SCALE, SCOPE AND KNOWLEDGE SPILLOVERS IN DRUG DEVELOPMENT: DO THEY MATTER FOR DRUG DEVELOPMENT SUCCESS OF BIOTECHNOLOGY FIRMS?

Zaur Rzakhanov

Suffolk University
Department of Economics
8 Ashburton Place
Boston MA, 02108-2770
USA
Tel: (617) 305-1768
Fax: (617) 994 - 4216
zaur@beaconhill.org

This version: June 2004*

Very preliminary draft, please do not quote.

* I would like to thank Darlene Chisholm, Jonathan Haughton, Inn-Mee Baek and Irina Tytell for useful comments and suggestions.

Abstract

The purpose of this paper is to gauge the degree to which firm structure plays a role in successful product development in an innovation-driven industry. Do firm scale, product variety and technological scope affect product development success rate and if so, through which channels? Pharmaceutical and biotechnology industries are naturally attractive subjects of such an inquiry. While basic facts about drug development success rate are well-known, its determinants of are poorly understood. Using comprehensive database of product development outcomes in the biotechnology industry the paper investigates the role of these factors in determining drug development success rate. The results indicate that some elements of structure of a biotechnology firm are correlated with probability of drug development success.

A number of empirical regularities are established. Measures of firm and project scale, choice of therapeutic and technological scope appear to be correlated with probability of drug development success. Evidence suggests that there is an inverse relationship between probability of development success and scale of drug development program. Another finding indicates that wider technological and therapeutic scope of drug development program contributes to higher probability of project development success. Experience with a chemical compound also matters: both successful and unsuccessful experience with a compound in one project affects the probability of success of other projects that use the same compound. Finally, research experience has weak adverse effect on the probability of drug development success.

1. INTRODUCTION

The objective of the study is to identify elements of firm structure that are conducive to successful product development in an innovation-driven industry. Structural characteristics of an innovation-driven firm such as scale and scope of its product development pipeline, choice and scope of technologies may influence product development success rate and development cost. Biotechnology and pharmaceutical industries provide natural context for such inquiry for a number of reasons. Both industries represent one of the leading R&D intensive industries in advanced economies. Developing successful and profitable medicine is exceptionally costly and hard, and while basic facts about drug development cost and success rate are well-known, the determinants of both are poorly understood. It takes about fifteen years of research and development to bring a drug to the market (Spilker, 1999) and roughly $600 million to come up with a single successful drug (Grabowski, Vernon and DiMasi 2002). Such arguably high cost of drug development is in part due to high failure rate of drug development projects. DiMasi (1995) estimates that only about 1 in 4 drugs make through the human clinical trials, yet all project, both eventual successes and failures require constant flow of cash adding to the final cost of successful projects. Despite recent advances in genomics, ethical drug development continues to be largely a trial-and-error process that relies on knowledge, experience and intuition of scientists (Robbins-Roth, 2000). The question arises as to what organizational and technological environment makes drug development less hard, more cost effective and eventually more profitable? Could one firm structure or configuration be more favorable to success of drug development than others? The answers to these questions have obvious implications for

management and design of innovation process in biotechnology and pharmaceutical industries.

The choice of firm structure should matter from a firm's point of view, as it may eventually affect profitability. A company must decide on the scale of its research and development program: how much to spend on drug development and how many drugs to work on. Larger scale of development efforts may enable a firm to spread its fixed cost over the greater amount of potential sales (Chandler, 1990), and additionally may increase the probability of product development success if scale enables more efficient drug development techniques. Firms must also decide on the therapeutic scope of its development effort: how many different diseases to work on. A firm could specialize in, say, cancer drugs or develop drugs in a wide variety of therapeutic areas. Choice of therapeutic scope is likely to be closely related to the characteristics of product market competition, desire to diversify sources of revenue and to the possible tradeoff between cost and success in drug development.

Likewise, a biotechnology firm must decide how many chemical compounds (which could be interpreted as technologies) to utilize in development effort. Should two different, but to an extent related, medical conditions be treated with two different chemicals, or should a firm leverage a single compound and use it treat both medical conditions. Such decision is the decision about technological scope of development program. The ability to use the same tangible or intangible asset for multiple purposes at little or no additional cost is important to achieve higher productivity of a given amount of effort (Panzar and Willig, 1981). In this setting, profitability could be enhanced as a single technology (chemical compound) can be "leveraged" or be used to develop a number of different products (medical treatments). The choice of technological scope

depends on cost and benefits of scope, specialization and knowledge spillovers that exist in pharmaceutical product development.

The direction of long-run adjustments in firm structure and the optimal structure will depend, among other things, on the marginal effect that changes in firm structure have on product development cost and on probability of development success (and thus on expected revenue). Modeling such adjustments in the context of dynamic R&D competition is beyond the scope of this paper, which is decidedly modest. The paper focuses more narrowly on empirical relationship between probability of development success and firm structure, rather than on the dynamic choice of firm structure.

The study describes various aspects of biotechnology firm structure and attempts to uncover the elements of firm structure that are correlated with probability of drug development success. Following previous studies, this relationship is investigated within an econometric setting that is based on estimation of probability of success function that takes as covariates various characteristics pertaining to firm structure such as firm scale, therapeutic or product scope, technological scope, measures of intrafirm knowledge spillovers and research experience. The main conclusion is that some aspects of firm structure such as the employment per project, technological and therapeutic scope, technological knowledge spillovers and development scale, appear to be relevant to drug development success. However, results also seem to indicate that choice of therapeutic field, knowledge spillovers across therapeutic areas and scale of firm's research activity have small or no impact on probability of development success.

The next section provides a review of relevant literature. The section 3 describes research hypotheses, while section 4 addresses measurement issues. Subsequently, data issues (section 5), descriptive statistics (section 6) and estimation approach (section 7) is

discussed. The final sections report and discuss results (section 8) and offer conclusions (section 9).

2. LITERATURE REVIEW

There has been a considerable amount of economic research on economies of scale in drug research or discovery[1], yet rather few researchers addressed such issues in the context of drug development. The two recent and noteworthy studies by Cockburn and Henderson (2001) and by Danzon, Nicholson and Pereira (2003) address productivity effects of firm scale, therapeutic scope and knowledge spillovers in drug development.

The first study by Cockburn and Henderson investigates the determinants of drug development success of large pharmaceutical firms. Working with comprehensive longitudinal data on 10 large pharmaceutical companies, authors investigate the impact of firm scale and scope on probability of success of individual drug development projects, using logit model in a panel data setting. Defining success as project's entry into regulatory stage (filing new drug application [NDA] with U.S. Food and Drug Administration), they conclude that:

"…In contrast to previous work on the discovery phase of pharmaceutical R&D we find strong correlation between the diversity of firms' development efforts and the success probability of individual projects, but no effect of scale per se. Large firms' superior performance in drug development appears to be driven by returns to scope rather than returns to scale. Scope is confounded with firm fixed effects, however, suggesting an important role for inter-firm differences in the organization and management of the development function" (Cockburn and Henderson, 2001).

[1] Concise review is provided by Cockburn and Henderson (2001).

This conclusion underscores the importance of firm-specific structural factors. Current paper expands the number and detail of explanatory variables and considers not only development scale and therapeutic scope, but also other possible determinants of drug development success: technological scope, scale of drug discovery activity, as well as intrafirm knowledge spillovers in drug development. All those factors could be responsible for differential success record across biotechnology firms. Considering such elements of firm structure explicitly could yield additional insights into firm-level determinants of development success.

