MANAGEMENT SCIENCE
Vol. 55, No. 10, October 2009, pp. 1638–1653
ISSN 0025-1909 | EISSN 1526-5501 | 09 | 5510 | 1638

DOI 10.1287/mnsc.1090.1055

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

Ashish Arora
Fuqua School of Business, Duke University, Durham, North Carolina 27708; and
National Bureau of Economic Research, Cambridge, Massachusetts 02138, ashish.arora@duke.edu

Alfonso Gambardella
Department of Management and KITeS, Università Bocconi, 20136 Milan, Italy, alfonso.gambardella@unibocconi.it

Laura Magazzini
Department of Economics, University of Verona, 37129 Verona, Italy, laura.magazzini@univr.it

Fabio Pammolli
IMT Lucca Institute for Advanced Studies, 55100 Lucca, Italy; and Fondazione CERM, 00187 Rome, Italy,
pammolli@gmail.com

This paper compares the innovation performance of established pharmaceutical firms and biotech companies, controlling for differences in the scale and scope of research. We develop a structural model to analyze more than 3,000 drug research and development projects advanced to preclinical and clinical trials in the United States between 1980 and 1994. Key to our approach is careful attention to the issue of selection. Firms choose which compounds to advance into clinical trials. This choice depends not only on the technical promise of the compound, but also on commercial considerations such as the expected profitability of the market or concerns about product cannibalization. After controlling for selection, we find that (a) even after controlling for scale and scope in research, established pharmaceutical firms are more innovative than newly entered biotech firms; (b) older biotech firms display selection behaviors and innovation performances similar to established pharmaceutical firms; and (c) compounds licensed during preclinical trials are as likely to succeed as internal compounds of the licensor, which is inconsistent with the "lemons" hypothesis in technology markets.

Key words: firm capabilities; drug development process; market for technology
History: Received March 1, 2007; accepted February 20, 2009, by Scott Shane, technological innovation, product development, and entrepreneurship. Published online in *Articles in Advance* August 3, 2009.

1. Introduction

Understanding the determinants of the innovative performance of firms is central to innovation and entrepreneurship research. Yet, the development of innovations by firms is not simply a matter of technical ability. Incentives matter as well. A classical example is Arrow (1962), whereby larger firms have lower incentives to innovate for fear of cannibalizing existing products. Similarly, if a firm faces a larger potential market or has established downstream commercialization capabilities, it will have higher innovation incentives. Therefore, in studying the drivers of innovation performance, it is important to distinguish between incentives and capabilities, as one may confound lack of incentives with lower technical capabilities or vice versa.

A key contribution of this paper is to develop and estimate a structural model that distinguishes between these two dimensions. We use observations on more than 3,000 drug research and development (R&D) projects initiated between 1980 and 1994, and followed through 2005. The pharmaceutical industry is ideal for our analysis because the drug innovation process follows well-defined steps (see, e.g., FDA 1999, DiMasi et al. 2003). New compounds are generated in the laboratory (initial discovery), tested on animals (preclinical research), and if the firm considers them sufficiently promising, both technically and economically, they are advanced into clinical trials on humans. This process has two important regulatory gates: at the beginning of clinical trials, when an investigational new drug application is filed, and at the end of clinical trials, when a new drug application may be filed. These two main stages play an important role in our analysis.

We argue that firms do not advance all technically promising compounds to clinical trials; economic considerations also affect this decision, creating differences in selection behavior across firms. Thus, if one measures innovation capability by the share of preclinical compounds that eventually become new

drugs, one may confound selection with capability (e.g., Shaver 1998).

The contribution of this paper to the literature goes beyond flagging a selection problem and developing a method to cope with it. We focus on two questions that can be studied particularly well in the context of the pharmaceutical industry. The first is the long-standing comparison between the innovation performance of large and small firms (e.g., Arrow 1983, Holmstrom 1989, Acs and Audretsch 1990, Henderson and Clark 1990, Chesbrough and Teece 1996, Levinthal and March 1993). One view holds that, whereas large firms have an advantage in commercialization, small firms have more flexible organizations that enhance creativity and attract more inventive minds, which makes them more productive in R&D. A different view holds that economies of scale and scope, and learning advantages, make larger firms better at innovation (e.g., Henderson and Cockburn 1996, Macher and Boerner 2006). In this view, the comparative advantage of small firms may lie in exploring technologies and markets that established firms are unwilling to explore (Christensen 1997, Klepper and Thompson 2006, Giarratana 2004). We cannot directly measure comparative advantage. However, because older, established firms and younger biotech firms each perform R&D, we can examine if older and younger firms differ in their ability to develop compounds successfully. This is our measure of innovation performance.

Our second question is whether markets for technology (Arora et al. 2001) are afflicted by a "lemons" problem (Akerlof 1970). Are compounds offered for license of lower quality than compounds developed in-house? The current evidence is mixed. Danzon et al. (2005) find that licensing has a positive effect on the probability of success of drug compounds, whereas Pisano (1997) and Guedj (2005) find that licensed drugs are less successful than those developed in-house. Our paper tests these contending views.

The next section reviews the main findings of the empirical literature on drug development in the pharmaceutical industry and places our contribution in context. Section 3 describes the model, and §4 develops our hypotheses. Section 5 describes our data and analysis. Section 6 discusses our estimates. Section 7 discusses the robustness of our results and concludes.

2. Innovation in Pharmaceuticals: Related Literature

One group of studies on drug innovation focuses on the internal economies of the firm (economies of scale, scope, spillovers, experience). Henderson and Cockburn (1996) study the determinants of research (drug discovery) performance, measured by the number of important drug patents. They use data at the level of the individual research program from the internal records of 10 major pharmaceutical firms, and find returns to scale at the level of both individual research programs and research expenditures of the firm as a whole, as well as evidence of economies of scope. Our paper investigates the stages after drug discovery, as do most of the studies described below.

Nerkar and Roberts (2004) examine the determinants of commercial success of new pharmaceutical products. They find that proximal technological experience (patents in the same therapeutic class) has a positive and significant effect on 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. Nerkar (2003) also studies the relationship between experience and the probability that a drug receives U.S. Food and Drug Administration (FDA) approval. He hypothesizes that experience may not lead to better performance if the experience is of the wrong sort (e.g., oriented toward drug discovery rather than drug development), and if feedback is inadequate or delayed. With drug development cycles stretching to 10 years or more, researchers may be rewarded for discovery (measured by patenting) rather than commercialization (measured by FDA approval). As a result, scientists may continue to work in areas that provide them with clumps of patents but do not lead to commercially useful drugs. As we discuss below, even small firms may suffer from a misalignment of incentives, as reflected in their selection behavior.

A second group of studies looks specifically into the drug development process by examining projects in clinical trials. Adams and Brantner (2003) and Abrantes-Metz et al. (2006) analyze drugs in clinical trials around the world between 1989 and 2002. They 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. Macher and Boerner (2006) also estimate time to complete clinical trials, but by looking at contract research organizations rather than integrated R&D projects in pharmaceutical firms. They find that scale and scope economies, and experience, are valuable.

Danzon et al. (2005) focus on the role of experience (both overall experience and experience in a particular therapeutic category) and alliances on the outcome of R&D projects. Their sample consists of R&D projects from more than 900 firms during 1988–2000.

They too find evidence of large, positive returns to a firm's overall experience in the larger and more complex late-stage trials. They recognize, but do not explicitly model, differences across firms in the quality of their drug candidates.

Our paper shares the approach and goals of these papers. However, none of them explicitly address the 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 preclinical trials, particularly late-stage clinical trials. Which compounds make it into clinical trials depends on whether managers think the compounds are promising. All else being equal, compounds with a higher likelihood of FDA approval are more likely to be advanced into clinical trials than compounds with a lower probability of success.

