The Value of Failures in Pharmaceutical R&D

Jing-Yuan CHIOU*
Laura MAGAZZINI†
Fabio PAMMOLLI‡
Massimo RICCABONI§

PRELIMINARY DRAFT¶
This version: April 15, 2011

Abstract

We build a cumulative innovation model where both success and failure provide valuable information for future research. To test this learning mechanism, we use a dataset covering outcomes of world-wide R&D projects in the pharmaceutical industry, and proxy knowledge flows with forward citations received by patents associated with each project. Empirical results confirm theoretical predictions that patents associated with successfully completed projects (i.e., leading to drug launch in the market) receive more citations than those associated to failed projects, which in turn are cited more often than patents lacking clinical or preclinical information. We therefore offer evidence of the value of failures as research inputs in (pharmaceutical) innovation.

Keywords: R&D competition, patent policy, pharmaceutical industry
JEL codes: D23; D83; O34

*IMT Lucca Institute For Advanced Studies
†University of Verona, Department of Economics
‡IMT Lucca Institute For Advanced Studies, and CERM Foundation
§University of Trento, DISA

¶Paper prepared for presentation at the 4th ZEW Conference on Economics of Innovation and Patenting (Mannheim, May 19-20, 2011)

1 Introduction

In this paper we look at the features of the learning mechanisms and R&D competition in the pharmaceutical industry. We build a cumulative innovation model where each research project may provide useful information for future research. Differently from the standard single-research-line model, such as Green and Scotchmer (1995) and O'Donoghue (1998), we consider a scientific space consisting of multiple research lines, or trajectories (Aghion *et al.*, 2008). Several trajectories are possible candidates to successfully reach a final invention (e.g. to develop and launch a drug in the market). But, for simplicity, we consider the case where at most one trajectory is the "right" approach. That is, trajectories compete with each other for success, and it may well happen that no trajectories can lead to the final invention. In this case, a piece of "positive" information that raises the successful likelihood of a trajectory will at the same time reduce that of the other trajectories. Similarly, "negative" information, e.g., when a research project on some trajectory fails, although reduce the successful probability of the experimented trajectory, will simultaneously raise the prospects of other trajectories. In a simple way, this captures the idea that both success and failure convey useful information for subsequent innovation.

Empirical analysis in the pharmaceutical domain reveals that, indeed, technological competencies are accumulated building both on successfully developed compounds as well as on failures (Magazzini *et al.*, 2009). On the one side, it is not surprising that marketed products play an important role in guiding subsequent research efforts of both the innovating firm and its rivals. On the other side, also failures substantially spur rival innovative efforts.

The model is tested using data from a comprehensive dataset about the innovative activity of pharmaceutical and biotechnology firms, including R&D project level data, patents, citations, and collaborations, allowing us to explore the nature of technological advances and of the underlying learning regime, shaping the industrial patterns of innovative activity (Nelson and Winter, 1982; Dosi, 1988; Winter, 1994; Malerba and Orsenigo, 1993; Breschi *et al.*, 2000). We exploit the information provided by patents and patent citations, as a proxy for research efforts and to identify the different approaches or trajectories in order to characterize the dynamic nature of the innovation process (Trajtenberg, 1990; Lanjouw and Schankerman, 1999; Harhoff *et al.*, 1999; Jaffe *et al.*, 2000; Trajtenberg *et al.*, 1997; ,Jaffe and Trajtenberg, 2002).

The pharmaceutical industry represents a unique framework for innovation studies, given

that: (i) the industry is a textbook example of a "science-based" sector (Pavitt, 1984): strong linkages exist between drug development and the scientific advances in the "Open Science", leading firms to dissect and analyze an increasing number of techniques, trajectories and exploration strategies (Orsenigo *et al.*, 2001); (ii) patents play an important role as means for protecting innovation – in exchange of the full disclosure of the characteristics of the innovation (Mansfield, 1986; Cohen *et al.*, 2000; Arundel and Kabla, 1998); (iii) the pharmaceutical R&D process is characterized by a large presence of knowledge spillovers, as rival research results are positively correlated with firm productivity (Henderson and Cockburn, 1994; Henderson and Cockburn, 1996); (iv) the innovation process is characterized by strong uncertainty and high failure rates in drug discovery and development (Munos, 2010; Pammolli *et al.*, 2011).

The paper is organized as follows. Section 2 describes the model and its main assumptions. Section 3 describes our empirical strategy, and the results are reported in Section 4. Section 5 concludes.

2 Theory

We consider scientific research as a cumulative process that consists of a series of experiments to discover the "right" or successful approach for product development (e.g., drugs). The technology space contains $I > 1$ paradigms. Within each paradigm, there are $J > 3$ candidate approaches or trajectories. Different trajectories in a paradigm are competing for success, but not different paradigms. That is, within one paradigm, there is at most one trajectory leading to success (e.g., a marketable drug); but whether a paradigm contain a successful trajectory is not affected by that of other paradigms. The true state of nature concerning a paradigm is described by which trajectory, if any, is the successful one.

A research project is an experiment on a trajectory. An experiment generates a piece of evidence concerning whether the experimented trajectory can eventually succeed. Let an experiment be run on trajectory j. Three outcomes may arise, $\tau^j \in \{s, f, n\}$, where a result $\tau^j = s$ means "success," while $\tau^j = f$ and n represent "failure" and "no result," respectively. With probability β, the experimental result coincides with the true state: the outcome is a good sign $\tau^j = s$ when trajectory j will succeed, and a bad sign $\tau^j = f$ when it will fail. With probability γ, the experiment generates no useful result: the occurrence of the signal

$\tau^j = n$ doesn't depend on the true state. And with the remaining probability $1-\beta-\gamma$, the experiment delivers the wrong result. We impose the following assumption so that a result $\tau_j \in \{s, f\}$ remains informative.

Assumption 1. *(Informative experiments)* $\beta > 1-\beta-\gamma > 0$.

