DRUID academy

Paper to be presented at the DRUID Academy conference in Rebild, Aalborg, Denmark on January

21-23, 2015

Generation and Diffusion of Technological Novelty in Biotechnology

Manuel Gigena
KU Leuven. VAT: BE 0419.052.173
Department of Managerial Economics, Strategy and Innovation
manuel.gigena@kuleuven.be

Ashish Arora
Duke University
The Fuqua School of Business
ashish.arora@duke.edu

Dennis Verhoeven
KU Leuven VAT: BE0419052173
Managerial Economics, Strategy and Innovation
Dennis.Verhoeven@kuleuven.be

Reinhilde Veugelers
KU Leuven
Managerial Economics, Strategy and Innovation
Reinhilde.Veugelers@kuleuven.be

Abstract

Generation and Diffusion of Technological Novelty in Biotechnology

Authors: Ashish Arora, Manuel Gigena(1), Dennis Verhoeven, Reinhilde Veugelers

(1)Submitting author
Affiliation: Managerial Economics, Strategy and Innovation, Faculty of Economics and Business, KU Leuven
Year of enrolment-finishing: 2013-2016 (currently in the second year)
e-mail address: manuel.gigena@kuleuven.be

State of the art
Inventive activity resulting in the creation and development of novel approaches to serve human purposes is widely acknowledged to be the key to Schumpeterian creative destruction and economic growth. Despite some anecdotal

evidence, prior literature does not provide a clear understanding of how the process of radical innovation unfolds.

Research gap
Technological development proceeds along trajectories, guided by technological paradigms (Dosi, 1982). The differences in the nature of technological activities and uncertainty of their benefits might translate into differences in the ability and incentives of actors to be active along the different stages. We are interested in the origins of technological novelty and its diffusion over time. In particular, we investigate the role of different types of players along the process.

Theoretical arguments
Large, incumbent firms are usually believed to have difficulties coming up with radical inventions (Henderson, 1993). They are claimed to have a disadvantage from the organizational viewpoint, as well as less economic motivation to come up with, and to introduce, novelty. Conversely, small firms are argued to suffer less from such disadvantagesâ??they are less prone to learning traps, and less concerned with cannibalizing well-established competences. Finally, universities, in theory, are not bounded by the competitive pressures that demotivate private firms to engage in research surrounded by high uncertainty. As such, their role in the process of radical innovation might be one of diversity-generating experimentation. Overall, this capabilities-based view leads to a division of labor wherein large firms are more efficient at commercializing novelty, whereas small firms may either be creators of novelty, or intermediaries between universities and large firms.

Methodology
We characterize three types of actors: universities, small firms and large firms. Moreover, we characterize inventions as introducing novelty or not. Within the group of novel inventions, we make a further distinction based on the level of diffusion of the novelty introduced. In addition, we investigate the role of different players as the technological trajectory is developed, i.e. in the early â?? intermediate â?? late stage of diffusion.
We measure inventive activity using patent documents. Our preliminary data includes all granted USPTO patent documents in biotechnology applied between 1988 and 1992. Patents that are classified in two IPC-codes (group level) which have never occurred together before, are used to proxy inventions introducing a new approach (Fleming, 2001). To track follow-up inventive activity to new approaches introduced by novel inventions, we identify â??followerâ?? patents, reusing the newly introduced IPC-pair. In order to characterize the actors involved in novelty generating/using patents, we use information on the sector and size of the applicant. Sectors are allocated using the algorithm developed at MSI which classifies each applicant as a company, university, or others. To identify firm size we use information on the possibility of fee reduction for small entities at the USPTO.

Results
Preliminary findings suggest a distinct role of universities, small firms and large firms in the generation and development of technological novelty. While the lionâ??s share of patents, novel patents and follow-up patents using novelty is to be attributed to large firms, it seems they have a relative disadvantage to generate novel patents that become highly diffused. Small firms are at a relative advantage to create novel patents, and within the group of novel patents, they have an advantage in generating the ones with high diffusion. Universities are not more likely to generate novel patents, but if they do, these patents are relatively more often among the most highly diffused ones.

References
Dosi, Giovanni. "Technological paradigms and technological trajectories: a suggested interpretation of the determinants and directions of technical change." Research policy 11.3 (1982): 147-162.
Fleming, Lee. "Recombinant uncertainty in technological search." Management science 47.1 (2001): 117-132.
Henderson, Rebecca. "Underinvestment and incompetence as responses to radical innovation: Evidence from the photolithographic alignment equipment industry." The RAND Journal of Economics (1993): 248-270.

