Paper to be presented at the DRUID Academy Conference 2017 at University of Southern Denmark, Odense, Denmark on January 18-20, 2017

What they don't know can hurt you: internal development of technological novelty and the informational structure of the market for technologies

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

Abstract

Research goal and motivation. This paper is about the relation between the informational structure of the market for technologies and the division of labor among firm types in the innovation process. A novel dataset tracing the trajectories of novel technologies in the biotech industry allows: 1) to track the involvement of different organizational types (large-incumbent and small-entrant firms, and university/research centers) along the stages of generation, development and commercialization of novel technologies in a large scale; and 2) to challenge established knowledge about how vertical specialization in the innovation process works. Data This study tracks for the biotech sector (i) the origination of technologically novel ideas, (ii) their further 'use' or development and (iii) their commercialization . To this end I use biotech patent data to identify novel ideas and their origin, identifying technological novelty as recombinant novelty, -I classify a patent as novel when it contains two previously uncombined IPC-codes (group level), reflecting its recombinant novelty -To track follow-up inventive activity to new approaches introduced by novel inventions, I identify patents reusing the newly introduced IPC-pair in the 10 years following the novel patent application. -I identify inventions leading to commercialization through drugs by linking the set of patents to the FDA's Orange Book (OB), which lists approved drugs and their related patents. I determine whether the patent applicant of the novel patent is a large or small company or a non-profit, research organization (university, research institute, university hospital, etc.). I also identify the technological age and experience of the applicant in the technology are to identify incumbency status of the applicant. I identify the organisational type of the novel patent applicant, but also of the patent applicants reusing the new combination used by the novel originator. And I identify the organisational type of the patent applicant that gets listed in the OB. This allows to trace the location of novel biotech inventions into different organisational types along the three different steps of creation, follow-on development and successful commercialization. Model and main results. I sketch a simple signaling model that relates the decision of firms to refine or develop internally the novel ideas they generate to the presence of asymmetric information about the commercial value of the innovations. I present the characteristics of the possible equilibriums and their implications for the configuration of the technological trajectories. The patterns observed in our data allow to map the involvement of different organizational types to different equilibriums. In particular, I shed light on two observed, and unexpected, empirical patterns: the substantial internal follow-on

development effort done by small firms to refine their own novel ideas, and the absence of large firms as prevalent followers of small firm novelty.

Jelcodes:O31,-

What they don't know can hurt you: internal development of technological novelty and the informational structure of the market for technologies

Manuel Gigena *

Managerial Economics, Strategy and Innovation, KU Leuven

December 2016

`Very early draft. Please do not distribute.`

Abstract

This paper is about the follow-on development of novel technologies and the creation of commercial value in the presence of asymmetric information in the market for technologies. A novel dataset tracing the trajectories of novel technologies in the biotechnology industry allows: 1) to track the involvement of different organizational types (large-incumbent and small-entrant firms, and universities) along the stages of generation, follow-on development and commercialization of novel technologies; and 2) to challenge established knowledge about how vertical specialization in the innovation process works. I sketch a signaling model that relates to the informational structure of the market for technologies firms' decision to refine/develop internally the novel ideas they generate. I present the characteristics of the possible equilibriums and their implications for technological trajectories and technical change.

*Email: manuel.gigena@kuleuven.be

1 Motivation

Schumpeter [28, 29] lays the foundations of a literature that alternatively supports small-entrant firms or backs large-incumbents as the main source of breakthrough innovation (Mark I vs Mark II). One strand of this literature considers small entrants endowed with an advantage when it comes to the generation of radical innovation, while large, incumbent firms are better in generating follow-on incremental innovations. Established firms are argued to have rigid organizational structures that raise difficulties in creating novelty, especially if that requires new combinations of existing capabilities [20] and search beyond familiar technologies [3]. Furthermore, as radical innovations might destroy their existing positions, incumbents have lower incentives to generate radical innovations [27]. As opposed to large, incumbent firms, small entrants are believed to be less prone to learning traps, being less concerned with cannibalizing incumbent competences, and better organized for radical innovation [2, 10]. Yet, there is also empirical evidence stressing the important role of large, incumbent firms in radical innovation [9, 24]. Moreover, a considerable stream of research in the ambidexterity perspective has investigated potential strategies for large firms to produce technological breakthroughs ([26, 3]).

