LSE RICAFE 2

Working Paper No. 034

# A Breath of Fresh Air? Firm Types, Scale, Scope and Selection Effects in Drug Development

**Ashish Arora**

**Heinz School, Carnegie Mellon University, USA**

**Alfonso Gambardella**
**Bocconi University, Milan, Italy**

**Laura Magazzini**
**University of Verona, Italy**

**Fabio Pammolli**
**University of Florence and IMT – Lucca Inst. of Adv. Studies, Italy**

**March 2007**

**RICAFE2 - Regional Comparative Advantage and Knowledge Based Entrepreneurship**

A project financed by the European Commission, DG Research
under the 'Citizens and governance in a knowledge-based society' (FP6) programme
Contract No : grant CIT5-CT-2006-028942.
Financial Markets Group, London School of Economics and Political Science, LSE, UK
Dipartimento di Scienze Economiche e Finanziarie Prato, Università di Torino, TORINO, Italy
Centre for Financial Studies, CFS, Germany
Haute Etudes Commerciales, HEC, France
Baltic International Center for Economic Policy Studies, BICEPS, Latvia
Amsterdam University, UVA, Netherlands
Neaman Institute for Advanced Studies in Science and Technology at Technion, TECHNION, Israel
Indian School of Business, ISB, India
Tilburg University, UTIL.CER, Netherlands
University of Southern Switzerland, USI, Switzerland
The Institute of Physics, IBP, Serbia

# A Breath of Fresh Air?
# Firm types, scale, scope and selection effects in drug development

Ashish Arora
Heinz School, Carnegie Mellon University, Pittsburgh, PA
ashish@andrew.cmu.edu

Alfonso Gambardella
Bocconi University, Milan, Italy
agambardella@unibocconi.it

Laura Magazzini
University of Verona, Italy
laura.magazzini@univr.it

Fabio Pammolli
University of Florence, and IMT – Lucca Institute for Advanced Studies, Italy
pammolli@gmail.com

This paper has benefitted from comments and suggestions from many people, including Giovanni Dosi, Sul Kassiech, Gregory Kruse, Atul Nerkar, along with seminar participants at the Sant'Anna School of Advanced Studies, Pisa, Italy, the Strategic Management Society 2006, Vienna, the Institute for Markets and Technology (IMT), Lucca, and the University of Verona. We also acknowledge financial support from the European Commission (grant CIT5-CT-2006-028942).

**Abstract**

This paper measures differences in the innovation performance of different types of firms in the pharmaceutical industry. We compare the innovation performance of incumbent firms with entrants, controlling for differences in the scale and scope of research, both at the firm level and at the project level. To do so, we develop a simple analytical framework of drug development, which we use to estimate a structural model, using data on 3,000 drug R&D projects in preclinical and clinical trials in the US during the 1980s-early 1990s. Key to our approach is a careful attention to the issue of selection – firms choose which compounds to advance into clinical trials. This choice depends upon the likelihood of success, but also upon economies of scale and scope, and strategic considerations about product cannibalization. It also depends upon how the costs of development and the rewards of success are shared within organizations and between alliance partners. After controlling for selection, we find that: a) incumbent pharmaceutical firms draw their compounds from better statistical distributions; b) over time, learning or environmental selection make entrants firms more similar to the established firms both in terms of selection behavior and research productivity; c) compounds licensed by pharmaceutical firms are at least as likely to succeed as internal developed projects, inconsistent with the "lemons" hypothesis; d) firm scale improves innovation performance but not scale at the project level.

*Key words*: firm capabilities, drug development process, market for technology.

## Non–technical abstract

This paper presents a model aimed at evaluating the empirical validity of theories dealing with inter-firm differences in firm performance. Entry is vital to any industry, being the source of new technology and new ideas about how production or research activities are organized and financed. In research-intensive industries, new entrants typically compete on the basis of superior innovation capabilities, while incumbents have the advantage of size and scope, relying upon established production and marketing capabilities. However, what happens if the innovation activity itself is characterized by scale and scope economies? How do entrants fare in comparison to established incumbents in terms of their innovation performance? Can one disentangle the role of size and scope from the intrinsic differences due to firm types? We explore these questions in the context of the pharmaceutical industry.

Apart from understanding and measuring the factors that affect pharmaceutical R&D performance, the industry is a good setting for studying the questions above given the evolutionary trends that have shaped its structure during the past thirty years.

First, since the advent of biotechnology, large established drug companies and smaller research-intensive biotech firms coexist within the industry. By estimating whether large pharmaceutical firms and smaller biotech companies draw their research discoveries (i.e., preclinical compounds) from different statistical distributions, our research contributes to the debate of the relative innovation performance of large and small firms. Unlike previous studies, we cast the debate in a proper context by emphasizing the importance of controlling for the potential selection bias arising from the different incentives for advancing compounds in the R&D process that characterize the different firm types.

Second, the rise of the biotech firms has encouraged a division of "innovative" labor and wide networks of collaborations have emerged. The literature has addressed the question of whether licensed compounds are more likely to succeed compared to integrated projects. The division of innovative labor perspective, as well as the network literature, suggest that firms benefit from joining resources and efforts because of scale, complementarities, or other advantages from interacting. By contrast, another strand of the literature posits that "lemons" problems can emerge due to asymmetric information in technology markets. The empirical evidence on the matter is mixed, and our model offers a sharp test of these contending views.

Finally, learning processes at the level of the firm, economies of scale and scope in R&D and innovation, and other similar issues, are central to the pharmaceutical industry. Our study contributes to this literature by understanding how the learning processes of firms, their experience in specific domains or in the R&D process as whole, or other factors, affect its innovation capabilities and the incentives to move compounds to later stages of the innovation process.

The model mimics the drug development process and explicitly deals with the issue of selection of compounds into human clinical trials, not only to avoid biases, but also because selection illuminates a number of important organizational and market processes that are of interest in and of itself. Preclinical studies aim at evaluating the toxic and pharmacological effects of the drug under study, providing the firm with a signal of the likelihood that the compound will pass FDA evaluation and will eventually be launched on the market. This characterizes the *technical dimension of the selection decision*. The firm

also compares the estimated net addition to profits the new product would make and the net addition to development costs, taking into account economies of scale and scope, strategic considerations about product cannibalization, and how the costs of development and the rewards of success will be shared within organizations and between alliance partners. The compound is then chosen for trials if the firm expects a net profit from its development, i.e. the selection decision also entails a *commercial or strategic dimension*. Though simple, our model is essential for disentangling the differences in outcome due to selection and those due to underlying differences in innovation performance.

Within this framework, a set of hypothesis is developed about selective and innovative capabilities as a function of scope and scale economies, both at the firm and project level, firm types and development arrangements (in-house versus jointly developed projects). The model is estimated by using a rich dataset about 3,000 drug R&D projects originated by pharmaceutical and biotechnology firms in preclinical and clinical trials in the US during the 1980s-early 1990s. As well, we develop extensive and fine grained controls for the therapeutic area and the compound being tested.

The results show significant differences in strategic behavior and performance of pharmaceutical and biotechnology firms. We believe that the firm types we consider reflect different models of organizational structures, learning processes, and strategic behavior. Incumbent pharmaceutical firms draw their compounds from better statistical distributions than entrants, even after controlling for differences in selection processes. Over time, learning or environmental selection make entrants firms more similar to the established firms both in terms of selection behavior and research productivity: older biotechs are closer to established incumbents than are the more recent entrants in both their selection threshold and overall performance. Finally, scale at the program level reduces the average probability of success whereas overall research scale at the firm level increases this probability. Even if significantly different from zero, the effect associated with scale at the firm level is negligible in magnitude, downplaying the importance of firm scale beyond the firm types we consider. Economies of scope appear to be unimportant.

As far as the debate about the effects associated with the division of innovative labor, our results provide no support to the “lemons” hypothesis: compounds licensed by pharmaceutical firms are at least as likely to succeed as internally developed projects.

The discovery of recombinant DNA was thought to have heralded the era of biotechnology in the pharmaceutical industry. Many thought that this would lead to a spate of entry which would unleash widespread changes in the structure of the pharmaceutical industry. Though much has changed in the pharmaceutical industry, entrants have been a breath of fresh air rather than the gale of creative destruction that some foresaw.

## 1. Introduction

Entrants to an industry are often the source of new technology, and new ideas about how production or research activities are organized, how they finance them, the market segments they focus. Though new, these ideas may not always be good; however, entrants can seldom hope to succeed by mimicking the established entrants who have the advantage of size and scope. When the industry is research-intensive, entry is typically on the basis of superior innovation capabilities, while incumbents are thought to compete on the basis of established production and marketing capabilities. However, what happens if the innovation activity itself is characterized by scale and scope economies? How do entrants fare in comparison to established incumbents in terms of their innovation performance? Can one disentangle the role of size and scope from the intrinsic differences due to firm types? We explore this question in the context of the pharmaceutical industry.

To do so, we have to deal with the issue of selection, not simply to avoid bias (e.g., Heckman 1979; Shaver 1998) but because selection illuminates a number of important organizational and market processes, and it is therefore interesting in and of itself. When innovation is measured, as is often done, by the number of innovations brought to market, selection processes are at work. For instance, large firms may shelve some inventions because they could cannibalize existing products (or more promising products under development), or simply because large firms do not find it profitable to operate in niche markets (e.g., Christensen 1997). A different type of selection process may reflect the nature of information feedbacks and channels for information flow in firms. For instance, managers in large firms may be less able than managers of smaller firms, with richer information flow between the research and management functions, to accurately assess the technical and commercial potential of a candidate compound (Arrow 1983). Researchers in the large firms act as champions for their own compounds (Nerkar 2003) or resist the introduction of new techniques (Thomke and Kuemmerle 2002) and this may lead to less rigorous scrutiny than the external markets. On the other hand, managers of startup firms, needing to satisfy impatient investors, may not have the luxury of time, and may have too much of a vested interest in the existing candidate compound (Guedj and Scharfstein 2005). To this we add the

conjecture that the larger firms will likely have smaller expected additions to net profits from a new compound because it may cannibalize sales of other products of the firm.

