A Deeper Understanding of Technology Evolution Using Cluster Destructiveness Index

Case Study of the Pharmaceutical Sector

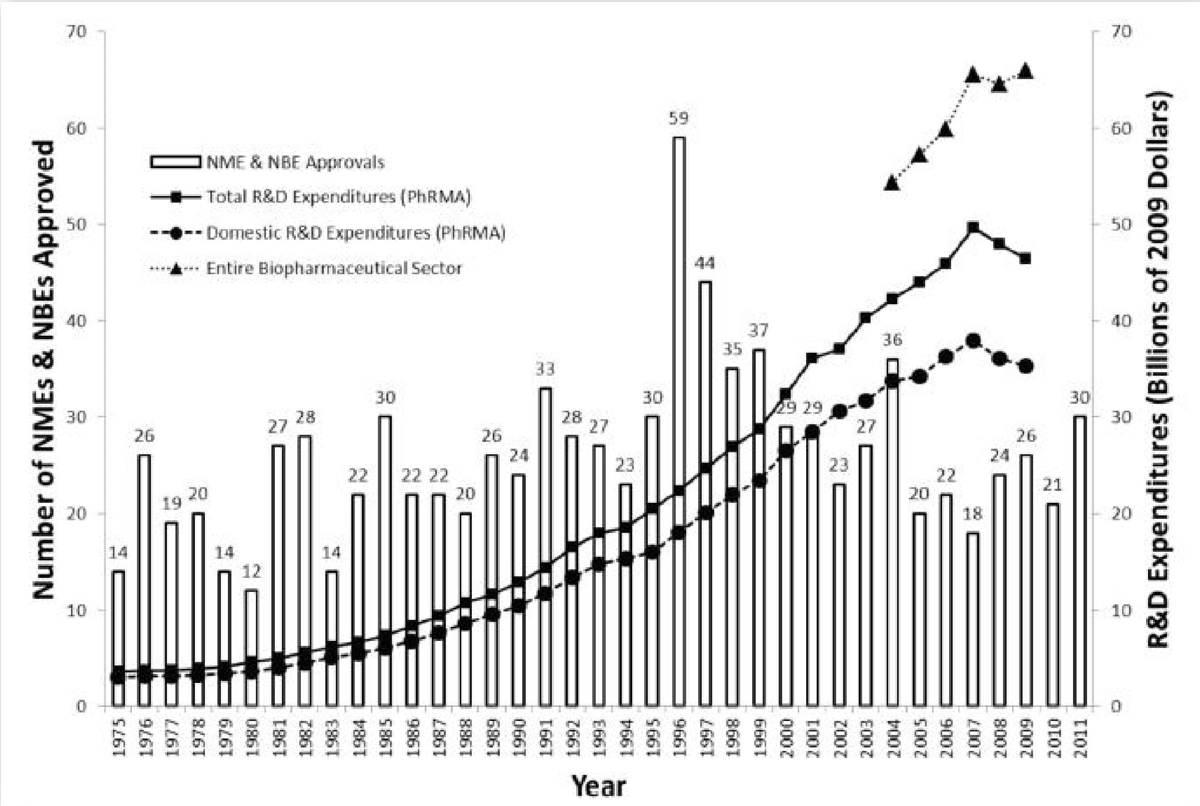
Yuan Gao
Emiliya Lazarova
School of Economics, University of East Anglia

Exploration is the engine that drives innovation. Innovation drives economic growth.

- Edith Anne Widder

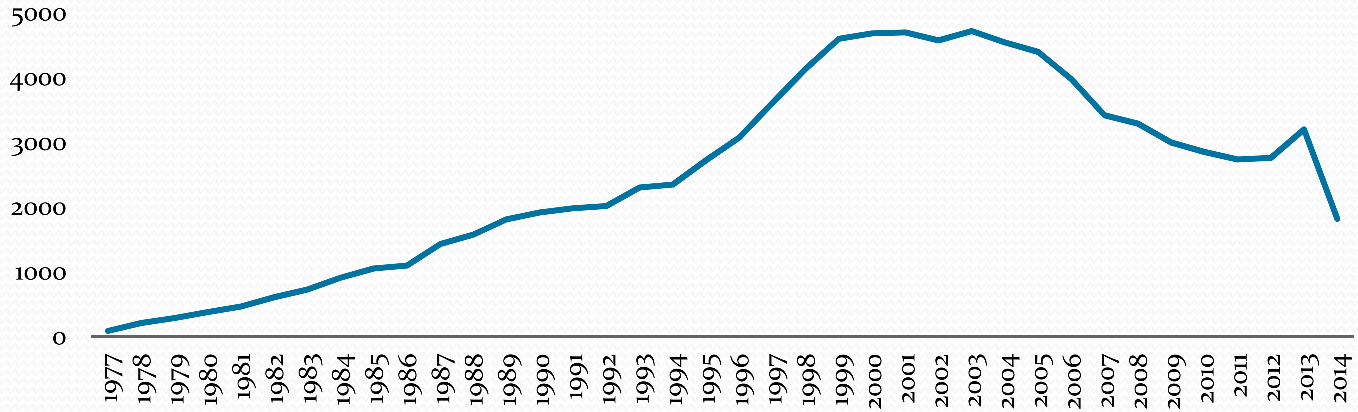
What's happening in U.S. bio-pharma innovation ecosystem?