Commercial considerations are also very important. Acemoglu and Linn (2004) present evidence that larger pharmaceutical markets see more innovation. Furthermore, a firm with downstream marketing capabilities in a particular market (such as Bristol-Myers Squibb in anticancer or Merck in cardiovascular) may be able to extract greater value from a drug than a firm that lacks such capabilities, and may therefore be willing to take a chance on a compound with a lower likelihood of success. Other considerations may also be important. A drug championed by a high-status research team may be selected for clinical trials even if it is below threshold (Nerkar 2003). Guedj and Scharfstein (2005) point to the agency problem between managers and investors. In biotech startups, managers typically have one or two compounds to bet on, and therefore try to push them into clinical trials. By contrast, larger firms 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 (Guedj and Scharfstein 2005). In our paper, we distinguish between preclinical and clinical trials, but do not separately analyze progress across the various clinical stages.

To appreciate further the contribution of our paper, it may be useful to compare it with Chandy et al. (2006), who examine the "conversion ability" of firms in translating drug-related patents into new drugs. They find that conversion ability is greatest in firms that develop an intermediate number of drug-related patents. Because firms patent compounds soon after discovery, Chandy et al. (2006), in essence, study the success rate of preclinical compounds. Our analysis "unpacks" this sequence into whether a compound is selected into clinical trials, and if selected, whether it successfully receives FDA approval, and models the drivers of the underlying processes. This enables us to move away from the assumption implicit in Chandy et al. (2006) that firms attempt to convert every compound. Instead, our model allows the stringency of selection to be driven by the expected profitability of the market, the scale and scope of the firm, and other technology, market, and firm characteristics, ensuring that the rate at which preclinical compounds succeed is determined not only by the factors that drive innovation performance but also by the factors that drive selection into clinical trials. In econometric terms, what allows us to identify selection is that the major resource commitments are made in clinical trials rather than in preclinical trials (DiMasi et al. 2003), so that we can assume that commercial considerations affect selection, but do not affect the outcome once a compound is into clinical trials.

Finally, the literature addresses the question of whether licensed compounds are more likely to succeed compared to in-house projects. The market-for-technology perspective (Arora et al. 2001) suggests that smaller, technology-specialist firms and established manufacturers have comparative advantages in different stages of the innovation process. In particular, technology specialists are relatively more effective in upstream innovation activities, which rely on creativity, rather than in downstream development and commercialization, which is intensive in resources and scale. Licensing permits the technology specialists to "cooperate" with downstream incumbents, increasing overall efficiency. Zeckhauser (1996) and Pisano (1997, 2006) argue that technology markets are inefficient and are potentially afflicted by a lemons problem. Pisano (1997) finds that in-house development is superior to licensed compounds, and Guedj (2005) finds that projects financed by pharmaceutical companies but developed by biotech firms are more likely to fail. 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. Note that Danzon et al. (2005) include in their sample alliances formed prior to the conclusion of each phase (up to phase III). Because these alliances are likely to include marketing agreements, which pharmaceutical firms often strike for successful compounds to enhance market access, there is the potential for an upward-biased estimate. To avoid this problem, we follow Guedj (2005) and only include licenses signed before phase I clinicals, and test whether licensed compounds differ from in-house compounds. However, comparing compounds discovered by biotech firms and licensed to pharmaceutical firms to compounds discovered in-house by pharmaceutical firms is not the correct way to test for a lemons problem.

A lemons problem would exist if firms systematically licensed inferior drugs and kept superior ones in-house. Thus, we estimate whether licensed compounds are drawn from a different distribution than those that firms develop in-house.

3. An Econometric Model of Selection and Success

To understand the interplay between selection and success, we provide a simple model. After a compound has been discovered, the firm starts preclinical trials to evaluate its properties. Preclinical research will provide the firm with an estimate of the probability that the compound will succeed, i.e., pass clinical trials and receive FDA approval. Let Pg be the firm's (unbiased) point estimate of the (uncertain) probability of success of a particular compound. We will sometimes refer to this as the quality of the compound. Formally, Pg_i, the probability that compound i of firm j will pass clinical trials, is a random variable drawn from a distribution with a firm specific mean, μ_j (the innovative capability of the firm).[1] Actual success or failure is only observed for compounds selected for clinical trials.

3.1. Drug Selection

Dropping the subscripts for compound and firm to avoid clutter, let V indicate the expected net additional revenue if the compound is eventually marketed, and let D indicate the net additional clinical trial costs, which we label development costs. The firm will take the compound into clinical trials 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. Note that Pg^* is not a random variable but rather a threshold value. The probability of a compound being selected depends on both the threshold, Pg^*, and the "quality" of the compound, Pg, where the latter is itself a random variable. Thus, the *average* probability that a compound is selected depends on Pg^* and μ.

The selection threshold, Pg^*, represents the economic and strategic dimensions of selection. For instance, firms with full pipelines of drugs under development will assess a higher net addition to development costs (D), and hence, have a high Pg^*. Market profitability matters as well. To anticipate our empirical analysis, competition in the product market will reduce the expected value, V, implying a higher Pg^*. Similarly, larger markets will increase V, implying a lower Pg^*. Also, public firms may have a lower cost of capital, implying a lower D, and hence, a lower Pg^*.

[1] In the empirical analysis, μ will also differ by therapeutic area (indication).

We assume that once a compound enters clinical trials, its future progress in clinical trials is determined only by its technical characteristics. Our interviews with industry executives suggest that over the course of development, the relative importance of technical characteristics increases. Early in drug development, the selection threshold is influenced by commercial considerations such as the size of the market, the extent of competition in the market, and the firm's relevant commercialization capabilities. Technical (therapeutic) characteristics gain influence as a compound progresses through trials, and its safety, effectiveness, and potential side effects become apparent. We simplify by assuming that technical promise, Pg, and commercial considerations, represented by Pg^*, both determine selection. However, once selected, the FDA approval takes place with probability Pg. Because we do not observe Pg, the expected probability of FDA approval is the expected value of Pg, conditional upon Pg being greater than Pg^*.[2]

3.2. Drug Development

Let $Y_1 = 1$ if the compound is selected for clinical trials, 0 otherwise. Therefore the probability of selection is $\Pr(Y_1 = 1) = \Pr(Pg > Pg^*)$. Let $Y_2 = 1$ if the project succeeds if selected for trial, and 0 if it fails in clinicals, so that $\Pr(Y_2 = 1) = E(Pg \mid Pg > Pg^*)$. To estimate the model, we assume that the log of the odds ratio, $\ln(Pg/(1-Pg))$, is normally distributed with mean μ and unit variance. Furthermore, we assume that μ 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. Because Pg^* must be between 0 and 1, we assume $Pg^* = \exp(X\beta)/(1+\exp(X\beta))$ so that $\ln(Pg^*/(1-Pg^*)) = X\beta$. We maximize the following log-likelihood function:

$$\begin{aligned} L(\beta,\gamma) = &\sum_{Y_1=0} \Pr(Y_1=0) + \sum_{Y_1=1,\,Y_2=0} \Pr(Y_1=1 \text{ and } Y_2=0) \\ &+ \sum_{Y_1=1,\,Y_2=1} \Pr(Y_1=1 \text{ and } Y_2=1) \\ &+ \sum_{Y_1=1,\,Y_2=\cdot} \Pr(Y_1=1 \text{ and } Y_2=\cdot), \end{aligned} \tag{1}$$

where the four terms represent, respectively, not selected, selected but failed, selected and succeeded, and selected and still in trial (log of probabilities is considered). Note that we have two binary dependent variables, and we are estimating the drivers of Pg^* and μ, rather than the more typical exercise of estimating the probability of selection and the probability

[2] A different model would start with some projects being "good" and others "bad," with a real valued signal, Pg, about whether the project is good. Instead, we consider each project to be like a coin toss, with the probability of success ("heads") equal to Pg. The firm observes Pg and then decides whether it wants to toss the coin.

of success conditional on selection. The probability of selection is

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

whereas the probability of success is

$$\Pr(Y_2=1\mid Y_1=1)$$
$$=E[Pg\mid Pg>Pg^*]=\int_{Pg^*}^{1} Pg\,f(Pg\mid Pg>Pg^*)\,dPg$$
$$=\frac{1}{1-\Phi(X\beta-Z\gamma)}\int_{Pg^*}^{1} Pg\,f(Pg)\,dPg. \quad (3)$$

Danzon et al. (2005) effectively estimate $\Pr(Y_2=1\mid Y_1=1)$. Although they acknowledge the selection issue by including the share of projects selected from preclinicals in that indication as a regressor, they cannot separately identify γ and β. Nerkar (2003) and Chandy et al. (2006) estimate $\Pr(Y_1=1 \text{ and } Y_2=1)\equiv \Pr(Y_2=1\mid Y_1=1)\Pr(Y_1=1)$. All else being equal, as a firm becomes more selective (the second term decreases), the compounds it advances into clinicals will be more likely to succeed (the first term increases), but the product of the two, the share of compounds in preclinicals that eventually succeed, may either fall or rise. Without modeling both $\Pr(Y_2=1\mid Y_1=1)$ and $\Pr(Y_1=1)$, and without explicitly recognizing the interdependence between them, one cannot distinguish innovativeness (i.e., μ) across firms.