Theory and evidence points to a division of labor along corporate types in the generation and further development of novelty. Breakthrough innovations occur in young, small firms, but improvements and wide-scale dissemination happens through large firms [8]. Further, a research stream initiated by Jensen & Thursby [21] documents corporate follow up of early-stage university-generated innovations. Support for a division of labor is strengthened when complementary assets are required for commercialization, or there exists a critical scale and stock of resources that gives large firms an absolute advantage in commercializing inventions [32].

Fundamental to the dynamics briefly described above is the existence of a market for technology [4, 16], as the trade in technologies facilitates vertical specialization according to comparative advantages [14, 19]. Firms without downstream assets can invest in innovation as they do not need to own, acquire or develop such assets. Small firms and universities can focus on the upstream segment, especially benefiting from technology trade [14, 5] while large firms use their relative advantage in downstream developing and commercialization [19]. As is the case with the operation of any market, the trade of technologies and the efficiency of the final allocations are affected by the informational structure of the market. Under perfect information, firms operating on the upstream segment can specialize on the generation of novel ideas. The latter will be sold to downstream firms that can unequivocally assess their quality. In presence of asymmetric information about the quality of novel ideas and technologies, agents might devote effort to signal the type of ideas they generate and wish to trade, even if such effort is unrelated

to productivity and is privately and socially inefficient [31].

This paper is about the follow-on development[1] of novel technologies and the creation of commercial value in the presence of asymmetric information in the market for technologies. Using a novel dataset on the generation and the follow-on development of novel technologies I document empirically patterns that do not conform to conventional theory about the involvement of different firm types. To understand them, I sketch a simple signaling model (based on Spence [30]) that relates the decision of firms to refine or develop internally the novel ideas they generate to the informational structure of the market in which those ideas are exchanged. Some implications of the equilibriums are discussed, and empirical patterns supporting the main predictions are introduced.

2 Stylized vs documented facts on novel technology development

The short review above leads to some expectations about the differential involvement of distinct organizational types along the trajectory of novel technologies. I depart from four (key) stylized facts:

1. **Novelty generation.** Large-incumbent firms have an advantage over small-entrants to create technological novelty with commercial value.

2. **Novelty development.** Small-entrant firms create commercial value indirectly (i.e. via follow-on improvement) rather than directly, and they are more likely to create indirect commercial value than large firms.

3. **Appropriation.** Large firms have an advantage over small firms in the appropriation of innovations, and are thus more likely to make internal follow-on of the novelty they create.

4. **Vertical specialization.** With efficient markets for technologies, small firms specialize on generation of novelty, being rather absent in the development stage. Thus, small firm novelty is more likely than large firm novelty to be developed externally (and into commercial success), especially by large firms.

[1] Throughout this document the expression 'development' is meant to represent the refinement, or follow-on improvement, of novelty, rather than the process that leads to the creation of novelty itself. Aligned with the ideas of Dosi [11], if technologies evolve along trajectories guided by paradigms, I see *novelty generation* as the starting point of that trajectory, and *development* as the subsequent, incremental, efforts along it.

These stylized facts cannot be confirmed empirically, however. In this paper I document a remarkable involvement of small firms in the technology development stage, which I characterize in three ways. First, I note that small, and not large firms, are the prevalent developers of novel technologies introduced by small firms. Second, I note that small firms are more prone to in-house development of novelty than large firms. Third, I find that follow-on improvement of novelty is an important source of commercial value, to which the contribution of small firms is also significant.

3 Empirical approach

Following scholars studying the evolution of technology [11, 6, 7] I look at three stages: the introduction of technological novelty, follow-on technological improvement and commercialization. Identifying technological novelty as recombinant novelty [18, 12, 6, 7], this study tracks the origination of novel ideas and their usefulness in two dimensions, namely technological and commercial. I trace further 'use' of technologically novel ideas in biotechnology using patent data, and their commercialization through their link to approved drug products.

Sample The dataset consists of granted USPTO novel patent documents in biotechnology (using the OECD definition, [25]), grouped at the level of the family using the DOCDB definition ([23]), with application priority years between 1982 and 1995. In order to limit the dataset to inventions with potential as approved commercial drugs, i.e. which could make it into the FDA's Orange Book (OB), only patents classified under main USPC classes 424, 514, 530 and 536 are included.[2] Patents having only one IPC6 code are excluded, for reasons that have to do with the measure of technological novelty. The final set comprises 1142 observations.