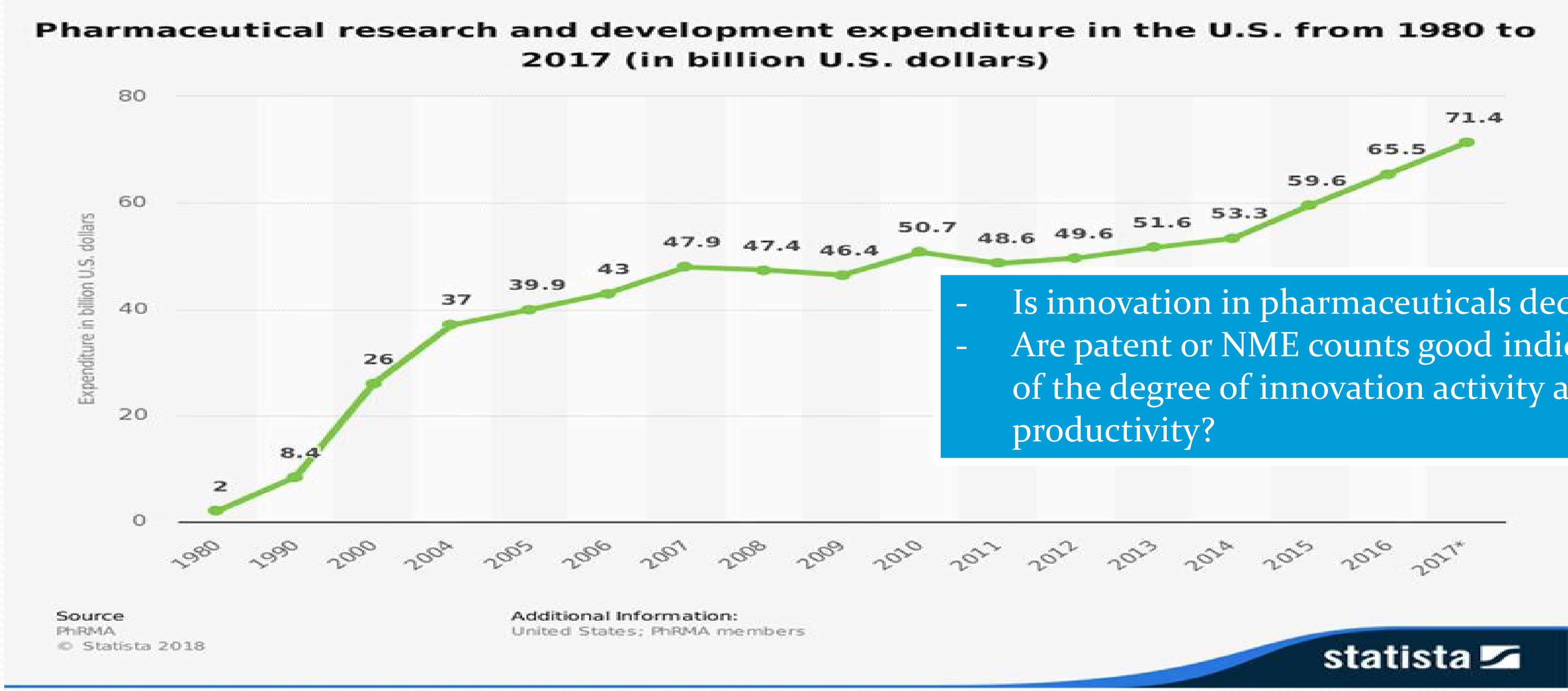
FIGURE 1. (a) Annual NME/NBE Approvals versus R&D Expenditures in 2009 Dollars.
(b) Estimates of Cost to Companies per New Molecular Entity in 2009 Dollars.



Source: President's Council of Advisors on Science and Technology (US). (2012). *Report to the president on propelling innovation in drug discovery, development, and evaluation*. Executive Office of the President, President's Council of Advisors on Science and Technology.

Number of pharmaceutical patent applications in the U.S.





- Is innovation in pharmaceuticals declining?
- Are patent or NME counts good indicators of the degree of innovation activity and productivity?