As well known, the drug R&D process follows well defined steps.[1] After preclinical research, in which new active compounds are discovered, potential new drugs enter multi-stage clinical trials, where their safety and efficacy is evaluated. At each stage compounds are either rejected or moved to the next step. After Clinical III the compounds are assessed by the FDA, and if successful, approved for sale. These stages provide natural milestones. Specifically, we model the decision to select a compound to move it from pre-clinical to clinical trials. This selection decision has a technical dimension, modeled here as the firm's estimate of the probability that the compound will successfully pass FDA evaluation. However, it also has a commercial or strategic dimension, namely the firm's estimate of the *net* addition to profits it would make relative to the *net* addition to development costs. We develop and estimate a structural model of drug development, using data on more than 3,000 drug R&D projects between 1980-1994. Our objective is to use this model to evaluate the empirical validity of various theories dealing with inter-firm differences in firm performance.

The pharmaceutical industry is an apt setting for our study. There is a great deal of interest in understanding and measuring the performance of pharmaceutical R&D. Over time, R&D costs have increased steadily and concerns have been expressed about how little return firms are getting on their R&D investments. Our results shed light on key firm characteristics that condition successful clinical development. However, there are theoretical reasons for choosing pharmaceutical industry as well.

First, since the 1980s, this industry has been populated by both large established drug companies and smaller research-intensive biotech firms (Gambardella 1995). The relative innovation performance of firms of different size has been the subject of a long debate (e.g., Arrow 1983; Holmstrom 1989; Acs and Audretsch 1990; Henderson and Clark 1990; Chesbrough and Teece 1996; Christensen 1997; Giarratana 2004; see also Levinthal and March 1993). Our analysis will shed light on this issue. We estimate whether

[1] See e.g. http://www.fda.gov/cder/handbook/. See also FDA (1999) or DiMasi et al. (2003).

large pharmaceutical firms and smaller biotech companies draw their research discoveries (i.e., pre-clinical compounds) from different statistical distributions. One contribution of our research is to cast the debate of the relative innovation performance of large and small firms in a proper context by emphasizing the importance of controlling for the potential selection bias.

Second, the rise of the biotech firms has encouraged a division of "innovative labor" (Arora et al. 2001), with widespread inter-firm collaborations (Powell et al. 1996; Pammolli and Riccaboni 2004). Typically, the biotech firms supply technology and compounds to large firms, which can better conduct costly clinical trials and commercialization activities. If there are "gains from trade", licensed compounds ought to be better than those developed in-house by the licensor. Offsetting this may be the "lemons" problem (Akerlof 1970), which Pisano (1997) has in particular highlighted as a danger of participating in the market for technology. Simply put, the prediction is that licensed compounds will fare worse than that compounds developed in-house by the licensor. The current evidence is mixed. Danzon et al. (2005) find a positive effect of licensing on the probability of success of drug compounds, while Pisano (1997) finds that licensed drugs are less successful than those developed in-house. Our model offers a sharp test of these two contending views.

Finally, learning, and economies of scale and scope in R&D, are central in the pharmaceutical industry (Henderson and Cockburn 1996, Nerkar and Roberts 2004; Macher and Boerner 2006). Our study contributes to this literature by understanding how the scale of research, in specific domains or in the R&D process as whole, affects research performance. In addition, we study how these factors affect the decisions of firms to select compounds into the later stages of the innovation process.

The next section reviews the main findings of the empirical literature that studied the drug development process in the pharmaceutical industry. Section 3 describes the model and Section 4 develops our hypotheses. Section 5 describes our data and analysis. Sections 6 presents our estimation results, section 7 explores the implications and limitations of the results, and section 8 summarizes our findings and concludes.

## 2. Innovation in Pharmaceuticals: Background Literature

An early study of innovation in the drug industry is the one by Henderson and Cockburn (1996), who use data at the level of the individual research program from the internal records of ten major pharmaceutical firms. They find returns to scale at the level of both individual research programs and the research expenditures of the firm as a whole, as well as evidence of economies of scope. Henderson and Cockburn (1996) focus only on the pharmaceutical firms, and do not include biotechnology firms. Moreover, their paper is best thought of as illuminating invention productivity rather than innovation performance, since their measure of output is "important patents", whereas our measure is the successful FDA approval.

Nerkar and Roberts (2004) examine the effects of nearby and distant technological or product market experience of pharmaceutical firms on the success of new products. They find that proximal technological experience (patents in the same therapeutic class) has a positive and significant effect on the first-year sales of a new product. Distal technological experience is positive and significant only when accompanied by a high level of distal product-market experience. Finally, they find that the interaction between distal and proximal technological experience is negative, which suggests that focused and diversified innovations may be alternative strategies. Macher and Boerner (2006) look at a similar set of questions, but with two differences. First, their measure of performance is the time to complete a drug development project. Second, they use data for drug development projects by contract research organizations (CRO), which are then passed on to other companies, rather than integrated R&D projects in pharmaceutical firms. They find that experience, as well as scale and scope economies matter. They also find that experience in one therapeutic category increases the ability of the firm to gain from knowledge spillovers in other therapeutic categories in which it may be involved. Simply put, experience in one area raises economies of scope or spillovers from another area, which is suggestive of the importance of absorptive capacity in drug R&D (Cohen and Levinthal 1990).

Adams and Brantner (2003) and Abrantes-Metz et al. (2006) analyze data on drugs in human clinical trials around the world between 1989 and 2002 and find that success rates and durations can vary substantially across observable characteristics of the drugs, including primary indication, originating

company, route of administration and chemistry. Danzon et al. (2005) focus on the role that experience (both overall experience and experience in a particular therapeutic category) and alliances have on the outcome of R&D projects. Their sample is composed of R&D projects by over 900 firms during 1988-2000. They too find that the probability that a compound succeeds in a particular stage of clinical trials varies across therapeutic categories. Danzon et al. (2005) also find evidence of large positive returns to a firm's overall experience for the larger and more complex late-stage trials. They find that products developed in alliances have a higher probability of success, at least for the more complex phase 2 and phase 3 trials, and particularly if the licensee is a large firm. They recognize, but do not explicitly model, differences across firms in the quality of their drug candidates.

Nerkar (2003) studies the relationship between experience and the probability that a drug receives FDA approval. He hypothesizes that experience may not always lead to better performance, if the experience is of the wrong sort (e.g., oriented towards drug discovery rather than drug development), and if feedback is inadequate or delayed (the drug development cycle can be fifteen years or more). Given the uncertainties and long delays, researchers may be rewarded for discovery (measured by patenting) rather than commercialization (measured by FDA approval). As a result, scientists may continue to work, building experience in areas that provide them with clumps of patents but do not lead to commercially useful drugs. In this paper, our unit of analysis is the firm, rather than the team. We measure how many compounds the firm develops, during our sample period, in a therapeutic area and overall. Though this is a measure of scale of operations, it is also plausibly a measure of experience.

None of these studies addresses explicitly our concern that firms may systematically vary in how selective they are in advancing compounds into clinical trials. As is well known, clinical trials are much more expensive than pre-clinical trials, particularly late stage clinicals. Which compounds make it into clinical trials depends upon whether managers think the compounds are promising. However, a variety of other factors play an important role. A drug discovered or developed in the pre-clinical stages by a research team that has high status in the organization may be moved into clinicals even if it is below threshold (Nerkar 2003). In contrast, Guedj and Scharfstein (2005) point to the agency problem between

managers and investors. In biotech startups, managers typically have one compound to bet upon, and therefore push it as much as they can into the clinical trials. By contrast, in larger firms managers have many compounds from which to choose, and are less likely to have a vested interest in any particular compound. As a result, biotech companies are more likely to advance products from phase I into phase II clinical trials, but these compounds are more likely to fail in later stages. In our paper, we distinguish between the pre-clinical and the clinical, but do not separately analyze the progress across the various clinical stages.

Finally, the literature has addressed the question of whether licensed compounds are more likely to succeed compared to integrated projects. The division of innovative labor perspective (Arora et al. 2001) suggests that established firms and biotech companies have a comparative advantage in different stages of the innovation process. Similarly, the network literature (Powell et al. 1996; Pammolli and Riccaboni 2004) posits that firms benefit from joining resources and efforts because of scale, complementarity, or other advantages from interacting. By contrast, Pisano (1997) argues that technology markets are affected by asymmetric information, leading to a "lemons" problem.

Pisano (1997) finds that in-house development is superior to licensed compounds. On the other hand, Danzon et al. (2005) find that compounds developed in alliances (roughly equivalent to licensed compounds) have a lower probability of failure in clinical trials. However, Danzon et al. (2005) include in their sample alliances that were formed prior to the conclusion of each phase (up to phase III). This is likely to include marketing agreements that pharmaceutical firms often strike for successful compounds to enhance market access. To avoid bias, we only include licenses signed before phase I, and test both whether licensed compounds are more likely to be selected into clinical trials and whether they are drawn from different distributions.

## 3. An Econometric Model of Selection and Success

To understand the interplay between selection and success we provide a simple model. After the compound has been discovered, the firm starts preclinical trials to evaluate its properties. Preclinical

research will provide the firm with a signal of the probability that the compound will pass the clinical trials and will be eventually launched on the market. For simplicity, we assume that the firm forms an estimate, Pg, of the probability that compound is safe and effective. In other words, Pg is the firm's point estimate of the (uncertain) probability of success. We assume that these estimates are unbiased.

Let V indicate the expected net additional revenue from the sales of the compound if it id eventually marketed, and D the net additional clinical trial costs, which we label development costs. The firm will take the compound into clinicals if it expects a net profit from its development i.e. if $PgV - D > 0$, or $Pg > D/V$. We denote D/V as Pg*, a threshold that is a function of V and D. The compound with expected probability greater than Pg* will be taken into clinical trials, otherwise the project is discontinued during the preclinical stage.

The selection threshold, Pg*, represents the economic and strategic dimension of selection. Firms with full pipelines of drugs under development will assess a higher net addition to development costs (D) and have a high Pg*. Similarly, firms with portfolios of similar products under development will likely assess a lower net addition to profits (V), and also have a higher Pg*. Since these measures are likely correlated with measures of firm scale as well, modeling this behavior is essential for estimating innovation performance.

Early in drug development, commercial considerations such as the size of the market, the extent of competition, and the firm's relevant commercialization capabilities, act as a threshold on which compounds to select. As the compound progresses, more and more information about its safety and efficacy becomes available, so that decisions become more deeply impregnated by technical characteristics. Thus, though selection has both technical and commercial aspects, we assume that once a compound enters clinical trials, its future progress in clinical trials is determined only by its technical characteristics. What the assumption allows us to do is to use the information on whether the compound is successful in passing clinical III as a measure of its underlying technical quality (conditional on selection) and use that information in the selection stage to distinguish between technical and commercial criteria at the selection stage.