Jelcodes:O33,-

Generation and Diffusion of Technological Novelty in Biotechnology

Ashish Arora – Manuel Gigena[1] – Dennis Verhoeven – Reinhilde Veugelers

Abstract

Despite anecdotal evidence, literature does not provide a clear understanding of how the process of radical innovation unfolds. A capabilities-based view leads to a division of labor wherein large firms are more efficient at commercializing novelty, whereas small firms may either be creators of novelty, or intermediaries between universities and large firms. Using patent-based indicators, we sketch a framework to study the role of different sectors in the generation of technological novelty, its diffusion, and approach its commercial application.

Introduction

Inventive activity resulting in the creation and development of novel approaches to serve human purposes is widely acknowledged to be the key to Schumpeterian creative destruction and economic growth. Despite some anecdotal evidence, prior literature does not provide a clear understanding of how the process of radical innovation unfolds.

Dosi (1982) describes how technological development proceeds along trajectories, guided by technological paradigms. A certain approach to solve a technological problem is proposed and follow-up inventive effort refines this approach, adhering to the paradigm introduced by the novel invention. This technological trajectory continues to evolve until a new approach is introduced, shifting the technological paradigm. Along the different stages of the trajectory the technological activities might be substantially different in nature. While novel approaches and their early development might result from experimentation with a wide variety of possible solutions, the activities in the later stage provide incremental improvements within the bounds of what is known about the technology. During the early stages it is far from clear to what extent the new approach will make it to a final application, while during mature development applications might already be present and the outcome of improvements are relatively clear.

[1] Contact details are included at the end of the document.

The creation and diffusion of novelty

The differences in the nature of technological activities and uncertainty of their benefits might translate into differences in ability and incentives of actors to be active along the different stages of a technological trajectory. Small firms contribute disproportionately to the class of breakthrough inventions, while large firms dominate the subsequent stages of their refinement through incremental innovation (Baumol, 2004). The literature finds some of the sources of this division of labor in firms' capabilities as well as in market structure. Large, incumbent firms are usually believed to have difficulties coming up with radical inventions (Henderson, 1993). On the one hand, they have a disadvantage from the organizational viewpoint. Introducing radical innovation requires severely different organizational capabilities compared to incremental improvements, as such, large firms are believed to suffer from learning traps (Ahuja & Lampert, 2001). These traps include searching in familiar technological fields (familiarity trap), searching in mature technological fields (maturity trap) and searching for solutions near existing ones (propinquity trap). On the other hand, they are claimed to have less economic motivation to come up with, and introduce radical innovation. Radical innovations might destroy existing capabilities of large firms because they might draw on unfamiliar knowledge domains. Since creating new capabilities is difficult and costly, firms are organized as to exploit existing capabilities and as such they might have less incentives to generate more radical innovations (Shane, 2001). Finally, if large firms have established, long-lived sales channels with spare capacity, then even if they do not have an organizational disadvantage in creating novelty, they have an absolute advantage in commercializing it (Henkel et al., 2014).

As opposed to large, incumbent firms, small ones (and entrants) are argued to suffer less from aforementioned disadvantages regarding coming up with radical innovations. They are believed to be less prone to learning traps as because they are less concerned with cannibalizing well-established competences. Moreover, they possess the entrepreneurial spirit necessary for radical improvements, while bureaucratic procedures in large firms might kill entrepreneurial effort.

Universities are claimed to play a major role in the generation and development of technological novelty. It has been argued basic science has shaped technological evolution by uncovering and providing deeper understanding of natural phenomena that might be essential for technological progress (Baumol, 2004). In theory, universities are not bounded by the competitive pressures that demotivate private firms to engage in research surrounded by high uncertainty. As such, their role in the process of radical innovation might be one of diversity-generating experimentation.

Overall, this capabilities-based view leads to a division of labor wherein large firms are more efficient at commercializing novelty, whereas small firms may either be creators of novelty, or intermediaries between universities and large firms. This study focuses on the origins of technological novelty and its diffusion over time. In particular, on the role of different types of players along the process. Currently, we characterize three types of actors: universities, small firms and large firms. Moreover, we characterize inventions as introducing novelty or not. Within the group of novel inventions, we make a further distinction based on the level of diffusion of the novelty introduced. In addition, we investigate the role of different players as the technological trajectory is developed, i.e. in the early – intermediate – late stage of diffusion.

RQ1: Is there a difference in involvement of different actors to generate technological novelty? Which of the actors are more likely to generate novelty that has a high diffusion?