A lower γ raises the likelihood that the experiment delivers some outcome (success or failure), and a higher β ensures that this outcome provides more information, namely, it is more aligned with the true state of nature. These parameters capture the "quality" of experiment and determine how informative an experiment outcome is.

Remark 1. (Other types of research). This type of research is "applied" research in the sense that the experiment is conducted only on one trajectory within on paradigm, despite its informational externality to be shown below. On the other hand, "basic" or "fundamental" research can be thought of as experiments delivering direct results about several trajectories, within or across paradigms according to how "fundamental" the research could be. This distinction between basic and applied research stresses not the timing of invention, but the contribution to the knowledge accumulation process. Another type of research is research tools, which can be modeled as an invention that increases the precision of applied research, i.e., a better research tool increases β. ‖

Given this information structure, we use Bayesian updating to calculate the successful probability in the research process. After t experiments conducted in the same paradigm, denote the profile of each trajectory's success probability as $\{\hat{\alpha}_t^j\}_{j=1,2,\ldots,J}$, with $0 \leq \sum_{j=1}^J \hat{\alpha}_t^j \leq 1$ due to mutually exclusive success of each trajectory. Let the $t+1$th experiment be conducted on trajectory j. The realization of the experiment outcome $\tau_{t+1}^j \in \{s, f, n\}$ causes new assessment of the success probability according to Bayesian updating. Let $k \in \{1, 2, \ldots, J\}$ and $k \neq j$,

- $\tau_{t+1}^j = s$: it occurs with probability $\hat{\alpha}_t^j\beta + (1-\hat{\alpha}_t^j)(1-\beta-\gamma)$, and

$$\hat{\alpha}_{t+1}^j = \frac{\hat{\alpha}_t^j\beta}{\hat{\alpha}_t^j\beta + (1-\hat{\alpha}_t^j)(1-\beta-\gamma)} > \hat{\alpha}_t^j, \quad \hat{\alpha}_{t+1}^k = \frac{\hat{\alpha}_t^k(1-\beta-\gamma)}{\hat{\alpha}_t^j\beta + (1-\hat{\alpha}_t^j)(1-\beta-\gamma)} < \hat{\alpha}_t^k; \quad (1)$$

- $\tau_{jt+1} = f$: it occurs with probability $\hat{\alpha}_t^j(1-\beta-\gamma) + (1-\hat{\alpha}_t^j)\beta$, and

$$\hat{\alpha}_{t+1}^j = \frac{\hat{\alpha}_t^j(1-\beta-\gamma)}{\hat{\alpha}_t^j(1-\beta-\gamma) + (1-\hat{\alpha}_t^j)\beta} < \hat{\alpha}_t^j, \quad \hat{\alpha}_{t+1}^k = \frac{\hat{\alpha}_t^k\beta}{\hat{\alpha}_t^j(1-\beta-\gamma) + (1-\hat{\alpha}_t^j)\beta} > \hat{\alpha}_t^k; \quad (2)$$

and

- $\tau_{jt+1} = n$: it occurs with probability $\hat{\alpha}_t^j\gamma + (1 - \hat{\alpha}_t^j)\gamma$, and

$$\hat{\alpha}_{t+1}^j = \frac{\hat{\alpha}_t^j\gamma}{\hat{\alpha}_t^j\gamma + (1 - \hat{\alpha}_t^j)\gamma} = \hat{\alpha}_t^j, \quad \hat{\alpha}_{t+1}^k = \frac{\hat{\alpha}_t^k\gamma}{\hat{\alpha}_t^j\gamma + (1 - \hat{\alpha}_t^j)\gamma} = \hat{\alpha}_t^k. \tag{3}$$

By mutual exclusivity, a trajectories's positive outcome "crowds out" the prospects of other trajectories in the same paradigm. More interestingly, a failed experiment reduces the success probability of that trajectory (and the whole paradigm), but at the same time *increases* the probability that the successful route may hide in other trajectories. In other words, both successful and failed experiments are informative.

To empirically test this insight, we use forward citations to assess the knowledge contribution of a patented technology. We consider a very simple R&D decision and patent citation generation process. Strategic interactions at the both the final market and R&D competition stages could be introduced to enrich the model. But as a first step, we ignore these concerns in order to focus on the learning mechanism outlined above. Assume that all experiments in a research field are made public and receive patents, regardless of the outcome of the experiment. In the patent application all the previous patents in the same fields are cited, but not in other fields, i.e., there is no strategic citation. The number of forward citations a patent receives, then, is positively correlated with subsequent inventors' incentive to enter into the same research field. Denote $\hat{\alpha}_t^* \equiv \max\{\hat{\alpha}_t^j\}$ as the highest successful probability in the paradigm after t experiments, and let the corresponding trajectory be j^*. For simplicity, assume that a firm's incentive to enter and start a research project in this paradigm is increasing in the probability $\hat{\alpha}_t^*$. Post entry, the firm, with the knowledge $\{\hat{\alpha}_t^j\}$, conducts experiment on the trajectory j^*. In sum, these assumptions imply that a patent will receive more forward citation if its outcome raises the highest successful probability of the paradigm by a larger magnitude.

We now derive two hypotheses. First, consider the firm's experiment on trajectory j^*. An uninformative result ($\tau_{t+1}^{j^*} = n$) does not change the knowledge stock and so the entry incentives, $\{\hat{\alpha}_{t+1}^j\} = \{\hat{\alpha}_t^j\}$. But a positive sign ($\tau_{t+1}^{j^*} = s$) raises the highest successful probability. A positive experimental result maintains the status of j^* as the most promising trajectory, and raises its successful probability, $\hat{\alpha}_{t+1}^{j^*} > \hat{\alpha}_t^{j^*}$. Therefore, a patent associated with a successful experiment should receive more forward citations than one with no informative outcome.