We denote $Y_1$ to be the variable that describes the selection process, i.e., $Y_1 = 1$ if the compound is taken into clinical trials, 0 otherwise. Therefore the probability of selection is $P(Y_1 = 1) = P(Pg > Pg^*)$. We use $Y_2$ to denote the success of the project, if taken into clinical trials. This is a random event whose probability depends on Pg and Pg*, viz. $P(Y_2 = 1) = E(Pg \mid Pg > Pg^*)$.

In order to estimate the model, we assume that Pg is a random variable such that ln(Pg/(1-Pg)) is normally distributed with mean $\mu$ and unit variance $\sigma^2$. Moreover, we assume that $\mu$ depends linearly on a set of independent variables Z, i.e. $\mu = Z\gamma$. The threshold Pg* is assumed to be a function of a set of independent variables X. Since Pg* is bounded to lie in the unit interval, we assume Pg* = exp(Xb)/(1+exp(Xb)). The parameters of interest are estimated via maximum likelihood. The log-likelihood function we maximize is

$$L(\beta,\gamma) = \sum_{Y_1=0} \Pr(Y_1 = 0) + \sum_{Y_1=1,Y_2=0} \Pr(Y_1 = 1)\Pr(Y_2 = 0) + + \sum_{Y_1=1,Y_2=1} \Pr(Y_1 = 1)\Pr(Y_2 = 1) + \sum_{Y_1=1,Y_2=.} \Pr(Y_1 = 1) \qquad (1)$$

where the four terms correspond to: ($Y_1=0$) not selected; ($Y_1=1,Y_2=0$) selected but failed; ($Y_1=1,Y_2=1$) selected and succeeded; ($Y_1=1,Y_2=.$) selected and still in trial. The probability of selection is:

$$\Pr(Y_1 = 1) = \Pr(Pg > Pg^*) = \Pr\left(\ln\frac{Pg}{1-Pg} > X\beta\right) = 1 - \Phi\left(\frac{X\beta - Z\gamma}{\sigma}\right), \qquad (2)$$

whereas the probability of success is :

$$\begin{aligned} \Pr(Y_2 = 1) &= E[Pg \mid Pg > Pg^*] = \int_{Pg^*}^{1} Pg\ f(Pg \mid Pg > Pg^*)\,dPg = \\ &= \int_{Pg^*}^{1} Pg \frac{f(Pg)}{\Pr(Pg > Pg^*)}\,dPg = \frac{1}{\Pr(Pg > Pg^*)} \int_{Pg^*}^{1} Pg\ f(Pg)\,dPg \end{aligned} \qquad (3)$$

## 4. Hypothesis development

We follow the existing literature and analyze the innovation capabilities of firms. Following Henderson and Cockburn (1996), we first consider three types of firm-level economies in research: i) economies associated with the scale of the individual research program; ii) economies associated with the research scale of the firm as a whole; iii) economies of scope across research programs of the firm (see also Macher and Boerner 2006; Nerkar and Roberts 2004). A large scale of research in development activity

in a given research area will increase the likelihood that the firm will successfully develop a drug in that area. There are a variety of sources of scale and scope economies. A firm with a larger or more diverse research scale may have a larger and more valuable chemical library (Thomke and Kuemmerle 2002), be more knowledgeable about the underlying biochemical processes, have better models for interpreting results from animal and small sample human studies, and be better able to modify the original compound to improve its efficacy and reduce its side effects. These benefits may be offset by organizations that are maladapted, with perverse incentives and inadequate feedbacks. There is, however, a major offsetting strategic effect as well. Firms with a number of compounds in trial for the same therapeutic area will likely not commercialize all of them, even if all are effective and safe. Instead, it will pick the most promising among them. Though our formal model assumes that once selected, compounds progress through clinical trials, this is a simplification. In reality, firms may choose to bring a number of compounds into trial for the same therapeutic area but pick only the most promising of the lot, terminating the rest even if they do not fail clinical trials. We label this effect the "portfolio effect".

Following our model, in all our hypotheses below we use the expression "higher innovation performance" to mean higher $\mu$, viz. a better distribution of the probability of success of compounds discovered in preclinical research.

*Hypothesis 1. The innovation performance of a firm in a given research area increase with a) the overall research scale of the firm; b) the span of the research across research areas (scope); c) Research scale at the level of the program. Program level economies of scale may be offset by the portfolio effect.*

*Hypothesis 2: Learning processes over time imply that the innovation performance of established firms will be higher than those of more recent entrants. Among the entrants, older, surviving entrants may perform better than newer ones.*

We now turn to our hypotheses about selection. In what follows we define a firm to be more selective if Pg* is higher. Recall that Pg* = D/V. That is, a key insight from the simple model developed in Section 3 is that the selection threshold depends upon the costs of development and the net revenue

obtained from a successful drug. It does not depend upon the probability that the compound will succeed in clinical trials.

Firms with a high scale in a given therapeutic area have an incentive to feed this specific capacity. Its costs of developing another potential drug will be lower because of superior ability to modify the lead compound to maximize effectiveness and minimize side effects, and because the firm may have closer links with hospitals that can carry out clinical trials in that area (lower D). By contrast, a firm with a larger number of compounds in clinical trials, i.e., a larger research scale at the firm level, will be more selective. Bringing an additional compound will imply higher incremental cost of development, and the firm will enter new compounds in clinical trials only if they are promising. Finally, scope, i.e. a wider diversification of projects across research areas, may increase value or reduce development costs because of spillovers across research programs, as Henderson and Cockburn (1996) suggested. This will make the firm less selective. We formally state the following hypothesis.

*Hypothesis 3. Research scale at the level of the individual research program and scope makes the firm less selective about the compounds to be moved from preclinical research to clinical trials. Research scale at the level of the firm makes it more selective.*

In large established firms new compounds are more likely to cannibalize existing products. There is a substantial economics literature that deals with the cannibalization effect (also called the replacement effect). Arrow (1962) argued that fear of cannibalizing the profits from existing markets and products would lead an incumbent firm to under-invest in research. The danger of cannibalizing existing products is also likely to be greater, the larger the sales and profits of existing products. Entrants, with no existing markets to protect, have no fear of cannibalization.

Organizational capabilities may reinforce the cannibalization effect. Interviews with industry participants reveal that large firms have well established routines and financial models for deciding how compounds are moved along through the various stages of pre-clinical research and clinical research. By contrast, many smaller entrants lack such discipline, and indeed, the founders may have strong biases that

favor compounds being selected. The need to provide "good news" to investors may also bias startups to push drugs into clinical trials (Guedj and Scharfstein 2005).

Finally, there may be differences in risk aversion. Ordinarily, one might imagine that large established firms may have a lower degree of risk aversion due to their superior financial reserves. However, as Stiglitz and Weiss (1981) showed in a seminal paper, limited liability laws may themselves make startups more risk loving, particularly in taking large risks that might result in bankruptcy. The intuition is simple. A large risk ("swinging for the fences") has high reward but also a high cost in terms of bankruptcy. However, with limited liability, the downside risk is limited whereas the upside can keep increasing. As a result, small startups may be inclined to swing for the fences, which implies that they would have a lower threshold. In sum, risk aversion further reinforces the cannibalization effect.

*Hypothesis 4. Established incumbents are more selective than recent entrants.*

Our final hypotheses focus on the licensed compounds. As noted, the literature provides mixed results, in part because selection issues have not been adequately addressed. We separately test whether licensed compounds are drawn from a better distribution (high $\mu$) and whether licensed compounds have to overcome a higher selection threshold (high Pg*). Following the earlier discussion about differences between established firms and recent entrants, we estimate the effect of licensing on selection and clinical success separately for the two types of firms. We expect that compounds licensed in by incumbents will face a higher threshold in order to proceed to clinical trials. The reason is straightforward: the incumbent firm will have to share the net revenues from the compound, if successful, with the licensor. However, it is unlikely that the licensor will share the development costs in the same proportion. If so, in terms of our model, this is as if V were lower, so that the selection threshold, Pg*(= D/V), is higher for licensed compounds than for internally developed compounds. However, when an entrant is the licensee, additional considerations may apply. The licensor may impose more stringent contractual conditions, making it more difficult for the licensee to terminate clinical development (Guedj and Scharfstein 2005). For instance, the licensee may be unable to terminate clinical trials on commercial grounds (unlike

established firms with much greater bargaining power), or it may be desperate to keep its development pipeline from becoming empty (Higgins and Rodriguez 2006). In sum, entrants will be less selective than incumbents regarding licensed compounds.

*Hypothesis 5. Licensed compounds have a higher selection threshold (higher Pg*) than internally developed compounds when the licensee is an established firm. When the licensee is an entrant, the selection threshold for licensed compounds is lower (lower Pg*) than for incumbent firms.*

As far as the performance of licensed compounds is concerned, the literature offers diametrically opposite predictions. Pisano (1997; 2006), following the economics literature spawned by Akerlof (1970), argues that markets for knowledge are imperfect, and characterized by information asymmetries, especially about the quality of the technology offered for license. The fear is that a licensor is likely to keep its most promising technologies for its own use and only offer inferior ones for license. If so, licensed compounds will be drawn from an inferior distributions (have lower $\mu$). Arora and Gambardella (1994), building on the insights of Cohen and Levinthal (1989, 1990) and Rosenberg (1990), argue that research capabilities also provide firms with the ability to evaluate external technologies. Thus, for sophisticated buyers, the "lemons problem" may not be serious, although inexperienced buyers, such as recent entrants, run the risk of licensing in a "lemon". This leads to the following hypothesis:

*Hypothesis 6. Entrants are likely to face a lemons problem in the licensing market but not incumbent firms. Licensed compounds will have the same distribution than in-house compounds when the licensee is an incumbent but a worse distribution when the licensee is an entrant.*

We have articulated these hypotheses about the underlying structural variables, namely the selection threshold, Pg*, and the average probability that the compound will be successful, $\mu$. What do these imply in terms of the observed probabilities of selection and success? The most obvious one is for hypotheses 5 and 6. Hypotheses 5 and 6 implies that incumbents will have a higher Pg* and the same $\mu$, respectively. This implies that the probability of selection is lower for compounds licensed by incumbents, but the probability of success is higher. Conversely, since $\mu$ is lower for compounds licensed by entrants, they will have a lower probability of success. The probability of selection is ambiguous: such

compounds have a lower Pg* but also a lower μ. Similarly, hypotheses 2 and 4 imply that incumbents have a higher μ and a higher Pg*, respectively, which implies a higher probability of success, but ambiguous implications for the probability of selection, compared to entrants.

