

Ownership vs. Contract: How Vertical Integration Affects Investment Decisions in Pharmaceutical R&D

Ilan Guedj[§]

The University of Texas at Austin

October 2005

Abstract

This paper explores the effect of vertical integration on investment behavior in the pharmaceutical industry. I study a detailed, project-level sample of 4057 drug candidates that were sponsored by 40 large pharmaceutical firms during the period 1984-2001. Of these projects, 447 were conducted through a contractual alliance with another, smaller company that had discovered the drug candidate. The existence of these two types of governance structures allows me to compare integrated and non-integrated projects within the same firm. Controlling for project and firm characteristics, I document that pharmaceutical firms are more selective in continuing their integrated projects than in continuing projects governed by contract. I hypothesize that this difference is caused by the rigidity of the contract that governs non-integrated projects, making them less flexible in adapting to changes in the firm's situation. In line with this hypothesis, I document that although more frequently continued, non-integrated projects have a lower probability of success. Moreover, investment in non-integrated projects is less sensitive to the firm's cash flow and to the existence of other projects in the firm.

[‡]I am especially grateful to David Scharfstein for his continuous support and guidance during this project. I would like to thank Dirk Jenter, Stewart Myers, Steve Ross, and Antoinette Schoar for their comments and advice. This paper also benefited from discussions with Pierre Azoulay, George Baker, Nittai Bergman, Robert Gibbons, Jiro Kondo, Josh Lerner, Oguzhan Ozbas, Jannette Papastaikoudi, and seminar participants at Chicago, Columbia, Cornell, Dartmouth, Harvard, Notre Dame, MIT, UCLA, USC, Wharton, and the universities of Washington at Seattle and Texas at Austin. Financial support from the Program on the Pharmaceutical Industry at MIT is gratefully acknowledged.

[§]Department of Finance, McCombs School of Business, University of Texas at Austin, Austin, TX 78712 (Phone: 512-471-5781; e-mail: guedj@mail.utexas.edu)

1 Introduction

The costs and benefits of vertical integration have been the focus of much research since Coase (1937) first raised the issue. Most of this research has been theoretical in nature¹. There is a smaller empirical literature – summarized in Whinston (2001) – that tries to test various theories by examining the causes of vertical integration. For example, researchers have examined whether firms that must make relationship-specific investments are more likely to integrate (Joskow 1985, Baker and Hubbard 2003) as suggested by a number of theories. There is, however, even less empirical research on the effects of vertical integration, i.e. on how integration affects firm behavior and performance. This literature has concentrated on the effect of vertical integration across firms, comparing integrated and non-integrated firms (Mullainathan and Scharfstein, 2001; Berger, et al., 2004). This paper is an attempt to understand the effects of integration by performing a comparison of integrated and non-integrated projects within the same firm. In particular, I examine R&D projects that are fully integrated within a biopharmaceutical firm to those that are governed through a contract with another firm.

I focus on the biopharmaceutical industry because large biopharmaceutical firms are vertically integrated in that they perform R&D to develop new drugs, manufacture the drugs, and market them. However, they are not completely integrated; while developing their own drugs internally, they often contract with smaller biopharmaceutical firms to develop drugs as well. This feature of the industry enables me to compare the management and performance of a set of projects that are governed by ownership to those undertaken by the same firm that are governed by contract. This allows me to examine the difference between transactions that occur within the firm

¹Including the transaction cost approach of Williamson (1975, 1985) and Klein, Crawford and Alchian (1978); the property rights approach of Grossman and Hart (1986), Hart and Moore (1990), and Hart (1995); and the incentives-based models of Holmstrom and Milgrom (1991, 1994), Holmstrom and Tirole (1991), and Holmstrom (1999).

to those that occur between firms, the central issue raised by Coase (1937) and the work of Williamson (1975), Grossman and Hart (1986), and Hart and Moore (1990).

The focus of the empirical analysis is on the clinical trials that are required for a drug to be approved by the Food and Drug Administration (FDA)². Drug candidates must go through three phases of clinical trials on human subjects: small Phase I trials designed in most cases to test a drug's safety; larger Phase II trials to test both its safety and efficacy; and finally very large Phase III trials with as many as a several thousand subjects. At each point along the way, a company must decide based on scientific, clinical, and financial information whether to continue to the next, more expensive phase of clinical trials. These are the biggest investments companies make in the R&D process.

The decision of whether to advance a drug candidate to the next phase of clinical trials is the critical governance challenge in drug development. Ideally, this decision would be based purely on scientific, clinical, and financial grounds. However, the agent who is most closely involved in a drug's development has a personal incentive to move it forward even when scientific, clinical, and financial factors are not so promising. The incentive to continue to the next phase is very strong in the case of a small, independent biopharmaceutical firm, because the agent owns the project. Continuation of a project may signal to an imperfectly informed capital market that the technology works, making it easier to raise additional capital. Discontinuation, on the other hand, sends a very negative signal and could destroy the possibility of raising funds in the future. Thus, the agent-owner has an incentive to continue a project even when the scientific, clinical, or financial factors do not fully justify it. By contrast, when an agent is running a project internally, the signaling value of continuation is likely to be much weaker. In this case, the firm finances any future projects of the agent. Unlike the external capital market, the firm itself is likely to

²The most recent estimate of the cost of getting a single drug approved is \$802 million (deflated to 2000). This estimate factors in the expected costs associated with failed attempts to develop a drug.

be well informed about the quality of the agent's other projects. Thus, I argue that these "private benefits" of continuation are greater for non-integrated projects than they are for integrated projects, and that this difference will give rise to differences in the management and performance of the two types of projects.

In Section 2 of the paper, I develop a simple theoretical framework to understand the effect of these differences. I compare the optimal contract between a firm and an employee (in the case of an integrated project) and the contract between the firm and an entrepreneur (in the case of a non-integrated project). The key assumption is that entrepreneurs derive greater private benefits from continuation than do managers of internal projects. Integrated projects will be continued only if they are valuable projects, while the contract governing non-integrated project will ex ante guarantee continuation even in cases when the continuation is not warranted ex post. This guarantee allows the entrepreneur to enjoy the private benefits of continuation. It also acts as an incentive mechanism; the agent who owns the project enjoys the continuation externalities and is thus motivated to exert more effort for the success of the project. The result is that the optimal continuation probability of a non-integrated project will be higher than that of an integrated project, while its average ex post quality will be lower. The model also predicts that the continuation decision for integrated projects will be more sensitive than the continuation decision for non-integrated projects to changes in the financial resources of the firm.

My empirical analysis compares the investment decisions of biopharmaceutical firms with respect to their own (internal) projects to their investment decisions with respect to alliance (external) projects. Alliance projects are projects that are developed and owned by a small firm, but a contract is signed with the large firm whereas the large firm finances the project. The empirical evidence in this study is consistent with the implications of the model. Controlling for firm and project characteristics, alliance projects are 21% more likely to move from Phase I to Phase II, compared to internal projects. However, alliance projects are less likely to eventually succeed.

I use two measures of success. One is whether the project moves from Phase II to Phase III. In prior work, it has been shown that this decision is highly correlated with the clinical results of Phase II trials (Guedj and Scharfstein, 2004). The other measure is whether the Phase III trials eventually receive FDA approval. On both counts alliance projects are less successful. Alliance projects are 20% less likely to move from Phase II to Phase III and they are 10% less likely to receive FDA approval. These findings are consistent with the view that firms use a higher continuation threshold on their internal projects than on their external projects.

Because internal and external projects could be fundamentally different in nature, I control for an array of project and firm characteristics: the novelty of the drug candidate (by looking at the number of drug projects that had used prior to it the same mechanism to attempt to cure the same disease), the potential market size, whether that drug candidate is of chemical or biological source, and the therapeutic class this drug candidate belongs to. Unsurprisingly, more novel drugs have a higher probability of advancing to the next phase and so do drugs with a higher potential market size. However, none of these controls alter the results that internal projects are less likely to advance to the next phase but eventually more likely to succeed.

Nevertheless, the results could be due to some unobserved heterogeneity in the type of project. In order to investigate this possibility, I look at small firms that had a contractual agreement with a larger firm, but were subsequently acquired by that large firm. If there is an unobserved heterogeneity in the projects originated by the small firm, this heterogeneity should continue to be present after the firm and its projects are fully integrated in the large firm. However, the probability of continuation of projects that had been originated by the start up but were subsequently developed as fully integrated projects by the acquiring firm, do not exhibit any differences neither in their continuation decisions nor in their probability of success, compared to fully integrated projects that had been initiated by the large firm. This result is another indication that the level of integration affects investment decisions and performance,

and once the level of integration changes, so do the decisions.

Finally, I investigate how the continuation decision depends on the firm's financial resources. I find that the continuation decision with respect to integrated projects depends on the amount of cash available to the firm, the existence of other advanced projects in the firm, and, to a lesser extent, the existence of similar projects in the firm. Non-integrated projects are less sensitive than integrated projects to the firm's financial resources and to the existence of other advanced projects in which the firm could invest. However, they are sensitive to other similar projects in the firm. This indicates that large biopharmaceutical firms seem to rank their internal projects relative to each other across therapeutic classes, and that they assess investment decisions taking into account their existing constraints. When assessing external projects, the firms seem to follow a more rigid method, which is much less affected by changes in their financial resources and other investment opportunities.