RQ2: What do the patterns of diffusion look like with regard to the actors involved? Are certain actors more likely to be active in the early – intermediate – late stage of diffusion of novelty?

Methodology

Sample

Our dataset consists of granted USPTO patent documents in biotechnology (using the OECD definition, see also Arts et al., 2013), applied between 1988 and 1992. Patents having only one IPC6 class are excluded from the set, for reasons that have to do with our measure of technological novelty. Part of the analysis excludes patents that did not reach full term because their renewal fees were not paid. Furthermore, we restrict the set to patents filed in at least one of the 695 IPC6 classes used by patents in the FDA's Orange Book. Using this subsample allows us to sketch the patterns in the data, without having to deal with truncation issues. Moreover, this period gives us a sufficient time span for which we can track diffusion of novelty. In general, patent data is sourced from EPO PATSTAT 2012-10. Patent renewal information is obtained from the USPTO dataset on maintenance fee events for granted patents (version 20141006).

Technological novelty

We measure inventive activity using patent documents. Patents that are classified in two IPC-codes (at the 6-digit level) which have never occurred together before, are used to proxy inventions introducing a new approach (adapted from Fleming, 2001). This measure is sensitive to the base of comparison used to identify novelty, and we use the following approach to minimize bias. Since we are interested is patents that are novel on a global scale (not only novel among USPTO patents), we take as a base of comparison all patents filed before 1988. Using this approach introduces the issue of equivalent documents in multiple offices. This is why we identify those US patents for which at least 1 of its family members (DOCDB definition) has a combination of IPC-codes that never occurred before.

Diffusion

To track follow-up inventive activity to new approaches introduced by novel inventions, we identify patents following the newly introduced IPC-pair. All patents granted after the focal patent which are assigned to at least one of the new IPC-pair introduced by a patent in 1988-1992 are seen as following patents. As such, for each novel invention between 1988 and 1992, we count the number of US patents following the focal patents. We distinguish between early – intermediate – late followers by considering the application filing date of the following patents. This approach only tracks diffusion within the US patent system. Hence, bias introduced because of office specificities (relating to the assignment of technological classes) is minimized.

Products

We provide more insight on the relationship to the commercial dimension of radical innovation. Integrating our dataset with the Orange Book (OB) maintained by the US Food and Drug Administration (FDA) we identify biotech patents and novelty reusers that are linked to drug releases. In the US, the grant of a patent on a substance, product, or method of application is not enough to bring a drug to market. The product must be approved by the FDA, which lists both drug and patent in the OB. This allows to identify where along the technological trajectory the novel technology turns into a successful commercial product. As with the diffusion analysis explained above, we locate where in the diffusion pattern these patents occur, which kind of applicant files them, and to what extent this differs for novelty introduced by different sectors.

Actors

In order to characterize the actors involved in novelty generating/using patents, we use information on the 'sector' and size of the applicant. Sectors are allocated using the algorithm developed at MSI (Magerman et al., 2006) which classifies each applicant as a company, university or other. To identify firm size we use information on the possibility of fee reduction for small entities at the USPTO. Small business (under 500 employees, including subsidiaries), individual inventors, and nonprofit organizations can claim small entity status and obtain a 50% reduction in application and maintenance fees at the USPTO. In our set, the category 'small firms' consists of company patentees granted such status.

Empirical analysis

Novelty and diffusion

In this section we first show descriptive statistics of the distribution of patent counts for different actors, and examine cross-tabulations of sector and novelty, and sector and diffusion. This part of the analysis excludes biotech patents that did not reach full term. Table 1 below shows that lapsed patents represent about half of the full dataset. While not very different to non-lapsed patents in terms of novelty, they are significantly less likely to be linked to a drug product. The novelty share among novel (non-novel) patents is 9.24% (7.52%), while that of drug products is 4.29% (0.59%). Thus, the share of non-lapsed patents that reach the OB is seven times that of lapsed patents.

Table 1. Patent expiration and novelty/products

SECTOR	Novel		Product		
	No	Yes	No	Yes	Total
NOT LAPSED	5,973	608	6,299	282	6,581
Expected	6,027	554.4	6,414.10	166.9	6,581
relative share in Novel (Product)		**1.10**		**1.69**	
row percentage	90.76	9.24	95.71	4.29	100
column percentage	52.18	57.74	51.7	88.96	52.65
LAPSED	5,474	445	5,884	35	5,919
Expected	5,420.40	498.6	5,768.90	150.1	5,919
relative share in Novel (Product)		**0.89**		**0.23**	
row percentage	92	7.52	99.41	0.59	100
column percentage	47.82	42.26	48.3	11.04	47.35
Total	11,447	1,053	12,183	317	12,500
row percentage	91.58	8.42	97.46	2.54	100
column percentage	100	100	100	100	100
Pearson chi2(1)	11.9583 Pr = 0.001		172.0169 Pr = 0.000		