4. Hypothesis Development

4.1. Selection Hypotheses

Our first set of hypotheses is about selection. We define a firm to be more selective if Pg^* is higher. Broadly, selection reflects the cost of development and the expected profit given a successful drug. These must be distinguished from factors that condition the likelihood of success—the drivers of μ.

We hypothesize that there are program-level economies of scale. For example, a company with long-standing arrangements with academic hospitals for trials in an indication will find it easier to develop another compound within that indication than in others, implying a lower development cost, D. Similarly, links with prescribing physicians in a therapeutic area will increase sales of an approved drug, increasing profits, implying a higher V. In both cases, Pg^* will be lower. Adding more programs or increasing the scale of other programs are unlikely to lower the development and clinical trial costs of a molecule or allow the firm to extract more profit from it if approved. Firm scale and scope will condition the likelihood of success, as we argue later. Accordingly, though we include measures of firm scale and scope as controls, we do not hypothesize how they condition Pg^*. Formally, we have the following:

Hypothesis 1. *Research scale at the level of the program makes the firm less selective* (*lower* Pg^*).

The pharmaceutical industry has two fundamentally different types of firms: the established pharmaceutical producers and the biotechnology companies. The established firms were founded early in the 20th century, if not earlier. A large gap separates the entry of the pharmaceutical firms and the first biotech company, Genentech, in 1976. Syntex, the youngest research-based U.S. pharmaceutical company before Genentech, was founded in 1958. Waves of biotech companies have entered after Genentech. Their main driver has been research, and the comparative advantages and specialization of these firms are clearly upstream compared to the drug manufacturers (Gambardella 1995). Such differences in origin, time of entry, and development underlie profound differences between these two company types.

For example, Arrow (1962) suggests that the fear of cannibalizing profits from existing markets and products leads an incumbent firm to underinvest in research. In established firms, new compounds are more likely to cannibalize existing products. By contrast, most biotech firms have no products on the market and have little to fear about cannibalization of existing products. Insofar as one cannot fully control for this effect in the empirical analysis, the cannibalization effect is an important reason to distinguish between different firm types. Difficult-to-measure differences in organizational capabilities are another reason to distinguish between biotech and pharmaceutical firms. Interviews with managers reveal that pharmaceutical firms have well-established routines and financial models for deciding how compounds are moved along through the various stages of preclinical and clinical research. By contrast, many of the smaller entrants lack such discipline, and 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).

In addition, there may be differences in risk aversion. Ordinarily, one might imagine that large, established firms would be less risk averse. However, as Stiglitz and Weiss (1981) showed in a seminal paper, limited liability laws may 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. With few assets to protect, and with outside equity investors frequently having a large share of the capital at risk

if the project (and the firm) fails, managers of small biotech firms may be inclined to swing for the fences, which implies that they would have a lower Pg^*. Indeed, using patent citation analysis, Owen-Smith and Powell (2006) find that biotech firms in the Bay Area, which has a greater involvement of venture capitalist firms, are more likely to initiate riskier projects than biotech firms in the Boston region, where universities (and National Institutes of Health funding) are more heavily represented. Interviews with R&D executives support the idea that biotech firms tend to push compounds into clinical development that firms with bigger product portfolios might hold back. When asked, the managers also believe that biotech firms are more inclined to swing for the fences.

Among the biotech firms themselves, there are distinctions to be made. Some of the oldest biotech firms that survive today, such as Genentech and Amgen, have become similar to the established drug companies. They have products in the market and they have to manage cannibalization. They have strong financial models for selecting projects, and also considerable wealth at risk from an expensive failure. This suggests that the Pg^* of the older biotech firms lies somewhere between that of established pharmaceutical firms and that of the smaller, more recent biotech entrants.[3] We can then formulate the following hypothesis:

Hypothesis 2. *Even after controlling for the scale and scope of research, established pharmaceutical companies are more selective (higher Pg^*) than biotech firms. The older and more established biotech firms are more selective than the younger biotech firms.*

4.2. Performance Hypotheses

Note that whereas Hypothesis 1 relates to how scale affects development costs and profits from successful compounds, here we focus on how scale and scope affect the probability of success, roughly analogous to the difference between incentives and innovation capabilities.

Henderson and Cockburn (1996) identify three types of economies in research: economies of scale at the level of the research program; economies of scale at the level of the firm; and economies of scope across research programs of the firm (see also Macher and Boerner 2006, Nerkar and Roberts 2004). There are a variety of sources of scale and scope economies. A larger firm with a diverse research portfolio may enjoy knowledge spillovers (Henderson and Cockburn 1996); may have a larger chemical library (Thomke and Kuemmerle 2002); may be more knowledgeable about the underlying biochemical processes; may have better models for interpreting results from animal and small sample human studies; or may be better able to modify the lead compound to improve efficacy and reduce side effects.

These benefits may be offset by organizations that are maladapted, with perverse incentives and inadequate feedbacks (e.g., Nerkar 2003). There is, however, another reason why measures of program-level economies may appear to result in lower μ. A firm with a number of compounds in trial for the same indication is unlikely to commercialize all of them, even if they could all win approval. Instead, the firm will pick the most promising among them. Though our formal model assumes that once selected, compounds progress through clinical trials based entirely on technical merits, this is a simplification. In reality, firms may choose to bring a number of promising compounds into trial for an indication, but pick only the best, terminating the rest (Dahan and Mendelson 2001). The cannibalization effect may work similarly. A firm with a product may choose not to commercialize a drug that is safe and effective but very similar in its target market to the existing product. Put differently, the portfolio and cannibalization effects do not reduce μ, but instead result in an estimated μ that is lower than the true μ.

Hypothesis 3. *The innovation performance of a firm in a given research area increases with* (a) *the overall research scale of the firm*, (b) *the span of the research across research areas* (*scope*), *and* (c) *research scale at the level of the program. The portfolio and cannibalization effects reduce the measured innovation performance.*

As with selection, there are other differences between established firms and biotech companies that may affect performance. There is a substantial literature arguing that smaller research-intensive companies are more creative than established firms. For example, smaller firms are said to have a more open environment, enhancing benefits from external knowledge (e.g., Chesbrough 2003), and are less prone to asymmetric information between managers and inventors inside the organization (Arrow 1983, Holmstrom 1989). Conversely, internal bureaucracies inside large firms may crush good ideas. Also, larger firms may tend to rest on one's laurels (Christensen 1997), and they may stress exploitation over exploration (Levinthal and March 1993).

However, an alternative hypothesis has the opposite prediction. The agency model in Guedj and Scharfstein (2005), and the idea that small biotech firms may have excessive incentives for risk, implies that biotech firms would advance even inferior compounds into clinicals. In our model, this implies a

[3] Clearly, selection effects are at work here as well. Biotech firms that entered early and survived are more efficient, reinforcing the argument that they have become more similar to the established companies.

lower *estimated* μ for biotech firms. Ultimately, the net effect of these two forces is an empirical matter. We then formulate two alternative hypotheses:

Hypothesis 4. *Biotech companies have higher innovation performance than established pharmaceutical firms because smaller organizations are more innovative.*

Hypothesis 4′. *Biotech firms have lower innovation performance because of misaligned incentives.*

4.3. Licensing Hypotheses

If the licensing agreement does not share costs and benefits in the same proportion across the licensee and licensor, this will bias the selection threshold. In the extreme case, if the licensee is responsible for the cost of development but has to pay a royalty from sales of a successful compound to the licensor, this effectively reduces V for the licensee, implying a lower Pg^*. However, knowing this, partners try to align incentives by sharing costs and benefits. We do not observe details of the licensing contracts, and accordingly, we do not formulate a hypothesis regarding the effect of licensing on Pg^*.