The second study of interest is by Danzon, Nicholson and Pereira (2003). Using data on both pharmaceutical and biotechnology firms the authors investigate the effect of scale (total experience in their terminology), therapeutic scope and therapeutic-specific experience on success probabilities for different drug development stages. These authors also adopt the logit model. Evidence presented suggests that there are positive effect of total experience for late stage clinical trials, as well as positive effect of therapy-specific experience and negative effect of wider therapeutic scope on the likelihood of successful completion of Phase 2 clinical trials.

Both studies suggest that some decisions about firm structure matter for successful drug development. Current paper builds upon these inquiries by systematically investigating the effects of development scale, technological and therapeutic dimensions of firm scope, intrafirm knowledge spillovers and research activity on drug development success in the biotechnology industry. This set of firm characteristics is more encompassing measure of firm structure than the explanatory variables used in the previous studies. While Danzon, Nicholson and Pereira investigate the probability of success separately for different human clinical trial stages (Phase 1, 2 and 3), this paper

focuses on the probability of bringing a project beyond the clinical trial stage (into regulatory stage). The difference in approach reflects the desire to evaluate the effect of firm structure on eventual technical success of a project and to generate results comparable to results obtained by Cockburn and Henderson (2001).

3. RESEARCH HYPOTHESES

The hypotheses presented below focus on the determinants of success of a drug development project which is defined in this study as a unique chemical compound – disease pair. This definition of a development project is warranted by the U.S. Food and Drug Administration's requirement of separate set of human clinical trials and separate regulatory review for each such compound-disease pair. For instance, a firm that uses a single chemical to treat two different diseases would be required to conduct two separate sets of clinical trials and two separate regulatory reviews for each compound-disease pair. Naturally, the notion of development success is clearly defined when such definition of development project is adopted. Hypotheses 1 through 5 presented below provide conjectures about the effects of firm structure on the probability of development success. Hypotheses 1 through 3 parallel, to some extent, Cockburn and Henderson (2001) hypotheses concerning scale and scope effects on productivity of big pharmaceutical companies; here, however, the focus is on biotechnology companies. Hypothesis 4 considers possible channels for knowledge spillovers across drug development projects, while the fifth hypothesis addresses the effect of cumulative discovery effort on probability of development success.

H.1: Ceteris paribus, larger scale of development effort has no effect on probability of development success.

Cockburn and Henderson (2001) found that there is no statistically significant partial effect of scale on drug development success, and here their finding is adopted as a null hypothesis. It is possible that zero effect found is simply a net effect of forces working in opposite directions, rather than a phenomenon reflecting genuine absence of scale effects. Beneficial effect of scale on the probability of success may work through learning-by-doing. A firm with twenty projects in development may know more about efficient and successful management of drug development projects than a firm with only two projects in the pipeline. Other things equal, firms with more projects in development could be expected to accumulate more knowledge about the science and art of pharmaceutical product development, by adapting more efficient techniques and learning from its more numerous past successes and failures. Additionally, larger scale may also lead to productivity-enhancing specialization. One, but by no means the only possible source of diseconomies of scale is managerial input. While supervisory and monitoring requirements for increasing number of projects could be met by hiring more managers or outside contractors (such as contract research organizations) coordination and agency issues (see Azolay(2003)) may become more important with increasing size. If such problems become severe enough they could outweigh benefits of scale in drug development that accrue through learning-by-doing and specialization.

H.2: Ceteris paribus, firm's technological specialization (narrower technological scope) lowers probability of project's success.

There are gains and cost to leveraging a single technology across different applications. The benefits of narrower technological scope may arise from cost sharing. Identifying the best compound to treat disease X is major and costly task of pre-clinical research. An alternative and cost-saving strategy is to use the compound that has already

shown promise in treatment of another and possibly related disease Y. When a single chemical compound is used to develop drugs to treat two or more different diseases, more products share the costs involved in compound discovery. Furthermore, passing safety clinical trials would indicate that a certain dose of chemical compound is likely to be safe no matter which disease is being treated. Hence, once safety is established for a compound treating disease X, a firm may not need conduct as large and as extensive a safety trial to treat disease Y or other diseases. Confirmation of safety for disease X, could therefore save some safety trial costs for other diseases.

The tradeoff in this decision is between saving on the discovery and development cost and finding the compound that is the best match to the disease in question. Using existing compound to treat another disease may lower discovery expenses, but at the same time may compromise the quality of match between a compound and target disease in terms of lower probability of development success. The second and costlier strategy is to search for the new compound from scratch, spending funds on identifying a new compound that is the best match to the disease in question. The implication for drug development outcome is that narrower technological scope compromises the quality of compound-disease match, resulting in a higher probability of development failure.

H.3: Ceteris paribus, firm's therapeutic specialization (narrower therapeutic scope) lowers probability of project's success.

Specialization in a narrow set of therapeutic areas may enable a firm to utilize similar clinical trial design across a number of projects, with likely savings in development cost. Yet such approach may also be detrimental to project success rate if such specialization limits managerial and scientific experience of a company. Managers and scientists in companies that are exposed to wide variety of therapeutic areas may

have better exposure to different clinical trials settings and challenges posed by drug development and hence may adapt better and faster in response to unfavorable trial results, by introducing appropriate changes in clinical trials design and conduct. Consequently, such exposure may lead to higher development success rate.

H.4: Development efforts are subject to the intrafirm knowledge spillovers. A chemical compound that succeeds in clinical trials for one medical indication is more likely to succeed in the clinical trial for another medical indication[2]. Similarly, success of a project in one therapeutic field is likely to increase the probability of success of other projects in firm's pipeline in the same therapeutic field.

Toxicity of a compound, when found in the clinical trial for one medical indication, may be serious enough to terminate other clinical trials that involve the same chemical compound. This type of technological spillover may be important rationale for avoiding technological specialization, despite its possible cost advantages. Correspondingly, within a firm a success (failure) of a project in one therapeutic area, may to some extent predict success (failure) of another project in the same therapeutic area. This correlation is likely to reflect acquired experience within a therapeutic area. For example, successful development of one anti-infective drug may provide necessary knowledge to develop successfully another anti-infective drug.

H.5: The size of firm's research program has no effect on the probability of development success.

The fifth hypothesis reflects an agnostic prior. For financially constrained firms discovery activities would tend to compete for resources with development activities. In this case, given the pool of resources, larger size of discovery program could harm

[2] Same applies to failure. A failure of a compound in one project could make more likely failure of other projects that use that compound.

development efforts. Conceivably, cohabitation of drug discovery and development activity within a firm could also have beneficial impact on development success, if tacit knowledge transmission is important. To an extent that knowledge about beneficial properties of discovered compound cannot be completely codified, all else equal, development of a drug discovered in-house may have higher success rate than a compound in-licensed from outside. In such a case the size of research (or discovery) program may have beneficial impact on the probability of drug development success. An implicit, and possibly questionable, assumption behind this beneficial effect is that transmission of knowledge within a biotechnology firm is easier than across firm boundaries. A priori it is not clear which effect dominates and thus the statement claiming no effect of research program size is adopted as the null hypothesis.