Table 2 below shows the cross-tabulation of novelty and type of assignee. It displays the number of novel and not novel patents for large companies, small companies, and universities that applied for a patent in biotechnology between 1988 to 1992. We also present the expected number of patents in each group under the assumption that the sector is not related to generating novelty (i.e. novelty is randomly distributed among sectors). In bold we include the relative share in novel patents (number of novel patents divided by the expected number). Nine percent of the patents in the sample are identified as novel. Of those, 66.6% are filed by large companies, 10.36% by small companies, and 22.86% by universities. Large firms are responsible for the bulk of technological novelty as defined here. They are as well, however, the largest player in the broader picture, that of overall innovative activity. When comparing these numbers to the expected values based on row and column independence, we see that large companies are only 3% more likely to introduce novelty compared to what is to be expected. Small companies are 57% more likely to do so, and universities are 20% less likely. Small companies, though not be the largest generator of novelty in absolute terms, are relatively the most specialized in generating novelty. Somewhat surprisingly, universities are not specialized in generating novel inventions.

Table 2. Novelty and assignee type

SECTOR	Not novel	Novel	Total
LARGE COMPANY	3,850	406	4,256
Expected	3,862.80	393.2	4,256
relative share in 'Novel'		**1.03**	
row percentage	90.46	9.54	100
column percentage	64.46	66.78	64.67
SMALL COMPANY	371	63	434
Expected	393.9	40.1	434
relative share in 'Novel'		**1.57**	
row percentage	85.48	14.52	100
column percentage	6.21	10.36	6.59
UNIVERSITY	1,752	139	1,891
Expected	1,716.30	174.7	1,891
relative share in 'Novel'		**0.80**	
row percentage	92.65	7.35	100
column percentage	29.33	22.86	28.73
Total	5,973	608	6,581
row percentage	90.76	9.24	100
column percentage	100	100	100
Pearson chi2(2) = 22.9139 Pr = 0.000			

In the next step, we singled out patents identified as novel and compare patents that have a considerable number of followers (>=10) versus patents with only few followers (<10). Table 3 below presents a cross-tabulation analogous to Table 2 above for novel patents with high follow-up inventions versus low follow-up invention. About 30% of the novel patents are followed in at least one of their new combinations by 10 or more patents. Within this group, small companies and universities are overrepresented, while other entities and large companies (to a lesser extent) are underrepresented. Small firms (universities) are 5% (40%) more likely to introduce highly-diffused novelty than the case when diffusion happens just by chance. In sum, large companies seem to have a considerable share in both generating novel and highly diffused novel inventions in absolute terms. Universities seem not to be specialized in generating novel inventions, but the novel inventions they do create are very likely to be amongst the most diffused ones. These figures do not shed light on an interesting side of the diffusion process: to what extent a novel technology ‘stays’ within a sector or with the introducer. This is addressed in the next section.

Table 3. High diffusion and assignee type

SECTOR	Low diffusion	High diffusion	Total
LARGE COMPANY	301	105	406
Expected	283.1	122.9	406
relative share in 'High diffusion'		**0.85**	
row percentage	74.14	25.86	100
column percentage	70.99	57.07	66.78
SMALL COMPANY	43	20	63
Expected	43.9	19.1	63
relative share in 'High diffusion'		**1.05**	
row percentage	68.25	31.75	100
column percentage	10.14	10.87	10.36
UNIVERSITY	80	59	139
Expected	96.9	42.1	139
relative share in 'High diffusion'		**1.40**	
row percentage	57.55	42.45	100
column percentage	18.87	32.07	22.86
Total	424	184	608
row percentage	69.74	30.26	100
column percentage	100	100	100
Pearson chi2(2) = 13.5674 Pr = 0.001			

Self- vs non-self-reuse

In this section we explore the identity of the assignees as introducers and reusers of novel technologies. As stated above, we study diffusion by tracking the reoccurrence of novel technology class combinations within a period of 15 years following the introduction of the new pair. Early, intermediate, and late adopters are identified according to whether the reusing patent was filed in the years 1-5, 6-10, or 11-15, respectively. We define own reuse as the case when the assignee of the reusing patent is the same as in the patent that introduced the novel combination.