Regarding the performance of licensed compounds, following Akerlof (1970), the economics literature predicts that markets for knowledge are characterized by information asymmetries about the quality of the technology offered for license (Zeckhauser 1996). Although these arguments are typically used to support the contention that licensing is not the preferred mode of commercialization, they can be extended to performance, as Pisano (1997) does in the pharmaceutical industry. The notion is that a licensor is likely to keep its most promising technologies for its own use and only offer inferior ones for license. Licensees therefore believe that compounds offered for licensing are on average worse than compounds not offered for license, conditional on observable characteristics. If so, the licensees must be offered a discount. But then the suppliers of good compounds are less likely to offer their compounds for license, thereby fulfilling the belief about compounds offered for license being inferior. If so, licensed compounds will be drawn from an inferior distribution (lower μ) than those retained for in-house development. Guedj (2005) also offers an asymmetric-information-based model that implies that licensed compounds not only have a lower Pg^* but also have a lower probability of success. He finds empirical support in a sample of projects selected for clinicals.

There are countervailing arguments suggesting that the market for lemons is not an important problem in technology markets. Arora and Gambardella (1994), building on the insights of Cohen and Levinthal (1989) and Rosenberg (1990), argue that research capabilities also provide firms with the ability to evaluate external technologies. Licensees in the drug industry are typically firms that perform research themselves, so that potential buyers are capable of evaluating the technical characteristics of the licensed compound, especially since strong patent protection in this industry facilitates the disclosure of relevant information. Furthermore, many licensing transactions are embedded in strategic research partnerships and involve minority investments or codevelopment deals, reducing potential information asymmetries. Of course, a potential licensor may suppress unfavorable test results or may not carefully probe avenues that might lead to uncomfortable findings. Therefore, the potential for asymmetric information and, consequently, for a lemons problem, always exists. Accordingly, we state the following:

Hypothesis 5 (Market for Lemons). *Licensed compounds exhibit lower innovation performance than internally developed compounds.*

5. Data

Our sample is drawn from the PHID (Pharmaceutical Industry Database) developed by the CERM Foundation, which combines proprietary data sets on the pharmaceutical industry. It reports information about more than 17,000 R&D projects all over the world carried on since the 1980s.[4] Following Danzon et al. (2005), we consider each indication for which the compound is being developed as a separate project. Also, countries have different institutions and procedures for advancing compounds to clinical trials and for approving them for sales. Even in our database, both the selection and success rates of European trials are higher than the United States, which suggests that the U.S. criteria are more stringent. To reduce this source of heterogeneity, we focus on U.S. trials. Our sample is composed of all projects with preclinical research conducted in the United States (whether by U.S. firms or not). We classify as "selected" those projects that start clinical I in the United Staes, and analogously for successful projects.[5] We exclude projects originated by hospitals, public sector labs, or universities, where selection processes may be different. We reduce the censoring problem posed by ongoing projects that are neither successes nor failures by including only those projects that entered preclinical

[4] Names of firms in the database are reported as they were in 2002. Therefore, we cannot consider separately the projects of firms that merged or were acquired during the 1990s.

[5] There are a few cases with missing information about the location of preclinical research, which we dealt with case by case, using other contextual information in the database.

trials between 1980 and 1994.[6] Our final sample is composed of 3,311 projects, with 329 distinct firms and 185 indications.[7]

We divide biotech firms into two types, depending on whether they were founded before or after 1988, the median founding year for the biotech firms in our sample. We also distinguish between the top pharmaceutical corporations worldwide (henceforth "*Established Pharma*"), and smaller pharmaceutical firms developing compounds that target predominantly infection and inflammation, asthma and diabetes, obesity, and sexual dysfunctions (henceforth "*Other Pharma*"), which are typically much less R&D-intensive. Specifically, we distinguish among four types of companies:

(i) *Pioneer Biotech*: Firms that apply biotechnological methods to the discovery and development of new drugs and were founded before 1988.

(ii) *Other Biotech*: All other firms using biotechnology for the discovery of new drugs.

(iii) *Established Pharma*: Pharmaceutical firms among the top 50 firms worldwide by sales for at least 10 years during the 15-year period 1983–1998 (excluding biotech firms and adjusting for mergers and missing data for 1995).

(iv) *Other Pharma*: All other pharmaceutical firms.

Table 1 reports the number of firms and projects related to the four types of firms considered in the analysis. Our empirical measures are summarized in Table 2. Measures that vary across firms and indications are noted as measured at the project level, whereas others vary only across firms, or only across indications.[8] Table 2 also shows the predicted sign of the variable. Some variables are excluded from μ (and the corresponding cell is left blank), whereas other variables are used as controls (and the corresponding cell has a check (✓) mark, rather than a plus (+) or minus (−) symbol).

Table 1 Firm Types and R&D Projects

Firm type	Number of projects (%)	Number of different firms (%)
Pioneer Biotech	1,031 (31.14)	123 (37.39)
Other Biotech	556 (16.79)	101 (30.70)
Established Pharma	1,512 (45.67)	31 (9.42)
Other Pharma	212 (6.40)	74 (22.49)
Total	3,311 (100)	329 (100)

Scale and Scope: We measure *Scale_program*, or the scale at the program level, as the cumulative number of past projects for that indication. *Scale_firm*, or scale at the level of the firm, is measured by the total number of ongoing projects in all indications. *Scope* is measured by the Herfindahl index of diversification of projects across indications, for the firm. These measures are based on the literature. Danzon et al. (2005) and Abrantes-Metz et al. (2006) use the number of drugs in development to measure overall firm scale. To measure program-level economies, Macher and Boerner (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 Herfindahl-index-based measure that we use, 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, which we cannot do because we lack R&D data at the program level.

Licensed Compounds: The database records the licensor and licensee(s).[9] We consider a project "licensed" only if the agreement was signed in preclinicals.

Controls: We develop binary, indication-level measures[10] to show (a) whether the disease is lethal (*lethal*), (b) whether it can result in organ damage or complications (*organ damage*), (c) whether the disease is chronic or acute (*chronic)*, (d) whether the disease

[6] Following Danzon et al. (2005), we classify as failures projects that remain in a phase, without any further reported events, for longer than the maximum number of years observed for completion of each phase in the noncensored sample. The maximum number of years is computed within each indication. When this is not practicable—due to missing observation for the indication—we consider the maximum over all the indications.

[7] By selecting trials started between 1980 and 1994, the sample reduces to 8,107 observations, of which 7,247 were started by pharmaceutical and biotechnology firms. We further exclude projects codeveloped by two or more institutions, leaving us with 6,659 observations. By focusing only on preclinical research projects undertaken in the United States and projects with identifiable outcomes, we end up with 3,311 observations.

[8] We use two complementary classification systems for identifying the relevant therapeutic market: namely, indication in standard term and the anatomical therapeutic classification (ATC). Indication refers to the target disease or clinical condition, focusing on the clinical symptoms for which the drug is being tested, whereas the ATC deals with the therapy, using anatomy as the main organizing theme. At the first ATC digit, compounds are grouped according to the organ or system on which they act, whereas at the second digit, classes are formed on the basis of pharmacological/therapeutic action. For example, class D comprises all dermatological compounds, with class D10 grouping all antiacne preparations. The corresponding indication is "acne." Class A refers to compounds targeted to the alimentary tract and metabolism, with diabetes drugs classified in class A10. The corresponding indication is "diabetes."

[9] Where multiple licensees were present, we assigned the development of the project to the largest U.S. firm, and manually verified that the firm was in charge of development.