Destructive technologies

Literature

- A perspective to observe technology evolution and evaluate innovation based on the degree of novelty
- Creative Destruction: a process that *incessantly* revolutionizes the economic structure *from within*, incessantly *destroying* the old ones, incessantly *creating* a new one (Joseph A. Schumpeter, 1942)
- Disruptive technologies: not distinguished by their complexity or novelty, but by a different *performance attributes package* not valued by existing customers and rapidly improving to penetrate the established market (Joseph L. Bower and Clayton M. Christensen, 1995)
- Technological novelty can be characterized by novelty in *recombination* and novelty in technological and scientific knowledge origins (Dennis Verhoeven et. al., 2016)

Disruptive technologies

Our theories

- *Recombination:* knowledge, method and application

  A disruptive innovation reshapes the technological framework either by
  - recombining the existing knowledge origins (new method), or
  - recombining a same objective with newly developed technologies or technologies not previously used for this purpose (new knowledge origins), or
  - recombining the same set of techniques or functions with a different market segment or purpose (new application).
- Schumpeter's "Innovation Trilogy": Invention -> Innovation -> Diffusion

  Our conjecture of *Technology Evolution Cycles*:

- Active in recombinant innovation
- Not necessarily productive
- Could be led by existing major market players or emerging startups

- Concentrate on trends identified and established through exploration
- Highly productive
- Big players often stand out

Patent data analysis

Literature

- A widely recognized data source in quantitative assessment of technological novelty and values (Zvi Griliches, Lee Fleming and other colleagues at NBER, from 1990s)

- The most used conventional measures are simple “counts” (e.g. numbers of applications, citations) and composite indices based on them (Mariagrazia Squicciarini et. al., 2013)

Patent data analysis

Our approach

- A *network* constructed by technologies (nodes) coexisting in the same patents (edges)
- Identify natural network clustering
- Measure the network partitions' recombination over time to evaluate the degree of *"destructiveness"*

A more generic approach which captures the emergence of technological recombinations.

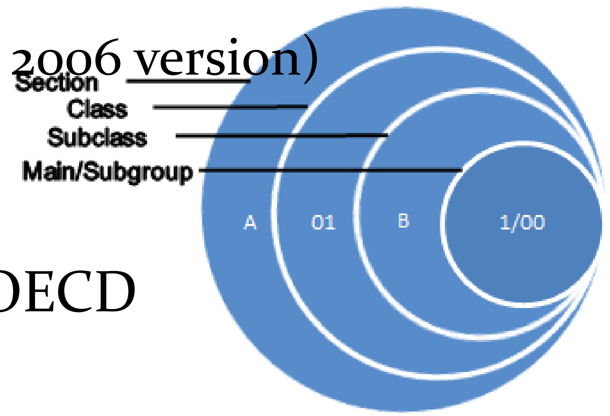
Innovation in U.S. pharmaceuticals

- The pharmaceutical patenting activity (relative to overall patenting activity) began to increase in the mid-1990s.
- The innovation ecosystem for public health is under significant stress and R&D productivity is declining. (PCAST, 2012, Danzon et. al., 2005; DiMasi et. al., 1991-2002)
  - Federal support and venture capital for startups are declining
  - Scientific knowledge gaps between basic research and commercial projects
  - Inefficiency in clinical trials
  - High attrition rates
  - Companies are exiting important fields of critical public health need
- While the output of new drugs has remained constant, total R&D investment by industry in drug discovery and development have grown exponentially, in inflation-adjusted terms.

Data

- Patent data from the OECD International Patent Dataset (Feb., 2019 release)

- The IPC technology classification system (WIPO, 2006 version)
  - 4-digit subclass codes used: 639 in total



- Leading R&D companies data from the JRC-OECD COR&DIP database 2017 and 2019

Method
I. Network Construction

Technology-Cohort network

- Based on different IPC subclasses assigned to the same patent
- Nodes: IPC subclasses
- Edges: number of patents every two different subclasses co-assigned to
- Aggregated with all the patent applications filed in the same year

Patent 1
- A01A
- A01B
- A01C

Patent 2
- A01A
- A01B

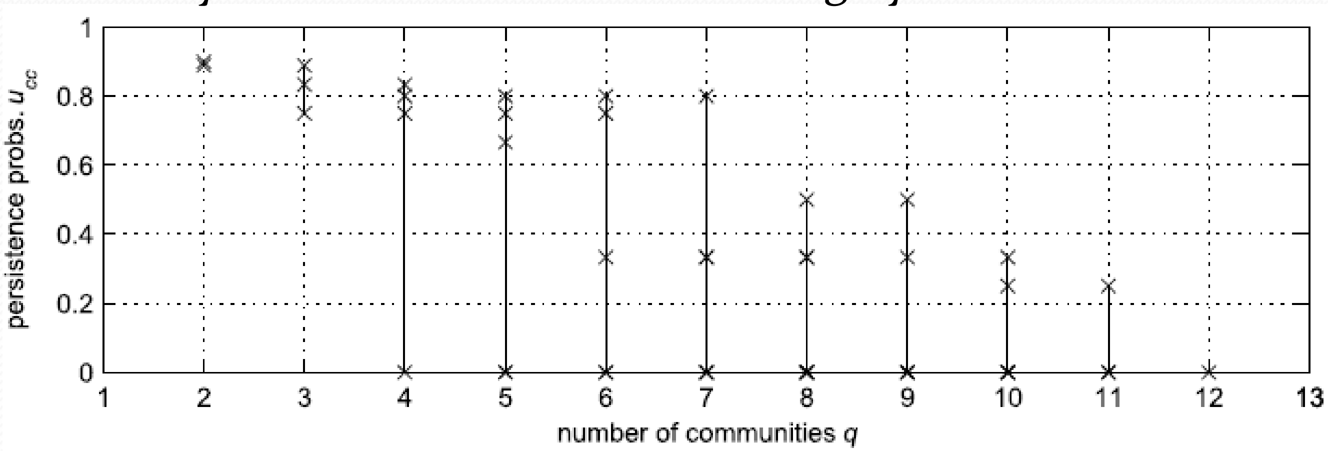
	A01A	A01B	A01C
A01A	0	2	1
A01B	2	0	1
A01C	1	1	0