## 5. Description of the data

Our sample is drawn from the ATA database (ATAdb), which combines proprietary datasets on the pharmaceutical industry. It reports information about more than 17,000 R&D projects all over the world carried over since the 1980s.[2] Following Danzon et al. (2005), we considered each indication for which the compound is being developed as a separate project. To reduce heterogeneity, we restricted our sample to projects where the first trials were conducted in the US, and excluded projects originated by hospitals, public sector labs or universities.[3] We reduced the censoring problem posed by ongoing projects by including only the projects that entered preclinical trials between 1980 and 1994.[4] Our final sample is composed of 3,311 projects.

Our sample includes both incumbents (pharmaceutical firms) and entrants (biotech firms). Biotech firms are far younger than pharmaceutical firms, as they were founded after the first biotech company Genentech in 1976. By contrast, the established pharmaceutical firms can trace their origins to the early part of the 20$^{th}$ century if not earlier. This suggests that there is a natural break between established pharmaceutical firms and biotech companies in terms of experience and long-term learning processes. However, biotech firms have been founded continuously since the late 1970s, and the early entrants who survived are, in our sample, about five times larger than newer entrants, and may have different organizational structures and financing models from newer entrants to the industry. Similarly,

[2] The names of the firms in the database are reported as they are in 2002. Therefore we cannot consider separately the projects of the major pharmaceutical firms that experienced merges and acquisitions during the 1990s.

[3] Universities and private firms differ in a substantial way in the incentive for innovation and in the way their innovative activity is organized (Dasgupta and David 1994). We also excluded the projects for which it is not possible to uniquely identify the originating company; i.e. the projects that are co-developed by two or more firms.

[4] For these projects we either observe the outcome, or impute the outcome --following Danzon et al. (2005) we classified as a failure a project which remains in a phase, without any further reported events, for longer than the maximum number of years observed for completion of each phase in the non-censored sample. The maximum number of years is computed within each indication. When this was not practicable – due to missing observation for the indication at hand – we considered the maximum over all the indications.

among pharmaceutical firms, a clear distinction emerges between the top pharmaceutical corporations worldwide, operating in a wide array of markets and with great experience in the development of new R&D based drugs for major therapeutic areas, and smaller and younger pharmaceutical firms developing compounds that target predominantly, cancer, infection and inflammation, asthma and diabetes, obesity and sexual dysfunctions.

Thus, we further subdivide incumbents into R&D intensive pharmaceutical firms (henceforth "Established Pharma") and those producing for niche markets (henceforth "Other Pharma"), which are typically much less R&D intensive. We also distinguish between entrants that are larger, and typically entered earlier and have survived for some length of time (henceforth "Pioneer Biotech") and other entrants (henceforth "Other Biotech"). Specifically, we distinguish among four types of companies:

i) Pioneer Biotech: firms, founded before 1990, that apply biotechnological methods to the discovery and development of new drugs.
ii) Other Biotech: All other firms using biotechnology for the discovery of new drugs founded;
iii) Established Pharma: companies among the top 50 pharmaceutical companies worldwide in terms of sales in 2002.
iv) Other Pharma: all other pharmaceutical firms.

**Table 1**: Firm types and R&D projects

| Firm Type | Number of Projects (%) | Number of different firms |
|---|---|---|
| Pioneer Biotech | 1,108 (33.51%) | 134 (40.73%) |
| Other Biotech | 476 (14.40%) | 90 (27.36%) |
| Established Pharma | 1,561 (47.22%) | 39 (11.85%) |
| Other Pharma | 161 (4.87%) | 66 (20.06%) |
| Total | 3,306 (100.00%) | 329 (100.00%) |

Information on the number and projects related to the four firm types we considered in the analysis are reported in Table 1. Table 2 reports the share of projects, over the total number of projects started by each firm type, classified according to the characteristics of the targeted disease or of the compound. The last row reports the p-value of the chi-squared test of independence in a contingency table obtained by tabulating the firm type variable and the variable in the column. Thus, for instance, biotech

firms are more likely to target rare diseases than average. Moreover, biotech firms seem to be much more involved in "first in class" compounds.[5]

*Measures of Scale and Scope:* We measure the scale at the program level by the number of projects already started by the firm for the given indication. Scale at the level of the firm is measured by the total number of projects for the firm. Scope is measured by Herfindahl index of diversification across indications, for the firm. Note that we are interested in knowledge based economies of scale and scope, for which these measures are reasonable. Danzon et al. (2005) and Abrantes-Metz (2003) use the number of drugs in development to measure overall firm scale. To measure program level economies, Macher and Berner (2006) use the number of projects completed and Danzon et al. (2005) use the total number of projects in the therapeutic area. Finally, Danzon et al. (2005) use the same herfindal index based measure as us, and Macher and Boerner (2006) use the number of therapeutic areas in which the firm is active to measure scope economies. Henderson and Cockburn (1996) use the concentration of R&D expenditures, something we cannot do because we lack R&D data at the program level. We also control for whether the firms is publicly traded (to control for differences in the cost of capital), and our firm type dummies also control for differences in overall size, to some extent.

**Table 2**: Project characteristics, % total number of projects by firm type

| Firm Type | Lethal | Organ damage | Multiple causes | Chronic | Rare | First in class | Unclassified. |
|---|---|---|---|---|---|---|---|
| Pioneer Biotech | 75.25 | 82.90 | 78.49 | 80.11 | 4.95 | 42.39 | 38.70 |
| Other Biotech | 73.43 | 72.79 | 81.43 | 79.71 | 3.20 | 22.09 | 35.72 |
| Established Ph. | 81.51 | 83.82 | 77.94 | 76.89 | 5.67 | 32.77 | 55.04 |
| Other Pharma | 76.54 | 74.69 | 85.80 | 79.63 | 2.47 | 17.90 | 33.33 |
| Total | 75.35 | 77.86 | 80.16 | 79.43 | 4.11 | 30.23 | 39.39 |
| p-value (chi-2) | 0.005 | 0.000 | 0.041 | 0.519 | 0.025 | 0.000 | 0.000 |

*Licensed Compound*: For licensed compounds, the database records the licensor and each licensee.[6] We considered a project as "licensed" only if the agreement was signed in preclinicals.

[5] This result must be interpreted with caution due to the different incidence of unclassified projects, (cf. last column of Table 2).
[6] When multiple licensees were present, we assigned the development of the project to the largest US firm. We then inspected this assignment to make sure that we consider the firm that was actually in charge of the R&D development of the project.

*Controls:* We developed a number of controls at the level of the indication.[7] We developed a set of binary indicators for whether the disease is lethal or not (*lethal*), whether it can result in organ damage or complications (*organ damage*), whether the disease is chronic or acute (*chronic),* and whether the disease has a single cause or multiple causes (*multiple causes).*[8] These factors condition both the cost of development and also the benefits if the drug is successful. They also condition $\mu$, the average probability of success. In addition, we include indicators for whether the disease is widespread or not (*rare)*.[9,10]

Finally, since it is possible that entrants and incumbents may choose projects with different levels of innovativeness or risk, we devised a proxy for the level of innovativeness (and therefore of risk) associated to the project. We used the information whether the compound was first-in-class or a follower molecule (*first in class*).[11] Other things being equal, we expect first-in-class compounds to exhibit a lower selection threshold. This is because there are fewer alternatives available, and hence they have to overcome a lower hurdle to improve upon existing remedies. We also expect them to be drawn from a worst distribution because there is lower learning about the indications for which they can be employed successfully. Unfortunately, in almost 40% of the cases it was not possible to classify the molecule under study because of the lack of information about its chemical characteristics. We create an indicator variable that codes for when this information is missing. Through the use of these fine grained controls at the level of indication, we are able to avoid having to use fixed effects for therapeutic areas. This is particularly important since we estimate a system of two highly non-linear equations through maximum likelihood, where having a lot of fixed effects is liable to create estimation problems and even bias.

[7] The main source for the disease information has been Braunwald et al. (2001). Other information has been found in e-medicine reviews from the disease database (http://www.diseasedatabase.com). For diffusion data, information has been drawn from the "rare disease database" referred to by the FDA, and available at the internet address: http://rarediseases.about.com/cs/orphandrugs/a/122103.htm.

[8] More precisely, we measure whether the disease has a multifactorial etiology versus an unknown or single factor etiology.

[9] An orphan or rare disease is generally considered to have a prevalence of fewer than 200,000 affected individuals in the United States. The "Orphan Drug Act" (1983) allows drug companies to take tax deductions for about three-quarters of the cost of the clinical studies (FDA 1999), with an implied reduction of cost for development.

[10] In previous specifications (not reported here) we also included a dummy variable indicating whether a pharmacological therapy was available for the targeted indication in 2003. However, a large share of projects (more than 95%) was targeted to indication with an existing therapy. As a result, this variable was dropped.

[11] The information about the chemical name of the projects and its Chemical Abstract Service (CAS) registry number (when available) has been employed to distinguish me-too and second-generation molecules from first-in-class compounds (see Reddy 2003).

In addition, to control for the economic characteristics of the targeted disease, we use information on the sales in the US at the therapeutic market level (at the Anatomic Therapeutic Classification at the 3rd digit level, ATC3), as a proxy for V, the net benefit from a successful drug. As well, we measure the degree of competition by the Herfindahl Index of concentration in the US market for the ATC3 class based on 1991 sales. Though we lack detailed firm controls because it is difficult to find data for the many small firms in our sample, we do control for the share of projects developed with universities, as a measure of the firm's links with research institutes, and its closeness to science (*closeness to science*). Further, we allow for inter-dependence across the observations for a firm by reported standard errors that are clustered at the firm level. Variables included in the estimated equations are summarized in Table 3 with the indication of the level at which they are measured (firm or project), whereas Table 4 presents descriptive statistics.