Technological novelty Patents classified in two group-level IPC codes (6-digit IPC, henceforth IPC6) which have never co-occurred before are used to proxy inventions introducing a new approach [12, 13, 33]. The base for comparison are all patent documents, from all patent offices, contained in the PATSTAT dataset. The issues posed by the existence of equivalent documents in multiple offices, as well as by continuation applications, are dealt with by working at the patent family level (DOCDB definition). I identify those US patents for which at least one of

[2]These are also the most frequent main USPC classes in OB patents, representing over 90% of the patents listed there. Classes 424 and 514 are for 'drugs, bio-affecting and body treating compositions'. Class 530 is for 'Chemistry, natural resins or derivatives; peptides or proteins; lignins or reaction products thereof'. Class 536 is for 'Organic compounds.'

its family members makes IPC combinations that never occurred before. To deal with simultaneous but independent development of novelty, all patents using a novel pair within 18 months of its introduction are taken as novel.[3]

Technological development To track follow-up inventive activity to new approaches introduced by novel inventions, I identify patents reusing the newly introduced IPC pairs in the 10 years following the application of novel patent. Patents granted after the novel patent which are assigned to at least one of the new IPC pairs introduced by a biotech patent from 1982-1995, and filed in the same main USPC class, are seen as following ('reusing') patents. I use disambiguated assignee identification data to distinguish 'internal' follow-on (made by the same agent that introduced the novel inventions) from external one. The total number of 'following' patents is 3423.

Commercial value: approved drugs I provide insight on the commercialization of biotech novelty by looking at one of its dimensions, approved-to-market drug products, thus moving away from purely patent-based indicator analysis. A New Drug Application (NDA) at the FDA includes information about the patents linked to the drug. By FDA rule, the subject matter of such patents can be a substance, a product, or a method of application. The patents are listed in the OB, which I link to the sets of novel patents, but also to followers, to identify where along the development trajectory new ideas turn into commercial products. It is important to note that while I flag inventions that reach the OB, the focus is (at the current stage, at least) on the *patent assignees* rather than the FDA sponsor (i.e. on the agent that creates the patent and not the one that buys, acquires or is licensed the patent and applies for regulatory permission).

Actors I use information on the type and size of the patent applicants to characterize the organizational type of the actors involved in novelty generation and development. The algorithm developed at ECOOM[22] is used to distinguish corporate from research or university applicants.[4]

[3]The use of such time window poses the question of considering the very early re-occurrence of a pair as follower or as novel inventions itself. The time lags between invention development and patent application, and to the publications by patent offices, to my judgment, justifies the approach. Different time spans were evaluated, comparing the characteristics of patents falling under both categories of this taxonomy and 18 months seems to provide the best division.

[4]Categories are: company, university, hospital, non-profit, government, and individual. Closer inspection reveals that categories like hospital, non-profit and government consist largely of research organizations (e.g. The Salk Institute, Sloan Kettering Institute, City of Hope, Institute of Cancer Research), so I cluster them together with universities. The algorithm assigns patents that do not list an institutional assignee to a class of 'Individual inventors'. This is often the

The USPTO grants a 50% application and maintenance fee reduction to small applicants (i.e. claiming 'small entity status', with less than 500 employees, including subsidiaries). I classify corporate patentees granted such status as 'small firms'. The final sample contains patents from about X large firms, Y small firms and Z universities.

Data sources The novel dataset is built by linking patent data from multiple sources. Basic patent data is obtained from EPO PATSTAT 2013-10, which I also use to construct the novelty indicator. Information on assignee size and payment of renewal fees is sourced from the USPTO Patent Maintenance Fee Events (v. 20150126). FDA data on approved drug products and their listed patents is obtained from June 2015 versions of Drugs@FDA and Orange Book datasets.

4 Patterns of technical change and commercial value creation

4.1 Empirical evidence

Table 1 shows the main variables driving the empirical analysis. The patterns observed are presented first analyzing descriptive statistics. About two thirds of the novel inventions have followers. University trajectories seem more likely than corporate ones to be built upon, both by internal and by external followers. Internal follow-on is less frequent than external, and is remarkably higher for small firm (20%) and university (26.7%) novelty than for large firms (13.5%). Large firms are prevalent in the sample (62%), and as external followers (54.6% of the novel inventions is followed at least by one large firm). They are not, however, and contrary to expectations prevalent followers of small firm novelty.