This paper relates to two empirical literatures. The first is the small empirical literature on the effects of vertical integration. Mullainathan and Scharfstein (2001) look at chemical producers that are integrated with the downstream users of their chemical, and compare them with producers that are not integrated. They find that the capacity of non-integrated producers is much more responsive to market conditions than is the capacity of integrated producers. In fact, integrated firms seem to hold more manufacturing capacity when it is least needed. Berger, et al. (2004) compare small to large (integrated) banks. They provide evidence consistent with small banks being better able to collect and act on soft information than large banks. In particular, they show large banks are less willing to lend to informationally "difficult" credits, such as firms with no financial records. These papers give an indication that the level of integration can have an important effect on the decisions firms make.

My findings are also related to the literature on the costs and benefits of internal capital markets. Much of this literature suggests that internal capital markets lead

to investment inefficiencies due to cross-subsidization of divisions in low-growth industries by those in high growth industries (Scharfstein and Stein, 2000, Scharfstein, 1998, Shin and Stulz, 1998 and Rajan, Servaes and Zingales, 2000). Another line of the empirical literature argues that internal capital markets enable firms to redeploy capital from unprofitable sectors to more profitable ones (Guedj and Scharfstein, 2004; Khanna and Tice, 2001; and Maksimovic and Phillips, 2002). This paper shows that in order for the firm to enjoy the benefits of an internal capital market, it needs to own the project, and that financing investments through contracts does not give the firm the flexibility to properly redeploy its capital.

The remainder of the paper proceeds as follows: in Section 2, I develop a simple model to structure the problem that arises when bringing a non-integrated project inside an internal capital market. In this section, I discuss the relevant empirical implications of the framework to the biopharmaceutical industry. Section 3 gives a short overview of alliances in the biopharmaceutical industry. Section 4 describes the data, and Section 5 presents the results. Section 6 concludes.

2 A Simple Model

This section outlines a simple model for comparing a firm's investment decisions with respect to a project when it is governed by outright ownership and when it is governed by a contract. Section 2.2 gives an interpretation the model in the context of biopharmaceutical development and develops the empirical implications of this framework.

The project has a life of two periods. At time $t = 0$ an investment I_0 is made. At time $t = 1$ uncertainty is resolved and the state of the world is observed. If the project is continued, an investment of I_1 is required. At time $t = 2$ the cash flows materialize. There are two states of the world, Good (G) or Bad (B). The ex ante probability of state G is θ and the ex ante probability of state B is $(1 - \theta)$. The

probability θ depends on the manager's effort. The state of the world is observable and verifiable. If the state of the world is B, the expected cash flow at time 2 is zero, and thus the project is terminated at time 1. If the state of the world is G, there are two possible outcomes. With probability α , a cash flow X is realized, and with a probability of $(1 - \alpha)$ no cash flow is realized. α can be seen as a measure of the quality of the project. α is random and can have one of two values, $\alpha \in \{\alpha_L, \alpha_H\}$, where $\alpha_H > \alpha_L$. The probability of $\alpha = \alpha_H$ is p , and the probability of $\alpha = \alpha_L$ is $(1 - p)$. α is observable at time 1, but it is not verifiable and thus non-contractible. See Figure 1 for a timeline of the model.

The project is initiated by a manager who is key to the development of the project. Thus, the manager needs to exert effort between time 0 and 1 in order to affect the probability of getting to the good state of the world. The project manager receives a payment w at time 2 if the project is successful. The manager chooses a level of effort, θ , which increases the probability that the project is successful. However, the manager incurs a personal, non-pecuniary cost, $\frac{1}{2}\beta\theta^2$, in doing so, where $\beta > 0$. This effort choice cannot be observed by anyone and thus contracts cannot be made contingent on it. It is also assumed that the manager has a "personal gain," b , if her project is continued at time 1, even if the project generates no cash flow at time 2. For example, b could be thought of as a shock to the manager's reputation, affecting the value of her other projects that share the same technology.

I will use this personal gain or externality, b , to differentiate between internal and external projects. When a project is internal, the manager has a low personal gain from the project. However, when the manager *owns* the project she has a high personal gain. When this personal gain is viewed as the externality received due to the information about the general value of the manager's projects, it is clear that if the manager does not own the project she does not gain much from that externality. Since the firm can observe the state of the world and the real value of the project, it doesn't need to continue a project in order to value the work of the manager. However,

when the manager owns the project, the informational value of continuation can be quite high, for example, if the manager owns a start-up company with two projects, one being developed and the other one needs funding. By having the first project continued, a positive signal would be sent to the market about the potential value of the second project, increasing the probability of securing funding for that project.

Regardless of the ownership of the project, it is assumed that the manager is risk neutral and does not finance the project outside of the firm. In the case of an independent entrepreneur, this model concentrates on the situation where the financially constrained entrepreneur has elected to develop the project in alliance with the firm. In order to simplify the analysis, it is assumed that, if possible, the manager prefers to have her project financed by the firm.

There are several potential investment decisions the firm can make. The firm can elect to always invest at time 1, to invest only if the state of the world is G, or only if the state of the world is G and the realization of the quality of the project is $\alpha = \alpha_H$. Given the that the state of the world is observable and contractible, the firm will not invest if the state of the world is B. However, if the state of the world is G, it can always invest or only if $\alpha = \alpha_H$. Since α is non-contractible, in this case the contract must be set up so that it is in the interest of the firm ex post to make an investment of I_1 only when α turns out to be α_H . By contrast, if the contract calls for the firm to always invest in state G regardless of α , the payments do not need to insure that it is ex post optimal for the firm to invest regardless of α .

In general, the contract can specify payments at time 0, time 1 conditional on G or B. It can be shown that the optimal compensation is at time 2 conditional on success as it fully aligns the manager's incentives with the firm's objective function. Therefore, the firm can offer two different payments dependent on whether it plans to always invest at time 1 or only if the realization of the quality of the project at time 1 is $\alpha = \alpha_H$.

2.1 Contracts

2.1.1 Contract 1 - Invest only if $\alpha = \alpha_H$

This contract stipulates that at time 1, the firm will invest in the project only if it observes that the state of the world is G and that the probability of a non-zero cash flow at time 2 is α_H . Thus, the firm will choose a payment w_H such that:

$$\max_{w_H} E[\pi^{Firm}] = E[\theta_H^*(\alpha(X - w_H) - I_1) - I_0 \mid \alpha = \alpha_H] \quad (1)$$

subject to

$$\theta_H^* \in \arg \max_{\theta_H} E[U(\pi^{Manager})] = E[\theta_H(\alpha w_H + b) - \frac{1}{2}\beta\theta^2 \mid \alpha = \alpha_H] \quad (ICC)$$

$$E[U(\pi^{Manager})] = E[\theta_H(\alpha w_H + b) - \frac{1}{2}\beta\theta^2 \mid \alpha = \alpha_H] \geq 0 \quad (IRC)$$

Therefore, the manager will optimize her effort level given the payment w_H and the firm will optimize the payment in order to increase the effort level θ_H . The result is:

$$\theta_H^* = \frac{p(\alpha_H X - I_1 + b)}{2\beta} \quad (2)$$

$$w_H^* = \frac{\alpha_H X - I_1 - b}{2\alpha_H} \quad (3)$$

The payment w_H^* meets the feasibility constraint:

$$w_H^* \leq X - \frac{I_1}{\alpha_H}$$

As one can see in equation (2), the manager's effort increases with b . The more the manager gains from the continuation of the project at time 1, the more the manager will exert effort in order to reach the state G and increase the probability of continuation. However, as equation (3) implies, the higher b is, the lower the payment w needs to be for the principal to induce the manager to exert the same level of effort. This is the basic insight of this model. When a manager owns her project, a lower remuneration is needed in order to induce exertion of effort by the

manager. However, when a manager who needs to exert effort for the success of the project does not own the project, the principal needs to pay a higher payment in order to induce the manager to exert effort.

2.1.2 Contract 2 - Always Invest

This contract stipulates that at time 1, the firm will invest in the project regardless of the realization α as long as the state of the world is G. Thus, the firm will choose a payment w_L such that:

$$\max_{w_L} E[\pi^{Firm}] = E[\theta_L^*(\alpha(X - w_L) - I_1) - I_0] \quad (1b)$$

subject to

$$\theta_L^* \in \arg \max_{\theta_L} E[U(\pi^{Manager})] = E[\theta_L(\alpha w_L + b) - \frac{1}{2}\beta\theta^2] \quad (ICC)$$

$$E[U(\pi^{Manager})] = E[\theta_L(\alpha w_L + b) - \frac{1}{2}\beta\theta^2] \geq 0 \quad (IRC)$$

Therefore, the manager will optimize her effort level given the payment w_L and the firm will optimize the payment in order to increase the effort level θ_L . The result is:

$$\theta_L^* = \frac{\bar{\alpha}X - I_1 + b}{2\beta} \quad (2b)$$

$$w_L^* = \frac{\bar{\alpha}X - I_1 - b}{2\bar{\alpha}} \quad (3b)$$

where $\bar{\alpha} = p\alpha_H + (1 - p)\alpha_L$. The payment w_L^* meets the feasibility constraint:

$$w_L^* \leq X - \frac{I_1}{\bar{\alpha}}$$

The basic results from the first contract hold in the same way with this contract. The effort level θ_L increases with b and the manager's payment decreases with b .