Identification of Pg* and μ: The system of equations we estimate is non-linear. To estimate it, we normalize the variance of the distribution of Pg to unity. As well, we impose three exclusion restrictions. We assume that the economic characteristics of the compound, i.e. the size and the level of concentration in the relevant final market (defined on the basis of the ATC class of the R&D project at the 3rd digit level) and whether the firm originating the compound is private or public, only affect the selection threshold and not distribution of the probability of success. This is how we identify selection (Pg*) as opposed to performance (μ). With more competition in the downstream market, the expected value V that the firm can hope to gain with a new compound is smaller. As a result, the threshold is higher. Similarly, larger markets are more attractive in terms of providing higher rewards. Also, for non-public firms it is more costly to find resources to develop their compounds. Since this implies higher D, they ought to exhibit a higher threshold. At the same time, these should not affect whether the firm draws compounds from a better or worst distribution. In the section exploring the robustness of these results, we show that our results do not critically depend on these identifying assumptions.

**Table 3**: Variables in the estimated equations

| Pg* | μ | Variable | Measured as |
|---|---|---|---|
| ✓ | ✓ | Incumbent versus entrant (firm) | Firm type dummies |
| ✓ | ✓ | Scale_Program (project) | Number of projects already started by the firm for that indication |
| ✓ | ✓ | Scale_Firm (firm) | Total number of projects by the firm |
| ✓ | ✓ | Scope (firm) | Herfindahl index of diversification of firm's projects across indications |
| ✓ | ✓ | Compound licensed to entrant (project) | If the compound licensed in preclinical, and the licensee is either a Pioneer Biotech or an Other Biotech |
| ✓ | ✓ | Compound licensed to incumbent (project) | If the compound licensed in preclinical, and the licensee is either an Established Pharma or an Other Pharma |
| ✓ | ✓ | Closeness to science (project) | Share of projects jointly developed with a public research organization for that indication |
| ✓ | | Public (firm) | Whether firm is public or not in 1994 |
| ✓ | | Market Size (project) | Log of 1991 of sales in the US market for the ATC3 class divided by 100 |
| ✓ | | Market Concentration (project) | Herfindahl Index of concentration in the US market for the ATC3 class (uses 1991 sales) |
| ✓ | | Miss Info ATC3 (project) | Dummy equal to 1 for projects with missing information about size and concentration in the ATC3 class |
| ✓ | ✓ | Disease Characteristics (project) | Lethal, organ damage, multi-causal, chronic, rare pathologies (see text for explanation) |
| ✓ | ✓ | First-in-class (project) | Dummy equal to 1 if the project is first-in-class and 0 otherwise |
| ✓ | ✓ | Unclassified (project) | Dummy equal to 1 if the information for first-in-class is missing |

**Table 4**: Descriptive Statistics

| Variable | Mean | S.D. | Variable | Mean | S.D. |
|---|---|---|---|---|---|
| Selection | 0.48 | 0.50 | License to biotech | 0.02 | 0.13 |
| Success | 0.34 | 0.48 | License to pharma | 0.06 | 0.24 |
| Pioneer Biotech | 0.34 | 0.47 | Closeness to Sc. | 0.05 | 0.16 |
| Established Pharma | 0.47 | 0.50 | Public Firm | 0.84 | 0.37 |
| Other Biotech | 0.14 | 0.35 | Size ATC3 | 0.10 | 0.05 |
| Other Pharma | 0.05 | 0.22 | Conc. ATC3 | 0.44 | 0.34 |
| Scale_Project | 2.67 | 3.53 | Miss Info ATC3 | 0.10 | 0.30 |
| Scale_Firm | 36.23 | 46.19 | First-in-class | 0.30 | 0.46 |
| Scope | 0.83 | 0.21 | Unclassified | 0.39 | 0.49 |

## 6. Empirical results

Table 5 presents probit estimates of the selection and success equations. Notice that compared to established pharmaceutical firms, biotech firms are less likely to take compounds into clinical trials (they have a lower probability of selection) and have a lower probability of success (although the difference is insignificant for Pioneer Biotech firms, when firm and market controls are added to the estimated equation). Firm scale increases the probability of success (though it leaves selection probability unchanged), whereas program scale decreases both selection and success probability, and scope reduces

selection but has no effect on outcome. Market concentration is associated with higher probability of selection. It is tempting to try to interpret these results in terms of the various theories about differences across firms or firm scale and scope. However, as we discussed in developing our model, the probability of success will depend upon not only the intrinsic quality of the project but also on the selection threshold. Conversely, the probability of selection depends upon the selection threshold but also on the intrinsic quality of the project. For instance, suppose that Pioneer Biotechs had the same selection threshold as Established Pharmaceutical firms but lower innovative capability. In this case, we would expect lower probability of selection and perhaps also lower probability of success conditional upon selection. However, the results are also consistent with Pioneer Biotechs having a lower selection threshold and the same level of innovation capability, or lower selection thresholds and lower innovation capability. In other words, we need to estimate the structural parameters. This will enable us to identify separately whether the behavior is driven by selection or success (or both).

**Table 5**: Probit estimates

| | Specification 1 | | Specification 2 | |
|---|---|---|---|---|
| | Pr Selection | Pr Success | Pr Selection | Pr Success |
| Pioneer Biotech | -0.32** (0.11) | -0.28* (0.15) | -0.55** (0.12) | -0.19 (0.16) |
| Other Biotech | -0.47** (0.14) | -0.93** (0.43) | -0.68** (0.17) | -0.71 (0.47) |
| Other Pharma | -0.20 (0.16) | 0.62** (0.23) | -0.41** (0.20) | 0.89** (0.30) |
| License (preclinical) to Pharma | 0.15 (0.17) | 0.42 (0.28) | 0.06 (0.17) | 0.34 (0.28) |
| License (preclinical) to Biotech | -0.35 (0.25) | -1.11** (0.53) | -0.38 (0.26) | -1.22** (0.59) |
| Scale_program. | | | -0.08** (0.02) | -0.13** (0.03) |
| Scope | | | -0.47** (0.18) | 0.31 (0.34) |
| Scale_firm. | | | .2E-3 (.1E-2) | .4E-2** (.1E-2) |
| Public | | | 0.23* (0.14) | 0.21 (0.24) |
| Closeness to Science | | | 0.54** (0.26) | 0.10 (0.26) |
| Market Concentration | | | 0.43** (0.14) | -0.12 (0.25) |
| Market Size | | | -0.04 (1.13) | 1.38 (2.40) |
| Miss Info Market | | | 0.33* (0.19) | -0.47 (0.47) |
| Lethal | 0.07 (0.09) | -0.16 (0.15) | 0.13* (0.07) | -0.16 (0.14) |
| Organ Damage | -0.17** (0.08) | 0.13 (0.1) | -0.20** (0.09) | 0.28** (0.10) |
| Multiple causes | -0.12 (0.08) | -0.22* (0.11) | -0.04 (0.08) | -0.17 (0.11) |
| Chronic | 0.03 (0.07) | -0.12 (0.15) | 0.13* (0.07) | -0.05 (0.15) |
| Rare | -0.14 (0.15) | 0.73** (0.2) | -0.14 (0.15) | 0.55** (0.16) |
| First in class | -0.35** (0.09) | -0.19 (0.15) | -0.35** (0.09) | -0.15 (0.16) |
| Unclassified (fc) | -1.47** (0.09) | -1.55** (0.24) | -1.41** (0.09) | -1.65** (0.21) |
| Constant | 0.96** (0.13) | 0.14 (0.21) | 1.07** (0.28) | -0.58 (0.55) |
| Obs. | 3,311 | 1,088 | 3,311 | 1,088 |
| Log-likelihood | -1847.25 | -610.94 | -1760.23 | -581.88 |

*Robust standard errors in parenthesis (clustered by firms). ** denotes p < 5%; * p < 10%.*

Table 6 reports the results obtained from estimating the model described in Section 3. We estimate two different specifications. In one, we only use the firm type dummies and disease controls, while in the second we also include measures of scale and scope and firm and market controls. We also estimated even more parsimonious specifications where, in addition to firm and market controls, and measures of scale and scope, we also excluded the licensing variables. The results were largely unchanged. Other specification issues are discussed in the robustness section.

**Table 6**: Structural Model, Maximum Likelihood Estimates

| | Specification 1 | | Specification 2 | |
|---|---|---|---|---|
| | Pg* | μ | Pg* | μ |
| Pioneer Biotech | -0.35 (0.28) | -0.68** (0.28) | -0.02 (0.34) | -0.58* (0.31) |
| Other Biotech | -1.55* (0.89) | -2.02** (0.88) | -1.20 (0.94) | -1.89** (0.92) |
| Other Pharma | 1.20** (0.45) | 1.00** (0.47) | 1.70** (0.29) | 1.28** (0.31) |
| License to Pharma | 0.68 (0.53) | 0.83 (0.58) | 0.69* (0.36) | 0.75* (0.42) |
| License to Biotech | -1.79* (1.04) | -2.14** (1.04) | -2.08* (1.11) | -2.47** (1.10) |
| Scale_Program | | | -0.20** (0.04) | -0.28** (0.05) |
| Scale_Firm | | | 0.01** (.1E-2) | 0.01** (.1E-2) |
| Scope | | | 1.04* (0.60) | 0.56 (0.59) |
| Closeness to science | | | -0.08 (0.36) | 0.47 (0.33) |
| Public | | | -0.22 (0.14) | |
| Market Concentration | | | -0.43** (0.14) | |
| Market Size | | | 0.18 (1.08) | |
| Miss Info Market | | | -0.34* (0.18) | |
| Lethal | -0.31 (0.23) | -0.24 (0.23) | -0.24 (0.25) | -0.11 (0.26) |
| Organ Damage | 0.32** (0.16) | 0.15 (0.16) | 0.49** (0.22) | 0.29 (0.20) |
| Multiple causes | -0.33* (0.17) | -0.45** (0.15) | -0.28* (0.17) | -0.33** (0.14) |
| Chronic | -0.22 (0.26) | -0.19 (0.27) | -0.11 (0.25) | 0.02 (0.25) |
| Rare | 1.33** (0.44) | 1.19** (0.47) | 1.15** (0.19) | 1.00** (0.25) |
| First in class | -0.16 (0.30) | -0.51* (0.28) | -0.15 (0.35) | -0.49 (0.33) |
| Unclassified | -2.32** (0.57) | -3.80** (0.60) | -2.44** (0.51) | -3.84** (0.54) |
| Constant | -0.99** (0.26) | -0.02 (0.29) | -1.96** (0.63) | -0.85 (0.59) |
| Obs. | 3,311 | | 3,311 | |
| Log-likelihood | -2458.68 | | -2348.29 | |

*Robust standard errors in parenthesis (clustered by firms). ** denotes p < 5%; * p < 10%.*

The results indicate that scale at the program level reduces the average probability of success but overall research scale at the firm level increases this probability. Economies of scope appear to be unimportant. Thus, hypothesis 1 is only partially supported. If we believe that program (indication) level research economies are in fact significant and positive, then our measure is confounding economies of scale with portfolio effects: because firms will ultimately commercialize only the most promising of

projects, a larger number of potential drugs under development results in a higher average failure rate. Similarly, the lack of statistical significance for the measure of scope may be because the scope effect may be captured by the scale variable at the level of the firm. After all, the latter captures effects across research programs, as the probability of success in a given indication is higher if the firm has more compounds in development for another indication. In sum, these results do reflect the importance of economies associated to the firm size and scope, even after controlling for the difference between biotech and established firms, and most importantly after controlling for potential selection biases in the introduction of new compounds onto clinical trials. However, our data may not be fine enough to discriminate between economies of scale or scope.