- - Table 1 about here - -

The creation of direct commercial value, here measured by the filing of a novel Orange Book (OB) patent, is a rare event (3.7%). Small firms seem to have a large relative advantage over large firms (7.4% vs 3.7%) and universities (1.1%). Further, follow-on development of novelty is a source of commercial value creation, with the indirect commercial reach being three times as large as the direct (11.3% vs 3.7%), and even across organization types. These figures mask an interesting pattern, however. Indirect internal development of commercial value by small

case of European university inventors who patent in their name rather than the institutions they work at. Extensive manual cleaning was done to correctly allocate such patents to the right organizational type.

firms (5.9%) is seven times larger than by large firm and by universities (very low for both types, 0.8% and 0.6%, respectively), and represents nearly half of the contribution of small firm commercial value creation. Indirect success of large firm and university novelty is achieved mostly externally, which reveals substantial external (and successful) exploration of technological trajectories by agents other than their initiators. This is despite the low expected probability of success, with technical search being an activity of high uncertainty. Also, in the case of university novelty, with a paucity of signals of direct commercial value.

Tables 2-3 extend the analysis by means of econometric analysis, controlling for invention characteristics related to the novelty indicator (count of IPC6 classes), the development indicator (patent family size, count of new combinations made), application year and main technology field (6-digit main USPC technology classes). Main patterns observed are confirmed. Thus, small firms seem to be not only actively involved in pure technical change, forging technological trajectories, but also in the refinement of novelty until it is commercially viable. Further, there is substantial evidence of the persistence of small firms within the technological and commercial trajectories they initiate.

- - Tables 2-3 about here - -

4.2 Turning novelty into a commercial success

Internal follow-on development seems prevalent among small firms and universities. This is somewhat unexpected in a world seen through the lens of a division of labor story, wherein the strength of these agents is experimentation rather than refinement. Further, where the existence of a mature market for technologies would allow trading early-stage ideas, incremental development that steers them into commercial applications can be undertaken by the agents involved in their commercialization. What drives their decision to develop novel technologies to a commercialization stage? What are the implications for technical change and for the structure of the market for technology?

The empirical evidence introduced above reveal that novel ideas differ in their direct value, as well as in their potential to be developed into value. Further, that that different actor types might be endowed with different power to carry a novel idea to commercial stage through internal development. Drawing from two-factor models of productivity [34, 15], consider that the probability of success through development of idea i is described by

$$p_i = \alpha_i + \beta_i g(e) \tag{1}$$

where p_i is the probability of success, α_i is the 'intrinsic' value of an idea, β_i represents potential to acquire value through development, $\mathbf{e}$ is a vector with the

different development effort types (i.e. internal and external) and their intensity, and $g : e \to \mathbb{R}$ represents the way in which effort types relate to success ($g_{e_i}(e)$), and their interdependence ($g_{e_i e_j}(e)$, when $i \neq j$)

To analyze the role of internal development, I estimate different specifications of equation 1 where e includes internal development only. Results are shown in Table 4. Internal development correlates positively, and strongly, with commercial success in the development trajectory (Model 1), particularly internal development by small firms (Model 2). The relationship with commercial success by external followers is weak, however (Model 3). If internal development is an important source of value creation, it is particularly so for small firms. Furthermore, appropriation of the proceeds of internal effort seems high (Model 4). The next section builds on this and attempts to answer the questions that open this sub-section.

5 A model of in-house development of novel technologies

The model is laid out to understand the relationship between the decision to do in-house technology development and the informational structure of the market for technologies. It is a loose adaptation of textbook signaling models typically used in labor economics to explain education choice under asymmetric information [30, 31]. It is laid out as a case when firms generate and trade technologies under asymmetric information, but it can somehow easily be rearranged to represent the interaction of agents in a technology landscape, modeling their choice to enter or enter a certain technology trajectory originated by different organizational types.