The data now consist of novel patents which are reused at least once, and this amounts to 844 observations. We create a dummy variable that takes the value of 1 if the original assignee is among the reusers in the time period under consideration (full; early, intermediate, late). The first column in part

(a) of Table 4 below shows that for the full set, between 25% and 28% of the combinations show some reuse by the original assignee. Own reuse is higher in the early stage and drops significantly during the subsequent phases. This is true for the three sectors, but the drop is largest for small firm novelty. Part (b) of the same table breaks down reuse by low- and high-diffused patents. High-diffusion patents are those whose combinations are reused 10 or more times. Those introduced by firms, both small and large, are twice as likely than low-diffusion ones to have the original assignee among reusers. This is not the case for university patents. Interestingly, for company patents own reuse in the early stage is higher in high-diffusion patents than in low-diffusion ones, while the opposite happens with university patents.

Table 4. Own reuse by sector and time

SECTOR	Full period	Early	Intermediate	Late	Freq.
(a)					
Full set					
Company-Large	0.252	0.277	0.122	0.050	556
Company-Small	0.253	0.299	0.069	0.028	91
University	0.284	0.324	0.091	0.034	197
(b)					
Low Diffusion					
Company-Large	0.180	0.255	0.051	0.004	383
Company-Small	0.196	0.263	0.027	0.026	56
University	0.280	0.392	0.079	0.028	118
High Diffusion					
Company-Large	0.410	0.309	0.220	0.114	173
Company-Small	0.343	0.345	0.114	0.029	35
University	0.291	0.257	0.103	0.039	79

Technological novelty and commercial products

We linked the set of biotech patents and reusers of novelty to a version of the OB that includes patents filed in the USPTO between 1978 and 2008, associated to approved drug releases. Table 5 shows the biotech patents with a cross-tabulation of sector and whether the patent has been listed in the OB for a drug product. Four percent of the set, 282 patents, is linked to drug products in the OB. Large companies originate the largest share or product-linked patents (75.89%) and are more likely to end up in the OB than other sectors' patents. For instance, small companies and universities account for 4.96% and

19.15% of the linked products, respectively, and are 25% and 33% less likely to reach the OB than in the case where products are randomly distributed.

Table 5. Biotech patents and the OB

SECTOR	No product	Product	Total
LARGE COMPANY	4,042	214	4,256
Expected	4,073.60	182.4	4,256.00
relative share in 'Product'		**1.17**	
row percentage	94.97	5.03	100
column percentage	64.17	75.89	64.67
SMALL COMPANY	420	14	434
Expected	415.4	18.6	434
relative share in 'Product'		**0.75**	
row percentage	96.77	3.23	100
column percentage	6.67	4.96	6.59
UNIVERSITY	1,837	54	1,891
Expected	1,810.00	81	1,891.00
relative share in 'Product'		**0.67**	
row percentage	97.14	2.86	100
column percentage	29.16	19.15	28.73
Total	6,299	282	6,581
row percentage	95.71	4.29	100
column percentage	100	100	100
Pearson chi2(2) = 16.3385 Pr = 0.000			

Novel inventions can be used in commercial products directly or through followers and refinements. Interestingly, the rate of novelty is strikingly similar among the biotech patents that are linked to an OB product and those that are not. This holds true across the different sectors as well (not reported here). Novel patents, thus, are just as likely as not-novel ones to end up in a commercial product, though they can also reach the market via the reusers of the technologies they lay out for the first time. Table 6 focuses on novely patents and their link to drug products. For the novel biotech patents in our set, shows cross-tabulations of sector, 'OB by patents' (the biotech patent is filed for a drug product), and

‘OB by reusers’ (the reuser of novelty is filed for a drug product). The first two columns distinguish among novel biotech patents that are used in a new drug application and those that are not. About 4.44% of the novel set makes it to the OB, and it is mostly the case of large company inventions (85%). The remainder (14.81%) is due to universities. None of the novel patents by small companies protects drug releases, and those by large firms are overrepresented in the set, being about 28% more likely to reach the OB. Columns three and four summarize, by sector, the shares of novel inventions listed in the OB through their reusers. Overall, 12.66% of the novel patents have at least one follower listed in the OB. Again, the largest group is that of large companies (62%), though the share of universities doubles (28.57%), and that of small firms is not null (9.09%).