4. MEASUREMENT

4.1 Product development scale

Development scale could be measured in a variety of ways. One choice is to use aggregate R&D expenditures or employment to measure the scale of development effort, which is a common practice in R&D productivity literature. Alternatively, scale could be measured along the extensive and intensive margins. Extensive margin reflects the total size of development effort and here is measured by the total number of projects a firm was engaged in from inception through 2000. These include active drug development projects, projects terminated prior to 2000, as well as the projects under regulatory review and on the market. This variable is intended to measure an overall experience of a firm in drug development. The intensive margin reflects the scale of effort at the level of individual project. Ideally, project-specific R&D expenditures or employment should be

used to capture project-level scale. Lack of project level cost and employment data forces me to define employment per project variable as the ratio of firm employment[3] to the total number of projects.

4.2 Technological and therapeutic scope: concept and measurement

Studies reviewed in section 2, focused on a dimension of firm scope that can be termed "therapeutic scope" of firm's drug development program. To illustrate the concept consider a firm with three development projects to treat three related diseases (say ovarian, breast and colon cancers). This firm is said to be narrow in its therapeutic scope, while a firm with three projects for three generally unrelated diseases (say AIDS, cystic fibrosis and Crohn's disease) has product development program of wide therapeutic scope. These two companies with the same number of development projects (or development scale) have very different therapeutic scope. The second dimension of scope could be called "technological scope". Again, consider a firm that decides to develop drugs to treat three different diseases. It can make a choice between using one, two, three or possibly more distinct chemical compounds to treat these diseases. If a firm chooses just one compound to treat all three diseases, its technological scope is narrow. On the other hand if a firm chooses to use a different chemical compounds for each disease (three different compounds in total) its technological scope is wider relative to the first case when only one compound is utilized.

Although therapeutic and technological scope could be correlated, a clear distinction in concept and measurement can be made between firm's therapeutic and technological scope. A firm with a given number of projects in development and a given therapeutic scope can have different technological scope depending on the number of

[3] Compustat data used in this paper do not differentiate between research and development employment or R&D expenditures

chemical compounds it decides to use in drug development. While therapeutic scope captures the breadth of firm's product market position, technological scope indicates firm's propensity to "leverage" its technology across different development projects. Therefore, the scope of drug development program can be viewed as having these two dimensions. As an example consider the following actual data

Table A: Therapeutic and technological scope of a firm, an example

Firm: Matrix Pharmaceuticals

Project#	Compound/Drug	Disease	Disease Class code	Project Phase
1	IntraDose	malignant melanoma	2	Lead
2	MPI 5017	bladder cancer	2	Pre-Clinical
3	MPI 5019	cancer	2	Pre-Clinical
4	IntraDose	primary liver cancer	2	Phase 2
5	IntraDose	metastatic liver cancer	2	Phase 2
6	IntraDose	prostate cancer	2	Phase 2
7	FMdC	colon cancer	2	Phase 2
8	IntraDose	head and neck cancer, solid tumors	2	Phase 3
9	IntraDose	recurrent or metastatic breast cancer	2	Phase 3
10	MPI 5020	breast tumor reccurence	2	Terminated
11	AccuSite	basal cell cancer	2	Terminated
12	AccuSite	squamous cell cancer	2	Terminated
13	AccuSite	psoriasis	12	Terminated
14	AccuSite	genital warts	1	Terminated

Table A lists all development projects that this firm was ever involved in. There are fourteen projects in development. Each project corresponds to a unique disease-compound combination because of FDA requirement to have separate set of clinical trials for each such combination. International Classification of Diseases, 9th edition (ICD-9, see Table B) was used to group all target diseases for this firm into three separate disease classes[4]: Neoplasms (2), Infections (1) and Diseases of the skin and subcutaneous tissue (12). There are also six distinct chemical compounds/drugs that are used to treat different diseases (IntraDose, MPI 5017, 5019, 5020, FMdC and AccuSite).

[4] Multum's Lexicon provides a table matching a list of roughly 38,000 diseases to a therapeutic class defined by ICD-9.

One way to measure therapeutic scope is to count the number of distinct disease classes. For Matrix Pharmaceuticals the number of distinct disease classes is three. Similarly, one can measure technological scope by the count of distinct chemical compounds/drugs in development. In this case there are six distinct chemical compounds. Furthermore, we can measure the scale of development effort, which is reflected in the total count of development projects (fourteen projects in total). Clearly, the firm's configuration we observe: therapeutic scope = 3, technological scope = 6 and scale = 14 is just one possibility. For a given number of projects this firm could have chosen a different technological and therapeutic scope. For instance a configuration of therapeutic scope = 14, technological scope = 14 and scale = 14 represents a firm with widest possible technological and therapeutic scope given the number of projects in development.

Table B: International Classification of Diseases, 9th edition

Disease/Therapeutic class	Class code
Infectious and Parasitic Diseases	1
Neoplasms	2
Endocrine, Nutritional, Metabolic Diseases, and Immunity disorders	3
Diseases of the blood and blood-forming organs	4
Mental disorders	5
Diseases of the nervous system and sense organs	6
Diseases of the circulatory system	7
Diseases of the respiratory system	8
Diseases of the digestive system	9
Diseases of the genitourinary system	10
Complications of pregnancy, childbirth and puerperium	11
Diseases of the skin and subcutaneous tissue	12
Diseases of the musculoskeletal system and connective tissue	13
Congenital anomalies	14
Certain conditions originating in perinatal period	15
Symptoms, signs and ill-defined conditions	16
Injury and Poisoning	17

4.3 Capturing the effects of knowledge spillovers

Success or failure of a chemical compound in one project may serve as a good predictor of outcomes of other projects that use the same chemical compound. As an example consider chemical toxicity. Toxicity of a chemical compound X that leads to failure of one project, may not affect other projects if for these projects the dose could be lowered without adverse therapeutic effects. However, if concerns about toxicity were serious enough, all projects that use compound X would be abandoned. Thus, learning about toxicity of a chemical in one development project is likely to affect the outcomes of other projects that use the same chemical. Conversely, successful development of a project involving chemical X, may increase the chance of successful progress of other projects that use the same chemical, for the knowledge gained in one such project could at little cost be shared across all chemical X projects in a firm's development pipeline.

Similar effects may exist within therapeutic areas. The likelihood of a successful project completion for a drug treating an infectious disease may be enhanced if there were other active projects to develop anti-infective therapies. Knowledge gained in the process of developing one anti-infective therapy is often useful in development of another anti-infective drug. Failures may also be informative. Inability to develop drugs in one therapeutic class may be a signal to firm's management that other projects within the same therapeutic class may now be less promising due to technical or other difficulties, and hence are more likely to be abandoned.

Dummy variables are used to test for the presence of these knowledge spillover channels. Suppose that a project *m* in question uses chemical *k* and treats a disease that falls into therapeutic class *t*, then the set of knowledge spillover variables consists of four dummy variables. The first variable $COF_{jm} = 1$ if there is at least one other *failed* project

in firm's j pipeline that used chemical *k* to treat a different disease, and $COF_{jm} = 0$ if otherwise. Also $COA_{jm} = 1$ if there is at least one other *active*[5] project in firm's j pipeline that uses chemical *k* to treat a different disease, $COA_{jm} = 0$ if otherwise. Observe that COF and COA are not mutually exclusive. A firm j can concurrently have other failed and active projects that use chemical *k* .

Similar dummies are defined for therapeutic classes. $TOF_{jm} = 1$ if there is at least one other *failed* project in firm's j pipeline that is in the same therapeutic class as given project *m*, and zero otherwise. $TOA_{jm} = 1$ if there is at least one *active* project in firm's j pipeline that is in the same therapeutic class as given project *m*, and zero otherwise. When all four dummies are equal to zero, the project *m* under consideration is a "stand-alone" project - that is a firm has no other active or failed projects in firm's pipeline that use the same chemical or are in the same therapeutic class as project *m*.