2.1.3 Which Contract Will be Preferred?

The second contract will be chosen if:

- 1) $E[\pi^{Firm} | \text{contract 2}] \geq E[\pi^{Firm} | \text{contract 1}]$ and
- 2) $E[\pi^{Manager} | \text{contract 2}] \geq E[\pi^{Manager} | \text{contract 1}]$.

This will happen if:

$$b \geq I_1 - \alpha_L X \quad (4)$$

If the second contract is chosen, the effort the manager will exert θ_H will be higher than the effort the manager would exert in case of the first contract θ_L :

$$\theta_H - \theta_L = \frac{(1-p)(\alpha_L X - I_1 + b)}{2\beta} \quad (5)$$

which is always positive if the second contract is preferred.

The second contract will be chosen regardless of b if the project has a positive NPV even when $\alpha = \alpha_L$ (i.e., when $\alpha_L X - I_1 > 0$). However, when the project has a negative NPV when $\alpha = \alpha_L$, the decision of which contract to choose depends on how large b is. If b is large enough, even projects with very negative NPVs will be developed at time 1. This situation is solely due to the existence of this external gain for the manager that allows in a sense to the firm to extract some of this gain by lowering the manager's payment.

2.2 Empirical Implications

This simple model generates several empirical implications regarding the investment decision of the firm when it has both integrated and non-integrated projects. This section develops the empirical implications in the context of the biopharmaceutical industry. In the context of the biopharmaceutical industry, the timeline of the model would be equivalent to engaging in Phase I (at time 0), deciding to move to Phase II (at time 1), and getting FDA approval (at time 2). α is the equivalent of the conditional probability of receiving FDA approval given that the project advanced to Phase II.

Prediction 1 *All else equal, the probability of a project advancing from Phase I to Phase II will be higher for a non-integrated (alliance) project than for an integrated (internal) project.*

Regardless of the chosen contract, the effort level exerted by the manager increases with b :

$$\frac{\partial \theta^*}{\partial b} > 0$$

Moreover, when $b \geq I_1 - \alpha_L X$, the second contract will be chosen for non-integrated projects, while the first contract will be chosen for integrated projects. In such a case, as equation (5) shows, the effort level will be higher since $\theta_H > \theta_L$.

Prediction 2 *Non-integrated projects will have ex post a lower success rate than integrated projects.*

When $b \geq I_1 - \alpha_L X$ the second contract will be chosen for alliance projects. The firm will always invest in those projects at Phase II. Due to that, the average probability of success of such projects will be equal to $\bar{\alpha} = p\alpha_H + (1 - p)\alpha_L$. However, internal projects will always have the first form of contract, resulting in a probability of success equal to α_H . The higher b , the greater the probability that a project with $\alpha = \alpha_L$ will be undertaken. The result is that cross sectionally, the average quality of an alliance project (its α) will be lower than the average quality of an internal project, since alliance projects will have a lower cutoff for continuation.

Prediction 3 *The firm's investment decisions regarding integrated projects compared to non-integrated projects will be more sensitive to financial and human capital constraints, and will be more sensitive to the existence of other projects in the firm's portfolio.*

The firm and the external entrepreneur have a binding contract. Until now I have assumed that all the agreement are not open for renegotiation. The agreement between the firm and its employee can be seen as somewhat different. This is not a

contract per se but an agreement where the firm commits to pay a certain payment. If the firm chooses so, it could pay the payment w_H and yet not continue a project, for example if it is financially constrained. This flexibility, is another aspect of ownership. As defined by Grossman and Hart (1986) this is one of the prerogative of the owner, to have the control on any aspect that has not been given away in a contract. The agreement between the owner and the employee does not stipulate continuation but rather a certain wage in which it is in the interest of the firm ex post to make that investment. However, if the firm's interest are altered, as the rightful owner, it has the flexibility not to invest.

3 The Pharmaceutical Industry and Alliances - Background

In this study I restrict my attention to the R&D of drug development in the pharmaceutical industry. There are a few reasons for this. First, due to regulation, detailed data is available on project level decisions. Second, it is possible to compare different projects inside the same firm, and across firms, since all drug development projects have to follow the same stages and procedures. This creates a relative homogeneity in the comparison of projects. Third, the outcomes of projects are measurable and comparable. A successful project is one that receives FDA approval. Fourth, there exists much heterogeneity in ownership structures of projects in this industry. Fifth, the biopharmaceutical industry is a major industry in R&D expenditure. In 2000, it accounted for 25% of all expenditure on R&D in the US.

This section gives a brief background on the industry and the mechanics of alliances.

3.1 The Drug Development Process

After a drug compound has been identified through pre-clinical research, the biggest investments that biopharmaceutical firms make are the clinical trials they conduct to prove the safety and efficacy of the drug candidate. In order to gain regulatory approval for market introduction, the FDA requires that biopharmaceutical companies provide substantial evidence of a drug's effectiveness through adequate and well-controlled clinical investigations. Proof of the drug's effectiveness must be provided by the results of randomized controlled trials. These are comprised of three main rounds of clinical trials on human subjects.

Phase I trials are typically conducted on fewer than 30 patients and are designed to determine a drug's safety. For most diseases, these trials are performed with healthy subjects. DiMasi, et. al. (2003), using a sample of 68 drug candidates undergoing trials at large biopharmaceutical firms between 1983 and 1994, estimate that the mean (median) out-of-pocket cost of a Phase I trial was \$15.2 million (\$13.9 million) deflated to year 2000 US Dollar terms.

Phase II trials are larger and more costly than Phase I trials. They are conducted on as many as a few hundred subjects, use patients with the disease, and are designed to test both the safety and efficacy of the drug agent. The mean (median) cost of a Phase II trial in the DiMasi, et. al. sample was \$23.5 million (\$17.0 million).

Finally, Phase III trials are typically very large studies, including possibly thousands of subjects. The mean (median) cost of a Phase III trial in the sample was \$86.3 million (\$62.0 million). After completing these trials, a drug sponsor can seek regulatory approval from the FDA by filing a New Drug Application (NDA).

3.2 Alliances

Alliances are a common way for pharmaceutical firms to augment their drug pipeline; they can be signed at any stage in the drug development process. Alliances signed at an early stage (preclinical) are usually longer-term relationships. The main aspect is

the financing of the entire process. These alliances usually have a broad scope and include most of the applications of an underlying technology. Alliances signed at mid stage (during clinical trials) are somewhat different. The disease the drug is supposed to cure is more defined and therefore the deal typically is narrower in scope. The rationale behind late stage alliances has less to do with financing (as most of the cost has already been incurred) and more to do with utilizing the marketing capabilities of large pharmaceutical firms. Nicholson, Danzon, and McCullough (2003) analyze the valuation of alliance deals showing that it is a vibrant and well performing market.

Figure 2 gives the distribution of alliances over time by the stage of signing. As one can see, the number of alliances has increased over time. Large biopharmaceutical firms are relying more on these alliances to supplement their own pipeline. Figure 3 details the alliance total cost by stages.

There are many ways to structure alliances. In general, however, most alliances have the same components. Most pre-commercial payments to licensors can be allocated to one or more of four categories: (1) upfront or licensing fees; (2) R&D payments; (3) milestone payments; and (4) equity³. Average upfront payments tend to be larger the more advanced the stage of the alliance at signing (this is a way to participate in the prior cost of research and development incurred by the licensor). For clinical-stage alliances, R&D payments usually depend on the development of the next generation or back-up compounds. “Milestone” is a general definition of all event-contingent payments, including achievement of benchmarks associated with a compound’s early or clinical development through first commercial sale, as well as patent issuance. Equity is one of the less common components of an alliance, as many do not include an equity investment.

Although each alliance is structured differently, the components are similar. However, alliances differ a lot depending on the stage at which they were signed. Therefore, in this paper I will concentrate only on early-stage alliances in order to reduce the

³See Lerner and Malmendier (2004) for a thorough analysis of alliance contracts.

variability of the different contracts and the effect it could have on the outcome of the drug development process.

4 Data

4.1 Data Source

The main data used in this study is obtained from PJB Publications' *PharmaProjects*. The PharmaProjects is a UK based commercial database that tracks drug compounds through their stages of development, from as early as pre-clinical laboratory studies to FDA approval. This database is typically licensed by major pharmaceutical, biotech, accounting, and law firms for the purpose of learning what the competition is doing. Since clinical trials are performed by medical institutions that publish the results of their research, the PharmaProjects is able to gather information on projects of both public and private companies. In addition to identifying the drug and its sponsoring organization, the PharmaProjects also gives information on the timeline of development (including the dates of Phase I, Phase II, Phase III, and FDA approval or discontinuation of the project). This database also provides detailed information on the drug candidates, including the indication it is aimed to cure, its pharmacological routes, its potential market value and of course the sponsoring firm. This data not only allows tracking of drug candidates over time, it also allows building a time series of each firm's R&D portfolio and its pipeline of drugs.