5.1 Model setup

The representative small firms generate novel technologies that may require additional development efforts to be commercially viable. The commercialization requires downstream assets only in the hands of representative large firms. There is a market for technologies that allows for the trading, licensing and negotiation of technologies disembodied of products [5, 17]. The intrinsic value of novel technology i, θ_i, is unobservable to the large firms and has a distribution with support over $[\underline{\theta}, \bar{\theta}]$. Firms can invest in further developing their novel technologies in-house, choosing an amount of development effort $e_i \in [0, \bar{e}]$. Development costs depend on the quality of the idea and the level of effort chosen by the firm, $c(\theta, e)$. Costs are continuous and twice differentiable, and I further assume $c(\theta, 0) = 0$, $c_e > 0$ and $c_{e\theta} < 0$. This is, the marginal cost of development is positive but is lower for ideas of greater quality.

The value of a novel technology to the firm that acquires it for commercialization is given by $v(\theta, e)$. Unable to observe θ_i, the large firm pays according to the development effort observed, $y(e)$, with a payment schedule derived from their own maximization behavior in a competitive market where the products are commercialized. Large firms take the value of technologies with given levels of development as given, not affected by the payment offered. Earning $y(e_i)$ by trading technology i, small firms are interested in their resulting net wealth, $n(\theta_i, e_i) = y(e_i) - c(\theta_i, e_i)$.

If large firms observed quality, they would offer a payment schedule based on θ_i, $y(\theta_i, e_i) = v(\theta_i, e_i)$. Technologies are traded according to their commercial value, and small firms can appropriate exactly the value of the technologies they create. If quality is unobservable, in-house development is not possible, and higher quality technologies remain in the market, the offer of large firms would be $y(\bar{\bar{\theta}})$, where, with $f(.)$ being the pdf of θ, $\bar{\bar{\theta}} = \int_{\underline{\theta}}^{\bar{\theta}} \theta f(\theta) d\theta$ is the expected quality. If technologies of all qualities are traded, firms offering $\theta_i > \bar{\bar{\theta}}$ would be better off in some alternative scenario in which their higher quality can be signaled so as to distinguish themselves from lower quality offers.

5.2 Solution

I am interested in a separating equilibrium where firms that generated higher-quality ideas find it optimal to signal it by means of in-house development effort. This equilibrium is defined by two conditions. First, small firms maximize their net wealth n by choosing a level of effort that equalizes marginal income to marginal cost,

$$\frac{\partial y(e)}{\partial e} = \frac{\partial c(\theta, e)}{\partial e}. \tag{2}$$

This must hold for all e. The solution is a maximum if the second-order condition holds,

$$\frac{\partial^2 y(e)}{\partial e^2} - \frac{\partial^2 c(\theta, e)}{\partial e^2} < 0. \tag{3}$$

In equilibrium, the signal transmits information if both absolute and marginal costs decrease as the unobserved trait increases.

The second condition is that the expectations of large firms in the technology markets must be confirmed. Their experience in the market is consistent with the schedule of payments offered, so that for all quality levels the offer matches the value of the idea, or

$$y(e) = v(\theta, e), \tag{4}$$

for all θ. This implies that beliefs of large firms about the quality of a technology must be correct, and

$$\theta = \hat{\theta}(e^*(\theta)), \tag{5}$$

where $e^*(\theta)$ is the level of effort that maximizes the small firm's net wealth. If $\theta > \hat{\theta}(e^*(\theta))$ the large firm would obtain returns in excess to its investment in the technology acquired, while if $\theta < \hat{\theta}(e^*(\theta))$ she would be better off by deviating and offering a smaller payment for that level of e.

In equilibrium $y(e) = v(\theta, e)$, for every level of θ. Differentiating we have that the marginal return to in-house development is

$$\frac{\partial y}{\partial e} = \frac{\partial v}{\partial e} + \frac{\partial v}{\partial \theta}\frac{\partial \theta}{\partial e}. \tag{6}$$

Since $v_\theta > 0$ and $\theta_e > 0$, the left hand side of equation 6 is larger than v_e. This implies that the private return to in-house development is higher than its direct contribution to productivity. The key here is the signaling effect, the part of the private return to development effort related to the unobserved quality level, captured by $v_\theta \theta_e$.

Further, combining equation 6 with the fact that in equilibrium the marginal income and cost of development are equalized, we have that $v_e - c_e < 0$. For all θ, in equilibrium the investment in development is higher than it would be with symmetric, perfect information. As stated above, if quality is observable, small firms choose development effort to maximize $v(\theta, e) - c(\theta, e)$ by setting $v_e = c_e$. Finally, if $v_\theta = 0$ for all θ, the signaling effect fully wanes. Even though the quality is unobservable, small firms make efficient development decisions.