As an illustration consider Matrix Pharmaceuticals' project #6 (Table A). Spillover dummy variables for this project take the following values [COA, COF, TOA, TOF] = [1,0,1,1]. Dummy variable COA = 1, as there are other active projects that use compound IntraDose and COF = 0 since there are no failed projects using IntraDose. Variable TOA = 1, because there other active projects that are in Neoplasms therapeutic class. Finally, TOF = 1, as there are failed projects that belong to Neoplasms therapeutic/disease class. Values of spillover dummies for other projects in the sample are determined in a similar fashion.

[5] An "active" project is defined as a project that is still in development or on the market.

5. DATA ISSUES

The data on drug development outcomes are assembled from the data made freely available by Recombinant Capital (ReCap) on its website[6] prior to October 2000. ReCap is a San Francisco-based consulting firm specializing in biotechnology alliances and capitalization. Its Clinical Trials Progress Database tracks the progression of more than 900 compounds in or near clinical development for which a biotechnology firm is involved in such compound's development and/or commercialization. The data covers the period between 1980 and 2000. The database provides the information on a company's drug pipeline, including the number of drugs in different phases of development; dates of entry into and exit from a particular phase of clinical trials; the number and development dates for projects whose development has been terminated; data on marketing partners for each drug development project.

The Recap database encompasses 292 existing and former public as well as private biotechnology companies involved in drug discovery and development. These companies encompass the entire spectrum of biotechnology firms: from boutique-scale companies with less than a dozen employees to biotechnology giants such as Biogen and Genentech. Matching Recap data with Compustat data on public companies yielded a sample of 220 public companies. Inspection of Compustat data uncovered 436 publicly traded U.S. companies that were involved in drug and diagnostics development prior to 1999 (excluding big pharmaceutical companies) of which 408 companies were listed in SIC 2833 to 2836 category[7] and 28 firms had 5122 as their primary SIC code. If these 436 companies represent the population of publicly traded U.S. biotechnology

[6] www.recap.com

[7] SIC 2833 (medicinal chemicals), SIC 2834 (pharmaceutical preparations), SIC 2835 (In vitro/in vivo diagnostics), SIC 2836 (biological products), SIC 5122 (drugs & proprietary)

companies, then 220 companies in the sample account for 50.5% of all public biotechnology companies in the United States.

In principle Recap database could be used to construct longitudinal dataset of drug development outcomes. However, since many development dates are missing it is impractical for this purpose unless one is willing to make imputations for project development dates. Despite this problem with missing dates, the eventual status and phase of the projects (as of 2000) was tracked fairly accurately. Tracking of product development outcomes by Recap for a random sub-sample of 30 companies was compared with the information available from company websites and with *Lexis-Nexis Medical* news indicated only a few discrepancies.

Compustat database was used to obtain measures of employment and other firm specific financial information for public biotechnology companies. Stock of patents for each firm was calculated using Hall, Jaffe and Trajtenberg (2001) patent data set. The result is reduction in sample size to 220 companies that went public prior to 1999 for which information on employment and patents was available.

The data sets present a number of potential measurement problems. Compustat database provides total R&D expenditures and firm employment without distinguishing between research and development stages. As a result, the effect of development scale on development success, if it exists, could be contaminated by likely correlation with research scale. If research scale also affects the chance of development success, the estimate of development scale effect will be biased. To fix this problem (and for other reasons) patent stock variable is included in productivity regressions, in effect mitigating omitted variable bias by picking up partial effect of the research scale.

The second potential problem concerns censored observations in the Recap database. Left censoring does not present a problem, because virtually all biotechnology companies were formed in the 1980's and 1990's[8]. Since the dataset starts in 1980 it captures end points of virtually all development projects that began in the 1980's and late 1970's. Right censoring is most probable for projects commenced in the late 1990's, as the end year of the dataset is 2000. Most of such projects are still in development and hence are neither terminated nor successfully brought to the market. This may present a problem as older and newer firms of comparable size may have different age structure of the development pipeline. For instance, a firm with 20 projects ever undertaken that has formed in the 1980's is likely to have more marketed products that otherwise similar firm that was born in the mid-90's. In this case while the scale of development effort may have the same effect on success for both firms, for the newcomer firm this effect will be censored. To control for this censoring effect entry date into Phase 1 clinical trial is used as another explanatory variable and a set of regressions is run on various sub-samples to elicit the effect of right censoring on parameter estimates.

6. DESCRIPTIVE STATISTICS

Table 2 (see appendix) provides a glance at the data. As of year 2000, an average biotechnology firm in the sample has been involved in 10.0 development projects. This number conceals considerable variation across firms. By year 2000, old and experienced firms have been involved in 40 to 60 projects, while the startups are usually involved in no more than a few projects. Another interesting observation concerns firm choice of technological and therapeutic scope. An average firm with 10.0 projects in development

[8] Genentech along with three others companis had an IPO in 1980. Only Biogen, Genentech and Genome Therapeutics incorporated prior to 1980 (Robbins-Roth, 2000).

is involved with roughly 6 distinct chemical compounds, and in 4 distinct therapeutic areas (Table 2A indicates that DC = 5.9 and TS = 4.3). Thus, there is some evidence of multiple applications of a chemical compounds in drug development: on average firms utilize a single compound to treat more than one medical condition. The same data also suggests that, on average, biotechnology firms have about two projects per therapeutic area.

Figures 1 and 2 provide information about the distribution of entry dates into Phase 1 clinical trials and initial public offering dates for biotechnology firms in the sample. Both clearly show biotechnology boom of early and mid-1990's, a time of entry for many new biotechnology firms. Correspondingly, we see many firms bringing new projects into human clinical trials around this time.

The success rate for drug development projects that entered Phase 1 of human clinical trials prior to 1990 is 25.7% (see Table 2C) and it is about equal to the success rate (23.5%) reported by DiMasi(1995) for the new chemical entities that entered clinical trials in early and mid-1980's. The number of projects still in development in 2000 rises sharply to 51.6% for projects that entered Phase 1 between 1990 and 1994, reflecting the effect of censoring. For the same cohort of projects the success rate is already about 25.5%, only slightly lower than for projects that entered Phase 1 prior to 1990. This could possibly reflect higher quality of 1990- 1994 cohort of projects. For the projects entering Phase 1 after 1994, percentage of projects still in development in 2000 is the highest - 85.1%, reflecting the proximity of censoring date (year 2000) to the Phase 1 date of entry.

As of 1999, an average biotechnology firm possessed 27.1 patents (Table 2A), yet only 10.0 projects were ever undertaken by 2000. If each patent reflects discovery of a new chemical entity, then at most 1 in 3 patented chemical compounds entered product

development and possibly even fewer considering that a compound could be involved in multiple drug development projects. The above observation is consistent with a notion that patented compounds are perceived to be different in value prior to development stage, and that only most ex-ante promising compounds enter product development (Hall, Jaffe and Trajtenberg, 2000).