Data on the alliances is obtained from the databases of Recombinant Capital (ReCap), a San Francisco-based consulting firm specializing in tracking the biopharmaceutical industry, and mostly the alliances between the companies in this industry. As mentioned earlier, one advantage of studying biopharmaceutical firms is the degree of disclosure in this industry. Publicly traded biopharmaceutical firms, like other concerns, are required by the US Securities and Exchange Commission (SEC) to file material documents. Biopharmaceutical companies tend to interpret this requirement

conservatively, and often file alliance contracts. This allows ReCap to gather information about alliance contracts between various biopharmaceutical companies even if one of them is a private company.

4.2 Data Sample

The empirical design of this study requires concentrating on firms that have both internally and externally developed projects. For this purpose, I concentrate on the 40 largest public biopharmaceutical firms. I focus only on public firms for two main reasons. The quality of information should be equivalent since they have to follow the same disclosure rules. It also allows me to link the drug development data to financial information from COMPUSTAT.

I extract from PharmaProjects all the drug development projects data that involves any of those 40 firms and that have valid information. This results in 3610 projects initiated by 40 companies between the years 1984-2001. I do not gather information on projects initiated after the year 2001 so as to have enough time to get information on movement to phase II.

I also extract from ReCap information on projects where there was an alliance that involved one of these 40 companies as the financing side of the alliance. I manually match these alliances with projects in PharmaProjects using the drug name, the originator's name and the start-up's name. In this study I concentrated on alliances signed before phase I. In particular, I eliminate alliances where:

- One of the parties is a university, medical center, non-profit organization, or government agency.
- One of the parties has a controlling interest in the other, either through a majority equity stake or through a purchase option (e.g., an alliance between a firm and one of its R&D limited partnerships).
- More than two firms are involved, making the analysis less clean and tractable.

- The agreement as filed contains no information on the stage it which it was signed.
- The alliance was signed after phase I had been initiated. The later the stage, the less the alliance is about research and development, and the more it is about combining marketing or manufacturing capabilities.

This results in 447 drug candidates. I will refer to projects originated by a firm as *internal* and those that were originated by another company and the firm signed an alliance on as an *external* project. All this results in 4057 drug development projects, of which 3610 were internally developed and 447 were externally developed.

4.3 Summary Statistics

4.3.1 Information on Clinical Trials

The main focus of my analysis is the decision a firm makes whether (and when) to advance a project from one phase to another. Table 1 Panel A reports the number of projects and the percentage of projects in each phase. Table 1 Panel B reports the distribution of the time between different phases. Since it takes some time to start a new clinical phase, in the cases where the time to Phase II is less than one year, it indicates that some or even all the preparations were undertaken before Phase I was finished. This implies that the decision was taken with no real regard to the clinical data generated by Phase I. As table 1 Panel B indicates, the average time from the beginning of a Phase I trial to the beginning of the first Phase II trial is 1.79 years. The mean time from Phase II to Phase III is 2.4 years. These results are similar to the results in DiMasi et. al. (2003).

Of course, not all trials move forward to the next phase. As table 1 Panel A shows, only 71% of the internally developed projects advance from the pre-clinical stage to Phase I. The figure regarding alliance projects is quite different, 93% of them advance to Phase I. However, this is not surprising. This is not an unbiased sample

of all pre-clinical projects developed by startups. This is the subset of those projects that generated enough interest for a large biopharmaceutical company to sign an alliance on. It is not unreasonable to expect that such an agreement is signed with the intent to advance to clinical trial. A caveat that must be noted is that since the most reliable source of information about drugs for *PharmaProjects* is clinical trial result, it is possible that if a drug was developed by a small private company that never reached clinical trial it may not have been recorded. In such a case there would have been no way to match it to an alliance agreement if one were available. Due to all that, I will not look at the decision to advance to Phase I, but only at the subsequent decisions that do not have those potential problems.

The other transition probabilities indicate the main results of this study. Internal projects advance to Phase II with a probability of 70%, while external projects advance with a probability of 75%. They also do so at a much higher pace: 1.57 years compared to 1.80 years. However, in the probability of advancing to Phase III the order is reversed. Internal projects have a probability of 54% while external projects advance only with a probability of 45%. The probability of success of an internal project from Phase I to FDA approval is 22% while external project have a success rate of 17%, even they got more chances by advancing to Phase II in higher numbers.

4.3.2 Control Variables

The analysis controls for important characteristics of the drugs in development.

Table 2 reports the distribution of therapeutic classes across the different compounds. A therapeutic class is the main disease the drug is aiming at. The therapeutic activity codes are as defined in the *PharmaProjects Therapeutic Classification System (PTCS)*. This classification is based on the original classification devised by the *European Pharmaceutical Market Research Association (EPHRA)*. The PTCS is divided into fourteen major sections covering broad therapeutic areas, such as blood and clotting products, anticancer agents or respiratory agents. This classification is

similar to the World Health Organization's (WHO) typology used by Danzon, Nicholson, and Pereira (2003). As Danzon et. al. have shown, different therapeutic classes carry different probabilities of success; it is thus important to control for those differences. As one can see in table 2, there is a wide dispersion between internal and external. Some of the smaller indications do not have alliance projects, mostly indicating that these indications either require some unique expertise, or that the cost cannot support the creation of a new company. In general, one can observe that firms use alliances in most of the indications they develop drugs for, making it a widely used tool for enhancing their pipeline.

Different compounds that aim at different diseases may also differ in their potential financial value. Table 3 Panel A reports the distribution of the estimated potential market size of each compound if eventually approved. These figures are estimated by the PharmaProjects. There doesn't seem to be a large difference in project types between alliance and internal projects.

In order to control for the novelty of the compound I use the pharmacology of a drug as described by Guedj and Scharfstein (2004). The pharmacology describes a drug's mechanism of action in the body, through which it exerts its therapeutic effect, i.e. it identifies the biological agent or process the drug stimulates or inhibits. The novelty ranking is based on the chronological use of a pharmacological mechanism for a certain therapeutic class. Using the entire universe of compounds developed in the past 20 years (both approved and discontinued compounds), I rank each compound by the chronological appearance of its pharmacological mechanism and therapeutic class. For example, compound A uses the mechanism "Phosphodiesterase V inhibitor" and its development started in 1990, and compound B uses the same mechanism, but its development started in 1991. Then, compound A will receive the rank 1 and compound B will receive a rank of 2 and so on. This methodology gives a basic sense of whether a compound uses a mechanism that is new or that has been previously used. Table 3 Panel B describes of mean comparison of novelty of drugs between

internal and alliance projects. The mean rank of a drug developed by an early stage drug is 24.07. While the mean rank of a drug developed by a mature company is 28.31. The difference is not statistically significant.

4.3.3 Information on Companies

Table 4 presents summary data (deflated to the year 2000 US Dollar terms) on the public companies sponsoring the trials in the sample. On average, the public companies are very large, with mean revenues of over \$9 billion, mean assets of almost \$13 billion, mean cash of close to \$2.2 billion and mean R&D of about \$1 billion. The average market capitalization is over \$32 billion and the mean Q is 3.45. Nonetheless, there is wide heterogeneity among the firms in the sample. This is mainly due to the long time series, whereas most of the companies have grown over time to the large size as indicated by the means. Some of the largest companies in the sample, are the largest world drug manufacturers. With sales of more that \$35 billion, R&D expenditure of almost \$3 billion and more than 100 thousand employees worldwide.

5 Empirical Analysis

In this section I compare the investment decisions firms make regarding projects that are vertically integrated (internal projects) and projects that are not integrated (external or alliance projects) but that are governed by a contract.

5.1 Basic Analysis

5.1.1 Phase I to Phase II Transition Probabilities

All the projects in this analysis were either developed inside the firm from the earliest stages (pre-clinical) or an alliance contract was signed prior to the beginning of Phase I. However, it is reasonable to assume that if a contract was signed prior to Phase I,

there might have been an ex ante intent to move to Phase I. Due to this possibility, in this work I will not concentrate on the decision to advance to Phase I, but rather on the decision of whether to move to Phase II given that Phase I was completed. As mentioned in section 4, Phase II is a large and expensive phase with a mean cost of \$23.5 million. Often, more than one Phase II is required, making it even more expensive. Therefore, due to its potential cost, the decision to move from Phase I to Phase II is an ideal setting to address the question of whether ownership has an effect on the investment decisions of the firm.

The following analysis looks at the factors that affect whether a firm advances a drug from Phase I to Phase II. This is aimed at testing Prediction 1 from section 2.2. The hypothesis of this prediction is that firms will be more likely to terminate an internal project than an alliance project. All the tables reported are of a linear probability estimation.