Table 6 shows that both Pioneer Biotech and Other Biotech have negative and significant impact on $\mu$, implying that entrants have lower innovative performance compared to incumbents. Within entrants, Pioneer Biotechs perform better than more recent entrants. This is consistent with hypothesis 2. The results also point to the underlying processes that increase similarity between entrants and incumbents, with the surviving entrants (Pioneer Biotechs) growing and becoming closer to the incumbents in innovation performance and selection behavior. This may stem from learning processes, and growth in size and number of products. This may also reflect selection. Pioneer Biotechs have survived over time, and hence they may be better firms.

Consistent with hypothesis 3, we find that research scale at the program level reduces the selection threshold whereas research scale at the level of the firm as a whole increases the selection threshold. As hypothesized, a potential explanation is that the firm has a larger capacity in the domain associated to a given indication, which reduces development costs associated to any new compound for that indication. By contrast, a firm with many compounds in development will be more selective about adding new compounds to clinical trials because capacity may be filled. As a result, they will only move to clinical trials new compounds with a relatively high Pg. Contrary to hypothesis 3, firm scope makes firms more selective, and the estimated effect is significant at the 10% level. This may either point to the absence of spillovers across research programs or potential overlap in our measure of scope and scale.

We find no significant difference in the selection process of the older biotechnology firms that first appeared into the industry (the Pioneer Biotech) and the Established Pharma. Other Biotech and Other Pharma, exhibit respectively a lower and a higher threshold than the Established Pharma. Thus, hypothesis 4 is supported.[12] Finally, the point estimate of our scope variable is consistent with the hypothesis.

Concerning licensed compounds, we see that compounds in-licensed by incumbents face a higher threshold than compounds developed internally and compounds in-licensed by entrants face a lower threshold, supporting hypothesis 5. We also find that incumbents have an innovation performance with licensed compounds that is no worse than their performance with internally developed compounds. However, entrants perform significantly worse with licensed compounds. These results imply that knowledgeable buyers, such as established pharmaceutical firms do not face a market for lemons when they seek external technology. They do rationally impose higher selection thresholds for licensed compounds, to offset the royalty payments they would have to make to the licensor. Alternatively, incumbent firms may find it even easier to terminate contracts with small biotech licensors than terminate early projects of politically powerful internal research groups (Guedj and Scharfstein 2005; Nerkar 2003). Entrants, perhaps because they are smaller and less experienced, may be buying lemons, and worse still, appear to be constrained in how selective they can be (Pisano 1997).[13]

Turning to the control variables, we see that more concentrated markets indeed lower the selection threshold: a successful drug could earn higher net profits (higher V). Similarly, the threshold is higher for a disease that is not widespread (because this implies a small market). Interestingly, the innovation performance in such cases is higher, for reasons that are not clear. We find that risky drugs have a lower innovation performance, though they do not face a higher selection threshold compared to "me-too" drugs. This suggests that first in class drugs have higher expected net revenues conditional on

[12] We interpret the high selectivity of Other Pharma, relative to established Pharma, as suggesting that the former have higher development costs than the latter, perhaps because innovation is not core to the Other Pharma. Other Pharma typically focus on non-R&D based drugs.

success, enough to offset the likely higher costs of development and the lower likelihood of success. Finally, the results on the public firms go in the right direction, since public firms are believed to face a lower cost of capital (implying a lower D), i.e. a lower selection threshold. However the result is not statistically significant.

## 7. Exploring the Results and their Robustness

The key benefit of estimating the structural coefficients is that they enable us to discriminate among different conjectures about the quality of the drug compounds that enter into the clinical trials. For instance, the higher probability of success and higher selectivity for incumbent projects could be due to higher selection thresholds, higher innovation performance or both, and structural estimates can sort this out. Moreover, having estimated structural parameters, we can use them for "what-if" scenarios to provide additional insights into the implications of our estimates. We begin by using the estimated coefficients from specification 2 in Table 6 to compute the probability of selection, the unconditional probability of success, the value of Pg* and $\mu$ for the firms in our sample.

**Table 7**: Estimated selection and success probabilities, Pg*, $\mu$, overall and by Originator type

| | Pr. Selection | Pr. Success | Pg* | $\mu$ |
|---|---|---|---|---|
| Overall | 0.48 | 0.25 | 0.11 | -2.92 |
| Pioneer Biotech | 0.42 | 0.30 | 0.10 | -3.15 |
| Established Pharma | 0.57 | 0.28 | 0.13 | -2.30 |
| Other Biotech | 0.31 | 0.06 | 0.03 | -5.05 |
| Other Pharma | 0.52 | 0.42 | 0.31 | -1.13 |

Figures 1 and 2 show the simulated the distributions of Pg for the whole sample and by originator types, using our estimates from specification 2 in Table 6. The distribution of Pg is skewed, with a low number of highly successful projects and a high number of projects with low probability of success. In general, the right tails of the four distributions become fatter as we move from the Other Biotech to Pioneer Biotech, Established Pharma, and Other Pharma. This is once again consistent with our discussion about

[13] We also experimented with a different definition of the licensing variable. In the selection equation we took into consideration licensed compounds up to preclinical, whereas in the success equation we included licenses up to Phase II Licensed compounds have a higher threshold and a higher probability of success.

learning processes and other characteristics across these firm types. The vast majority of the projects started by the Other Biotech have a very low probability of success.

**Figure 1**: Estimated distribution of Pg

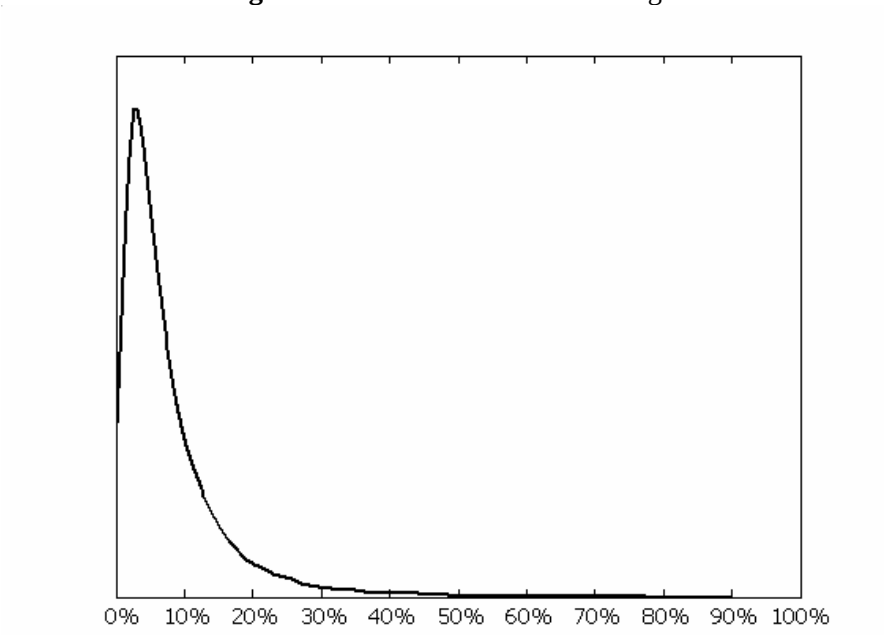


**Figure 2**: Estimated distribution of Pg, by Originator type

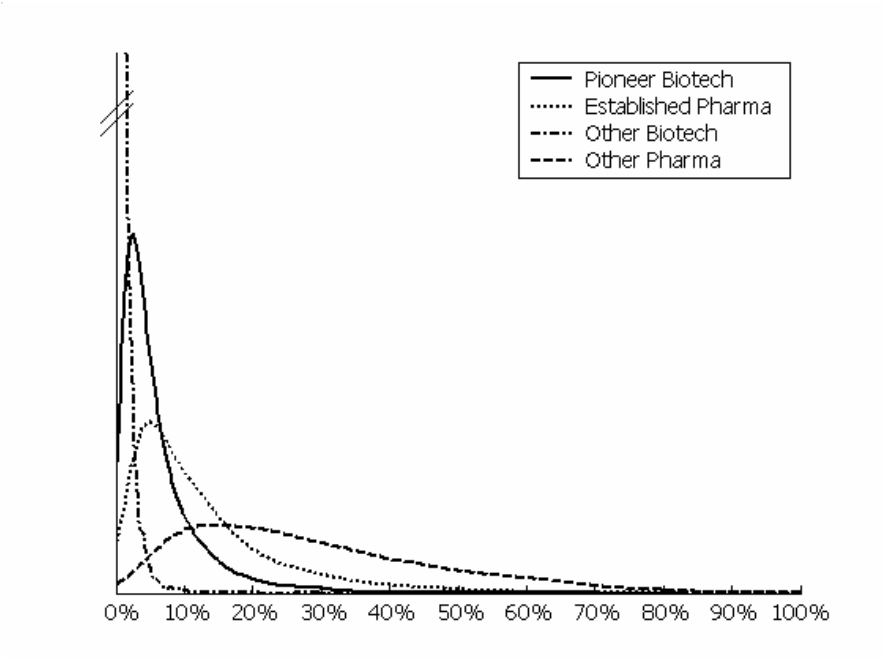

**Table 8:** Simulating an increase in the research scale of the firm (Scale_Firm)

| | Pioneer Biotech | | Established Pharma | | Other Biotech | | Other Pharma | |
|---|---|---|---|---|---|---|---|---|
| | Baseline | +1 σ | Baseline | +1 σ | Baseline | +1 σ | Baseline | +1 σ |
| Pr Selection | 0.42 | 0.42 | 0.57 | 0.57 | 0.31 | 0.31 | 0.52 | 0.52 |
| Pr Success | 0.14 | 0.15 | 0.23 | 0.29 | 0.04 | 0.04 | 0.38 | 0.38 |
| Pg* | 0.10 | 0.11 | 0.13 | 0.18 | 0.03 | 0.03 | 0.31 | 0.32 |
| μ | -3.15 | -3.05 | -2.30 | -1.83 | -5.05 | -4.99 | -1.13 | -1.08 |