There are comparatively few biotechnology firms that have good record of success in drug development and those firms that do succeed consistently have superior access to capital, large employee base, extensive laboratory and equipment capital. At the same time most biotechnology firms are quite small firms characterized by relatively few projects in development, small number of employees, as well as relatively small R&D expenditures and cash funds (Robbins-Roth, 2000). This configuration is a well-known feature of biotechnology industry and is clearly reflected in the summary statistics presented in Table 2A. For all firm characteristics the mean values considerably exceed their respective median values and the distributions of these characteristics in the sample are highly skewed and exhibit large kurtosis.

Lastly, the mean values of the spillover dummy variables COA, COF, TOA and TOF indicate that there is a potential for intrafirm knowledge spillovers across projects and within therapeutic/disease areas (Table 2B). The mean value of COA is 0.393 and mean value of COF is 0.106 indicate that for 39.3% of projects in the sample there is at least one other active project within a firm that utilizes the same chemical compound and for 10.6% of projects in the sample there is at least one other failed project within a firm that utilized the same chemical compound. Similarly, TOA = 0.66 and TOF = 0.21 indicate that for 66% of projects in the sample there is at least one other active project

within a firm in the same therapeutic class and for 21% of projects in the sample there is at least one other failed project within a firm in the same therapeutic class.

7. EMPIRICAL MODEL

The empirical model is based on a single estimating equation that relates a measure of project's success to various variables that describe firm structure and project-specific characteristics. The unit of observation is project which is a patented chemical compound designated to treat a particular disease (medical indication). A single chemical compound could be used to treat several diseases and a disease could be treated by a number of different compounds. Each chemical compound-disease pair (here called project) is required by U.S. Food and Drug Administration to have separate clinical and pre-clinical development, as well as regulatory review. For example, a firm using one chemical compound to treat two different medical conditions would be required to conduct two separate sets of trials. Thus, the development pipeline of such a firm will consist of two projects. In this setting, the probability of development success of a project *m* belonging to a firm j becomes,

$$\Pr(SUCCESS_{jm}) = \beta_0 + \beta_1 \cdot EMPP_j + \beta_2 \cdot EMPP_j^2 + \beta_3 \cdot TOTPROJ_j + \beta_4 \cdot TOTPROJ_j^2 +$$

$$+ \beta_5 \cdot PATSTOCK_j + \beta_6 \cdot DC_j + \beta_7 \cdot TS_j + \beta_8 \cdot COA_{jm} + \beta_9 \cdot COF_{jm} +$$

$$+ \beta_{10} \cdot TOA_{jm} + \beta_{11} \cdot TOF_{jm} + \beta_{12} \cdot Controls_{jm} + \sum_{TC=1}^{17} \lambda_{TC} + \theta_j + \varepsilon_{jm}$$

The dependent variable ($SUCCESS_{jm}$) is a dummy variable that is equal to 1 if project *m* belonging to firm j reaches an NDA filing stage (i.e. a firm files an NDA for a project by year 2000) and equal to zero otherwise. The independent variables are firm

level and project level variables. Firm level variables include employment per project (EMPP) and EMPP squared; the total number of development projects undertaken between 1980 and 2000 (TOTPROJ) and total number of projects squared; patent stock of firm j (PATSTOCK); measure of technological scope (DC_j) and a measure of therapeutic scope (TS_j). Intrafirm knowledge spillovers are captured by the set of project level spillover dummy variables (COA, COF, TOA and TOF). Seventeen therapeutic class dummies are included for two reasons. One is to control unobserved heterogeneity in development difficulty as drugs in some therapeutic classes may be harder to develop than in others. Another reason for inclusion of therapeutic dummies is to alleviate potential omitted variable bias. Firms with wide therapeutic scope may also be the ones that select therapeutic specialties where developing drugs is relatively easy. In such a case positive effect of wider therapeutic scope on the probability of success would reflect unmeasured difficulty of therapeutic areas. Finally θ_j is firm effect, and ε_{jm} is a project specific disturbance term. Table 1 provides detailed definitions of variables. The logit and random effects logit models are estimated by maximum likelihood. For the logit model standard errors are computed correcting for possible heteroskedasticity (White or robust standard errors) and clustering on firm[9].

Partial effect of scale will show up in $\beta_1, \beta_2, \beta_3$ and β_4. Effects of technological and therapeutic scope will be captured by β_6 and β_7, respectively. If there are intrafirm spillover effects $\beta_9 < 0$, that is the failure of other projects that use the same chemical *k* decreases the probability of success for project *m*. Correspondingly, $\beta_8 > 0$, would indicate that presence of other active projects in firm's pipeline that use the same chemical *k* increases the chance of success for project *m*. Similar interpretation pertains to

[9] Estimation is implemented in Stata 8.2 using logit and xtlogit commands.

the signs of therapeutic coefficients: $\beta_{12} < 0$ implies smaller chance of success if at least one other project within the same therapeutic class has failed, and $\beta_{11} > 0$ implies greater chance of success if at least one other project within the same therapeutic class is active. Control variables include a set of dummies reflecting entry date into Phase 1 clinical trials. This is intended to account for different age structure of projects across firms. Finally, coefficient on PATSTOCK (β_5) would capture the effect of firm's cumulative *research* effort on the probability of development success.

8. RESULTS AND DISCUSSION

8.1 Main specifications

The results confirm that scale and scope of firm's development effort along with intrafirm knowledge flows play important role in determining the probability of success of a project. The following points are notable; they indicate partial effects of various variables on the probability of success of a development project,

(1) Greater employment per project has positive but weakly diminishing effect on the probability of success.

(2) Greater total number of projects has *negative, but diminishing in absolute value* effect on the probability of success.

(3) Greater patent stock (which is a measure of firm's research output) has weak negative or no effect (in one specification) on the probability of development success.

(4) Wider technological and therapeutic scope increases probability of project's success. However, in some specifications the effect of therapeutic scope is statistically insignificant.

(5) Therapeutic area of a project *does not affect* the probability of successful completion of a project.

(6) Technological knowledge spillovers across projects within a firm do exist and have substantial and statistically significant effect on the probability of success, indicating that both positive and negative development experience with same chemical compound in other projects affect the probability of success of a given project. The evidence of knowledge spillovers via therapeutic channel is less decisive.

The positive effect of greater scale on probability of project's success seems to work to large extent through an "intensive margin": greater effort per project (measured as employment per project) leads to higher probability of success. However, along the "extensive margin" the relationship between the total number of projects and probability of success is negative. Having one more project in the pipeline decreases the probability of success (Table 2, coefficient on TOTPROJ) even after holding employment per project, scope, spillover and research effort measures constant.

All else equal, it appears that involvement in the greater number of projects is associated with lower probability of success. Negative effect is the strongest for small firms, remaining negative but declining in magnitude with greater firm size. This inverse dependence may reflect unequal relationship between learning-by-doing and specialization effects favoring positive effect of firm size (that may start playing some role for relatively large and experienced firms) and effects that favor negative effect of firm size, such as managerial coordination problems, lack of focus and agency problems that firms may face when they decide to expand the size of product development program. One simple explanation is that monitoring and agency problems may compound

with development program size, while benefits of acquired practical knowledge and input specialization may not be strong enough to counteract such forces.