I do not estimate Probit models, because Maximum likelihood estimators in the presence of fixed effects suffer from the 'incidental parameter problem' as was first analyzed by Neyman and Scott (1948). There is a vast literature such as Heckman (1981) who shows that in a Probit regression with a small sample, the presence of fixed effects might generate an upward bias in the estimation of β . This upward bias is exacerbated the smaller the sample is compared to the number of fixed effects. Although this sample is quite large, I report a linear model in order to have consistent estimates. The results also hold very similarly when using a Probit model and when using a Conditional Probit Model as suggested by Chamberlain (1980) and a Cox proportional hazard model.

In table 5, I report the results of the estimation of the following regression:

$$\begin{aligned}
 Pr[\textit{Advancing from Phase I to Phase II}] = & \\
 = \alpha + \beta_1 \textit{Alliance} + \beta_2 \textit{Novelty} + \beta_3 \textit{Bio} + \beta_4 \textit{LGMKT} + \beta_5 \textit{SMMKT} + & \\
 + \sum_{i \in \textit{Company}} 1_i + \sum_{j \in \textit{Therapy}} 1_j + \sum_{k \in \textit{Year}} 1_k + \epsilon & \quad (6)
 \end{aligned}$$

The dependent variable gets the value of 1 if the drug moved to Phase II within two years of initiating Phase I and zero otherwise. I use a two year cutoff on Phase II decisions for two main reasons. First, without a cutoff, Phase I trials that were begun in the early part of the sample would be more likely to be taken forward. If there is an over-representation of one type of ownership structure in the early period, this would bias the findings. The second reason to use a time cutoff is to measure the aggressiveness with which firms move forward in the clinical trials process. As one can see in table 1, the mean time to move from Phase I to Phase II is 1.69 years and the median is 1.82 years. To avoid making seemingly arbitrary cutoffs, a Cox proportional hazard model has been estimated (not reported) and yields similar results. The results are robust to using different time cutoffs, for example Table ?? reports results using a 3 year cutoff.

The independent variables are: *Alliance* is a dummy variable that gets the values of 1 if the project was originated by another company and is undertaken in alliance with that company, and gets the value of 0 if it is an internal project. *Novelty* is a measure whether the technology of the drug is new in treating that specific disease. It is constructed by looking at all the compounds developed over time and their pharmacological description (their mechanism of action). Each compound is ranked chronologically by the group of compounds that use the same pharmacology for the same therapeutic class. *Bio* is a dummy variable that receives the value of 1 if the project is about a compound that is biologic. If the compound is of chemical origins the dummy gets the value of 0. Two variables are used to control for the potential market size of the drug. The *PharmaProjects* estimates the potential market size of a drug. From this assessment I generate three dummies for Small, Medium, and Large market size (0-2\$b, 2-5\$b, and +5\$b respectively). *LGMKT* is a dummy that gets the value of 1 if the drug is potentially targeting a large market share; and *SMMKT* receives a value of 1 if the drug is targeting a small market. Three fixed effects are added: company fixed effects, therapeutic class fixed effects, and year fixed effects.

The first column of Table 5 is a simple comparison of means. An alliance project has a probability of moving from Phase I to Phase II in two years that is higher by 15.9% than an internal project. All the t-statistics in the table are based on robust standard errors corrected for firm-level clustering.

One potential explanation for this result is that alliance projects are inherently different. Those projects are developed by young start-up companies. For an entrepreneur to start a start-up and receive venture funding, the idea behind the start-up must have a high potential value. Moreover, one could speculate that the idea should be original and novel enough to have a potential for competing with products from well established firms. Therefore, one could reasonably hypothesize that alliance projects are more novel and therefore it is not unreasonable for those companies to be more aggressive regarding their decision to advance from Phase I to Phase II. Column 2 of Table 5 adds a measure of novelty. This control does not change the basic result, yet shows that indeed more novel drugs will have a higher propensity to move to Phase II. Those start-ups are usually generically referred to as 'Biotech' since a large majority of them use a biologic source (as opposed to a chemical source) for their compounds. Since biologic and chemical compounds could be different, column 3 adds a dummy for whether the compound is of biologic source. Since biologic compounds are relatively new, this adds another measure of novelty.

A second possible explanation for this finding could be that drugs developed by start-ups aim at higher payoff markets (increasing Y_i in the model). Column 4 adds controls for the potential market size. Drug candidates with higher potential payoffs have a higher propensity to move to Phase II, however, this doesn't hinder on the basic result that alliance projects are more likely to move to Phase II.

In order to control for other potential explanations for this result I add three sets of fixed effects: firm fixed-effects, therapeutic class fixed-effects, and year fixed-effects. Company fixed effects are added in column 5. Different companies may have different policies and different degrees of aggressiveness in pursuing clinical trials. I

add therapeutic class fixed-effects since different diseases carry different probabilities of success (see for example Danzon, Nicholson, and Pereira (2003) for an analysis of success probability by therapeutic class). Column 6 gives the results of the analysis with therapeutic fixed effects, not materially altering the result. Year fixed effects are added in column 5. As argued by Lerner, Shane, and Tsai (2003), in different years there are different market conditions that may lead to renegotiations of alliance contracts or cancelations of those contracts.

Column 7 summarizes the results including all the controls and fixed effects. The estimated marginal effect of an alliance is 0.2104, indicating that an alliance project is 21% more likely than an internal project to move forward from Phase I to Phase II. The results of the same regression but with a dependent variable that uses a cutoff of 3 years are similar and unreported.

5.1.2 Performance of the Projects that Advanced to Phase II

As Prediction 2 suggests, non-integrated projects should have a higher probability of advancing from Phase I to Phase II, but the eventual success of these projects should be lower than the success of integrated projects. I use two measures for the probability of success of the projects: the probability of moving from Phase II to Phase III, and the actual approval of the drug by the FDA.

There are several reasons I use the decision to advance from Phase II to Phase III as a measure of success although it is a choice the firm makes. First, the mean cost of a Phase III is estimated at \$86.3 million, but the actual cost can escalate to hundreds of millions of dollars depending on the therapeutic class and the number of times it needs to be undertaken. Most Phase clinical trials III are done more than once and in different countries in order to receive regulatory approval from each separate country. This high cost reduces the incentive of a firm to advance a low quality compound. Second, as shown by Guedj and Scharfstein (2004), the decision to advance from Phase II to Phase III is closely linked to clinical results at Phase II. This makes this

decision a good proxy for the quality of a compound. However, since FDA approval is the ultimate measure of success of any drug development project, I will use it as a second measure of quality.

Table 6 reports the results of the regression of the probability of advancing from Phase II to Phase III within three years, given that the drug candidate was in Phase II.

$$\begin{aligned}
 Pr[\textit{Advancing from Phase II to Phase III}] = & \\
 = \alpha + \beta_1 \textit{Alliance} + \beta_2 \textit{Novelty} + \beta_3 \textit{Bio} + \beta_4 \textit{LGMKT} + \beta_5 \textit{SMMKT} + & \\
 + \sum_{i \in \textit{Company}} 1_i + \sum_{j \in \textit{Therapy}} 1_j + \sum_{k \in \textit{Year}} 1_k + \epsilon & \quad (7)
 \end{aligned}$$

Similarly to the Phase II analysis, I use a three-year time cutoff on Phase III decisions. The results are robust to different time cutoffs and to a non reported Cox proportional hazard model analysis. Column 1 is a simple comparison of means. An alliance project that advanced to Phase II has about 9% less chances to advance to Phase III. The other columns represent the regression of this probability on the same controls that were described in section 5.1.1. Controlling for the novelty of the drug has quite a strong effect. More novel drugs have a 15% higher probability of advancing to the next phase just due to their novelty. At that stage the potential market size does not seem to play a large role in the firm's decision. However, adding the therapeutic class fixed effects makes a dramatic change. As mentioned earlier, different diseases carry different degree of complication in the drug development and also to some extent different costs in the clinical trials. By controlling for those disparities one gets a clear picture of the success of internal versus external projects. After controlling for project, company, and year characteristics, the probability of advancing a project from Phase II to Phase III is 20.7% lower for an alliance project than for an internally developed project. These results seem to indicate that having reached a more costly and complicated decision point, there was a reassessment of the real quality of the projects leading to the discontinuation of lower quality projects.

If one accepts that the decision to advance from Phase II to Phase III is a proxy of the quality of a project, then these results corroborate the prediction that the average quality of alliance projects is lower than the average quality of internal projects.

The ultimate definition of quality of a drug in the biopharmaceutical industry is whether a drug is approved by the FDA. In order to corroborate the above results, I run a regression of the probability of getting eventual FDA approval for a drug candidate that had a Phase III done. Table 7 presents the regression analysis. Indeed, not only do non-integrated projects advance with a lower probability to Phase III, but even those that do advance to Phase III have a lower probability of eventually getting FDA approval. Column 1 gives a simple comparison of means. An alliance project has a 9% lower chance of getting FDA approval than an internal project. This result holds using the standard controls; corroborating Prediction 2 that the average quality of an alliance project, as defined by its probability of success, is lower than the average quality of an internal project.