Using the estimated coefficients of the pooled model in Table 6 we also performed some simulation exercise. We first studied the effects of changes in firm size. Table 8 reports the estimated effects of a one-standard-deviation increase in Scale_Firm. When this variable increases Pg* increases. Since μ also increases, the probability of selection increases too. Interestingly, the effects of the firm research scale are rather small. Thus, while Scale_Firm is statistically significant in both the selection and performance equations in Table 6, ultimately its overall effect is small relative to the impact of firm type. As discussed, we believe these firm types reflect different models of organizational structures, learning processes, and strategic behavior. Our empirical results than suggest that such differences are the crucial ones to understand their strategic behavior about innovation or their innovation performance.

**Table 9**: Simulating the licensing effects by originator type

| | Pioneer Biotech | | | Established Pharma | | |
|---|---|---|---|---|---|---|
| | No License | License to Pharma | License to Biotech | No License | License to Pharma | License to Biotech |
| Pr. Selection | 0.43 | 0.45 | 0.31 | 0.57 | 0.59 | 0.46 |
| Pr. Success | 0.13 | 0.21 | 0.02 | 0.23 | 0.32 | 0.04 |
| Pg* | 0.09 | 0.16 | 0.01 | 0.13 | 0.21 | 0.02 |
| μ | -3.17 | -2.43 | -5.64 | -2.29 | -1.54 | -4.76 |
| | Other Biotech | | | Other Pharma | | |
| | No License | License to Pharma | License to Biotech | No License | License to Pharma | License to Biotech |
| Pr. Selection | 0.31 | 0.32 | 0.21 | 0.53 | 0.55 | 0.41 |
| Pr. Success | 0.03 | 0.06 | 0.00 | 0.39 | 0.49 | 0.10 |
| Pg* | 0.02 | 0.05 | 0.00 | 0.32 | 0.45 | 0.07 |
| μ | -5.05 | -4.30 | -7.52 | -1.04 | -0.29 | -3.51 |

Table 9 reports on the impact of licensing. It shows that when the licensee is a biotech firm, the drug compounds are drawn from a worst distribution than internal projects. By contrast, drug compounds are drawn from a (weakly) better distribution when the licensee is a pharmaceutical firm. There may be

two explanations, which are not mutually exclusive. Incumbent pharmaceutical firms are experienced in the domain – they have high absorptive capacity, and considerable bargaining power. Alternatively, the pharmaceutical buyers may have better development capabilities and thus they can make the same compound more successful than if it stayed with the originating firm. This may reflect the greater advantages and downstream resources of the pharmaceutical firms downstream – e.g. greater ability to organize the clinical trials; learning and experience from such trials; or ability to fertilize one domain from knowledge and expertise in other domains. Even when the licensor is another pharmaceutical firm, the licensee may be more competent in that therapeutic category than the licensor.

**Table 10a**: Probability of selection, by originator and developer type

| | | licensed to | | | | | |
|---|---|---|---|---|---|---|---|
| Originated by | in-house | PB | EP | OB | OP | Tot. Lic. | Total |
| Pioneer Bio (PB) | 0.40 | 0.50 | 0.61 | 0.00 | 0.64 | 0.54* | 0.42 |
| Established Pharma (EP) | 0.57 | 0.75 | 0.58 | 0.86 | 0.25 | 0.62 | 0.57 |
| Other Bio (OB) | 0.38 | 0.25 | 0.29 | 0.13 | 0.42 | 0.30 | 0.37 |
| Other Pharma (OP) | 0.53 | 1.00 | 0.50 | 0.50 | 0.57 | 0.59 | 0.54 |
| Total | 0.48 | 0.52 | 0.47 | 0.31 | 0.53 | 0.47 | |

** Chi-squared test of association between license (preclinical) and success statistically significant at 5% level*

**Table 10b**: Probability of success, by originator and developer type

| | | licensed to | | | | | |
|---|---|---|---|---|---|---|---|
| Originated by | in-house | PB | EP | OB | OP | Tot. Lic. | Total |
| Pioneer Bio (PB) | 0.31 | 0.00 | 0.36 | | 0.89 | 0.47 | 0.34 |
| Established Pharma (EP) | 0.37 | 0.00 | 0.50 | 0.00 | 0.00 | 0.70 | 0.36 |
| Other Bio (OB) | 0.17 | 0.00 | 0.00 | | 0.00 | 0.00 | 0.16 |
| Other Pharma (OP) | 0.49 | 0.67 | 1.00 | 0.00 | 1.00 | 0.82** | 0.54 |
| Total | 0.34 | 0.22 | 0.38 | 0.00 | 0.75 | 0.41 | |

*** Chi-squared test of association between license (preclinical) and success statistically significant at 5% level. However, this cell has fewer than 5 observations.*

The poor performance of biotech licensees, even relative to their in-house projects, does raise the possibility that these buyers face a lemons problem. An alternative hypothesis is that such licensees lack the absorptive capacity to use externally developed technology. It is possible to sort through these hypotheses with our data. The lemons theory suggests that some licensors, with low quality projects, misrepresent the quality of their projects to be higher than the true level. The implication is that the projects retained in-house would have a higher quality. However, this is not borne out in our data. Tables 10a and 10b show that there is no significant difference in the probability of success of in-house projects

and licensed-out projects at the aggregate level, even though licensed projects are characterized by a higher probability of success, both conditioning on selection and unconditionally. This is confirmed by re-estimating our model where we control for the source of the license. The results are qualitatively unchanged and are available upon request. All in all, we find little support for the lemons hypothesis.

In some cases, we have multiple observations on the same compound, when the same compound undergoes trials for multiple indications. In particular, if the compound has already been advanced to clinical trials for one indication, this may raise the probability of the compound being advanced for another indication. To test whether this significantly biases our results, we included in the selection equation, a dummy variable indicating whether a compound undergoing multiple trials for different indications had already been moved into clinicals for a different indication. This would decrease the *net* additional cost for undertaking the trial (since safety, assessed during Clinical I, is already established). As expected, the estimated coefficient of the dummy variable is negative and significant, but none of the other estimated coefficients change substantially.[14]

**Check for external validity**

To validate our estimation results, we compared our estimated Pg* thresholds, which correspond to the ratio of development costs and expected revenues, with estimates of costs and revenues provided by the existing literature. In Table 11, we report the estimates of cost and net present value, and compute Pg* as the ratio of cost to present value deflated by the GDP deflator. As a proxy for D, we use the cost estimates produced by DiMasi et al. (2003), who report average cost expenditure for projects developed by ten major pharmaceutical companies. As a proxy for V, we use a variety of different estimates. We use net present value estimates from OTA (1993) and Grabowski et al. (2002). They report respectively value estimates based on the new chemical entities (NCE) launched in 1981-1983, and 1990-1994. Grabowski et al. (2002) consider different scenarios for the estimation of the present value of new chemical entities launched between 1990 and 1994. As an alternative estimate, we used the average sales of products that

[14] The results are not reported here but available upon request.

entered the US market in 1992 over the 10 years following their launch.[15] The estimates of Pg* obtained from these calculations range from 0.07 to 0.19, compared to our estimated value of 0.11, which is very close to the median.

**Table 11**: Estimates of Pg* from existing literature

| Cost Estimate | Present Value Estimate | Pg* |
|---|---|---|
| DiMasi, Hansen, Grabowski (2003) | Sales in 17 countries for 10 years | |
| $39,854,193 | $210,814,761 | 0.19 |
| DiMasi, Hansen, Grabowski (2003) | OTA (1993; NCE launched 1981-83) | |
| $39,854,193 | $321,326,923 | 0.12 |
| OTA (1993) | OTA (1993; NCE launched 1981-83) | |
| $24,218,269 | $321,326,923 | 0.08 |
| DiMasi, Hansen, Grabowski (2003) | Grabowski, Vernon, DiMasi (2002; NCE launched 1990-94) | |
| $39,854,193 | baseline $502,131,971 | 0.08 |
| $39,854,193 | at 40% margin $430,043,718 | 0.09 |

**Limitations and future research**

Our results are from the pharmaceutical sector drug development. This sector is extraordinarily research intensive, and over the last three decades, its innovation process has been transformed, with a much greater role for genetics and molecular biology. These transformations have opened up opportunity for entry. Thus, it is a good test bed for exploring differences in innovation capability between incumbents and entrants. Even so, we focus only on one part of the innovation process, drug development, and arguably the part most favourable to incumbents. A fuller analysis, which also incorporates the more upstream research and discovery process, is left for further research.

Another avenue for additional research has to do with differences in risk aversion. We have discussed the impact risk aversion on the selection process. However, such differences may also be a source of unobserved heterogeneity in conditioning success. In the analysis we assume that our measure of market size and concentration, and our controls for whether the molecule is first-in-class or not, as well as the other disease characteristics such as lethality, chronicity, and multiplicity of causes, control for differences in the potential net revenues. However, it is possible that incumbents show superior performance by trying to hit singles while entrants are swinging for the fences. In economic terms,

[15] Since we limit sales to 10 years we are underestimating the net present value of the launched products, whose revenue can extend beyond 10 years.

entrants are willing to take bets with smaller probabilities of success but much higher rewards if successful. By contrast, incumbents have more financial reserves, which also makes them more cautious in terms of how much they are willing to bet on any particular project. From a societal perspective, this is an efficient division of labor. Society enjoys the benefits of experimentation by entrants without unduly risking the valuable capabilities in research, development and marketing of the established incumbents. On the other hand, the conventional wisdom is that even the incumbents are searching for blockbusters; their large downstream capabilities mean that unless the compound has a significant market potential, incumbent pharmaceutical firms are uninterested in it. We cannot resolve this issue here and leave it for further research.