If the experiment outcome is negative, ($\tau_{t+1}^{j^*} = f$), then the successful probability of

trajectory j^* becomes lower. But the information brought by a failed experiment may sufficiently boost the successful probability of other trajectories, such that the highest successful probability after incorporating the new information is larger than $\hat{\alpha}_t^{j^*}$. That is, there may exist $k \neq j^*$ such that

$$\hat{\alpha}_{t+1}^k = \frac{\hat{\alpha}_t^k \beta}{\hat{\alpha}_t^{j^*}(1-\beta-\gamma) + (1-\hat{\alpha}_t^{j^*})\beta} > \hat{\alpha}_t^{j^*} \Leftrightarrow \beta \left(\hat{\alpha}_t^{j^*} - \frac{\hat{\alpha}_t^{j^*} - \hat{\alpha}_t^k}{\hat{\alpha}_t^{j^*}} \right) > (1-\beta-\gamma)\hat{\alpha}_t^{j^*}. \quad (4)$$

When this condition holds, a patent associated with failed outcome will also receive more forward citation than one with uninformative outcome.[1] In general, this condition requires that the difference in successful probability between the most promising and the second most promising trajectory is not too large. For instance, if at the initial state, i.e., before any experiments are run, the prior belief is characterized by uniform distribution, $\hat{\alpha}_0^j$ is the same for all j, then the condition holds. Or, if the number of candidate trajectories J is large enough, and if past experiments are all failures, then there is always some "untested" trajectory k. The successful probability of trajectory k is the same as $\hat{\alpha}_t^{j^*}$; the condition also holds. These two scenarios seem fit the pharmaceutical context well, where most patents either have no informative results in their backward citation, or only cite past failures.[2]

Second, we compare positive and negative experimental outcomes. Fixing $\hat{\alpha}_t^{j^*}$, compare the highest successful probability after an experiment is conducted on trajectory j^*. For $k \neq j^*$,

$$\frac{\hat{\alpha}_t^{j^*}\beta}{\hat{\alpha}_t^{j^*}\beta + (1-\hat{\alpha}_t^{j^*})(1-\beta-\gamma)} > \frac{\hat{\alpha}_t^k \beta}{\hat{\alpha}_t^{j^*}(1-\beta-\gamma) + (1-\hat{\alpha}_t^{j^*})\beta} \quad (5)$$

$$\Leftrightarrow \quad \hat{\alpha}_t^{j^*}[(1-\hat{\alpha}_t^{j^*}) - \hat{\alpha}_t^k]\beta > [\hat{\alpha}_t^k(1-\hat{\alpha}_t^{j^*}) - \hat{\alpha}_t^{j^*2}](1-\beta-\gamma), \quad (6)$$

because $\beta > 1-\beta-\gamma$, and because

$$\hat{\alpha}_t^{j^*}[(1-\hat{\alpha}_t^{j^*}) - \hat{\alpha}_t^k] > \hat{\alpha}_t^k(1-\hat{\alpha}_t^{j^*}) - \hat{\alpha}_t^{j^*2} \quad \Leftrightarrow \quad \hat{\alpha}_t^{j^*} > \hat{\alpha}_t^k, \quad (7)$$

where $\hat{\alpha}_t^{j^*} + \hat{\alpha}_t^k \leq 1$. A positive result must raise the highest successful probability by a larger amount than a negative result, and so must receive more forward citations. Note that this is robust to the history of research, as summarized in $\{\hat{\alpha}_t^j\}$.

[1] If we assume that patents with informative outcomes receive more forward citations than those with uninformative outcomes, then this relationship holds directly from the information spillover of failed experiments.

[2] In our data, the majority of patents (80.34%) do not contain any informative outcomes in their backward citation, and only 15.31% of patents contain one known outcome in the backward citation.

Last, note that as β increases, i.e, as each experiment has a higher probability to deliver an informative result, then both a success and failure becomes more informative:

$$\frac{\partial}{\partial\beta}\left(\frac{\hat{\alpha}_t^{j^*}\beta}{\hat{\alpha}_t^{j^*}\beta+(1-\hat{\alpha}_t^{j^*})(1-\beta-\gamma)}\right)>0, \quad \text{and} \quad \frac{\partial}{\partial\beta}\left(\frac{\hat{\alpha}_t^{k}\beta}{\hat{\alpha}_t^{j^*}(1-\beta-\gamma)+(1-\hat{\alpha}_t^{j^*})\beta}\right)>0. \tag{8}$$

In the empirical analysis, we use the number of scientific publication a patent cites (in backward citation) to capture the "quality" of the experiment associated with the patent. Citing more scientific publications implies, say, absorbing more knowledge and so leads to a more thorough experiment design, which raises β. We then expect a higher number of forward citations.

To summarize, we will test the following two hypotheses:

Hypothesis 1. (Information spillover) A patent associated to a successful project receives more citation than a patent associated to a failed project or a patent whose project is inconclusive. When condition (4) holds, a patent corresponding to a failed project also receives more forward citation than one corresponding to project that is inconclusive.

Hypothesis 2. (Scientific publication) A patent receives more forward citations when it cites more scientific publications.

3 Empirical strategy

The empirical analysis builds on a comprehensive dataset on the innovative activity undertaken within the pharmaceutical industry, including R&D project level data, patents, citations, and collaborations maintained by the CERM Foundation, Rome, Italy. The database contains information about all pharmaceutical and biotechnology patents granted by the USPTO since 1965, including backward and forward citations.[3] Firm data at the level of specific R&D projects worldwide in the last 30 years are also available. The database tracks the development history of more than 22,000 R&D projects, starting from patent application to the latest stage of drug development through preclinical, clinical development and commercialization. In case of aborted projects, the database reports the time when the firm

[3]US patents are selected in the database on the basis of the the International Patent Classification (IPC) and US classification. Namely, pharmaceutical patents are defined as those in IPC classes A61K and A01N (Lanjouw and Cockburn, 2001) and we further include patents in US classes 424, 435, 514, and 800.

announces that the research around the compound has been terminated. By exploiting the information about the patents protecting the compound, the project data have been linked to patent data (number of forward citations up to May 2004, the application date, and the name of the assignee(s)).[4] Patent history is available for 49 per cent of the projects included in the database.