Even if one does not accept that the decision to advance from Phase II to Phase III is a proxy for quality, these results seem to indicate that the average quality of alliance projects is lower than the average quality of internal projects. When the firm decides at Phase II whether to advance to Phase III, it has more information than when it makes the decision to advance from Phase I to Phase II. Thus, one should expect this more informed decision to yield a better outcome, i.e. a higher probability of getting FDA approval. The fact that the probability of getting FDA approval for a non-integrated project is lower, indicates that indeed non-integrated projects are of lower quality.

5.2 Alternative Explanations

The empirical results are in line with the implications of the model in Section 2. However, there could be alternative explanations that would still be in line with the results. The results could be due to two different sorts of alternative

explanations. First, it could be due to problems in the data; either because of an endogeneity or because of a selection problem. Second, there are three main competing alternative explanations: learning, diversification, or real-options. In this section I will describe those explanations and compare them to the hypothesis developed in Section 2.

A. Endogeneity

The most basic criticism of this work is whether there exist an endogeneity problem. For that purpose all the regressions use a set of control to try and account for such a possibility. The *Novelty* variable is an attempt to capture differences in novelty and technology of the projects. The *Bio* variable account for differences due to different formulations. The *MarketSize* variables try to deal with the potential idea that different organizational forms target different parts of the distribution of payoffs of drugs. The results would make sense if smaller firms targeted only "blockbuster" drugs, and thus were willing to endure lower probability of success. The fixed effects help account for potential differences in companies, diseases, and years that could affect the results.

However, there could still be two sources of endogeneity in the data. First, there could be some unobserved heterogeneity that is not captured with the control variables. Second, the regression framework may not fully account for the differences in the characteristics of projects. In order to deal with those two potential problems I perform the following tests:

5.2.1 The Problem of Causal Inference

The problem of causal inference is a problem due to the fact that only one realization of events is observed. In the case of this study we observe for a certain drug only one realization, it is either developed internally or in alliance. If we had both realizations we could compare them directly and answer unequivocally what is the affect or integration investment decisions. Since we only observe one possibility there is a concern

that there might be an endogeneity problem. In order to address that I perform a matching analysis.

Table 8 reports the results of such a matching estimator as specified by Heckman, Ichimura, and Todd (1998) and Abadie and Imbens (2002). Observations are matched based on all observable characteristics: Novelty, Bio, Market Size, Therapeutic Class, and Year. I use 4 matchings per observation in order to improve the statistical reliability. Then the probability of advancing from Phase I to Phase II is compared between the matched observations. All the estimators are bias corrected following Rubin (1973a and 1973b) and Abadie and Imbens (2002). Panel A reports average treatment effects. The average difference of the probability of advancing from Phase I to Phase II for the matched sample is 0.2944, statistically significant at the 1% level. This shows that once observations are matched by their characteristics the differences in the probability of advancing from Phase I to Phase II still hold. Panel B reports average treatment effects for the treated. Since there are many more observations of internal projects than of alliance projects restricting ourselves only to the treated is in order. In essence by doing so one matches for every alliance projects an internal projects, this results with an even panel. The result still holds, this weakens the results by reducing the coefficient from 0.2944 to 0.2034, and the z-statistic from 9.81 to 6.40, but this is the same result. Matched by observables, the probability of advancing from Phase I to Phase II is higher for alliance projects compared to internal projects.

5.2.2 Technology Differences - Acquisitions

One potential criticism of the results in section 5.1 is that the controls do not account for something unique in the projects originated by the start-up companies. For this criticism to hold, one should expect that whatever makes those projects unique should still hold after the start-up is acquired by a larger firm.

Therefore, in order to corroborate the above results and in order to alleviate this potential criticism, I track start-up firms that had an alliance with a mature firm and

subsequently were acquired by that mature firm⁴. I define an acquired project as a project that was originated (or its technology was originated) by a start-up company. The company was acquired and the clinical trials started *after* the company was acquired. Table 9 reports the result of the regression of the probability of a project advancing from Phase I to Phase II. All the projects in that sample are fully owned by the firm, but a subset of them was originated by companies that eventually were acquired by the firm. Column 1 is a comparison of means. The mean probability of a project moving from Phase I to Phase II is not statistically significantly different if the project was originated by a large firm or if it had been originated by a startup but is now owned by the large firm. All these results hold with the regular sets of control used in section 5.1. This seems to indicate that what drives the above results is the organizational form governing the management of the project. Non-integrated projects seem to be treated differently by the firm, a situation that seems to change once the firm fully owns the project, making it fully integrated.

These results indicate that there seems to be something quite different in the decision mechanism, when a firm makes an resource allocation decision relating to integrated and non-integrated projects.

B. Other Alternative Explanations

There are three main compelling stories that potentially could explain the results in the same way the contracting story put forward does:

5.2.3 Learning

A possible explanation for the results could be that large firms *knowingly* choose to continue a project even if it is not that good since they might gain spillover knowledge and learn just from the exposure to a technology that the firm is not an expert on. Jaffe (1986) and Gomes-Casseres, Jaffe, and Hagedoorn (2004) show that spillover of knowledge is very strong when a project is inside a firm, even when the firm does

⁴See Higgins and Rodriguez (2003) for an analysis of post alliance acquisitions.

not necessarily own the project. Thus, if a firm wants to "acquire" knowledge or to let it trickle down in the organization, it might be willing to continued a project it knows does not have high probability of success but generates value in another way. In such a case the firm might be willing to do that between Phase I and Phase II as the cost is not too high and the project is valuable since it hasn't been long in the organization. The firm might be reluctant to continue do that after Phase II exactly for the opposite reasons: high costs and most of the knowledge should have been acquired by then.

Such a theory does not seem to fully fit the fact that even the decision to advance from Phase II to Phase III does not seem to be fully efficient. Such a theory would predict a very harsh discontinuation decision at Phase II. However, one could always argue that since does not disprove this idea directly. In order to test this idea I develop a measure to how much the form could "learn" from having a project internalized in the firm. I define a variable, *Expertise*, as a variable that receives the value of 1 if: The firm had at least one drug approved in the preceding 3 years, or if the firm has at the time of its decision, at least 2 other projects in Phase III in the same therapeutic class. This dummy variable give a sense of whether the firm has experience in a certain therapeutic class, which should indicate that it has *less* need to learn from a project and thus should be more reluctant to continue a mediocre project since it has less to gain from its continuation.

Table 10 reports the result of the regression of the probability of advancing a project from Phase I to Phase II controlling for the *Expertise* of a firm. The *Expertise* is statistically significant, indicating that there is indeed an impact to how much can a firm benefit indirectly from having a project. It shows that there is an impact on the firm's decision by how much it is likely to gain from the project. We can see however that the impact from internal and alliance projects is different. Internal projects will be more likely to be discontinued if a firm is an expert in a field while internal projects will be more likely to be continued. This is indeed in line with

the idea of spillover. An alliance projects will be less likely to be continued if a firm knows the area well, and has less to gain from mere continuation. However, internally it probably is an expert and thus the value of its internal projects is probably higher. The interesting fact is that these result do not change the fact that alliance projects are more likely to be continued, this seems to indicate that expertise and spillover, thought influential do not drive the above results but only supplement it. This indeed seems to show that spillover does not contradict the idea of the impact of contracts but is complementing it.

5.2.4 Diversification

Another alternative potential explanation is that large firms sign alliances in order to diversify their pipeline of drugs. In such a case they would be more likely to sign an alliance if they are weak in a certain class of drugs. If indeed this is the case, firms should be less likely to discontinue an alliance drug simply because it might hurt an already weak division. This idea is similar to the expertise idea, but is different in the sense that it stresses not that the firm could learn from the drug but that it may need it just so as to "satisfy" the stock market. Thus, if a firm is weak in a certain class it may be willing to continue a lower quality project just in order to maintain (even if it is for the short run) a "diversified" drug pipeline.

In order to test this idea I develop a *Concentration Index*. This is essentially the ratio of the number of projects in a specific therapeutic class to the total number of projects the firm is undertaking. This is an attempt to measure to what extent the results can be explained by the need to maintain a project in a therapeutic class where the firm is weak. This index is calculated at the date a Phase I of a project is initiated.

Table 11 Panel A gives the summary statistics of this *Concentration Index*. There does not seem to be any difference between the concentration of the firm in a specific therapeutic class when it sign an alliance compared to initiating an internal project.

However, the median is lower for internal projects and the standard deviations is much higher. This indicates that although on average there is no difference the distribution of the concentration at the time of initiation is somewhat different. Firms tend to sign alliances when they are weak in a certain class but not when they do not have many projects. This is probably due to the facts that smaller firms, that expect to learn from the larger firms in the process of the alliance, would be reluctant to sign a contract with a firm that has no real experience in that class. It also seems that when a large firm has many project in a class it will be less likely to need to sign an alliance with a small firm.

Table 11 Panel B gives the results of a regression of the probability of advancing from Phase I to Phase II controlling for the concentration of projects in the firm. High concentration has a *negative* effect on the firm's decision. The more concentrated a firm is in an area the less likely it is to take a project forward in that area. However, this reluctance is not different for internal and alliance projects. There is no statistically significant difference between the decision with respect to internal and alliance projects. This indicates that although there is an effect to diversification, this effect is not different for internal and alliance projects, showing that this cannot explain the results the way that a contract theorem can.