Our results on licensed compounds point to the benefits of the division of labor between entrants and incumbents. From an econometric viewpoint, we treat licensing as exogenous, because estimating a separate licensing equation (to account for potential endogeneity) will require much more detail about potential licensors and licensees than we currently have. The available evidence suggests that projects that are out-licensed and that are retained in-house appear to be drawn from very similar distributions. This suggests that endogeneity is unlikely but we cannot be more definitive at this stage.

Another extension would be to allow for interactions across projects carried out by a single firm. In developing our hypotheses we discussed some such interactions, most notably the portfolio effect. A firm with multiple projects for the same indication will likely commercialize the most promising and abandon others, even perhaps some that could have successfully passed clinical trials. In our estimation, such interactions are indirectly captured through our measures of scale and scope, leading to potentially confounding results, and partly accounted for by reporting estimates where errors are clustered at the level of the firm. Fully allowing for interactions would imply a very complex likelihood function that would be very difficult to estimate given our data.

## 8. Conclusions

The discovery of recombinant DNA was thought to have heralded the era of biotechnology in the pharmaceutical industry. Many thought that this would lead to a spate of entry which would unleash widespread changes in the structure of the pharmaceutical industry. Though much has changed in the pharmaceutical industry, entrants have been a breath of fresh air rather than the gale of creative destruction that some foresaw. Two decades later, we are beginning to understand better the nuances of this evolution.

In particular, in this paper, we have focused on understanding differences across firms in the performance in drug development. We do so by developing a simple model, wherein we explicitly model how a firm decides to select a compound tested in animals (pre-clinicals) into human clinical trials. This selection decision depends upon economic factors, such as the size and profitability of the market, but also technical factors, including the likelihood that the compound will successfully pass the clinical trials and receive FDA approval. Though simple, our model is essential for disentangling the differences in outcome due to selection and those due to underlying differences in innovation performance. Further, by modeling selection, we encompass theories that directly speak to the issue of how firms manage R&D, instead of merely treating selection as a nuisance problem plaguing the measurement of performance. We bring this model to a rich dataset, where unlike much of the prior literature, we also observe the compounds that are rejected and never brought into clinical trials. As well, we develop extensive and rich controls for the therapeutic area, and for the compound itself.

We find that incumbents demonstrate greater success in taking projects successfully through clinical trials compared with entrants, even after very extensive and fine grained controls for the type of therapeutic area, the type of compound being tested, and, most importantly, after controlling for differences in selection processes. Though incumbents are indeed more selective, their superior performance is not due to that alone, nor to economies of scale and scope, nor firm size. Not is it an artifact of incumbents pruning their portfolio of candidate compounds to leave only the most promising or

other strategic considerations. Instead it appears that incumbents represent a collection of experience and expertise, born of learning and honed by selection.

Entrants start by being different but over time may become more similar to incumbents; older biotechs are closer to established incumbents than are the more recent entrants in both their selection threshold and overall performance. Entrants are more aggressive than incumbents in moving compounds into clinical trials, perhaps because, paradoxically enough, they are better situated to take on risks. Even though these compounds fail at a much higher rate than do those of incumbents, the social benefits of entry are not entirely negated. In part, this is because entrants can supply incumbents with innovative drugs to be tested and developed. Though the market for licensed technology can be marred by problems of valuation, transaction costs and asymmetries of information, it appears that incumbents in this industry can avoid the problem of licensing lemons.

## References


Abrantes-Metz, R., C. Adams, A. Metz. 2006. Pharma Development Phases: A Duration Analysis. *Journal of Pharma Finance, Economics and Policy* **14(4)** 19-41.

Acs, Z., D. Audtretsch. 1990. Innovation and Small Firms. Cambridge MA. MIT Press.

Adams, C., V. Brantner. 2003. New Drug Development: Estimating Entry from Human Clinical Trails. FTC Working Paper #262.

Akerlof, G.A. 1970. The Market for 'Lemons': Quality Uncertainty and the Market Mechanism. Quarterly Journal of Economics 84 488-500.

Arora, A., A. Fosfuri, A. Gambardella. 2001. Markets for Technology: The Economics of Innovation and Corporate Strategy. Cambridge MA. MIT Press.

Arora, A., A. Gambardella. 1994. Evaluating Technological Information and Utilizing it. Journal of Economic Behavior and Organization 24 91-114.

Arrow, K.J. 1962. Economic welfare and the allocation of resources for invention, in: R.R.Nelson (ed.), The rate and direction of inventive activity, Princeton University Press: Princeton, 609-624.

Arrow, K.J. 1983. Innovation in Large and Small Firms. In Ronen J. (ed.) Entrepreneurship. Lexington MA. Lexington Books.

Braunwald, E., A.S. Fauci, D.L. Casper, S.L. Hauser, D.L. Longo, J.L. Jameson. 2001. Harrison's Principle of Internal Medicine. 15th edition. Mc Graw-Hill.

Chesbrough, H., D.J. Teece. 1996. When is Virtual Virtuous? Organizing for Innovation. Harvard Business Review 74(1) 65-72.

Christensen, C. 1997. The Innovator's Dilemma. Boston: Harvard Business School Press.

Cohen, W.M., D. Levinthal. 1989. Innovation and Learning: the Two Faces of R&D. Economic Journal 99(397) 569-596.

Cohen, W., D. Levinthal. 1990. Absorptive Capacity: A New Perspective on Learning and Innovation. Administrative Science Quarterly 35 128-152.

Danzon, P.M., S. Nicholson, N.S. Pereira. 2005. Productivity in Pharma-Biotechnology R&D: the Role of Experience and Alliances. Journal of Health Economics 24 317-339.

Dasgupta, P., P.A. David. 1994. Toward a new economics of science. Research Policy, 23(5) 487-521.
DiMasi, J.A., R.W. Hansen, H.G. Grabowski. 2003. The Price of Innovation: New Estimates of Drug Development Costs. Journal of Health Economics 22(2) 151-186.
Gambardella, A. 1995. Science and Innovation in the US Pharma Industry. Cambridge UK: Cambridge University Press.
Giarratana, M.S. 2004. The Birth of a New Industry: Entry by Start-Up and the Drivers of Firm Growth. The Case of Encryption Software. Research Policy 33 787-806.
Grabowski, H.G., J. Vernon, J.A. DiMasi. 2002. Returns on R&D for 1990s New Drug Introductions. www.dklevine.com/archive/grabow-randd_returns.pdf.
Guedj, I., D. Scharfstein. 2005. Organizational Scope and Investment: Evidence from the Drug Development Strategies and Performance of Biopharmaceutical Firms. NBER Working Paper 10933. Cambridge MA.
Heckman, J.J. 1979. Sample Selection Bias as a Specification Error. Econometrica 47(1) 153-161.
Henderson, R., K. Clark. 1990. Architectural Innovation: The Reconfiguration of Existing Product Technologies and the Failure of Established Firms. Administrative Science Quarterly 35(1) 9-30.
Henderson, R., I. Cockburn. 1996. Scale, Scope, and Spillovers: The Determinants of Research Productivity in Drug Discovery. RAND Journal of Economics 27(1) 32-59.
Higgins, M.J., D. Rodriguez. 2006. The outsourcing of R&D through acquisition in the pharmaceutical industry. Journal of Financial Economics 80 351-383.
Holmstrom, B. 1989. Agency Costs and Innovation. Journal of Economic Behavior and Organization 12(3) 305-327.
Levinthal, D., J. March. 1993. The Myopia of Learning. Strategic Management Journal 14 95-112.
Macher, J.T., C.S. Boerner. 2006. Experience and Scale and Scope Economies: Trade-Offs and Performance in Development. Strategic Management Journal 27 845-865.
Nerkar, A. 2003. The Folly of Rewarding Patents and Expecting FDA Approved Drugs. Working paper.. http://bschool.washington.edu/departments/mo/PDF/Nerkar_UW.pdf., Retrieved 27 Feb, 2007.
Nerkar, A., P. Roberts. 2004. Technological and Product-Market Experience and the Success of New Product Introductions in the Pharma Industry. Strategic Management Journal 25 779-799.
Pammolli, F., M. Riccaboni. 2004. Technological Competencies in Networks of Innovators. In Cantwell, J., A. Gambardella, O. Granstrand (eds.). The Economics and Management of Technological Diversification. London and New York: Routledge.
Pisano, G. 1997. R&D Performance, Collaborative Arrangements, and the Market for Know-How: A Test of the "Lemons" Hypothesis in Biotechnology". Mimeo. Harvard Business School, Cambridge MA.
Pisano, G. 2006. Can Science Be a Business? Lessons from Biotech. Harvard Business Review 84(10) 114-124.
Powell, W., K. Koput, L. Smith-Doerr, L. 1996. Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology. Administrative Science Quarterly 41 116-145.
Reddy, K.A. 2003. The agony and ecstasy of Drug Discovery. CRIPS 4(1) 10-12.
Rosenberg, N. 1990. Why Do Firms Do Basic Research (With Their Own Money)?. Research Policy 19 165-174.
Shaver, J.M. 1998. Accounting for Endogeneity When Assessing Strategy Performance: Does Entry Mode Choice Affect FDI Survival?. Management Science 44(4) 571-585.
Stiglitz, J. E., A. Weiss. 1981. Credit Rationing in Markets with Imperfect Information. *The American Economic Review* 71(3) 393-410.
Thompke, S., W. Kuemmerle. 2002. Asset Accumulation, Interdependence And Technological Change: Evidence From Pharmaceutical Drug Discovery. Strategic Management Journal 23(7) 619-635.

U.S. Congress, Office of Technology Assessment (OTA). 1993. Pharmaceutical R&D: Costs, Risks and Rewards. U.S. Government Printing Office, Washington, DC.
U.S. Food and Drug Administration (FDA), Center for Drug Evaluation and Research. 1999. From Test Tube to Patient: Improving Health Through Human Drugs (special report). U.S. Government Printing Office, Washington, DC.