The matching of the different sets of data proved to be a formidable, large-scale task, that tied up a great deal of our research efforts for a long time, providing us a unique dataset that monitors R&D activities of pharmaceutical and biotechnology firms from patenting to commercialization (if any) of the protected compound.

We further selected the patents associated to R&D projects of candidate drugs whose outcome was a success (s), i.e. a new product launched on the market, or a failure (f), i.e. the project was stopped due to the emergence of toxicological effects or to lack of effectiveness. The final database encompasses information about 2,000 R&D projects with informative outcome and their associated patents, entering into clinical trials from 1977 to 2002. Henceforth, we refer to successful (failed) patents as the patent associated to successful (failed) R&D projects.

Citation patterns of our sample of patents is compared with patents whose protected compound has not entered preclinical or clinical development. For each (successful or failed) patent in the original sample, a patent has been randomly matched from the set of bio-pharmaceutical patents with the same application year, publication year, and IPC class but no information about preclinical or clinical development (n, no result). By comparing the citation pattern of failed and successful patents against the sample of "no info" patents, we will be able to ascertain the level of knowledge utilization and diffusion associated with each trajectory and the related R&D competition dynamics. The three groups resemble the three research outcomes considered in the model (successful: s; failed: f; no information on preclinical or clinical development: n).

Since the focus is on the pharmaceutical industry, only citations from subsequent patents

[4]For projects listing a patent granted by a patent office other than the USPTO, we considered the US patent in the same family as the one listed in the database. In case no US patent is identified, the project is not considered in the analysis. This choice has been driven by the fact our data sources only provide citation data for the US patents. Moreover, different patent examination procedures characterizes the the US and European patent offices, leading to large differences in the average number of citations per patent (Breschi and Lissoni, 2004; Michel and Bettels, 2001). Focusing only on US citations avoids the emergence of spurious results driven by different institutional settings. We further excluded old molecules and/or natural products, which do not have any associated patent.

in the pharmaceutical domain have been taken into account. We disregard self-citations and only consider citations made by other companies, proved to be a good proxy for knowledge spillovers (Jaffe *et al.*, 2000). On the contrary, self-citations are considered to be indicators of the cumulative nature of the technology and a measure of the extent to which innovators are able to reap the benefits of their own research (Hall *et al.*, 2001).

First, we employ the double-exponential function to model the citation lag distribution for successful and failed patents, against biopharmaceutical patents with no information about development. The model provides a flexible framework for studying the process underlying the generation of citations, where an exponential process by which knowledge diffuses is combined with a second exponential process by which knowledge becomes obsolete (Jaffe and Trajtenberg, 1996; Caballero and Jaffe, 1993).

The analysis allows us to characterize knowledge dynamics of successful and failed trajectories (as identified trough patents) with respect to the uninformative patents issued within the biopharmaceutical domain (Hypothesis 1).

Next, we test Hypothesis 2 and whether condition (4) is satisfied within the pharmaceutical domain. In particular, we run a regression where the dependent variable is the number of citations received by the patents adjusted on the basis of the estimated citation lag distribution, in order to reflect life-time citations. Variables on the right hand side measures the importance of scientific references and the level of knowledge accumulated within the relevant trajectory, controlling for characteristics of the technological classes (defined on the basis of the International Patent Classification, henceforth IPC), assignees and the patent-innovation itself.

The two sets of results will provide a comprehensive view of the dynamics underlying R&D competition in the biopharmaceutical domain, providing a clearcut test of the model developed in Section 2.

4 Results

□ **The value of failures.** Figure 1 compares the observed and estimated citation lag distribution functions for successful and failed patents, taking as a benchmark the citation lag distribution function of patents with no information about preclinical or clinical development. The x-axis reports the citation lag, i.e. the difference between the citing and the cited patent

grant year. It represents the time elapsed from the grant date. The vertical line drawn 8 years after patent grant corresponds to the average length from patent application to termination of the project.[5] The y-axis depicts the (average) observed and estimated citation intensities, i.e. the likelihood that any patent will be cited by the patents granted x years apart (Jaffe and Trajtenberg, 1996).

The observed citation lag distribution is computed as the ratio between the number of citations received by patents granted in year t from patents granted in year T and the theoretical number of potential citations ($T - t$ is the citation lag):

$$p(t, T, \tau) = \frac{C(t, T, \tau)}{N(t, \tau)\, N(T)}$$

where t indicates the grant year of the cited patent, T is the grant year of the citing patent, and τ represents the outcome of the associated R&D project ($\tau \in \{s, f, n\}$). The potential number of citations is given by the number of citations that would have been observed if all patents granted in year T would cite all patents granted in year t with outcome τ, that is equal to the product of the number $N(T)$ of patents granted in the citing year and the number $N(t, \tau)$ of patents granted in the cited year with a known outcome τ.

In order to estimate the theoretical citation lag distribution, we considered the following specification of the double-exponential function, modeling the likelihood that a patent granted in year T will cite a patent granted in year t, with $T > t$ (Jaffe, Trajtenberg, 1996):

$$p(t, T, \tau) = \delta_0(t, T, \tau) \exp[-\delta_1(\tau)(T - t)](1 - \exp[-\delta_2(T - t)]).$$

where δ_0 is linked to the overall likelihood of receiving a citation, whereas δ_1 and δ_2 are indicators of, respectively, the rate of obsolescence of knowledge (i.e., the rate at which new knowledge replaces the existing one) and the rate of diffusion of the knowledge related to the invention protected by the patent. We claim that the grant year of the citing and the cited patent only affect the average citation intensity δ_0, while the outcome of the project affect both the average citation intensity δ_0 and the rate of obsolescence δ_1.[6] Due to identification

[5] This is actually few months longer for marketed compounds, being equal to 7.8 years for failed R&D projects and to 8.3 years for marketed R&D projects. The value is coherent with previous studies analyzing average duration of the drug development process (Abrantes-Metz *et al.*, 2004)

[6] As in previous empirical literature dealing with this model, convergence problem forbids the estimation of the model where all the cited-year effects are considered. The problem is solved by introducing the cited-year effects defined on the basis of 5-year time periods.