Proposition 1. There exists a separating equilibrium if the signaling schedule is strictly increasing in the development effort, and small firms find it optimal to do some non-zero level of in-house development, which is greater than what is done under symmetric information.

Proposition 2. In this equilibrium, for high-quality-novelty-generating small firms the private return to in-house development is higher than its contribution to productivity.

6 Discussion

The empirical patterns observed reveal a very skewed distribution of success in the generation of commercially viable novelty. They reveal as well a substantial involvement of small firms in the development that follows the introduction of novel technologies. The signaling model proposed somewhat explains the incentives of small firms to undertake such development efforts as a function of a feature of the

informational structure of the market where technologies are traded. The separating, signaling equilibrium proposed involves inefficient levels of signaling. To avoid being offered in payment an amount that corresponds to the value generated by the lowest quality of novel technologies, small firms other than those generating the lowest-quality ideas must make an otherwise unnecessary investment in the signal. In this regard, all firms but those generating the lowest quality ideas will do strictly worse in the asymmetric information scenario relative to one with full information. However, due to signaling, in equilibrium the private return to in-house development is higher than its contribution to productivity. It remains in future development of this work in progress to identify the overall implications of both effects.

7 Conclusions

This paper is about the development of novel technologies and the creation of commercial value in the presence of asymmetric information in the market for technologies. A novel dataset tracing the trajectories of novel technologies in the biotechnology industry allowed to track the involvement of different institutional types along the stages of generation, development and commercialization. Further, it challenges established knowledge about how vertical specialization in the innovation process works. The patterns of technical change documented here suggest that stylized facts about the development of novel technologies might be confirmed empirically. In particular, this work documents a remarkable involvement of small firms in the technology development stage, which is hereby linked to the existence of asymmetric information about the quality of novel ideas in the market for technologies. A simple signaling model was introduced that relates the decision of firms to refine the novel ideas they generate in-house to the informational structure of the market for technologies. As signaling is costly, it is suggested that the informational structure of the market of technologies might bring about efficiency losses. Small, entrant firms are enticed to go beyond their role of experimenters, in which they have a relative advantage, and enter the development stage, typically the domains of large, incumbent firms.

References

[1] Akerlof, G. (1970). The market for lemons: Quality uncertainty and the market mechanism. Quarterly Journal of Economics, 84(3), 488-500.

[2] Acs, Z. J., Audretsch, D. B. (1990). Innovation and small firms. Mit Press.

[3] Ahuja, G., Morris Lampert, C. (2001). Entrepreneurship in the large corporation: A longitudinal study of how established firms create breakthrough inventions. Strategic Management Journal, 22(6?7), 521-543.

[4] Arora, A., Fosfuri, A., Gambardella, A. (2004). Markets for technology: The economics of innovation and corporate strategy. MIT press.

[5] Arora, A., Ceccagnoli, M. (2006). Patent protection, complementary assets, and firms' incentives for technology licensing. Management Science, 52(2), 293-308.

[6] Arthur, W. B. (2007). The structure of invention. Research Policy, 36(2), 274-287.

[7] Arthur, W. B. (2009). The nature of technology: What it is and how it evolves. Simon and Schuster.

[8] Baumol, W. J. (2004). Entrepreneurial enterprises, large established firms and other components of the free-market growth machine. Small Business Economics, 23(1), 9-21.

[9] Chandy, R. K., Tellis, G. J. (2000). The incumbent? curse? Incumbency, size, and radical product innovation. Journal of marketing, 64(3), 1-17.

[10] Cohen, W., Klepper, S. (1996). Firm Size and the Nature of Innovation within Industries: The Case of Process and Product RD. The Review of Economics and Statistics, 78(2), 232-243.

[11] Dosi, G. (1982). Technological paradigms and technological trajectories: a suggested interpretation of the determinants and directions of technical change.Research policy, 11(3), 147-162.

[12] Fleming, L. (2001). Recombinant uncertainty in technological search.Management science, 47(1), 117-132.

[13] Fleming, L., Mingo, S., & Chen, D. (2007). Collaborative brokerage, generative creativity, and creative success. Administrative Science Quarterly, 52(3), 443-475.

[14] Gans, J. S., Stern, S. (2003). The product market and the market for "ideas": commercialization strategies for technology entrepreneurs. Research policy, 32(2), 333-350.