Method

II. Network cluster identification

- Cluster identification based on lumped Markov chains random walking algorithm (Carlo Piccardi, 2011)
  - Clusters are identified based on a configurable threshold, the *Persistence Probability* U_{CC} associated to a cluster C not smaller than α (0< α<1) $U_{CC} = \frac{\sum_{i,j \in C} \pi_i p_{ij}}{\sum_{i \in C} \pi_i}$, *$\pi_i$ is the probability of being in node i (at time t)*

    p_{ij} is the probability of random walking from node i to j
  - A sudden drop of U_{CC} indicates the breaking of a significant "natural" community – alternative threshold setting by number of clusters



Method

III. Computation of Destructiveness

- The Cluster Destructiveness Index (CDI) measures the extent of technological recombination of each cluster in a given time window based on the cluster's nodes diversification of nodes in terms of different communities in the previous time window.

  We denote CDI of time window *t-s* to *t* by C_t:

$$C_t = \sum_{i=1}^{n_t} \frac{P^t_{t-s}(i)}{\underbrace{\sum_{j=1}^{n_{(t-1)}} \left(\frac{|N^{t-1}_{t-s-1}(j) \cap N^t_{t-s}(i)|}{|N^t_{t-s}(i)|} \right)^2}_{\text{Destructiveness of cluster } N^t_{t-s}(i)}}$$

n_t: number of identified clusters in time window *t-s* to *t*

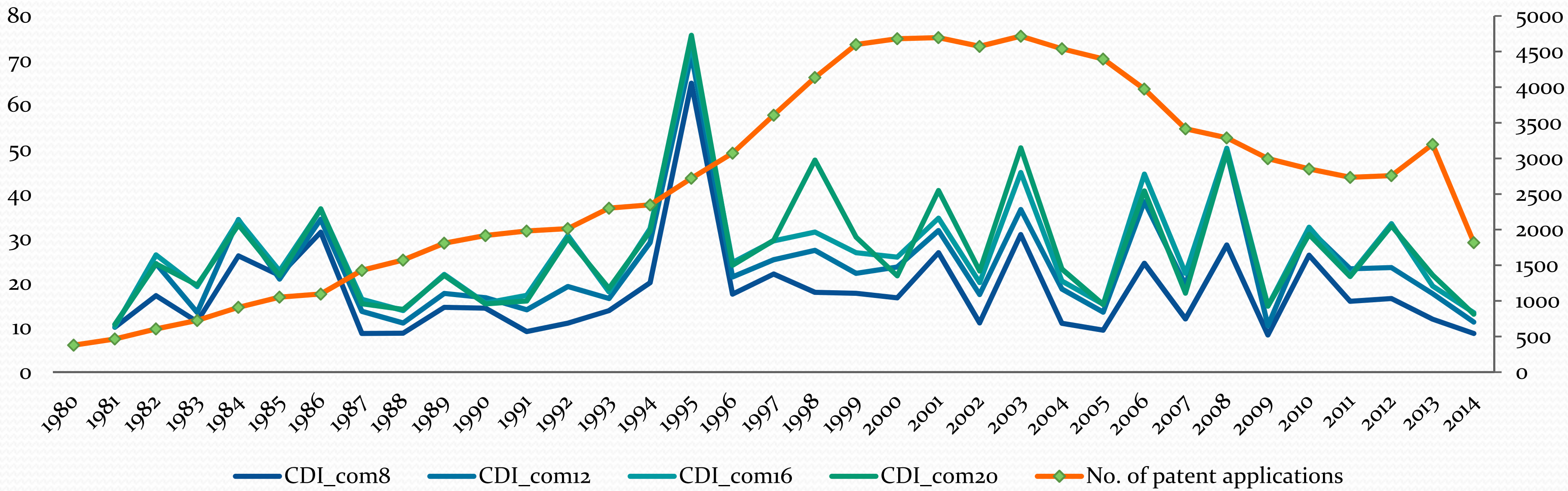
$N^t_{t-s}(i)$: the i_{th} cluster of time window *t-s* to *t*

$P^t_{t-s}(i)$: Persistence Probability of cluster $N^t_{t-s}(i)$

- We used time window length of *1 year* in our analysis and calculated CDI with different cluster numbers