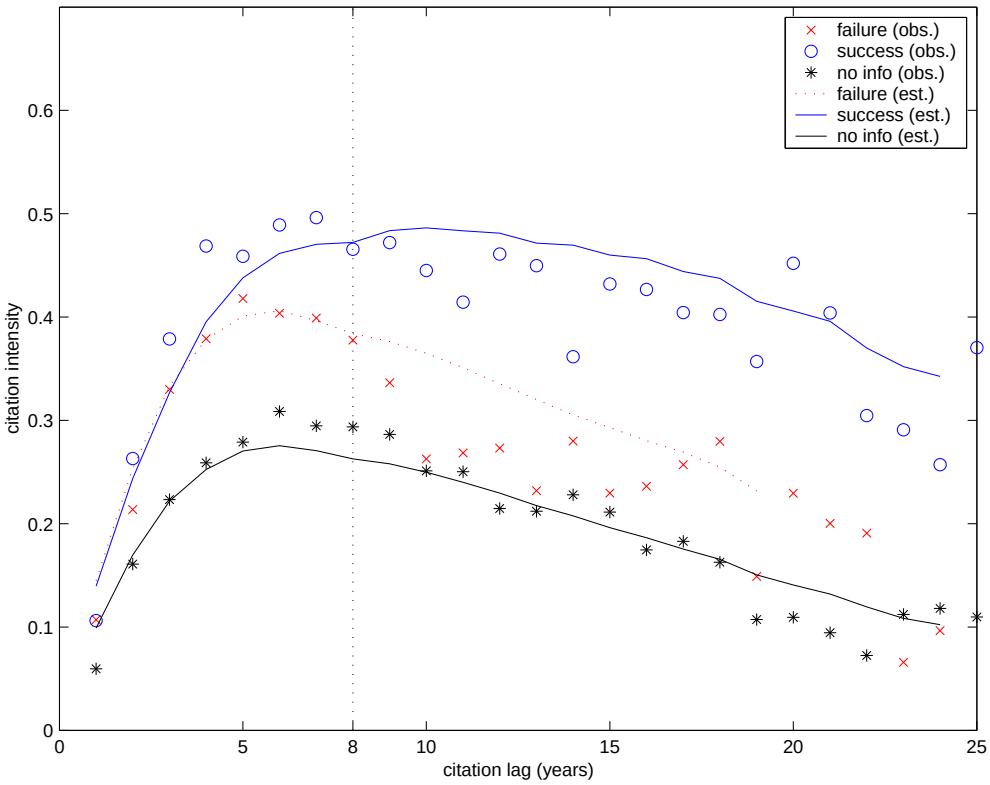

Figure 1: Observed and estimated citation lag distribution

Coefficients	Model 1	Model 2
δ_0^f	1.264 *(.182)**	1.329 *(.111)**
δ_0^s	1.309 *(.158)**	1.511 *(.108)**
δ_1	0.107 *(.008)**	0.084 *(.013)**
δ_1^f	0.827 *(.135)**	1.049 *(.108)**
δ_1^s	0.554 *(.087)**	0.604 *(.077)**
δ_2	0.114 *(.014)**	0.248 *(.042)**
Cited year effects	no	yes
Citing year effects	no	yes
R-squared	.609	.641

* statistically significant at 5% level.

Table 1: Double-exponential function, results of estimation, dependent variable: citation intensity

problems, the rate of diffusion (δ_2) is considered constant over time and across the three sets of patents.

According to our model, we expect that successful patents receive on average a higher number of citations with respect to failed patents. The model specification however allows us also to dig further into the dynamics of knowledge utilization and diffusion and to analyze the speed and extent to which existing knowledge is "picked up" in the case of failures and successes, as a proxy for the diffusion and utilization of the associated trajectory. Results, reported in Table 1, are obtained by nonlinear least squares estimation, weighting each observation by $[N(t,\tau)\,N(T)]^{1/2}$. The lines depicted in Figure 1 are obtained by taking the average,[7] for each lag, of the fitted values from Model 2.

Coherently with our model and previous literature showing that the number of citations received by a patent is positively associated to its value,[8] citations turn out to be related to the outcome of the project. The observed and estimated distributions indicate that, on average, failed patents receive a number of citations that is lower than the number of

[7]Both in the case of observed and estimated citation lag distributions, weighted averages are considered, where the weights are the same as the ones used in the estimation process.

[8]See Trajtenberg (1990); Lanjouw and Schankerman (1999); Harhoff *et al.* (1999); Jaffe *et al.* (2000); Trajtenberg *et al.* (1997); Jaffe and Trajtenberg (2002).

citations received by patents associated to successful projects (compare the estimated values of the δ_0^f and δ_0^s coefficients). Furthermore, coherently with Hypothesis 1, both sets receive a higher number of citations than the average patent in the biopharmaceutical domain with no information about clinical or preclinical development ($\tau = n$). The estimate of the δ_0 coefficients associated to failed and successful patents are higher than 1, indicating that patents associated to preclinical or clinical development, irrespective of their outcome, are more likely to receive a citation than biopharmaceutical patents with no informative results (taken as the reference category). The analysis shows the existence of a value associated also with failed patents: even though the compound associated with the patent will never reach the market (e.g., due to the emergence of toxicological problems or lack of effectiveness); the opened research trajectory is subsequently exploited by firms other than the original innovator.

