* [0.049]	−0.0952** [0.040]	−0.1295*** [0.039]	−0.0923** [0.042]
Experience diversity	0.2013*** [0.038]	0.1982*** [0.037]	0.1910*** [0.037]	0.2072*** [0.028]	0.2039*** [0.046]	0.1619*** [0.046]	0.1910*** [0.037]	0.2156*** [0.044]	0.1787*** [0.039]
Network size	−0.0269 [0.029]	−0.0268 [0.029]	−0.0294 [0.029]	−0.0348 [0.022]	−0.0196 [0.034]	−0.0020 [0.035]	−0.0294 [0.029]	−0.0097 [0.029]	−0.0228 [0.031]
Patent references	0.1835*** [0.020]	0.1910*** [0.020]	0.1842*** [0.020]	0.2051*** [0.016]	0.1652*** [0.022]	0.1901*** [0.023]	0.1841*** [0.020]	0.1861*** [0.022]	0.1627*** [0.021]
Non-patent references	0.0848*** [0.015]	0.0817*** [0.016]	0.0851*** [0.016]	0.0903*** [0.012]	0.0936*** [0.020]	0.0776*** [0.020]	0.0851*** [0.016]	0.0881*** [0.015]	0.0816*** [0.016]
Number of classes	0.3136*** [0.043]	0.2427*** [0.046]	0.2459*** [0.046]	0.2714*** [0.033]	0.2790*** [0.057]	0.1818*** [0.055]	0.2461*** [0.046]	0.3534*** [0.055]	0.2505*** [0.047]
Number of subclasses	0.1146*** [0.033]	0.1288*** [0.032]	0.1316*** [0.032]	0.0959*** [0.024]	0.1254*** [0.039]	0.1078*** [0.037]	0.1316*** [0.032]	0.1854*** [0.040]	0.1178*** [0.034]
Single subclass	0.2626*** [0.097]	0.3725*** [0.102]	0.3266*** [0.096]	0.2850*** [0.072]	0.4064*** [0.119]	0.2711** [0.108]	0.3282*** [0.098]		0.3120*** [0.101]
Newest subclass	−0.0106 [0.014]	−0.0018 [0.017]	0.0175 [0.018]	0.0155 [0.013]	0.0386* [0.021]	0.0268 [0.020]	0.0173 [0.018]	0.0486* [0.026]	0.0196 [0.019]
New subclass							−0.0190 [0.122]		
Familiarity components used in new combinations								0.0289* [0.017]	
Year fixed effects	Inclusive	Inclusive	Inclusive	Inclusive	Inclusive	Inclusive	Inclusive	Inclusive	Inclusive
Technology fixed effects	Inclusive	Inclusive	Inclusive	Inclusive	Inclusive	Inclusive	Inclusive	Inclusive	Inclusive
Log likelihood	−5643.1052	−5612.6824	−5603.4974	−9606.3937	−3446.8549	−4222.2669	−5603.49	−3564.5516	−5209.7608
Observations	84,119	84,119	84,119	84,119	84,119	62,628	84,119	40,521	53,586

Variables used in interaction terms are centered, Robust standard errors in brackets, Clustered at assignee level, $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$

Empirical analysis of the link between search and breakthrough

- **Results**
  - Role of technology brokering, i.e. the creation of new inventions by combining formerly disparate technology components, as a key search process leading to more useful inventions
  - Breakthroughs most likely originate from creatively recombining familiar technologies in new ways
  - Exploratory search process leading to the discovery of breakthroughs involves both recombinant novelty and the reuse of familiar technology.
  - Recombinant novelty allows to avoid the familiarity trap associated with using very familiar components.
- **Limits**
  - The identification of brokering does not explicitly take into account whether the formerly uncombined technologies come from otherwise disconnected industries and markets
  - Limited to industry using patents and specifically biotechnology industry

Summary and open questions

- Technological progress is based on recombination of partly familiar and partly unfamiliar technological components
- Technological progress is based on a search process based on search scope and search depth
- Radical inventions/ Breakthrough inventions are based on rather uncommon and original recombinations
- **Open questions and PHD Subjet :**
  - Do some organizations have more facility or opportunities to generate radical inventions
    - Small versus large firms ; Diversified firms ; Firms engaged in national and international collaborations ; Firms occupying brokerage positions within networks ?

Readings

- Katila, R. and Ahuja, G., (2002), Something old, something new : a longitudinal study of search behavior and new product introduction, Academy of Management Journal, 45(6), 1183–1194
- Fleming, L. (2001), Recombinant Uncertainty in Technological Search, Management Science, 47(1), 117-132, 2001
- Fleming, L. and Sorenson, O. (2004), Science as a map in technological search, Strategic Management Journal, (2004)
- Arts S. and R. Veugelers (2014), Technology familiarity, recombinant novelty, and breakthrough invention, Industrial and Corporate Change, 24(6), 1215-1246
- Strumsky, D., Lobo, J. (2015), Identifying the sources of technological novelty in the process of invention, Research Policy, 44, 1445 - 1461
- Verhoeven D., Bakkera J., Veugelers R. (2016), Measuring technological novelty with patent-based indicators, Research policy, 707-723
- Tushman M. and Anderson P. (1986), Technological discontinuities and organizational environment, Administrative Science Quarterly, 31 (3), 439-465