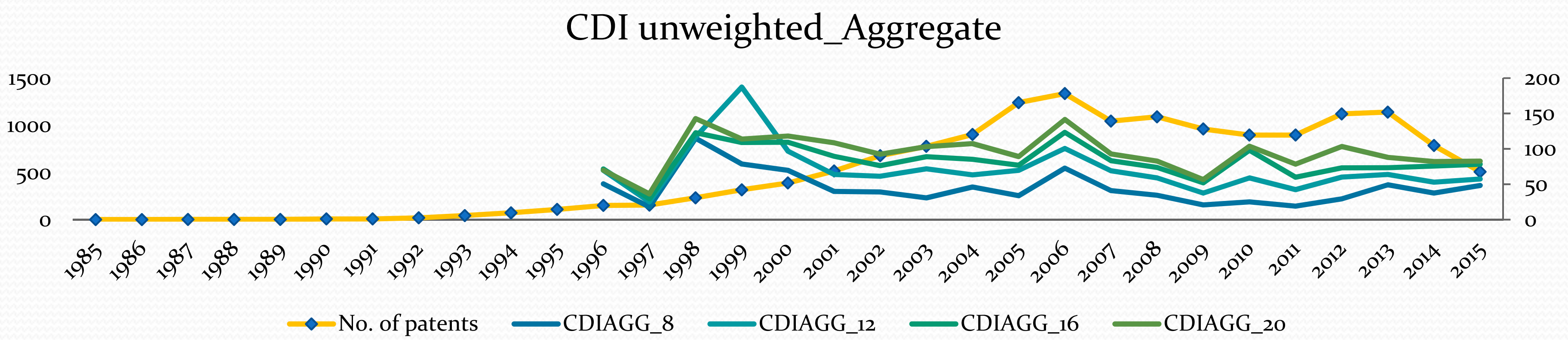
Results at country level

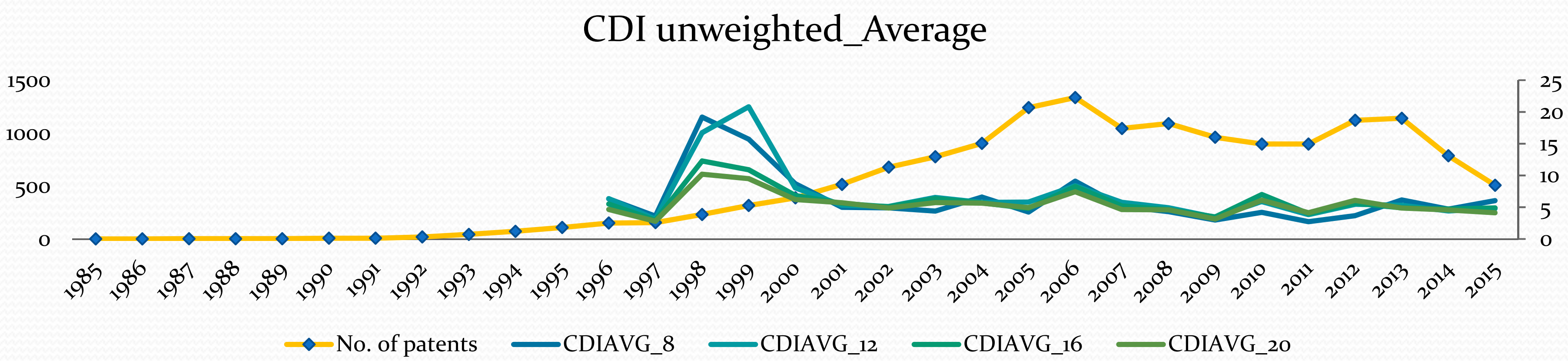
U.S. 1981-2014



Results at company level

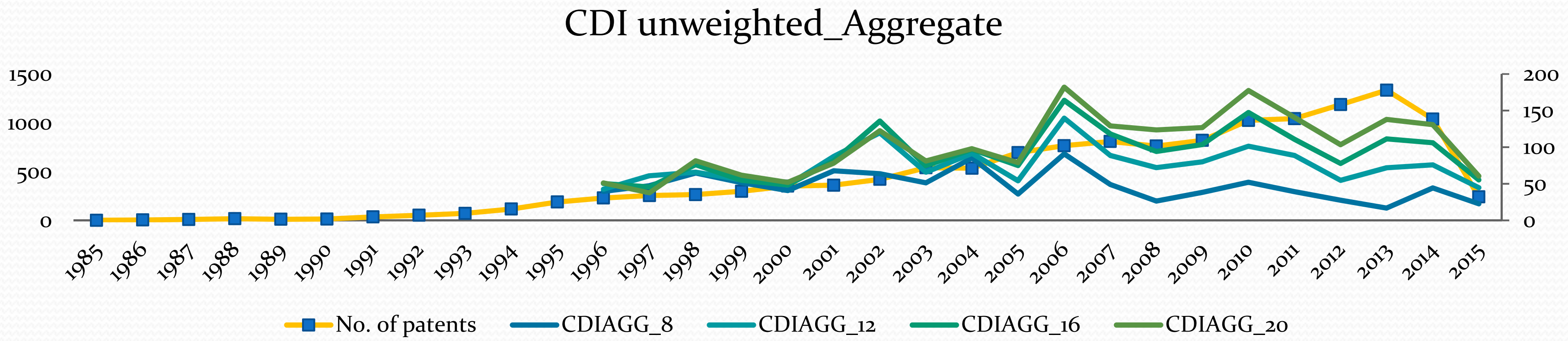
Johnson & Johnson_1996-2015

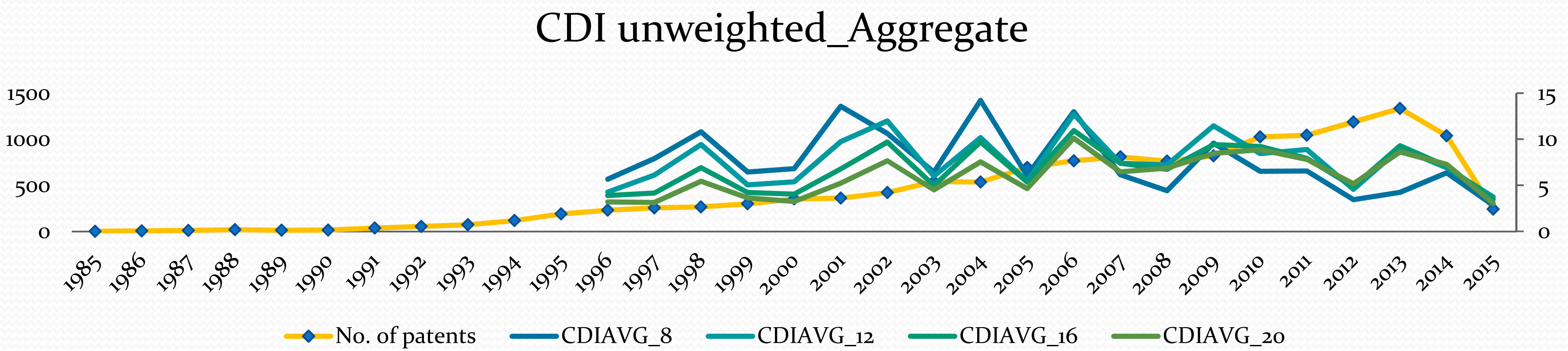