It is interesting to note that within the first 5 years from patent grant, no significant difference is detected between failed and successful patents, whereas starting from year 5 the two series start to diverge in a significant way. The analysis of the estimates of δ_1 reveals an important difference between successful and failed compounds in terms of knowledge obsolescence. Under this perspective, failed and the uninformative patents exhibit very similar dynamics (the δ_1 associated with associated with failed patents is very close to 1, pointing to no differences between failed patents and the uninformative patents). On the contrary, the knowledge embedded in patents protecting marketable compounds becomes obsolete less quickly than the other patents in the biopharmaceutical domain (the δ_1 associated with successful patents is lower than 1). Indeed, the citation intensity of marketed compounds is rather stable after commercialization, whereas the citation intensity of failed patents decreases substantially.[9]

The maximal citation frequency for failed patents is earlier in time than the maximal citation frequency of successful patents.[10]

[9]Also note that the larger departures between the estimated and observed citation lag distribution in the case of failed patents is registered right after the average time when the project is stopped. This might point to the fact that the termination of the research around a compound/mechanism of action is a major signal for rival firms that nonetheless regain interest after few years from the time of discontinuation and the citation intensity of failed patents is still higher than the citation intensity of the patents with no preclinical or clinical development (no info), also many years after discontinuation. On this issue, we asked a pharmacologist to extensively inspect the patents citing failed projects in search of a reason for the citation, finding no instance of "negative" citations, rather citations refer to pharmacological action or the structure of the compound (i.e., to the trajectory / paradigm opened up by the original innovator).

[10]Our estimates are coherent with the estimates of the Drugs and Medical sector presented by Hall et al.

□ **When are failures valuable?** The disclosure of the information about the compound under study in patents and the advances in science sets the ground for a "race" for reaching the market, where competitors start exploring the new research arena pursuing parallel research trajectories even though the outcome is still highly uncertain. Competition on the R&D side in the pharmaceutical industry is substantial and firms entering the new research arena build both on failures and successes (Figure 1).

We now build a regression model that looks at failures and successes separately and at the drivers of their citation intensity. The analysis will allow us to understand the characteristics of innovations and trajectories that are subject to a higher degree of exploration by rival firms. In particular we will test Hypothesis 2, explicitly considering in the regression the share of scientific publications cited over the total number of citations (patent and non-patent references).

The dependent variable is the number of citation received during the life time of the patent, where the observed citation frequency has been adjusted using the estimated coefficient of the citation lag distribution function (on the basis of estimated coefficients in Table 1).

The independent variables are listed in Table 2 and aim at capturing the characteristics of the cited patent (pt), of the IPC class (ipc) and of the patenting firm (asg).

Patents characteristics are measured using the indicators developed by Trajtenberg *et al.* (1997) on the basis of backward citations. The importance of scientific sources with respect to technological ones within the patent is captured by science, which is the ratio between the non-patent references and the total number of references (previous patents or previous scientific literature) listed in the patent. The closer to 1, the larger the scientific underpinnings of the research, relying more heavily on the scientific literature rather than on previous patents. Hypothesis 2 posits that this variable has a positive effect on the number of forward citations received. We also consider the share of self-citations in the patents (selfc)

(2001), with the exception of the estimated δ_2 which is lower. This might be explained by the fact that we only consider citations by institution other than the original assignee, which can require a longer time span with respect to self-citations to manifest. Moreover, an interesting pattern emerges in their results when comparing Drugs and Medical to other sectors. The citation lag distribution for this sector is more flat, whereas the citation lag distribution functions for the sectors of Computers and Communications, Electrical and Electronics, Chemical, and Mechanical have higher peak earlier in time. Knowledge in the Drugs and Medical sector diffuses less rapidly and takes a longer time to become obsolete. Important information about the protected compounds in terms of toxicological effects and effectiveness are revealed over time leading to a lengthier process of citation within this industry.

Variable	**Char.**	**Description**	**Mean**	σ
science	(pt)	Science Index*	0.34	0.35
cite-sf	(pt)	Number of backward citations with known outcome (either failure or success)	0.25	0.59
selfc	(pt)	Share of self-citations of the patent*	0.14	0.29
orig	(pt)	Index of originality of the patent*	0.43	0.37
timeb	(pt)	Average time lag*	5.56	4.50
importb	(pt)	Importance of cited patents*	122.4	640.7
nimp	(ipc)	Number of firms operating in the same IPC class	110.21	126.32
conc	(ipc)	Concentration of the IPC Class (Herfindhal index)	0.10	0.21
coree	(asg)	Share of firm patent within the same technology class (IPC)	0.14	0.23
dbf	(asg)	Dummy equal to 1 if the originating firm is a dedicated biotechnology company	0.19	0.39
pro	(asg)	Dummy equal to 1 if the originator is a public research organization	0.09	0.29

* defined as in Trajtenberg *et al.* (1997).

Char.: characteristic of (pt) patent; (ipc) IPC class; (asg) assignee.

Table 2: Description of the variables

that measures the extent to which benefits from research antecedents are appropriated by the firm and help in understanding whether the patent belongs to a research trajectory strongly rooted within the company. The index of originality of the patent (orig) measures the breadth of its technological roots,[11] whereas the importance of the previous patents cited by the patent under investigation is measured by importb which takes into account the number of backward citations in the patents and the number of citations they receive.[12] Besides this variable, we also include the number of citations with known outcome in the regressions to proxy the level of information that characterizes the patent's research trajectory (cite-sf). Finally, timeb measures the time distance between the citing and the cited patents. The higher timeb, the older the sources the patent builds upon.[13]

As far as the characteristics of the technological class of the patent, we consider the number of firms active in the IPC class, and the Herfindahl index of concentration computed at the technology class level on the basis of patent counts.

As for the assignee, we take into account the share of firm patent within the same technology class (IPC), and two dummy variables indicating whether the patentee is a dedicated biotechnology company (DBF) or a public research organization (PRO). The largest share of patents in our sample are assigned to pharmaceutical companies: 19 per cent of patents are assigned to DBF, and 9 per cent to PRO. Finally, cited year dummies are included in all the specifications.