[15] Garen, J. (1984). The returns to schooling: A selectivity bias approach with a continuous choice variable. Econometrica: Journal of the Econometric Society, 1199-1218.

[16] Geroski, P. (1995). Markets for technology: knowledge, innovation and appropriability. Handbook of the economics of innovation and technological change, 90-131.

[17] Graham, S. J., Merges, R. P., Samuelson, P., Sichelman, T. (2009). High technology entrepreneurs and the patent system: Results of the 2008 Berkeley patent survey. Berkeley Technology Law Journal, 1255-1327.

[18] Hargadon, A. (2003). How breakthroughs happen: The surprising truth about how companies innovate. Harvard Business Press.

[19] Henkel, J., Rønde, T., Wagner, M. (2015). And the winner is—Acquired. Entrepreneurship as a contest yielding radical innovations. Research Policy,44(2), 295-310.

[20] Henderson, R. M., Clark, K. B. (1990). Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. Administrative science quarterly, 9-30.

[21] Jensen, R., Thursby, M. (2001): Proofs and Prototypes for Sale: The Licensing of University Inventions. The American Economic Review 91, 240–259.

[22] Magerman, T., Van Looy, B., & Song, X. (2006). Data production methods for harmonized patent indicators: Patentee Name Harmonization. EUROSTAT Working Paper and Studies, Luxembourg.

[23] Martínez, C. (2011). Patent families: when do different definitions really matter?. Scientometrics, 86(1), 39-63.

[24] Methe, D., Swaminathan, A., Mitchell, W., Toyama, R. (1997). The underemphasized role of diversifying entrants and industry incumbents as the sources of major innovations. Strategic Discovery: Competing in New Arenas, 99-116.

[25] ORGANIZATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). (2005). A framework for biotechnology statistics.

[26] Rosenkopf, L., Nerkar, A. (2001). Beyond local search: boundary?spanning, exploration, and impact in the optical disk industry. Strategic Management Journal, 22(4), 287-306.

[27] Shane, S. (2001). Technology regimes and new firm formation. Management science, 47(9), 1173-1190.

[28] Schumpeter, J. A. (1934). The theory of economic development: An inquiry into profits, capital, credit, interest, and the business cycle (Vol. 55). Transaction publishers.

[29] Schumpeter, J. A. (1942). Socialism and democracy.

[30] Spence, M. (1973). Job market signaling. The quarterly journal of Economics, 355-374.

[31] Spence, M. (2002). Signaling in retrospect and the informational structure of markets. The American Economic Review, 92(3), 434-459.

[32] Teece, D. J. (1996). Firm organization, industrial structure, and technological innovation. Journal of Economic Behavior Organization, 31(2), 193-224.

[33] Verhoeven, D., Bakker, J.,&Veugelers, R. (2016). Measuring technological novelty with patent-based indicators. Research Policy, 45(3), 707-723.

[34] Willis, R. J., Rosen, S. (1979). Education and Self-Selection. The Journal of Political Economy, S7-S36.

Tables

	All org. types N=1142 100%		Company-Large n=714 62.5%		Company-Small n=179 15.7%		University n=249 21.8%	
	mean	sd	mean	sd	mean	sd	mean	sd
No. IPC6	10.585	5.401	10.560	5.597	10.950	5.592	10.394	4.646
No. of new pairs	4.029	7.031	3.723	7.069	4.654	7.339	4.458	6.662
Has followers	0.639	0.480	0.627	0.484	0.626	0.485	0.683	0.466
Has ext followers	0.614	0.487	0.604	0.489	0.581	0.495	0.667	0.472
Has internal followers	0.169	0.375	0.134	0.341	0.207	0.406	0.241	0.429
High follow-on(1*MAD)	0.338	0.473	0.290	0.454	0.393	0.491	0.429	0.496
followed by ext Large firm	0.516	0.500	0.522	0.500	0.436	0.497	0.554	0.498
followed by ext Small firm	0.352	0.478	0.308	0.462	0.369	0.484	0.466	0.500
followed by ext Uni	0.385	0.487	0.366	0.482	0.352	0.479	0.466	0.500
Direct CS	0.038	0.190	0.029	0.169	0.112	0.316	0.008	0.089
Indirect CS	0.107	0.309	0.105	0.307	0.128	0.336	0.096	0.296
Indirect CS-int	0.018	0.131	0.006	0.075	0.084	0.278	0.004	0.063
Indirect CS-ext	0.093	0.290	0.099	0.299	0.061	0.241	0.096	0.296
Dir OR Ind CS	0.127	0.333	0.127	0.334	0.162	0.369	0.100	0.301