Confirming predictions of hypotheses 2 and 3, wider technological and therapeutic scope improves chances of project's success. Consistent with the finding of Cockburn and Henderson (2001), experience in a greater variety of therapeutic areas appears to have beneficial effect on the probability of success. The positive effect is present in all specifications, however it is statistically significant only in random effects model. A number of interpretations are possible. Better exposure of management and scientists to varied challenges posed by drug development in different therapeutic areas may improve the quality of response to unfavorable trial results. Conversely, firms developing drugs in only one therapeutic area may not have accumulated necessary knowledge and skill to deal with product development setbacks. It is also possible that estimated effect of therapeutic scope may be due to omitted ability bias. Firms that are involved in a variety of therapeutic areas may be the once that employ more talented top management team able to work on projects in disparate, yet profitable therapeutic areas and at the same time manage to develop drugs that are more likely to succeed. Whatever the ultimate explanation, firms with wider therapeutic scope have on average a better chance to develop a successful drug.

Wider technological scope has a similar positive and statistically significant effect on probability of success (see coefficient on DC variable in Table 3). The chance of multiple project failure is lessened if more chemical compounds are used to treat a given number of diseases. Employing a single chemical compound with proven safety record to treat multiple diseases may reduce development cost, but may also increase the risk of failure in efficacy trials. One possible explanation is that this leverage strategy

compromises the quality of compound-disease match with a resulting negative effect on probability of success. The alternative strategy of looking for the best compound from scratch may improve the probability of development success but also likely to be more costly as each new compound would require separate discovery expenses and probably larger safety trials. Clearly, biotechnology firms have to consider such detrimental effects on probability of success if they decide to adopt narrower technological scope as a cost saving strategy.

Joint significance test on therapeutic dummies indicates that once we account for the effect of structural characteristics of a firm, the variation in development difficulty across therapeutic areas does not affect the probability of successful completion of a project. This result is consistent with Cockburn and Henderson (2001) finding that therapeutic dummies have insignificant joint effect on probability of success for large pharmaceutical companies. Inspection of individual statistical significance of therapeutic dummies reveals that with exception of two therapeutic areas, the differences between areas are statistically insignificant.

Technological spillovers have substantial and statistically significant effect on the probability of success, confirming the first part of conjecture outlined in hypothesis 4. Working and succeeding with a chemical compound in treatment of one disease, substantially increases the probability that the same compound will be proven to be successful in treatment of another disease. Negative experience also matters: failure of a chemical compound to pass a clinical trial tends to spill over onto trial outcomes for other diseases that use the same chemical compound. Marginal impacts of COA and COF dummies in Table 3 on probability of success indicate that both positive and negative experience with a compound in other projects has about the same effect (only with

opposite sign) on the probability of development success. Consequently, the knowledge gained in one project seems to be a good predictor of an outcome of other development projects that utilize the same chemical compound.

Knowledge spillovers within therapeutic areas appear to have smaller impact on the probability of success than technological spillovers (see marginal impacts of TOA and TOF in Table 3), and insignificant in random effects model. Having another active project in the same therapeutic class, increases the probability of success of a given project, while having another failed project in the same therapeutic class has no effect on the probability of success. This result only partially supports the evidence presented by Danzon, Nicholson and Pereira (2003) suggesting positive effect of therapy-specific experience. It could still be that the therapeutic area definitions used in the present study are simply too inclusive to detect strong spillover effects within therapeutic areas that would have been possibly observed over the narrower individual disease channels.

Another point worth emphasizing is that cumulative research experience (measured by patent stock) has weak negative effect on the probability of development success, once development scale, scope and spillovers are taken into account. The coefficient estimate on PATSTOCK variable is consistently negative, but significant only in some specifications. Relocating a development project from a firm with substantial research experience to an otherwise similar firm with less research experience, positively affects the probability of project's success. As a practical matter this result seems to suggest that specializing in product development by licensing compounds from other companies may have some beneficial effect on the probability of project's development success.

8.2 Robustness

An important concern for the stability of the estimates is right censoring. Two firms with the same measures of scale, scope and other characteristics may, nonetheless, have different age structure of their development portfolios. An older of two firms may naturally have older projects, some of which could have resulted in NDA by year 2000; an otherwise similar younger firm may have no successful projects by 2000, even though both firms could be equally productive. Hence, if right censoring is present, lack of success may reflect a project's young age rather than a genuine failure. Main specifications reported in Table 3 address this issue by introducing dummy variables reflecting the date of entry into Phase 1 clinical trials. An alternative approach is to restrict the sample to observations that are less likely to be right censored. Table 4 reports the estimates for the sub-sample of projects that have had entered Phase 1 prior to 1990 (Eq.4) and for a larger sub-sample of projects that entered Phase 1 prior to 1994 (Eq.5, 6 and 7). Qualitatively, with exception of jointly significant therapeutic dummies and wrong and statistically significant sign on TOF dummy in Eq.4, almost all results obtained for the entire sample of projects (Table 3) also remain true for these sub-samples, suggesting little relevance of right censoring for the signs of estimates of the logit model. Interestingly, the quantitative effects are the strongest for the sub-sample that includes only the projects that entered Phase 1 prior to 1990 (compare Eq.4 to Eq.5 in Table 4, and to Eq.1 in Table 3).

Another issue is possible survivorship bias. If the sample of firms includes only existing public companies, they may represent competent survivors. In this case, negative effect of scale on development success could be an artifact of survival or "startup fratricide" (Robins-Roth, 2000), because new but unsuccessful companies that have

ceased to exist are not in the sample, while successful ones are. As more companies went public in the 1990's than in the 1980's, more intense competition could have selected more efficient winners in the 1990's than in the 1980's. Therefore, even if there is no scale effect it would appear that firm size is inversely related to probability of success, as recent surviving startups are simply more efficient than older and larger firms.

The likelihood of survivor effect warrants a closer look at the data. The match between Compustat, Recap and patent databases yielded 220 companies that had an initial public offering between 1980 and 2000. Of this number, 33 companies or 15% were no longer public in 2000 (Compustat lists them in the research database). These 33 companies were involved in 250 development projects out of 2208 projects in the sample, or 11.3% of the total. Four of the 33 companies went public prior to 1990, and all four were involved in more than 10 projects prior to year 2000 (an average for these companies was 15.8 projects). Another 14 companies of 33 had an IPO after 1994. Finally, 25 of these 33 companies were acquired prior to 2000, while the fate of the other 8 is unclear. Including a dummy for these companies in the logit regressions (not reported) produced negative but statistically insignificant effect on the probability of success. It appears that the survivor effect is likely to be small, if not absent altogether.

A number of auxiliary regressions were run to determine whether estimates are sensitive to the choice of specification. Including aggregate firm-level R&D expenditures instead of the number of projects and employment per projects as an independent variable produces small positive but statistically insignificant coefficient on R&D input. This change in specification had no effect on the sign and statistical significance of scope and spillover measures. This result is consistent with Cockburn and Henderson (2001) finding of no effect of firm scale, which they also measure as total firm development spending.

Inclusion of an interaction effect between technological and therapeutic scope produced no statistically significant effect on the probability of development success.

Sensitivity of results to definition of variables was also tested. Redefining total number of projects to be projects that are currently in development (rather than the number of projects ever developed) had no substantial effect on sign or statistical significance of any of the independent variables, with exception of PATSTOCK. The negative coefficient on patent stock increased substantially in absolute value, remaining negative but becoming statistically significant at 1% level.

9. CONCLUSIONS

Evidence from biotechnology industry suggests that some key structural aspects of innovation-intensive firms such as firm scale, effort per project, therapeutic and technological scope, along with technological knowledge spillovers appear to be correlated with product development success. However, results also seem to indicate that choice of therapeutic field, knowledge spillovers within therapeutic areas and size of research activity have no or little impact on probability of drug development success.