5.3 Allocation of Resources

Section 5.1 presents evidence that pharmaceutical firms are more prone to discontinue integrated projects than non-integrated projects when deciding whether to advance from Phase I to Phase II. It also presents evidence that those projects are of lower quality, both defined by the probability of advancing to Phase II and getting FDA approval. The model in Section 2 suggests that the reason that firms would promote further non-integrated projects that would yield lower quality projects is the rigidity of the contract governing them. Therefore, as Prediction 3 suggests, changes in the firm's financial condition should have a larger effect on internal projects than on

external projects.

In order to test this prediction, I use several measures for factors that can affect the decision of a firm whether to continue an R&D project. First, I use the amount of cash (as measured by COMPUSTAT, in US Dollars deflated to the year 2000) the company has at the end of the year it started undertaking Phase I of the project. I normalize the amount of cash by the number of projects at that time in the firm. The results are very robust to different measures on normalization such as a weighted sum of projects, weighted by the expected cost of each project. Second, I use the number of other projects the firm has at the moment of decision. I normalize the number of projects by sales (in US Dollars deflated to the year 2000) of the firm that year so that they will proxy for the opportunity of the firm and not to its size. Again, the results using this variable are robust to using other deflators. All the regressions using the number of projects are robust to using only internal or internal and external projects. Third, I use not only the number of projects in the firm in general but also the number of projects in the firm targeting the same therapeutic class as the project I am looking at. This number is normalized in the same way as is the number of projects in the firm.

Table 12 presents the results of the regression of the probability of advancing from Phase I to Phase II on cash and the number of projects in the pipeline of the firm, using the same controls used in section 5.1. The proxy for the firm's financial constraint, the amount of cash the company has, affects the decision regarding internal projects but not external projects. The effect, though statistically significant, does not have a large economic significance. This is probably due to the fact that if they want, those large biopharmaceutical companies can raise added cash. These firms are sensitive in their decision to the cash at hand, indicating that they do not promote projects unless they have the resources to do so, and since they do not have endless resources they probably promote those projects that have the higher probability of success. When adding to the regression the overall number of projects in the firm we

get a similar picture. The decision whether to advance from Phase I to Phase II is negatively related to the product pipeline. If the firm has many other projects (scaled by sales) it will be less likely to advance a specific project to the next phase. This result coupled with the cash result seems to indicate a behavior similar to "winner picking" by the firm. Projects are more likely to advance if there is more cash, less projects overall in the pipeline and if that specific project is more novel or targets a larger potential market. Non-integrated projects, on the other hand, don't seem to exhibit the same behavior. They don't seem to be very sensitive to either the amount of cash in the firm or the firm's pipeline. As prediction 3 suggests, this can be explained by the fact that non-integrated projects are governed by a pre set contract. If the projects meet their goals, it is very difficult to renegotiate (or break) the contract simply because the firm's opportunity set has changed. This rigidity seems to be at the heart of the above results. It must be noted that both integrated and non-integrated projects show sensitivity to the existence of a pipeline of projects in the same therapeutic class. This is not explicitly predicted by the model. This could be explained by the fact that if the firm has a similar project at the same stage it is easier for it to terminate an external project. This could be due to the fact that this public knowledge of such projects reduces the asymmetric information that is involved in the signal perceived by the termination of an alliance. In such a case, the market or other firms will know that the reason for termination of the contract is not bad results but a clear change in the firm's priority, mitigating any reputation problem that may arise generally from the discontinuation of an alliance.

6 Conclusion

In this paper I compare two different financial arrangements: projects financed by internal funding, and Projects financed via a contractual alliance. Those two different arrangements govern integrated and non-integrated project respectively. I document

that biopharmaceutical firms use different criteria in their allocation of resources when assessing integrated and non-integrated projects. The probability of a non-integrated project that is governed by a contract to advance from Phase I to Phase II is 20% higher. However, the probability of those projects to advance from Phase II to Phase III is 20% lower, and eventually they are 10% less likely to get FDA approval if they undertake Phase III. I interpret this result of having a lower probability of success as an indication of the project being of lower quality. These findings are robust to a set of project and company characteristics and controls. These results disappear once the non-integrated company is acquired by the larger firm.

It seems that the decision criteria employed by a firm when considering integrated and non-integrated projects are different. Integrated projects are sensitive to the amount of cash on hand and to the number of other unrelated projects in the firm. Non-integrated projects, however, are much less sensitive to cash and are not sensitive to the existence of other unrelated projects. All these results seem to indicate that biopharmaceutical firms make their resource allocation decisions regarding integrated projects in a manner reminiscent of “winner picking.” Given their financial constraint, they seem to pick the projects that have a higher probability of performing better. However, when dealing with non-integrated projects, it seems that the contract reduces the flexibility and latitude the firm has to terminate or even postpone a decision regarding the continuation of a project. This lack of flexibility results in promoting further non-integrated projects at the expense of better integrated projects. These projects eventually do not perform as well, resulting in a lower success rate.

There are a number of ways in which I hope to build on this research. First, since integrated projects are of higher quality than non-integrated project, it would be very interesting to concentrate on the small company’s portfolio. How do projects that the company out-licenses compare to projects the company decided not to out-license. Do they differ in characteristics or in quality. If so, is the company selling “lemons” in the contracts it out-licenses? This is a potential serious issue of adverse

selection. Conversely, the firm could elect to out-license only its better projects in order to create a good reputation. In such a case the firm would be selling "lemons" not to the licensor but to the market.

Second, The theoretical framework in the model implies that in return for guaranteeing continuation the firm pays *less* than the cost of developing a project internally. This raises the question of comparing the cost of developing an R&D project. Is integration more costly? If so, what are its other values?

Third, Guedj and Scharfstein (2004) have shown that young companies are excessively prone to continue projects when they have the financial resources to do so. This paper shows that having such a project in an alliance with a larger firm does not resolve this over-investment problem. Since it seems difficult to discipline a start-up's decision, there must be something valuable in those companies that entices either the market or more established companies to invest in them. Therefore, these start-ups have much lower success rates. Understanding the origin of the innovation of these companies and what makes them seem valuable (at least *ex ante*) and yet makes them not as successful is a very valuable question.

References

- [1] Abadie, Alberto and Guido Imbens, 2002, Simple and BiasCorrected Matching Estimators for Average Treatment Effects, *National Bureau of Economic Research Technical Working Paper 283*.
- [2] Baker, George P., and Thomas Hubbard, 2003, Make v. Buy in Trucking: Asset Ownership, Job Design and Information, *American Economic Review* 93, no. 3.
- [3] Berger, Allen N., Nathan H. Miller, Mitchell A. Peterson, Raghuram G. Rajan, and Jeremy C. Stein, 2004, Does Function Follow Organizational Form? Evidence From the Lending Practices of Large and Small Banks, *forthcoming, Journal of Financial Economics*.
- [4] Chamberlain, Gary, 1980, Analysis of Covariance with Qualitative Data, *Review of Economic Studies*, 47, 225-238.
- [5] Coase, Ronald H., 1937, The Nature of the Firm, *Economica* 4,386-405.
- [6] Danzon, Patricia M., Sean Nicholson, and Nuno S. Pereira, Productivity in Biotech-Pharmaceutical R&D: The Role of Experience and Alliances, *National Bureau of Economic Research Working Paper 9615*.
- [7] DiMasi, Joseph A., Ronald W. Hansen and Henry G. Grabowski, 2003, The Price of Innovation: New Estimates of Drug Development Costs, *Journal of Health Economics*, 22, 151-185.
- [8] Gomes-Casseres, Benjamin, Adam B. Jaffe, and John Hagedoorn, 2004, Do alliances promote knowledge flows?, *forthcoming, Journal of Financial Economics*.
- [9] Grossman Sanford J. and Oliver D. Hart, 1986, The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration, *Journal of Political Economy* 94, 691-719.
- [10] Guedj, Ilan, and David Scharfstein, 2004, Organizational Scope and Investment: Evidence from the Drug Development Strategies of Biopharmaceutical Firms, *National Bureau of Economic Research Working Paper 10933*.
- [11] Hart, Oliver D., 1995, Firms, Contracts, and Financial Structure, *Oxford University Press, Oxford*.
- [12] Hart, Oliver D. and John Moore, 1990, Property Rights and the Nature of the Firm, *Journal of Political Economy* 98 1119-1158.