[10] The main source for the disease information is Braunwald et al. (2001). Other information comes from e-medicine reviews from http://www.diseasedatabase.com. For diffusion data, we use the "rare disease database" cited by the FDA, and available at http://rarediseases.about.com/cs/orphandrugs/a/122103.htm.

Table 2 Variables Definitions

Pg^*	μ	Variable	Measured as	Measured at the level of
−	+	*Scale_program*	Number of projects already started by the firm for that indication	Project
✓	+	*Scale_firm*	Total number of projects by the firm	Firm
✓	+	*Scope*	Herfindahl index of firm's projects across indications	Firm
✓	−	*Licensed compound*	If the compound is licensed in preclinical	Project
✓	✓	*Closeness to science*	Share of projects jointly developed with a public research organization for that indication	Project
✓	✓	*Presence in ATC2 product market*	Equals 1 if firm has a product in the ATC2 class, equals 0 otherwise	Project
−		*Public*	Equals 1 if firm is public when the project is started	Firm
−		*Market size*	Worldwide sales in the ATC2 class	Indication
+		*Market competition*	1. Log of the number of firms with a product in the ATC2 class 2. Share of established pharma with a product in the ATC2 class	Indication
✓	✓	*Disease characteristics*	Lethal, organ damage, multi-causal, chronic, rare, first-in-class (see text)	Indication

has multiple causes (*multiple causes)*,[11] and (e) whether the disease targeted is rare (*rare*).[12] We also use a dummy variable indicating whether the firm already has a product in the same market, as defined by the anatomical therapeutic classification at the 2nd digit (ATC2), to proxy for the firm's downstream assets and experience in that market.[13] These variables are used to control for differences across indications in the likelihood of success, μ, as well as differences in profitability and cost of clinical trials, which condition Pg^*, the selection threshold.

Finally, because pharmaceutical and biotech firms may choose projects with different levels of innovativeness or risk, we devise a proxy for the level of innovativeness (and therefore of risk) associated with the project. We determine if the compound is first in class or a follower molecule (*first-in-class*).[14] All else being equal, we expect first-in-class compounds to exhibit a lower selection threshold. This is because there are fewer available alternatives, making it easier to improve upon existing remedies. We also expect first-in-class compounds to be drawn from an inferior distribution because there is less information about potential side effects. Unfortunately, in almost 40% of cases it is impossible to classify the molecule under study because its chemical characteristics are not available. To properly distinguish first-in-class compounds from followers, we create an indicator variable, *first-in-class unknown,* for when this information is missing. Through the use of these fine-grained controls at the level of indication, we can avoid using indication-fixed effects, which can bias nonlinear estimates, a possibility confirmed by (unreported) Monte Carlo simulations.

Economic Characteristics of Therapeutic Area: We use information on the worldwide sales (at the ATC2 level) in the year the project is started as a proxy for V.[15] We measure the degree of competition by the number of firms operating in that indication (worldwide) at the time the project is launched. We also employ the share of established pharmaceutical firms as an additional control for the intensity of competition. We use 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*). These are hypothesized to affect only Pg^*, because they affect the profitability of a project rather than its likelihood of success.

Specification and Identification of Pg^* and μ: The system of equations we estimate is nonlinear. We

[11] *Multiple causes* equals 1 if the etiology is multifactorial and equals 0 if it is unknown or single factor.

[12] An orphan or rare disease affects fewer than 200,000 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.

[13] We also tried the total number of molecules in the same ATC2 with no change in the results.

[14] We use the chemical name of the projects and its CAS (Chemical Abstract Service) registry number (when available) to distinguish me-too and second-generation molecules from first-in-class compounds (see Reddy 2003).

[15] We use 1983 figures for projects started in 1980–1982 as well, because data for those years are missing.

Table 3 Descriptive Statistics

Variable	Mean	SD	Min	Max	Variable	Mean	SD	Min	Max
Selection	0.48	0.50	0	1	*Closeness to science*	0.05	0.16	0	1
Success	0.34	0.48	0	1	*Public firm*	0.84	0.37	0	1
Pioneer Biotech	0.31	0.46	0	1	*Market size*	2.38	2.47	0	14.40
Established Pharma	0.46	0.50	0	1	*Competitors* (log)	3.09	1.02	0	5.23
Other Biotech	0.17	0.37	0	1	*Share of established pharma in market*	0.28	0.18	0	1
Other Pharma	0.06	0.24	0	1	*First-in-class*	0.30	0.46	0	1
Scale_project	2.67	3.53	0	31	*First-in-class unknown*	0.39	0.49	0	1
Scale_firm	36.23	46.19	0	586	*Lethal*	0.75	0.43	0	1
Scope	0.83	0.21	0	0.98	*Organ damage*	0.78	0.42	0	1
Presence in ATC2 product market	0.18	0.39	0	1	*Multiple causes*	0.80	0.40	0	1
License (*preclinical only*)	0.08	0.27	0	1	*Chronic*	0.79	0.40	0	1
					Rare	0.04	0.20	0	1

also impose three exclusion restrictions. We assume that the economic characteristics of the compound—i.e., the size of the market, the level of competition in the market, and whether the firm is private or public—only affect the selection threshold and not the distribution of the probability of success. This is how we identify selection (Pg^*) as opposed to performance (μ). Second, recall that we normalize the variance of $\ln(Pg/(1-Pg))$ to unity. We do allow for interdependence across the observations for a firm by clustering standard errors at the firm level.

Table 3 presents the descriptive statistics. It shows that about half the projects were selected for clinical trial, and of those, roughly one third were successful. Reflecting the dominance of established firms (see Table 1), Table 3 shows that, on average, a project was associated with over 2.5 projects in the same indication, and 8% of projects were licensed in preclinicals.

6. Empirical Results

Table 4 presents simple probit estimates of the selection and success equations estimated separately. Compared to established pharmaceutical firms, biotech firms have a lower probability of selection and a lower probability of success, though the difference is not statistically significant. Firm scale increases the probability of success (though it leaves the selection probability unchanged), whereas program scale decreases both selection and success probability. Scope reduces selection but increases success. Projects for more profitable markets (larger size and lower competition) have higher probability of selection and success. The presence of downstream assets (a product already in the market) makes selection more likely but reduces success.

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 also depend on the selection threshold, and vice versa. For instance, is scope associated with higher success because it is associated with lower selection probability? If so, should not program scale (associated with lower success probability) have been associated with greater selection probability? The results in Table 4 are difficult to reconcile with a simple Heckman-selection model, in which the same underlying process drives both selection and success. In other words, we need to estimate the structural parameters of the full model described in §3, where we jointly model selection decisions and performance. Table 5 reports the results of pseudomaximum likelihood obtained using STATA. Note that Table 4 reports the drivers of the probability of selection and success, and Table 5 reports the drivers of the selection threshold and the innovative performance. For instance, the coefficients reported under probability of selection in Table 4 are $\beta - \gamma$, whereas Table 5 reports β (see §3). We estimate two specifications. In one, we use firm-type dummies (*Established Pharma* is the reference group) and disease controls only, whereas in the second we also include measures of scale and scope, and firm and market controls.