Results at company level

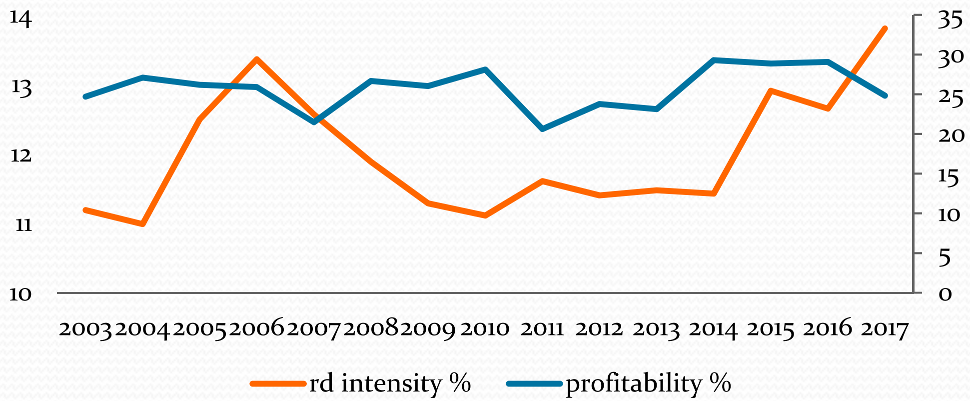
Roche_1996-2015

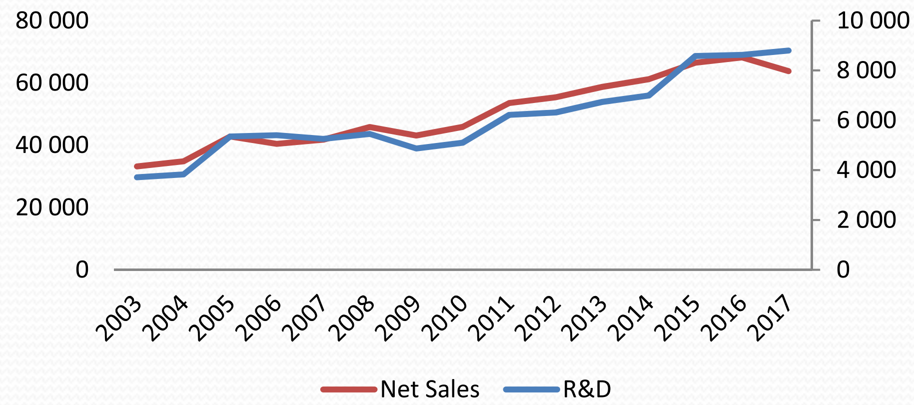




Results at company level: Johnson & Johnson
Correlation with R&D expenditure and profitability