An important empirical question for future inquiry is how the choice of firm structure affects cost of clinical development. Effects of firm structure on R&D cost and probability of success, along with a firm's competitive environment may determine the choice of scale and scope in both the short and long run. This brings to the forefront deeper theoretical and modeling questions about privately and socially optimal structure of innovation-driven firms and more generally about the dynamic nature of competition in drug development.

References

Azoulay, P. (2003). A Portfolio Approach to Transaction-Costs Economics: Evidence from Clinical Development. Columbia University.

Chandler, A. (1990). *Scale and Scope*. Cambridge, MA: The MIT Press.

Cockburn, I. and Henderson, R. (2001). Scale and scope in drug development: unpacking the advantages of size in pharmaceutical research. *Journal of Health Economics*, 20, 1033-1057.

Danzon P. M., Nicholson S. and Pereira N. S. (2003). Productivity in Pharmaceutical-Biotechnology R&D: The Role of Experience and Alliances, NBER Working Paper No. 9615.

DiMasi, J.A. (1995). Success rates for new drugs entering clinical testing in the United States. *Clinical Pharmacology & Therapeutics*, **58**(1), 1-13.

Grabowsky, H., Vernon, J. and DiMasi, J.A. (2002). Returns on R&D for 1990s new drug introductions. *Pharmacoeconomics*, 20(3), 11-29.

Hall, B.H., Jaffe, A.B., and Trajtenberg, M. (2000). Market Value and Patent Citations: A First Look, NBER Working Paper No. 7741.

Hall, B.H., Jaffe, A.B., and Trajtenberg, M. (2001). The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools, NBER Working Paper No. 8498.

Panzar, J., and Willig, R. (1981). Economies of scope. *American Economic Review*, **71** (2), 268-272.

Robbins-Roth, C. (2000). *From Alchemy to IPO: The Business of Biotechnology*. Cambridge, MA: Perseus Publishing.

Spilker, B. (1999). The Drug Development and Approval Process. In Pharmaceutical Research and Manufacturers of America (Ed.) (1999). *New Drug Approvals in 1998*, Washington D.C.

Appendix

Table 1: Description of variables

Variable	Measure of…	Definition
SUCCESS	development success	Dummy: equals 1-if a project has filed new drug application (NDA) by year 2000, and 0 otherwise
TOTPROJ	firm scale/total experience	Number of projects that were ever in development (including failed ones and on the market) between 1980 and 2000
EMPP	Development project scale	Average number of employees per project (1980- 2000)
DC	Technological scope	Number of distinct chemical compounds ever used in development portfolio between 1980 and 2000
TS	Therapeutic scope	Number of distinct therapeutic areas ever involved in between 1980 and 2000
PATSTOCK	Research Scale	Number of patent applications filed by 1999
COA	Positive technological spillover across projects	Dummy: COA = 1 if there is at least one other *active* project in firm's j pipeline that uses the same chemical k to treat a different disease, and COA = 0 if otherwise
COF	Negative technological spillover across projects	Dummy: COF = 1 if there is at least one other *failed* project in firm's j pipeline that used the same chemical k to treat a different disease, COF = 0 if otherwise
TOA	Positive spillover within therapeutic area	Dummy: TOA = 1 if there is at least one other *active* project in firm's j pipeline that is in the same therapeutic class, and 0 otherwise
TOF	Negative spillover within therapeutic area	Dummy: TOF = 1 if there is at least one other *failed* project in firm's j pipeline that is in the same therapeutic class, and 0 otherwise

Table 2: Summary statistics

(A) Firm level – 220 biotechnology firms, year 2000.

Variable	Mean	Std	Min	Max	Median	Skewness	Kurtosis
TOTPROJ	10.0	8.9	1	60	5	2.5	11.5
EMPP	43.6	292.0	0.33	4850	11.9	16.1	265.1
DC	5.9	5.7	1	40	4	2.3	10.5
TS	4.3	3.2	1	15	3	1.2	4.0
PATSTOCK	27.1	60.0	1	601	11	6.7	57.7

(B) Project level – 2208 drug development projects, year 2000

Variable	Mean	Std	Min	Max
Success dummy	0.128	0.35	0	1
COA	0.393	0.49	0	1
COF	0.106	0.31	0	1
TOA	0.660	0.47	0	1
TOF	0.210	0.41	0	1

(C) Project outcomes

Project status	Sample: Projects that entered Phase 1 before 1990	Sample: Projects that entered Phase 1 between 1990 and 1994	Sample: Projects that entered Phase 1 after 1994	Sample: All Projects
Success (Filed NDA by 2000)	25.7%	25.5%	8.6%	12.8%
Still in clinical development in 2000	12.6%	51.6%	85.1%	72.9%
Failure	61.7%	22.9%	6.3%	14.3%

Figure 1: Frequency distribution of project entry dates into Phase 1 clinical trial