Results of the estimation of a Poisson regression model are reported in Table 3 where we distinguish the dynamics characterizing successes and failures. The Poisson modeling has been preferred to simple regression of a log-linearized equation, as recent research shows that, under heteroskedasticity, OLS estimation of log-linearized models lead to biased estimates of the true elasticities (Santos Silva and Tenreyro, 2006). Furthermore, we record 20.77%

[11]The index is computed as an Herfindahl index of diversification, considering the share of backward citation in each IPC class. The closer orig is to 1, the broader are the technological roots of the underlying research, i.e. they span many different IPC classes. The index is zero when all backward citations contained in the patent are classified within the same IPC class.

[12]The higher is the value of importb, the higher is the number of backward citations contained in the patent and the citations they receive.

[13]As compared to the descriptive statistics reported in Trajtenberg *et al.* (1997), no difference emerges with respect to the value of selfc. On the contrary, the average value of timeb in our sample is lower, indicating younger sources for our sample of patents, whereas the values of orig, science, and importb are higher. One important difference with the sample in Trajtenberg *et al.* (1997) relies in the fact that we only consider pharmaceutical patents, and citations are counted only within the pharmaceutical technological classes.

Variable	Failure	Success
science	.9689 *(.1729)**	.4841 *(.1744)**
cite-sf	-.2704 *(.0895)**	.0174 *(.0794)*
conc	-.2452 *(.3553)*	.0997 *(.2850)*
nimp	-.0009 *(.0005)*	.0005 *(.0005)*
selfc	-.2626 *(.1863)*	-.1983 *(.1909)*
orig	-.8878 *(.1368)**	-.5642 *(.1534)**
timeb	-.0030 *(.0106)*	-.0253 *(.0130)*
importb	.0002 *(.4E-4)**	.0003 *(.6E-4)**
coree	.1886 *(.2292)*	-.0587 *(.1630)*
dbf	.7057 *(.1085)**	.6450 *(.1289)**
pro	.6216 *(.1914)**	.3617 *(.1775)**
constant	2.672 *(.2422)**	3.247 *(.2853)**
Obs.	1,554	725
Log lik.	-21123.2	-11679.7

* statistically significant at 5% level.

Cited year (application) included in all regressions.

Table 3: Regression results (Poisson). Dependent variable: number of adjusted forward citations).

patents receiving zero citations.

Even though the magnitude of the coefficients differ between successes and failures, statistical significance of the chosen regressors does not change. The one exception is the coefficient associated with the number of citations with known outcome (cite-sf) that is negative for failures and seems to exhibit no effect for successes. The lower the number of informative outcomes, the higher the citations to failed patents. Put it differently, failures receive a higher number of citations (i.e. are the basis for subsequent innovations) in nascent fields as compared to fields where larger knowledge has been accumulated, leading us to claim that condition (4) is actually satisfied within the pharmaceutical domain. On the contrary, citation to successes does not seem to be driven by the level of knowledge accumulated along the trajectory.

Coherently with Hypothesis 2, patents with predominance of scientific sources over technological ones contribute more heavily to subsequent research.

The estimated coefficient of orig shows that patents with sparse technological roots receive a lower number of citations by other firms. These are likely patents within narrow fields of application, therefore being relevant only to the firms and institutions working within the same technological domain. Patents building on an important (in the sense of highly cited) knowledge base are more often subsequently cited.

As far as the characteristics of the patent assignee are concerned patents by DBF and PRO receive on average a higher number of citations. The former result is consistent with Hall *et al.* (2000) who find that in the pharmaceutical sector smaller biotechnology firms are more likely to average a higher citation rate. As a tentative explanation for this phenomenon, we propose that this is due the growing division of innovative labor and the wide network of collaborations among the different actors involved in the drug development process that has come to characterized the pharmaceutical industry (Arora and Gambardella, 1994; Powell *et al.*, 1996; Orsenigo *et al.*, 2001). The small biotechnology firms are highly specialized in the early stages of drug development, but they lack the resources and capabilities that are needed for the large clinical trials, therefore they are more likely to license-out their compounds to the large pharmaceutical companies with significant expertise with clinical trials. This is also true for PROs, strongly oriented toward basic science and the early stages of the innovation process in pharmaceuticals. As a result, DBF and PRO patents are more likely to be the object of an alliance and the basis of subsequent research by firms other than the original innovators, leading to knowledge transfer from the original innovator to the company that licensed-in the compound and continues the research around it.

The characteristics of the IPC class of the patent do not seem to exert a significant effect on the number of subsequent citations received by the patent.

5 Concluding discussion

The paper has explored the learning processes and R&D competition in the pharmaceutical domain. We developed a cumulative innovation model analyzing R&D competition and learning mechanisms along different research trajectories, that posits a role for successes as well as (under suitable conditions) failures as inputs to the innovative activities of the

innovation process. The model has been tested in the context of the worldwide pharmaceutical industry, a unique framework for studying issues related to innovation and innovative activities.

By relying on a comprehensive dataset of innovative activities in the pharmaceutical industry, we have provided evidence of the value of failures in the pharmaceutical innovation process. The analysis suggests the existence of a social value associated to failed projects and to the disclosure of the associated information.

Within this scenario, the discussion about patent scope becomes crucial in this industry, where research is highly cumulative in nature and firms enjoy knowledge spillovers spanning from internal and, to some extent, from external R&D projects, pointing to a trade-off that cannot be easily resolved. This poses problems for the optimal design of patent law. On the one side, it is necessary to fully reward early innovators for the technological foundation they provide to later innovators, but also later innovators should be rewarded adequately for their improvements and new products.[14]

References

[1] Abrantes-Metz, R.M., Adams, C.P., and Metz, A. (2004), Pharmaceutical Development Phases: A Duration Analysis, *Bureau of Economics, US FTC*.

[2] Aghion, P., Dewatripont, M., and Stein, J. (2008), Academic Freedom and Private Sector Focus, *The RAND Journal of Economics*, forthcoming.

[3] Arora, A., and Gambardella, A. (1994), The Changing Technology of Technological Change: General and Abstract Knowledge and the Division of Innovative Labour, *Research Policy* **23**(5), 523-532.

[4] Arundel, A., and Kabla, I. (1998), What Percentage of Innovations are Patented? Empirical Estimates for European Firms, *Research Policy* **27**(2), 127–141.