6.1. Selection Results

We find that research scale at the program level reduces the selection threshold consistent with Hypothesis 1. However, firm scale and scope increase the selection threshold. *Other Biotech* and *Other Pharma* exhibit, respectively, a lower and a higher threshold than the *Established Pharma*. *Other Pharma* may have higher development costs than *Established Pharma*, perhaps because the former lack the strong links with reputed academic centers enjoyed by established pharmaceutical firms such as Lilly and Merck; they may also have limited production and marketing assets. By contrast, there is little difference in the selection process between *Pioneer Biotech* and *Established Pharma*. However, the coefficient of the *Other*

Table 4 Probit Estimates: Factors Affecting Selection (Pr Selection), and Factors Affecting Success for Selected Projects (Pr Success)

	Model 1		Model 2	
	Pr selection	Pr success	Pr selection	Pr success
Pioneer Biotech	−0.31** (0.11)	−0.22 (0.15)	−0.37** (0.12)	−0.08 (0.17)
Other Biotech	−0.47** (0.13)	−0.79** (0.36)	−0.54** (0.16)	−0.63 (0.41)
Other Pharma	−0.18 (0.17)	0.53** (0.20)	−0.33 (0.21)	0.78** (0.26)
License (preclinical)	0.03 (0.15)	0.14 (0.28)	−0.06 (0.14)	−0.02 (0.27)
Scale_program			−0.09** (0.02)	−0.14** (0.03)
Scale_firm			−0.1E−4 (0.1E−2)	0.01** (0.2E−2)
Scope			−0.47** (0.19)	0.54† (0.32)
Closeness to science			0.59** (0.25)	0.4E−2 (0.23)
Presence in ATC2 class			0.29** (0.12)	−0.30** (0.12)
Public			0.17 (0.11)	0.04 (0.20)
Competitors (log)			−0.14** (0.04)	−0.15** (0.07)
% *Established pharma competitors*			0.36 (0.22)	0.18 (0.24)
Market size			0.02 (0.02)	0.10** (0.02)
Lethal	0.07 (0.09)	−0.13 (0.14)	−0.02 (0.09)	−0.21 (0.15)
Organ damage	−0.17** (0.08)	0.13 (0.09)	−0.15† (0.08)	0.09 (0.11)
Multiple causes	−0.12 (0.08)	−0.20† (0.11)	−0.02 (0.08)	−0.13 (0.11)
Chronic	0.02 (0.07)	−0.13 (0.15)	0.3E−2 (0.07)	−0.04 (0.15)
Rare	−0.12 (0.15)	0.73** (0.20)	−0.11 (0.14)	0.58** (0.17)
First-in-class	−0.36** (0.10)	−0.20 (0.15)	−0.35** (0.09)	−0.17 (0.15)
First-in-class unknown	−1.47** (0.09)	−1.54** (0.24)	−1.39** (0.09)	−1.59** (0.20)
Constant	0.97** (0.14)	0.09 (0.21)	1.69** (0.24)	−0.24 (0.42)
No. of observations	3,311	1,088	3,311	1,088
Log-likelihood	−1,849.98	−616.08	−1,752.67	−578.12

Note. Robust standard errors are in parentheses (clustered by firms).
**$p < 5\%$; †$p < 10\%$.

Biotech dummy is sizable (ranging from −1.25 to −1.15), implying that younger biotech firms have lower selection thresholds than older ones, consistent with theories about greater risk profiles, agency problems, and weaker management models. Nonetheless, *p*-values of the tests comparing *Pioneer Biotech* and *Other Biotech* provide only marginal support to Hypothesis 2.[16]

Among the measures of profitability, the number of competitors increases Pg^*, implying lower profitability. Similarly, previous commercial experience in the ATC2 class reduces Pg^*, which suggests higher economic value from successful compounds in markets where the firm has commercialization capability. Market size also increases profitability, but the estimated coefficient, though negative, is not statistically significant.

6.2. Performance Results

Table 5 shows that program scale reduces μ but firm scale increases it. The effect of scope is positive, but imprecisely measured. Thus, Hypothesis 3 is only partially supported. If we believe that there are significant program-level economies of scale, then our measure is confounding economies of scale with the portfolio effect. Consistently, we find that if a firm has a product in the market, μ is lower, which is surprising if one believes that past experience should improve innovative performance. The alternative explanation appeals to the portfolio effect. To ensure that its downstream assets are fully utilized, a firm with downstream assets in a market will develop many compounds in clinical trials for the therapeutic target (e.g., Higgins and Rodriguez 2006, Dahan and Mendelson 2001), resulting in lower estimated Pg^*. However, the firm will commercialize at most one of the compounds, even if more could gain FDA approval. The higher "failure" when there are more projects under development for the same indication reflects such portfolio effects.

Table 5 shows that both *Pioneer Biotech* and *Other Biotech* have lower μ, implying that they have lower innovative performance compared to *Established Pharma*. Formally, Hypothesis 4 is rejected in favor of the alternative Hypothesis 4′. Table 5 also shows that *Pioneer Biotech* perform better than *Other Biotech*, even after controlling for scale, scope, and downstream assets.[17] There are three interpretations, which are not mutually exclusive. The first is that

[16] We tested the null hypothesis that the two selection thresholds are equal, against the alternative hypothesis that $Pg^*(\textit{Other Biotech}) < Pg^*(\textit{Pioneer Biotech})$. The *z*-statistics for the test equal 1.33 in Model 1 and 1.34 in Model 2. In both cases, the *p*-value is approximately 0.09.

[17] The *p*-values of the test statistic for a one tail test is 0.060 in Model 1 and 0.057 in Model 2.

Table 5 Structural Model, Maximum Likelihood Estimates

	Model 1				Model 2			
	Pg^*		μ		Pg^*		μ	
Pioneer Biotech	−0.26	(0.32)	−0.56†	(0.31)	−0.16	(0.31)	−0.55†	(0.29)
Other Biotech	−1.25†	(0.71)	−1.72**	(0.70)	−1.15	(0.73)	−1.71**	(0.72)
Other Pharma	1.04**	(0.40)	0.86**	(0.42)	1.37**	(0.28)	1.02**	(0.36)
License (in preclinical)	0.24	(0.41)	0.27	(0.45)	0.13	(0.38)	0.08	(0.42)
Scale_program					−0.20**	(0.04)	−0.29**	(0.04)
Scale_firm					0.01**	(0.6E−4)	0.01**	(0.1E−2)
Scope					1.11**	(0.53)	0.62	(0.54)
Closeness to science					−0.20	(0.26)	0.39	(0.30)
Presence in ATC2 class					−0.71**	(0.21)	−0.42**	(0.18)
Public					−0.17	(0.11)		
Competitors (log)					0.13**	(0.04)		
% *Established pharma competitors*					−0.30	(0.21)		
Market size					−0.01	(0.02)		
Lethal	−0.26	(0.25)	−0.19	(0.24)	−0.04	(0.20)	−0.04	(0.21)
Organ damage	0.31†	(0.17)	0.14	(0.17)	0.38†	(0.20)	0.23	(0.18)
Multiple causes	−0.30	(0.22)	−0.42**	(0.18)	−0.30†	(0.18)	−0.32**	(0.15)
Chronic	−0.23	(0.27)	−0.20	(0.28)	0.05	(0.23)	0.07	(0.24)
Rare	1.34**	(0.30)	1.22**	(0.33)	1.06**	(0.22)	0.94**	(0.26)
First-in-class	−0.19	(0.29)	−0.54†	(0.28)	−0.20	(0.28)	−0.54**	(0.26)
First-in-class unknown	−2.32**	(0.54)	−3.79**	(0.58)	−2.39**	(0.51)	−3.78**	(0.54)
Constant	−1.07**	(0.38)	−0.10	(0.37)	−2.51**	(0.57)	−0.82	(0.55)
No. of observations	3,311				3,311			
Log-likelihood	−2,466.54				−2,345.18			

Note. Robust standard errors are in parentheses (clustered by firms).
**$p < 5\%$; †$p < 10\%$.

biotech firms are in fact less innovative than established pharmaceutical firms. The second is that biotech firms advance into clinical trials less promising compounds because their managers have perverse incentives, or because they are incapable of correctly evaluating the potential of the compound, or because they are willing to take more risk. The third interpretation is that the regulatory hurdles for the average biotechnology product are more difficult to meet than those for the average pharmaceutical product, and our controls do not adequately control for these differences.

6.3. Licensing Results

Recall that the lemons hypothesis is based on the seller systematically offering a lower-quality good, keeping the superior-quality good for internal use. This implies that licensed compounds should have a lower μ than compounds developed in-house by the licensor. Table 5 shows that we do not observe any market for lemons. Licensed compounds face a higher selection threshold, Pg^*, as might be expected if the licensee has to share revenues (but not costs) with the licensor. However, they also have a higher μ, although neither coefficient is statistically significant. Thus, Hypothesis 5 is not supported. The countervailing effects discussed in §3 appear strong enough to avoid a lemons problem.