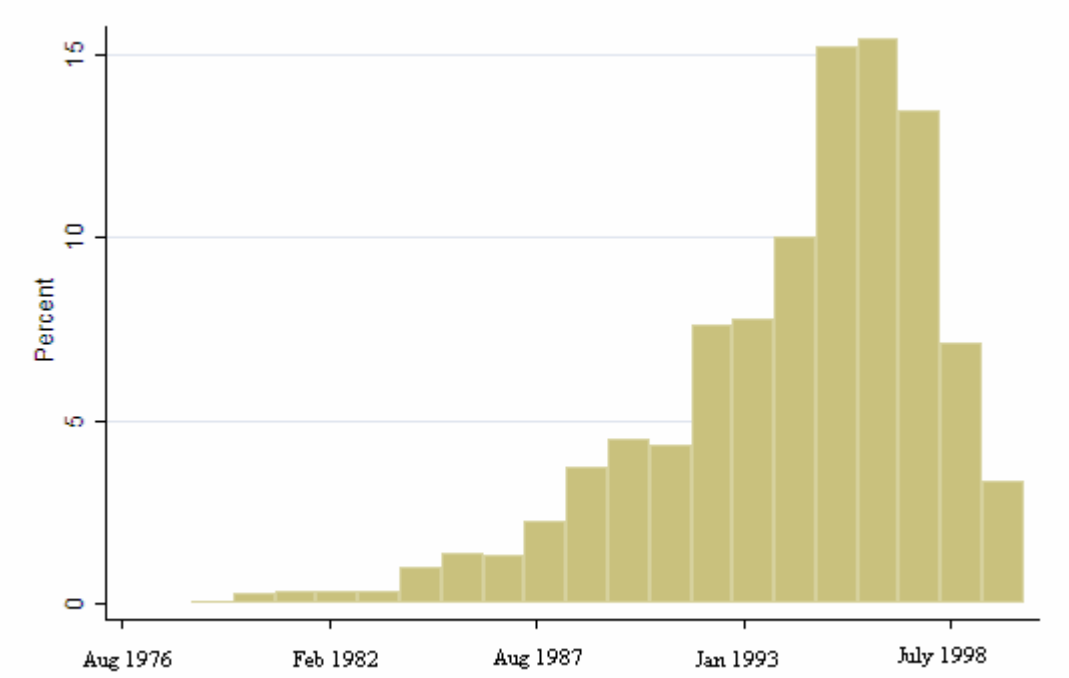

Figure 2: Frequency distribution of initial public offering dates

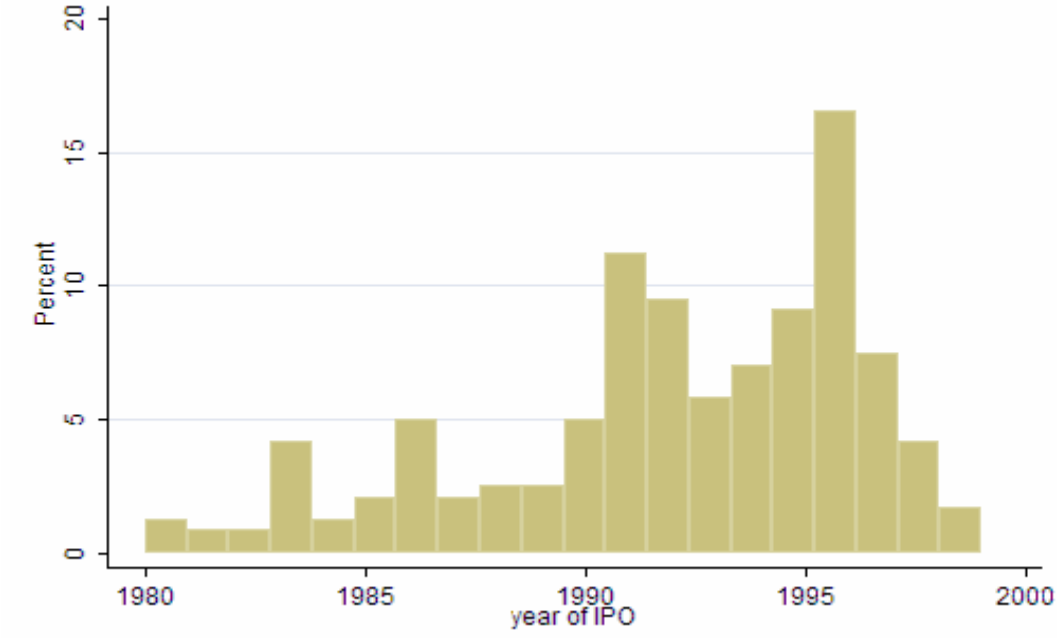

Table 3: Probability of development success regressions, Year 2000.

Unit of observation: project (chemical compound-medical indication pair), N = 2208 (entire sample)
Dependent variable: Dummy = 1 if NDA filed by year 2000, 0 - otherwise

Independent variables	Eq.1 Logit	Eq.2 Logit	Logit marginal impact	Eq.3 Random effects (RE) logit
Employment per project (EMPP)	0.0072*** (.0022)	0.0075*** (.0023)	0.00052*** (.00019)	0.0073** (.0037)
EMPP squared	--	-1.31×10^{-6}*** (4.87×10^{-7})	-9.0×10^{-8}** (3.73×10^{-8})	-1.22×10^{-6} (1.35×10^{-6})
TOTPROJ (Number of projects)	-0.187*** (.060)	-0.186*** (.060)	-0.013*** (.004)	-0.228*** (.060)
TOTPROJ squared	0.0015** (.0007)	0.0015** (.0007)	0.00011** (.00004)	0.0021** (.0008)
DC (technological scope)	0.118*** (.041)	0.117*** (.041)	0.0081** (.0032)	0.147*** (.044)
TS (therapeutic scope)	0.114 (.083)	0.114 (.083)	0.0078 (.0056)	0.126* (.072)
PATENT STOCK	-0.0019* (.0011)	-0.0019* (.0011)	-0.00013* (.00008)	-0.0028 (.0025)
Entered Phase 1 before 1990 dummy	1.93*** (.25)	1.93*** (.25)	0.253*** (.053)	1.98*** (.23)
Entered Phase 1 after 1994 dummy	-2.15*** (.27)	-2.15*** (.27)	-0.208*** (.028)	-2.45*** (.21)
COA	0.72*** (.20)	0.72*** (.20)	0.053*** (.016)	0.73*** (.18)
COF	-1.46*** (.26)	-1.46*** (.26)	-0.065*** (.012)	-1.48*** (.30)
TOA	0.44** (.22)	0.44** (.22)	0.028** (.013)	0.36 (.23)
TOF	0.28 (.33)	0.28 (.33)	0.020 (.026)	0.17 (.22)
Therapeutic area dummies	Jointly insignif. (p-value = 0.21)	Jointly insignif. (p-value = 0.21)	--	Jointly insignif.(p-value = 0.71)
% correctly predicted	89.3	89.3	--	89.0
Log-Likelihood	-618.1	-618.1	--	-603.0

Notes: (1) For logit regressions (eqs.1,2) standard errors in parenthesis corrected for heteroskedasticity and clustering on firm, (2) ***, **, * - significant at 1, 5 and 10 percent level, respectively, (3) Omitted category – entered Phase 1 in 1990 – 1994. (4) Constants are not reported, (5) marginal impacts calculated for eq. 2 and evaluated at sample means.

Table 4: Effects of right censoring

Dependent variable: Dummy = 1 if NDA filed by year 2000, 0 - otherwise

Independent variables	Eq.4 Sample: entered Phase 1 before 1990 only (Logit)	Eq.5 Sample: Entered Phase 1 before 1994 only (Logit)	Eq.6 Sample: Entered Phase 1 before 1994 only (Logit)	Eq.7 Sample: Entered Phase 1 before 1994 only (RE Logit)
Employment per project (EMPP)	0.012* (.007)	0.0049* (.0025)	0.0050* (.0026)	0.0049 (.0031)
EMPP squared	--	--	-6.54×10^{-7} (6.60×10^{-7})	--
TOTPROJ	-0.529*** (.151)	-0.229*** (.059)	-0.229*** (.059)	-0.229*** (.050)
TOTPROJ squared	0.006*** (.001)	0.0025*** (.0007)	0.0025*** (.0007)	0.0024*** (.0006)
DC (technological scope)	0.085* (.048)	0.069** (.034)	0.069** (.034)	0.069** (.029)
TS (therapeutic scope)	0.621** (.248)	0.178*** (.067)	0.178*** (.067)	0.178*** (.062)
PATENT STOCK	-0.0053** (.0023)	-0.0017 (.0011)	-0.0017 (.0011)	-0.0017 (.0015)
Entered Phase 1 before 1990 dummy	--	1.93*** (.26)	1.93*** (.26)	1.93*** (.21)
Entered Phase 1 after 1994	--	--	--	--
COA	1.84*** (.51)	0.87*** (.23)	0.87*** (.23)	0.87*** (.20)
COF	-1.50*** (.52)	-1.37*** (.27)	-1.37*** (.27)	-1.37*** (.29)
TOA	0.35 (.67)	0.23 (.28)	0.23 (.28)	0.23 (.26)
TOF	1.02** (.50)	-0.003 (.30)	-0.003 (.30)	-0.003 (.23)
Therapeutic area dummies	Jointly signif. (p-value = 0.001)	Jointly insignif. (p-value = 0.45)	Jointly insignif. (p-value = 0.45)	Jointly insignif. (p-value = 0.58)
% correctly predicted	73.6	77.4	77.4	77.4
Log-Likelihood	-95.6	-386.1	-386.1	-386.1

Notes: (1) For logit regressions (eqs.4,5 and 6) standard errors in parenthesis corrected for heteroskedasticity and clustering on firm (2) ***, **, * - significant at 1, 5 and 10 percent level, respectively, (3) Sample size: Eq. 4: N = 201 , Eqs.5, 6 and 7: N =774 , (4) Constants are not reported.