[5] Breschi, S., Malerba, F. and Orsenigo, L. (2000), Technological Regimes and Schumpeterian Patterns of Innovation, *The Economic Journal* **110**(463), 388–410.

[14] See Scotchmer (1991) for a detailed discussion of the optimal patent scheme in the case of cumulative knowledge.

[6] Breschi, S. and Lissoni, F. (2004), Knowledge Network From Patent Data: Methodological Issues and Target Research, CESPRI, Milan, Working Paper.

[7] Caballero, R.J. and Jaffe, A.B. (1993), How High are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth, NBER Working Paper.

[8] Cohen, W.M., Nelson, R.R. and Walsh, J. (2000), Protecting Their Intellectual Assets: Appropriability Conditions and Why US Manufacturing Firms Patent (or Not), NBER Working Paper.

[9] David, P.A., Mowery, D. and Steinmueller, W.E. (1992), Analysing The Economic Payoffs From Basic Research, *Economics of Innovation and New Technology* **2** (1), 73–90.

[10] Dosi, G. (1988), Sources, Procedures, and Microeconomic Effects of Innovation, *Journal of Economic Literature* **26**(3), 1120–1171.

[11] Green, J. and Scotchmer, S. (1995), On the Division of Profit in Sequential Innovation, *The RAND Journal of Economics*, **26**, 20–33.

[12] Hall, B.H., Jaffe, A.B. and Trajtenberg, M. (2000), Market Value and Patent Citations: A First Look, NBER Working Paper.

[13] Hall, B.H., Jaffe, A.B. and Trajtenberg, M. (2001), The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools, NBER Working Paper.

[14] Harhoff, D., Narin, F., Scherer, F.M. and Vopel, K. (1999), Citation Frequency and the Value of Patented Inventions, *The Review of Economics and Statistics* **81**(3), 511–515.

[15] Henderson, R. and Cockburn, I. (1994), Racing to Invest? The Dynamics of Competition in Ethical Drug Discovery, *Journal of Economics and Management Strategy* **3**(3).

[16] Henderson, R. and Cockburn, I. (1996), Scale, Scope, and Spillovers: The Determinants of Research Productivity in Drug Discovery, *The RAND Journal of Economics* **27**(1), 32-59.

[17] Jaffe, A.B. and Trajtenberg, M. (1996), Flows of Knowledge from Universities and Federal Labs: Modeling the Flowof Patent Citations Over Time and Across Institutional and Geographic Boundaries, NBER Working Paper.

[18] Jaffe, A.B., Trajtenberg, M. and Fogarty, M.S. (2000), The Meaning of Patent Citations: Report on the NBER/Case-Western Reserve Survey of Patentees, NBER Working Paper.

[19] Jaffe, A.B. and Trajtenberg, M. (2002), Patents, Citations and Innovations: a Window on the Knowledge Economy, MIT Press.

[20] Lanjouw, J. and Schankerman, M. (1999), The Quality of Ideas: Measuring Innovation with Multiple Indicators, NBER Working Paper.

[21] Lanjouw, J.O. and Cockburn, I.M. (2001), New Pills for Poor People? Empirical Evidence after GATT, *World Development* **29**(2), 265–289.

[22] Magazzini, L., Pammolli, F., Riccaboni, M. and Rossi, M.A. (2009), Patent Disclosure and R&D Competition in Pharmaceuticals, *Economics of Innovation and New Technology* **18**(5), 467-486.

[23] Malerba, F. and Orsenigo, L. (1993), Technological Regimes and Firm Behavior, *Industrial and Corporate Change* **2**(1), 45-71.

[24] Mansfield, E. (1986), Patents and Innovation: An Empirical Study, *Management Science* **32**(2), 173-181.

[25] Michel, J. and Bettels, B. (2001), Patent Citation Analysis. A Closer Look at the Basic Input Data from Patent Search Reports, *Scientometrics* **51**(1), 185–201.

[26] Munos, B (2010), Lessons From 60 Years of Pharmaceutical Innovation, *Nature Reviews Drug Discovery* **8**, 959–968.

[27] Nelson, R.R. and Winter, S.G. (1982), An Evolutionary Theory of Economic Change, Belknap Press.

[28] O'Donoghue, T. (1998), A Patentability Requirement for Sequential Innovation, *The RAND Journal of Economics* **29**, 654–79.

[29] Orsenigo, L., Pammolli, F. and Riccaboni, M. (2001), Technological Change and Network Dynamics: Lessons from the Pharmaceutical Industry, *Research Policy* **30**(3), 485-508.

[30] Pammolli, F., Magazzini, L. and Riccaboni, M. (2011), The Productivity Crisis in Pharmaceutical R&D, *Nature Reviews Drug Discovery*, forthcoming.

[31] Pavitt, K. (1984), Sectoral Patterns of Technical Change: Towards a Taxonomy and a Theory, *Research Policy* **13**(6), 343–373.

[32] Powell, W.W., Koput, K.W. and Smith-Doerr, L. (1996), Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology, *Administrative Science Quarterly* **41**(1).

[33] Santos Silva, J.M.C., and Tenreyro, S. (2006), The Log of Gravity, *Review of Economics and Statistics* **88**(4), 641-658.

[34] Scotchmer, S. (1991), Standing on the Shoulders of Giants: Cumulative Research and the Patent Law, *The Journal of Economic Perspectives* **5**(1), 29–41.

[35] Trajtenberg, M. (1994), A Penny for Your Quotes: Patent Citations and the Value of Innovations, *The RAND Journal of Economics* **21**(1), 172–187.

[36] Trajtenberg, M., Henderson, R. and Jaffe, A. (1997), University vs. Corporate Patents: A Window on the Basicness of Innovations, *Economics of Innovation and New Technology* **5**(1), 19–50.

[37] Winter, S.G. (1994), Schumpeterian Competition in Alternative Technological Regimes, *Journal of Economic Behavior and Organization* **5**, 287–320.