7. Discussion and Conclusions

A key takeaway from our analysis is that in comparing the innovation performance of firms, one must pay careful attention to differences in behavior, and to factors, such as incentives, that may drive such differences. We explicitly model one aspect of behavior, namely, the decision to select a compound into clinical trials. However, there are other aspects of our results that illustrate the same point, such as our finding no program-level economies of scale, or that a firm with past experience in a market (and that, by inference, has downstream assets) is less successful in innovating in the market. These apparently anomalous findings can be reconciled by considering the incentives of firms to ensure the full utilization of downstream assets by selecting a number of compounds into clinical trials, at most one of which is marketed, thereby reducing estimated innovative performance. The problem lies with the measure of program-level economies of scale (commonly used in the literature), which confounds the economies of scale with the portfolio effect.

Our results imply that even after controlling for scale and scope of the firm, there are significant differences across different types of firms. These differences reflect both performance and behavior, most notably, selection. The different results in Tables 4 and 5 suggest that explicitly modeling selection is important. To

see this, compare the results about the firm type dummies for specification 2 in the two tables. In Table 4, the dummies for *Pioneer Biotech* and *Other Biotech* are negative and significant in the selection equation, but in the success equation, the estimates are insignificant. Table 5 shows that they have significantly lower μ but also lower Pg^*. Thus, we conclude that the reason biotech firms are less likely to advance compounds into clinical is because *on average* they have lower-quality compounds compared to average pharmaceutical firms, not because biotech firms are more selective.

7.1. Robustness Checks

We conduct a series of tests to explore the robustness of our results to changes in the sample, alternative measures of market profitability, and alternative assumptions about the distribution of Pg. For brevity, we summarize our results, with details available from the authors upon request. First, our results are largely unchanged if we expand our sample to include projects initiated in Europe, or by universities and other public research organizations. Similarly, defining an established pharmaceutical firm to be one that appears on the list of the top 50 pharmaceutical firms by sales for 3, 5, or 10 years does not materially change our results. Second, although the nonlinear specification of our likelihood function contributes to identification, our results do not depend on it. We also estimate a specification where the log-odds ratio of Pg is uniformly distributed over some subset of the unit interval, with little qualitative change in our results. Finally, our results are robust to some alternative empirical measures. For instance, using the Herfindahl index of market shares in an indication (for the year 2001) to measure competition in the market leads to similar results. Dropping *first-in-class*, for which there are many missing values, does not change our results. Similarly, using sales at the ATC3 level in an indication (albeit for a single year, 2001) to measure market size does not affect our findings.

To validate our estimation results, we also compared our estimated Pg^*, which corresponds to the ratio of development costs and expected revenues, with estimates of costs and revenues provided by the existing literature (DiMasi et al. 2003, OTA 1993, Grabowski et al. 2002). The estimates of Pg^* obtained from these calculations range from 0.07 to 0.19, compared to our estimated value of 0.12, which is very close to the median.[18]

7.2. Further Explorations

Because we estimate 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 5 to compute the probability of selection, the unconditional probability of success, and the value of Pg^* and μ for the firm types in our sample (Table 6).

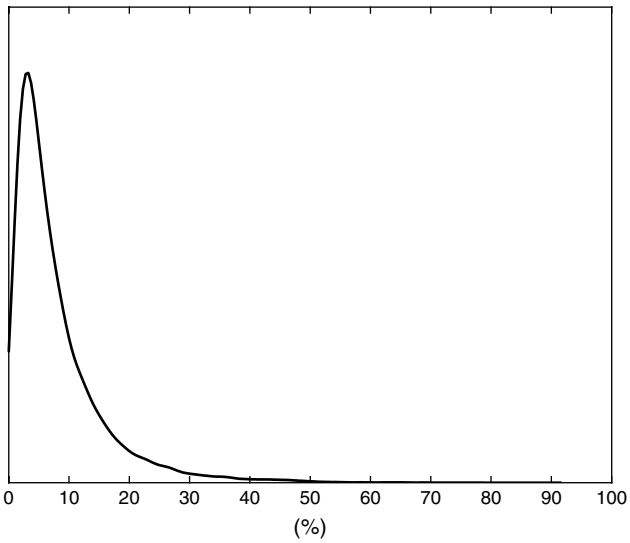
Table 6 **Estimated Selection and Success Probabilities, Selection Threshold (Pg^*), and Average Performance (μ), by Originator Type**

Originator type	Pr selection	Pr success	Pg^*	μ
Overall	0.48	0.25	0.12	−2.87
Pioneer Biotech	0.44	0.19	0.11	−2.97
Established Pharma	0.57	0.27	0.13	−2.35
Other Biotech	0.30	0.07	0.03	−4.79
Other Pharma	0.55	0.43	0.31	−1.04

Figures 1 and 2 show the simulated distributions of Pg for the whole sample and by originator types, using our estimates from specification 2 in Table 5. 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*. Using the estimated coefficients of specification 2 in Table 5, we also perform some simulation exercises. We first study the effects of changes in firm size.

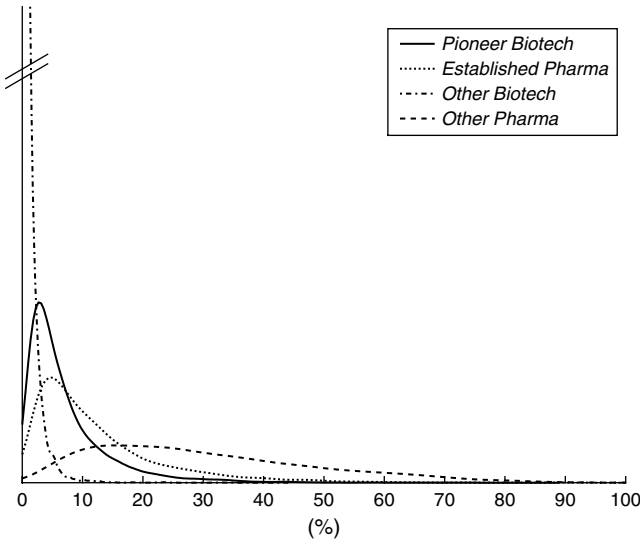
Table 7 reports the estimated effects of a one-standard-deviation increase in *Scale_firm*. When *Scale_firm* increases, Pg^* increases. Because μ also increases, the probability of selection increases too. Interestingly, the effects of the firm research scale are rather small. Thus, although *Scale_firm* is statistically significant in both the selection and performance equations in Table 5, its overall effect is small relative to the

Figure 1 **Estimated Distribution of *Pg***



[18] Details of these calculations are available upon request.

Figure 2 Estimated Distribution of *Pg*, by Originator Type



impact of firm type. As discussed, we believe these firm types reflect different organization structures and strategic behavior. Our empirical results suggest these differences are important for understanding differences in innovation performance.

7.3. Limitations and Future Research

Our results are from the pharmaceutical sector. This sector is extraordinarily research intensive, and its innovation process over the last three decades 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 on only one part of the innovation process, drug development, and arguably the part most favorable 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 discussed the impact of risk aversion on the selection process. However, such differences may also be a source of unobserved heterogeneity in conditioning other outcomes. We assume that market size, measures of competition, the presence of downstream assets, whether or not the molecule is first-in-class, as well as the other disease characteristics such as lethality, acuteness, and multiplicity of causes, together 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, entrants are willing to take bets with smaller probabilities of success but much higher rewards if successful. By contrast, incumbents have more financial reserves, making them reluctant to bet heavily on a 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. Though plausible and consistent with our results, this story is inconsistent with the conventional wisdom, which holds that even the incumbent pharmaceutical firms are focusing on blockbusters. We leave this issue for further research. One implication of these results is that the division of innovative labor (Arora et al. 2001) in the pharmaceutical industry must rely on differences in comparative (rather than absolute) advantage across biotech firms and pharmaceutical firms. Alternatively, the advantage of biotech firms may lie in researching unexplored avenues, rather than in doing better what pharmaceutical firms already do.