Table 6. Novel biotech patents and the OB

	OB by patent		OB by reusers		
SECTOR	No	Yes	No	Yes	Total
LARGE COMPANY	383	23	358	48	406
Expected	388	18	354.6	51.4	406
relative share in ‘Product’		**1.28**		**0.93**	
row percentage	94.33	5.67	88.18	11.82	100
column percentage	65.92	85.19	67.42	62.34	66.78
SMALL COMPANY	63	0	56	7	63
Expected	60.2	2.8	55	8	63
relative share in ‘Product’		**0.00**		**0.88**	
row percentage	100	0	88.89	11.11	100
column percentage	10.84	0	10.55	9.09	10.36
UNIVERSITY	135	4	117	22	139
Expected	132.8	6.2	121.4	17.6	139
relative share in ‘Product’		**0.65**		**1.25**	
row percentage	97.12	2.88	84.17	15.83	100
column percentage	23.24	14.81	22.03	28.57	22.86
Total	581	27	531	77	608
row percentage	95.56	4.44	87.34	12.66	100
column percentage	100	100	100	100	100
Pearson chi2(2) =	5.1619 Pr = 0.076		1.6547 Pr = 0.437		

If the intensity of reuse of the novel patents is weighted in (Table 7), the universities lead the links to OB via followers, with a share of products among total reusers of 0.92%. This is nearly twice as large as large and small companies (0.51% and 0.38%, respectively). This is a very skewed phenomenon, though, and at this level of disaggregation the number of observations is quite small, so these results might better be taken with caution.

Closing this section, Table 8 introduces the timing of drug products among the novelty reusers. Part (a) shows the minimum, average, and maximum number of years between the application of a novel patent and the application of reusing patents linked to drug products. Results are averaged across novelty-makers. This is used to proxy the time it takes a for a novel invention to 'reach' the market. More precisely, to reach another invention that has itself commercial application. On average, such 'commercial applications' arrive at the intermediate stage of reuse, 6-7 years after the original novel invention. Novelty by small firms and universities, though, reach a commercial phase already in the early stage of reuse. In the case of small firms (universities) the first product-by-reuser shows up, on average, 4 (4.95) years after the novel patent; this is two thirds of the time it takes in the case of large firms.

Part (b) of the same table requires more explanation. For each novel patent associated to a product via its reusers, we identify the full set of reusers, locate the products within it, and arrange them by application time—this defines their relative position in a reuse timeline. A value of 0 means the product is the same novel patent whereas a value of 1 means the product is in the last reuser. Again, small firms and universities reach the market earlier in the reuse profile than large firms. All this is consistent with a scenario in which small firms and universities specialize in producing high-quality novelty, but not its commercialization.

Table 7. Reuse intensity of novel patents by sector

SECTOR	Number of reused patents (a)	Average reuse (b)	Full reuse count (a*b)	# of reusers linked to products (c)	Share of products among reusers (c/(a*b)*100)
Large Company	406	23.13	9390	48	0.51
Small Company	63	29.59	1864	7	0.38
University	139	17.29	2403	22	0.92

Table 8. Timing of drug products among the reusers of novelty

SECTOR	Min	Avg	Max	Freq
(a)				
Years-to-product by novelty reusers				
Company-Large	6.02	7.21	8.67	48
Company-Small	4.00	7.05	10.71	7
University	4.95	6.05	7.05	22
(b)				
Relative order of product among novelty reusers				
Company-Large	0.33	0.47	0.65	48
Company-Small	0.14	0.38	0.75	7
University	0.20	0.35	0.51	22

Multivariate analysis

In this section we estimate models to predict novelty and diffusion on a set of covariates, with sector dummies as the main variable of interest. In this section too the analysis is done at the level of DOCDB families. In model (1) of Table 9 the dependent variable is a dummy that takes the value of 1 if there is at least one patent making a novel combination in the biotech patent's DOCDB family, and zero otherwise. Models (2)-(5) are conditional on novel patents, and the dependent variables are dummies which measure different levels of diffusion (whether the number of times the new combinations are reused exceed a certain threshold). All models include dummies for small companies and universities, large companies are the reference group, and a count of the patent's IPC6 classes. Variable *first_us* is a dummy assuming value of 1 if the first patent in the family was filed in the USPTO. In order to maximize the number of observations for the regression analysis we return to the full set including expired patents, and include a dummy to control for patents not reaching full term: *lapsed* is a dummy that takes the value of 1 if all the US patents in the family expired before full term due to lack of payment of maintenance fees. For the patents-families making new IPC6 combinations, variable *nr_newpairs* gives a count of the number of new pairs made. Models include dummies for application year. Correlations and descriptives are shown in Table 8.