Arbitrarily set 5 years as the time needed for R&D expenditure to take effect: 2008-2015





	ptqty	cdiagg_8	cdiavg_8	rd intensity	profitability
ptqty	1				
cdiagg_8	-0.2694	1			
cdiavg_8	-0.2237	0.9711	1		
rdintensity	0.2627	-0.4169	-0.4621	1	
profitability	-0.5563	0.3913	0.3201	-0.6932	1

	ptqty	cdiagg_8	cdiavg_8	rd	ns
ptqty	1				
cdiagg_8	-0.2694	1			
cdiavg_8	-0.2237	0.9711	1		
rd	-0.148	0.2752	0.1992	1	
ns	-0.5272	0.6589	0.7233	0.5561	1

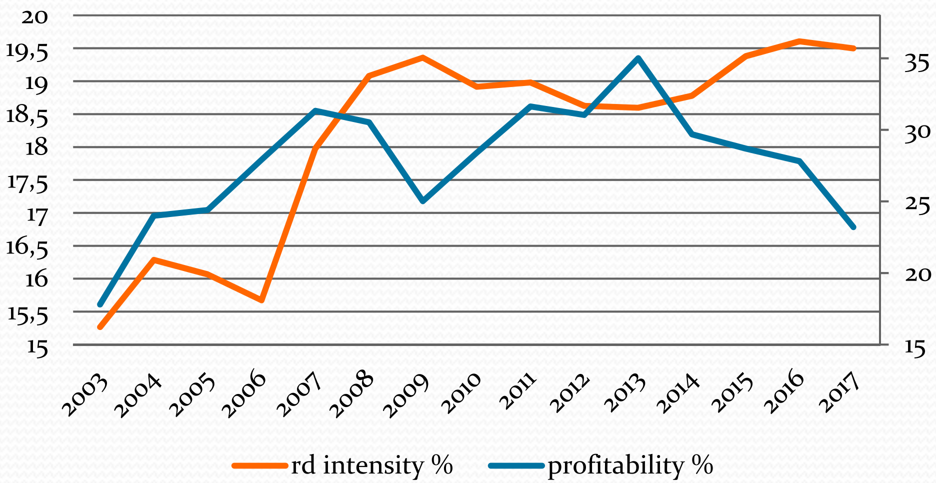
rd: R&D expenditure rd intensity: rd/ns *use data from 2003-2010 for these 2 variables

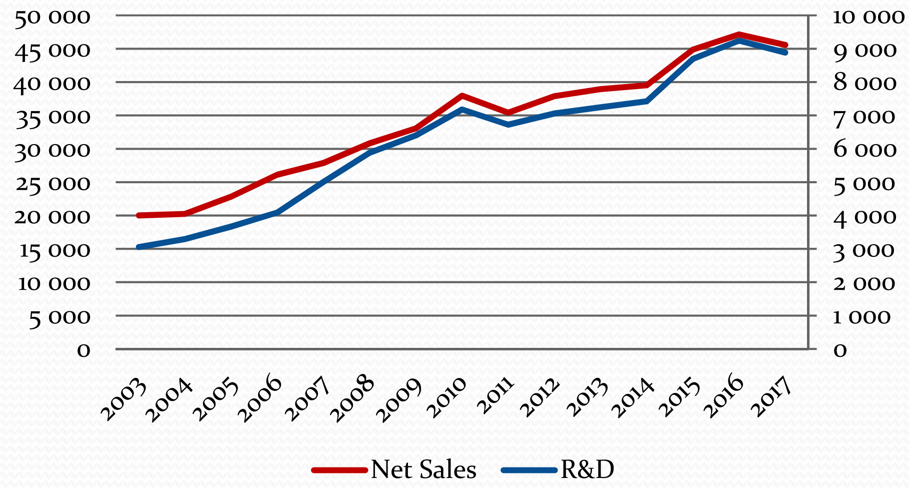
ptqty: patent application quantity

ns: net sales profitability: operation revenue/ns * in MM$

Results at company level: Roche
Correlation with R&D expenditure and profitability

Arbitrarily set 5 years as the time needed for R&D expenditure to take effect: 2008-2015





	ptqty	cdiagg_8	cdiavg_8	rd intensity	profitability
ptqty	1				
cdiagg_8	0.1638	1			
cdiavg_8	0.1251	0.6748	1		
rdintensity	0.0395	-0.4294	-0.365	1	
profitability	0.5416	-0.5895	-0.5235	0.2996	1

	ptqty	cdiagg_8	cdiavg_8	rd	ns
ptqty	1				
cdiagg_8	0.1638	1			
cdiavg_8	0.1251	0.6748	1		
rd	-0.1965	-0.5474	-0.3687	1	
ns	-0.2765	-0.5025	-0.1721	0.9033	1

rd: R&D expenditure rd intensity: rd/ns *use data from 2003-2010 for these 2 variables

ptqty: patent application quantity

ns: net sales profitability: operation revenue/ns * in MM$

Conclusion

We propose:

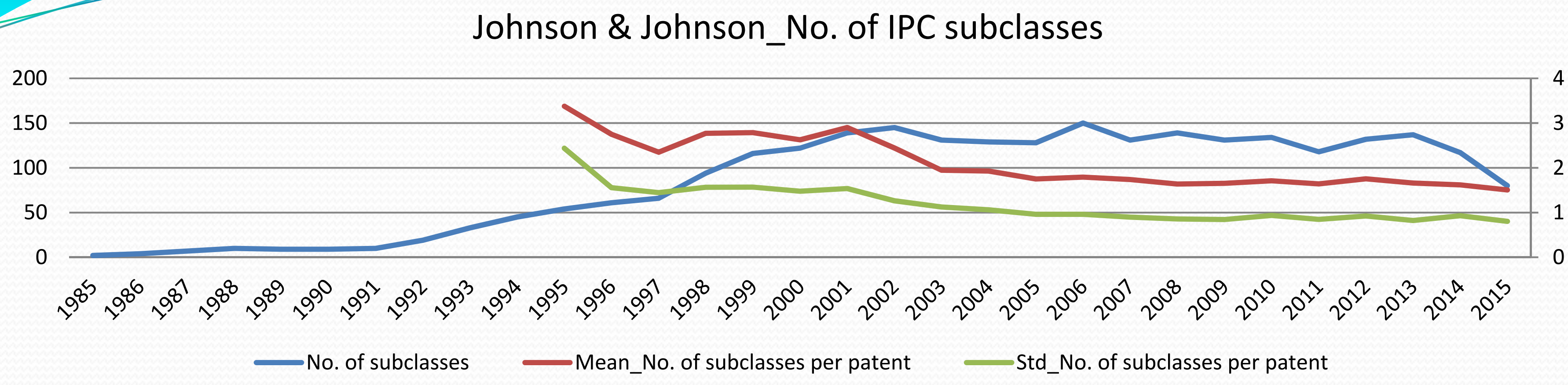
- A refined definition of disruptive technologies with a focus on *recombination*
- A study approach based on network analysis with patent data and an index to measure technology cohort destructiveness
- Conjectures of technology evolution cycles consisting of *exploration and intensification* stages

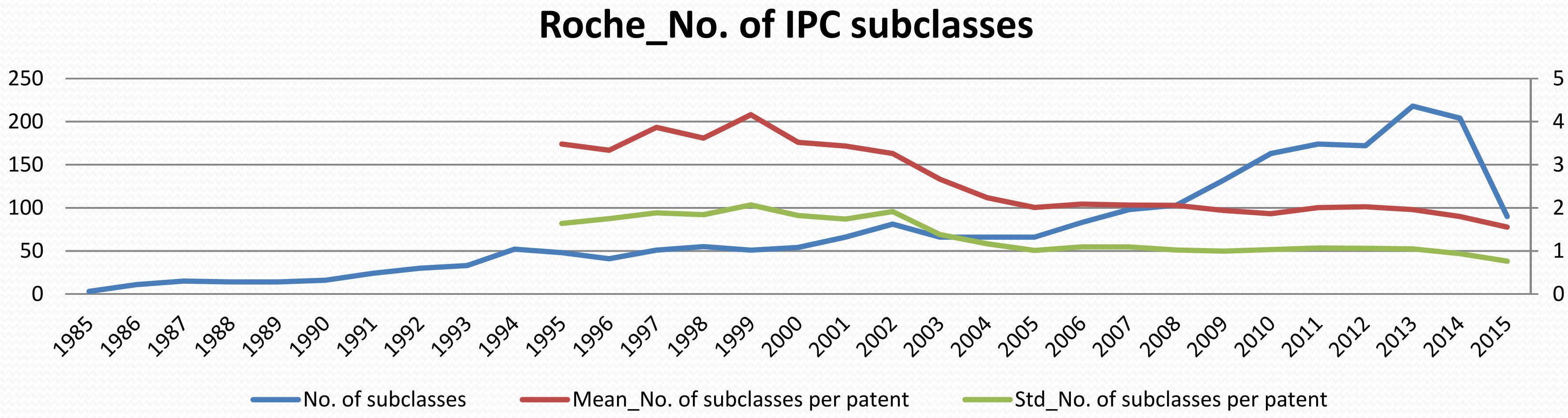
Through a case study of the pharmaceuticals sector in the U.S. and of 2 leading R&D companies:

- CDI tends to increase as patent applications increase, and starts to drop as patent quantity continues to approach the peak.
- At country level, CDI helps explain the inconsistency in conventional patent indicators and aligns with NME and industry expert opinions.
- At company level, the relationships among CDI, R&D investment and profit are not yet clear, mainly due to the lagging effect from investment to output in drug developments, and the extra complications with clinical trials and drug approvals.

Future work will involve more countries and look to better define the company level variables. We will also include patent citations as a quality indicator.

Understanding of the technology evolution cycles helps governments and market players anticipate the innovation trajectory, identify potential disruptive technologies, and be more informed in strategic planning.


Johnson & Johnson_No. of IPC subclasses
200
150
100
50
0
4
3
2
1
0
1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
No. of subclasses
Mean_No. of subclasses per patent
Std_No. of subclasses per patent


Roche_No. of IPC subclasses
250
200
150
100
50
0
5
4
3
2
1
0
1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
No. of subclasses
Mean_No. of subclasses per patent
Std_No. of subclasses per patent