Table 7 Simulating an Increase in the Research Scale of the Firm (*Scale_Firm*)

	Pioneer Biotech		Established Pharma		Other Biotech		Other Pharma	
	Baseline	$+1\ \sigma$	Baseline	$+1\ \sigma$	Baseline	$+1\ \sigma$	Baseline	$+1\ \sigma$
Pr selection	0.44	0.44	0.57	0.57	0.30	0.30	0.55	0.55
Pr success	0.15	0.15	0.22	0.23	0.04	0.04	0.38	0.39
*Pg**	0.11	0.11	0.13	0.13	0.03	0.03	0.31	0.31
μ	−2.97	−2.95	−2.35	−2.33	−4.79	−4.77	−1.04	−1.01

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 projects that are outlicensed and those that are retained in-house appear to be drawn from similar distributions. This suggests that endogeneity is unlikely, but we cannot be more definitive at this stage.

An extension of our analysis would be to explicitly model 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, we cluster the errors at the level of the firm, but are unable to cluster at the firm-cum-indication level. As a result, such interactions may be captured through our measures of scale and scope, leading to potentially confounding results. Future research may explicitly model how a firm chooses among the portfolio of compounds in clinical trials for a particular indication, and how this choice varies with whether it has relevant downstream assets, and with potential for product cannibalization.

Despite these limitations, our paper makes two important contributions. It provides a simple model wherein firms choose which compounds to advance into clinical trials, taking into account the technical promise of the compound, as well as economic and commercial considerations. The model also specifies how the observed innovation performance is conditioned by this choice. Our second contribution is to analyze the factors driving this choice, and in particular, how differences in the underlying innovativeness between established firms and entrants, as well as differences across them in incentives to advance compounds to trials, drive both their behavior and their innovation performance.

Acknowledgments

The authors gratefully acknowledge comments and suggestions from the department editor, three anonymous reviewers, and seminar participants in the Sant'Anna School of Advanced Studies, Pisa, Department of Economics at the University of Verona, the Strategic Management Society, Wien, November 2007; however, the authors retain full responsibility for the contents of this paper. Alfonso Gambardella thanks the Italian Ministry of Education (PRIN, 2006–2008), the CNR–ESF Project "Science and Technology in Large and Small Firms," and Bocconi University for financial support.

References

Abrantes-Metz, R., C. Adams, A. Metz. 2006. Pharma development phases: A duration analysis. *J. Pharma Finance, Econom. Policy* **14**(4) 19–41.

Acemoglu, D., J. Linn. 2004. Market size in innovation: Theory and evidence from the pharmaceutical industry. *Quart. J. Econom.* **119**(3) 1049–1090.

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

Adams, C., V. Brantner. 2003. New drug development: Estimating entry from human clinical trials. FTC Bureau of Economics Working Paper 262, Federal Trade Commission, Washington, DC.

Akerlof, G. A. 1970. The market for "lemons": Quality uncertainty and the market mechanism. *Quart. J. Econom.* **84**(3) 488–500.

Arora, A., A. Gambardella. 1994. Evaluating technological information and utilizing it. *J. Econom. Behav. Organ.* **24**(1) 91–114.

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

Arrow, K. J. 1962. Economic welfare and the allocation of resources for invention. R. R. Nelson, ed. *The Rate and Direction of Inventive Activity*. Princeton University Press, Princeton, NJ, 609–624.

Arrow, K. J. 1983. Innovation in large and small firms. J. Ronen. ed. *Entrepreneurship*. Lexington Books, Lexington, MA, 15–28.

Braunwald, E., A. S. Fauci, D. L. Kasper, S. L. Hauser, D. L. Longo, J. L. Jameson. 2001. *Harrison's Principles of Internal Medicine*, 15th ed. McGraw-Hill, New York.

Chandy, R., B. Hopstaken, O. Narasimhan, J. Prabhu. 2006. From invention to innovation: Conversion ability in product development. *J. Marketing Res.* **43**(3) 494–508.

Chesbrough, H. 2003. *Open Innovation*. Harvard Business School Press, Boston.

Chesbrough, H., D. J. Teece. 1996. When is virtual virtuous? Organizing for innovation. *Harvard Bus. Rev.* **74**(1) 65–72.

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

Cohen, W. M., D. Levinthal. 1989. Innovation and learning: The two faces of R&D. *Econom. J.* **99**(397) 569–596.

Dahan, E., H. Mendelson. 2001. An extreme-value model of concept testing. *Management Sci.* **47**(1) 102–116.

Danzon, P. M., S. Nicholson, N. S. Pereira. 2005. Productivity in pharma-biotechnology R&D: The role of experience and alliances. *J. Health Econom.* **24**(2) 317–339.

DiMasi, J. A., R. W. Hansen, H. G. Grabowski. 2003. The price of innovation: New estimates of drug development costs. *J. Health Econom.* **22**(2) 151–186.

Gambardella, A. 1995. *Science and Innovation in the U.S. Pharma Industry*. Cambridge University Press, Cambridge, UK.

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. *Res. Policy* **33**(5) 787–806.

Grabowski, H., J. Vernon, J. DiMasi. 2002. Returns on R&D for 1990s new drug introductions. *Pharmacoeconomics* **20**(Supplement 3) 11–29.

Guedj, I. 2005. Ownership vs. contract: How vertical integration affects investment decisions in pharmaceutical R&D. Working paper, University of Texas, Austin. http://www.aeaweb.org/annual_mtg_papers/2006/0106_0800_0202.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, National Bureau of Economic Research, Cambridge, MA.

Henderson, R., K. Clark. 1990. Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. *Admin. Sci. Quart.* **35**(1) 9–30.

Henderson, R., I. Cockburn. 1996. Scale, scope, and spillovers: The determinants of research productivity in drug discovery. *RAND J. Econom.* **27**(1) 32–59.

Higgins, M. J., D. Rodriguez. 2006. The outsourcing of R&D through acquisition in the pharmaceutical industry. *J. Financial Econom.* **80**(2) 351–383.

Holmstrom, B. 1989. Agency costs and innovation. *J. Econom. Behav. Organ.* **12**(3) 305–327.

Klepper, S., P. Thompson. 2006. Submarkets and the evolution of market structure. *RAND J. Econom.* **37**(4) 861–886.

Levinthal, D., J. March. 1993. The myopia of learning. *Strategic Management J.* **14**(S2) 95–112.

Macher, J. T., C. S. Boerner. 2006. Experience and scale and scope economies: Trade-offs and performance in development. *Strategic Management J.* **27**(9) 845–865.

Nerkar, A. 2003. The folly of rewarding patents and expecting FDA approved drugs. Working paper, Columbia University, New York. http://bschool.washington.edu/departments/mo/PDF/Nerkar_UW.pdf.

Nerkar, A., P. Roberts. 2004. Technological and product-market experience and the success of new product introductions in the pharma industry. *Strategic Management J.* **25**(8–9) 779–799.

Owen-Smith, J., W. Powell. 2006. Accounting for emergence and novelty in Boston and Bay Area biotechnology. P. Brauner-hjelm, M. Feldman, eds. *Cluster Genesis: The Emergence of Technology Clusters and Their Implications for Government Policy*. Cambridge University Press, Cambridge, UK, 61–85.

Pisano, G. 1997. R&D performance, collaborative arrangements, and the market for know-how: A test of the "lemons" hypothesis

in biotechnology. Working paper, Harvard Business School, Cambridge, MA.

Pisano, G. 2006. Can science be a business? Lessons from biotech. *Harvard Bus. Rev.* **84**(10) 114–124.

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)? *Res. Policy* **19**(2) 165–174.

Shaver, J. M. 1998. Accounting for endogeneity when assessing strategy performance: Does entry mode choice affect FDI survival? *Management Sci.* **44**(4) 571–585.

Stiglitz, J. E., A. Weiss. 1981. Credit rationing in markets with imperfect information. *Amer. Econom. Rev.* **71**(3) 393–410.

Thompke, S., W. Kuemmerle. 2002. Asset accumulation, interdependence and technological change: Evidence from pharmaceutical drug discovery. *Strategic Management J.* **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.

Zeckhauser, R. 1996. The challenge of contracting for technological information. *Proc. Natl. Acad. Sci. USA* **93**(23) 12743–12748.