Table 8. Correlation matrix and descriptive statistics

Correlations								Full set		Novel only	
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	Mean	Std. Dev.	Mean	Std. Dev.
Large Company	1.00							0.64	0.48	0.66	0.47
Small Company	-0.50	1.00						0.07	0.26	0.11	0.31
University	-0.76	-0.19	1.00					0.29	0.45	0.22	0.42
First_us	-0.15	0.03	0.15	1.00				0.66	0.47	0.59	0.49
Nr IPC6 classes	0.01	0.00	-0.01	-0.04	1.00			4.85	3.31	8.68	4.90
Lapsed	-0.01	0.03	-0.01	-0.02	-0.06	1.00		0.47	0.50	0.42	0.49
Nr new pairs	0.00	0.00	0.00	-0.01	0.45	-0.04	1.00	0.33	2.51	3.96	7.78
N								12500		1053	

The results presented in Table 9 are consistent with those already exposed above. Small firm (university) patents are more (less) likely to introduce novelty than those from large firms. Conditional on novelty, however, both small firms and universities are more likely to be among the high diffused novelty, and this effect is reinforced for higher levels of reuse. Expired novel patents, expectedly, are on average less likely to reach higher reuse, which suggests they are both commercially and technically less valuable. Or inversely, that this measure of high-diffusion may correlate with value.

Table 10 estimates the same models as in Table 9, this time including as dummies for patents' technology classes (TC) at the IPC3 level. For each case (novelty, and diffusion) the table includes both the empty model (IPC3 and application year dummies only) and the model with covariates. In all models the R-squared increases significantly, indicating that novelty is local to some technology classes. This is expected, given the nature of the novelty indicators employed. The sector dummies in model (2) are attenuated relative to the model without TC in Table 9, but that of small firms is still large and significant. The difference in the generation of novelty among large and small firms can be explained only partially by them occupating different areas of the technology space. This is the case with universities, though, since their sector dummy vanishes if TC are controlled for. The sector coefficients in the diffusion models are also attenuated after the inclusion of TC, though in a similar fashion to the models in Table 9. Thus, the differences in the diffusion cannot, too, be solely attributed to the fact that the different sectors specialize on distinct technologies.

Table 9. Logit estimates of novelty and diffusion

	(1) novel	(2) diff_1	(3) diff_5	(4) diff_10	(5) diff_15
main					
comp_small	0.635***	-0.130	0.088	0.263	0.409
	(0.116)	(0.255)	(0.216)	(0.235)	(0.264)
university	-0.308***	0.225	0.325*	0.449*	0.654***
	(0.085)	(0.210)	(0.162)	(0.176)	(0.197)
first_us	0.105	0.442**	0.170	0.111	0.180
	(0.073)	(0.165)	(0.137)	(0.154)	(0.178)
nr_class_ipc6	0.276***	0.158***	0.080***	0.092***	0.105***
	(0.009)	(0.028)	(0.017)	(0.017)	(0.018)
biotech_exp		-0.319*	-0.074	-0.345*	-0.544**
		(0.162)	(0.134)	(0.153)	(0.180)
nr_newpairs		0.083*	0.121***	0.081***	0.074***
		(0.036)	(0.022)	(0.018)	(0.017)
Year dummies	yes	yes	yes	yes	yes
Intercept	-3.937***	-0.227	-1.744***	-2.413***	-3.127***
	(0.119)	(0.298)	(0.247)	(0.275)	(0.322)
N	12500	1053	1053	1053	1053
r2	0.161	0.086	0.091	0.102	0.136
Recall	9.687	99.289	41.612	21.603	17.822
Precision	54.545	80.115	66.783	66.667	70.588
% correct	91.712	79.677	65.527	75.689	82.811
test_p	0.000	0.231	0.328	0.477	0.394

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 10. Logit estimates of novelty and diffusion controlling for technology class