Table 1: Descriptive statistics for novel patents

	(1) Has followers	(2) Has external followers	(3) Has internal followers	(4) Followed by external large firm	(5) Followed by external small firm	(6) Followed by external uni
Comp-Small	-0.025	-0.040	0.062	-0.098*	0.128*	-0.033
	(0.537)	(0.325)	(0.073)	(0.042)	(0.021)	(0.545)
University	0.053	0.060	0.099**	-0.027	0.174***	0.042
	(0.138)	(0.091)	(0.001)	(0.435)	(0.000)	(0.358)
No. IPC6	0.022***	0.024***	0.007**	0.009***	0.007	0.002
	(0.000)	(0.000)	(0.007)	(0.000)	(0.068)	(0.599)
No. of new pairs	0.001	0.001	0.003	0.000	0.003*	0.005*
	(0.687)	(0.705)	(0.052)	(0.909)	(0.037)	(0.024)
Constant	0.541***	0.535***	0.044	0.778***	0.272*	0.498***
	(0.000)	(0.000)	(0.647)	(0.000)	(0.028)	(0.000)
Observations	1142	1142	1142	701	701	701
r2	0.099	0.113	0.070	0.100	0.119	0.094
Sample	Novel patents	Novel patents with followers	Novel patents with followers	Novel patents with ext followers	Novel patents with ext followers	Novel patents with ext followers

*p values in parentheses. + $p < 0.1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$*
includes application year dummies and controls for a set of 6-digit main uspc

Table 2: OLS estimates of follow-on

	(1) Direct CS	(2) Indirect CS	(3) Indirect CS-int	(4) Indirect CS-ext
Comp-Small	0.084***	0.074^{+}	0.385***	-0.029
	(0.000)	(0.093)	(0.000)	(0.452)
University	-0.021*	0.005	-0.034	0.009
	(0.040)	(0.873)	(0.418)	(0.803)
No. IPC6	0.005*	0.005	0.007	0.003
	(0.011)	(0.118)	(0.193)	(0.379)
No. of new pairs	-0.001	-0.001	-0.001	-0.000
	(0.432)	(0.467)	(0.632)	(0.910)
Constant	0.003	0.242*	-0.050	0.266**
	(0.961)	(0.015)	(0.750)	(0.005)
Observations	1142	730	193	701
r2	0.089	0.093	0.396	0.092
Sample	Novel patents	Novel patents with followers	Novel patents with int followers	Novel patents with ext followers

*p values in parentheses. + $p < 0.1$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$*
includes application year dummies and controls for a set of 6-digit main uspc

Table 3: OLS estimates of OB (I)

	(1)	(2)	(3)	(4)
	Indirect CS	Indirect CS	Indirect CS-ext	Indirect CS-ext
Internal follow-on	0.118***	0.044	0.056^{+}	0.031
	(0.000)	(0.300)	(0.087)	(0.497)
Comp-Small		-0.055		-0.056
		(0.246)		(0.234)
University		-0.004		-0.008
		(0.932)		(0.845)
Internal follow-on=1 × Comp-Small		0.367***		0.078
		(0.000)		(0.388)
Internal follow-on=1 × University		0.016		0.042
		(0.829)		(0.577)
No. IPC6	0.005	0.005^{+}	0.002	0.002
	(0.098)	(0.068)	(0.427)	(0.407)
No. of new pairs	-0.001	-0.001	-0.000	-0.000
	(0.456)	(0.538)	(0.839)	(0.881)
Constant	0.177*	0.197**	0.195**	0.201**
	(0.014)	(0.006)	(0.006)	(0.005)
Observations	730	730	701	701
r2	0.107	0.135	0.094	0.096
Sample	Novel patents with followers	Novel patents with followers	Novel patents with external followers	Novel patents with external followers

*p values in parentheses. + p < 0.1 * p < 0.05, ** p < 0.01, *** p < 0.001*
includes application year dummies and controls for a set of 6-digit main uspc

Table 4: Internal follow-on and indirect commercial success