	(1) novel	(2) novel	(3) diff_1	(4) diff_1	(5) diff_5	(6) diff_5	(7) diff_10	(8) diff_10	(9) diff_15	(10) diff_15
main										
comp_small		0.544***		-0.045		0.102		0.239		0.364
		(0.140)		(0.304)		(0.244)		(0.271)		(0.308)
university		-0.028		0.026		0.109		0.284		0.480*
		(0.096)		(0.234)		(0.181)		(0.202)		(0.230)
first_us		-0.003		0.537**		0.197		0.092		0.199
		(0.083)		(0.186)		(0.153)		(0.173)		(0.203)
nr_class_ipc6		0.298***		0.132***		0.048*		0.074***		0.111***
		(0.013)		(0.036)		(0.021)		(0.021)		(0.023)
biotech_exp				-0.232		0.052		-0.140		-0.363
				(0.182)		(0.150)		(0.173)		(0.207)
nr_newpairs				0.146**		0.177***		0.105***		0.095***
				(0.047)		(0.029)		(0.023)		(0.022)
Year dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes	Yes
IPC3 dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
_cons	-4.425***	-4.592***	0.937**	0.321	-1.652***	-1.974***	-2.776***	-2.976***	-3.535***	-3.916***
	(0.149)	(0.170)	(0.345)	(0.408)	(0.286)	(0.339)	(0.338)	(0.393)	(0.401)	(0.475)
N	12500	12500	991	991	1044	1044	1053	1053	1053	1053
r2	0.236	0.331	0.111	0.168	0.119	0.186	0.163	0.226	0.183	0.279
Recall	22.697	31.434	97.970	97.335	49.333	55.111	31.359	38.328	23.267	38.614
Precision	67.135	69.979	81.866	82.120	64.162	73.373	76.271	71.895	77.049	71.560
% correct	92.552	93.088	81.130	81.029	66.284	72.031	78.632	79.107	83.951	85.280
test_p		0.000		0.837		0.980		0.882		0.732

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Discussion and research agenda

In general, these preliminary results unveil some interesting patterns. The main take-away in this stage of the research is that the patterns we see seem to make sense. Hence, it seems valuable to further investigate the dynamics of technological trajectories according to the framework introduced. Preliminary findings suggest a distinct role of universities, small firms and large firms in the generation and development of technological novelty. While the lion's share of patents, novel patents and follow-up patents using novelty is to be attributed to large firms, it seems they have a relative disadvantage to generate novel patents that become highly diffused. Small firms are at a relative advantage to create novel patents, and within the group of novel patents, they have an advantage in generating the ones with high diffusion. Universities are not more likely to generate novel patents, but if they do, these patents are relatively more often among the most highly diffused ones.

In order to address the second research question, which we only lay out here, we will focus on novel patents only and explore the shares of the different actors in the using patents. Moreover, we can distinguish between different stages of the diffusion process and try to disentangle the role of different players over different stages. Further, we can break down by type of player introducing the technological novelty. This allows us to show descriptively which actor might play which role in the generation and diffusion of technological novelty in the biotechnology industry.

As an avenue to gaining more insight on the commercial dimension of radical innovation, as well as the role played by market competition, we have integrated our dataset with data from the FDA's Orange Book. We analyze the links to drug releases and have uncovered some interesting patterns, but more questions lay ahead. Where in the diffusion pattern do these patents occur? Are they more likely to filed by large companies? What is the moment of 'entry' by large companies close to patent related to drug? Are more important drugs to be found earlier in the diffusion pattern? Do large firms face disadvantages for the 'new drugs'? Moreover, we can explore whether the drug introduces a New Chemical Entity (NCE) or was considered for the 'fast track' (drugs with high therapeutic relevance), and how this relates to novelty.

References

Ahuja, Gautam, and Curba Morris Lampert. "Entrepreneurship in the large corporation: A longitudinal study of how established firms create breakthrough inventions." Strategic Management Journal 22.6-7 (2001): 521-543.

Arts, Sam, Francesco Paolo Appio, and Bart Van Looy. "Inventions shaping technological trajectories: do existing patent indicators provide a comprehensive picture?." Scientometrics 97.2 (2013): 397-419.

Baumol, William J. Education for innovation: entrepreneurial breakthroughs vs. corporate incremental improvements. No. w10578. National Bureau of Economic Research (2004).

Dosi, Giovanni. "Technological paradigms and technological trajectories: a suggested interpretation of the determinants and directions of technical change." Research policy 11.3 (1982): 147-162.

Fleming, Lee. "Recombinant uncertainty in technological search." Management science 47.1 (2001): 117-132.

Henderson, Rebecca. "Underinvestment and incompetence as responses to radical innovation: Evidence from the photolithographic alignment equipment industry." The RAND Journal of Economics (1993): 248-270.

Henkel, Joachim, Thomas Rønde, and Marcus Wagner. "And the winner is—Acquired. Entrepreneurship as a contest yielding radical innovations." Research Policy (2014).

Magerman, Tom, Bart Van Looy, and Xiaoyan Song. "Data production methods for harmonized patent statistics: Patentee name harmonization." DTEW-MSI_0605 (2006): 1-88.

Shane, Scott. "Technology regimes and new firm formation." Management science 47.9 (2001): 1173-1190.

Contact details

Manuel Gigena
Faculty of Business and Economics
Dept. of Managerial Economics, Strategy and Innovation
Naamsestraat 69
3000 Leuven
Phone: +32 (0) 16 37 90 54
e-mail: manuel.gigena@kuleuven.be