- [13] Heckman, James J. , 1981, The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time - Discrete Data Stochastic Process. Structural Analysis of Discrete Data with Econometric Applications, Manski, C. and McFadden D. (eds.). *MIT Press: Cambridge*.
- [14] Heckman, James J., Hidehiko Ichimura, and Petra E. Todd, 1998, Matching as an econometric evaluation estimator, *Review of Economic Studies*, 65(2), 261-294.
- [15] Higgins, Matthew J. and Daniel Rodriguez, 2003, The Outsourcing of R&D Through Acquisitions in the Pharmaceutical Industry, *working paper*.
- [16] Holmstrom, Bengt, 1999, The Firm as a Subeconomy, *Journal of Law, Economics, and Organization* 15: 74-102.
- [17] Holmstrom, Bengt and Paul Milgrom, 1991, Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design, *Journal of Law, Economics, and Organization* 7:24-52.
- [18] Holmstrom, Bengt and Paul Milgrom, 1994, The Firm as an Incentive System, *American Economic Review* 84: 972-91.
- [19] Holmstrom, Bengt and Jean Tirole, 1991, Transfer Pricing and Organizational Form, *Journal of Law, Economics, and Organization* 7:201-28.
- [20] Jaffe, Adam B., 1986, Technological opportunity and spillovers of R&D: Evidence from firms patents, profits and market value, *American Economic Review* 76, 984-1001.
- [21] Joskow, Paul L., 1985, Vertical Integration and Long-term Contracts: The Case of Coal-burning Electric Generating Plants, *Journal of Law, Economics, and Organization* 1:1, 33-80.
- [22] Khanna, Naveen and Sheri Tice, 2001, The Bright Side of Internal Capital Markets, *The Journal of Finance*, Vol. 56, pp. 1489-1528.
- [23] Klein, Benjamin, Robert G. Crawford, and Armen A. Alchian, 1978, Vertical integration, appropriable rents, and the competitive contracting process, *Journal of Law and Economics* 21, 297-326.
- [24] Lerner, Josh, and Ulrike Malmendier, 2004, Contractibility and the Design of Research Agreements, *working paper*.
- [25] Lerner, Josh, Hilary Shane and Alexander Tsai, 2003, Do Equity Financing Cycles Matter? Evidence from Biotechnology Alliances, *Journal of Financial Economics*, 67, 411-446.

- [26] Vojislav Maksimovic and Gordon M. Phillips, 2002, Do Conglomerate Firms Allocate Resources Inefficiently Across Industries? Theory and Evidence, *The Journal of Finance*, April, 721-767.
- [27] Mullainathan, Sendhil, and David Scharfstein, 2001, Do Firm Boundaries Matter?, *American Economic Review*, 91: 2, 195-199.
- [28] Neyman, J. and Elisabeth L. Scott, 1948, Consistent Estimates Based on Partially Consistent Observations, *Econometrica*, 16, 1-32.
- [29] Nicholson, Sean, Patricia M. Danzon, and Jeffrey McCullough, 2003, Biotech-Pharmaceutical Alliances as a Signal of Asset and Firm Quality, *forthcoming, Journal of Business*.
- [30] Rajan, Raghuram; Servaes, Henri and Zingales, Luigi, 2000, The Cost of Diversity: The Diversification Discount and Inefficient Investment, *The Journal of Finance*, 55(1), pp. 35-80.
- [31] Rubin, Donald B., 1973a, Matching to Remove Bias in Observational Studies, *Biometrics*, 29, 159-183.
- [32] Rubin, Donald B., 1973b, The Use of Matched Sampling and Regression Adjustments to Remove Bias in Observational Studies, *Biometrics*, 29, 185-203.
- [33] Scharfstein, David, 1998, The Dark Side of Internal Capital Markets II: Evidence from Diversified Conglomerates, *National Bureau of Economic Research Working Paper 6352*.
- [34] Scharfstein, David and Jeremy Stein, 2000, The Dark Side of Internal Capital Markets: Divisional Rent-Seeking and Inefficient Investment, *The Journal of Finance*, Vol. 55, 2537-2564.
- [35] Shin, Hyun-Han and Rene Stulz, 1998, Are Internal Capital Markets Efficient?, *Quarterly Journal of Economics*, 113, 531-553.
- [36] Whinston, Michael D., 2001, Assessing the Property Rights and Transaction-Cost Theories of Firm Scope, *American Economic Review*, 91: 2, 184-188.
- [37] Williamson, Oliver E., 1975, Markets and Hierarchies: Analysis and Antitrust Implications *Collier Macmillan, New York*.
- [38] Williamson, Oliver E., 1985, The Economics Institutions of Capitalism *Free Press, New York*.

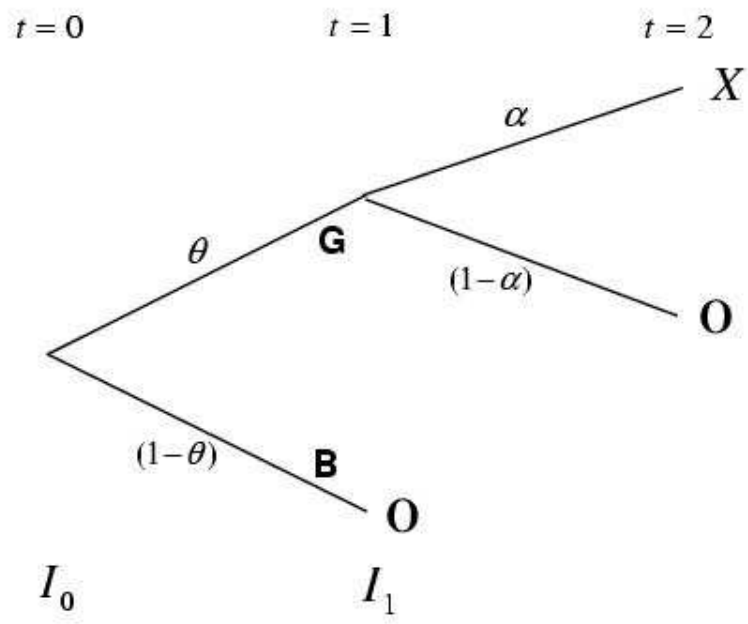


Figure 1: Timeline of the model.

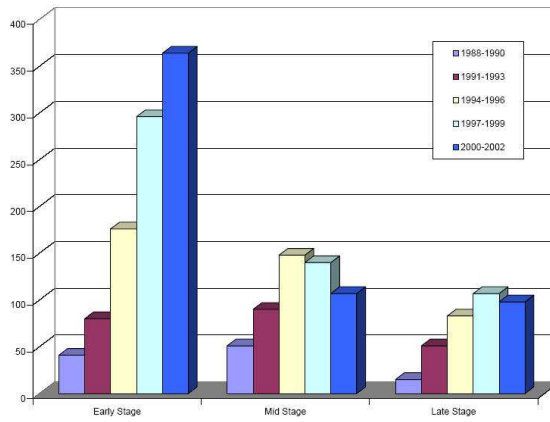


Figure 2: Number of alliances in the BioPharmaceutical industry in the years 1988-2002, grouped by the stage the alliance was signed at.

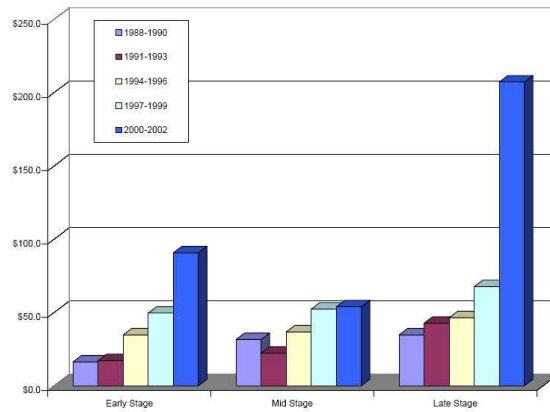


Figure 3: Average total size of an alliance (in US Dollars) in the BioPharmaceutical industry in the years 1988-2002, grouped by the stage the alliance was signed at.

Table 1: Time and Probability of Advancing from One Phase to Another
This table reports the time and probability of moving from one development phase to another. Panel A reports the number of drugs in the sample that moved from pre-clinical, to Phase I, Phase II, Phase III, and eventual FDA approval. Panel II reports the mean, median, and standard deviation of the time of moving from one phase to another. Time is measured in years.

	All		Internal		External	
	Number	Advance Rate	Number	Advance Rate	Number	Advance Rate
Number of Projects	4057		3610		447	
Went to Phase 1	2985	73.6 %	2567	71.1 %	418	93.5 %
Went to Phase 2	2108	70.6 %	1794	69.9 %	331	79.2 %
Went to Phase 3	1112	52.8 %	970	54.1 %	142	42.9 %
FDA Approval	671	60.3 %	597	61.5 %	74	52.1 %

		All	Internal	External
Time to Phase 1	Mean	0.70	0.68	0.82
Time to Phase 1	Median	1.10	1.10	1.08
Time to Phase 1	Std	0.53	0.53	0.48
Time to Phase 2	Mean	1.69	1.70	1.47
Time to Phase 2	Median	1.82	1.90	1.50
Time to Phase 2	Std	1.17	1.14	1.39
Time to Phase 3	Mean	2.40	2.38	2.51
Time to Phase 3	Median	2.60	2.60	2.60
Time to Phase 3	Std	1.10	1.10	1.16
Time to Approval	Mean	2.91	2.88	3.10
Time to Approval	Median	2.74	2.74	2.74
Time to Approval	Std	1.02	1.00	1.13
Time to cancellation	Mean	2.85	2.78	3.75
Time to cancellation	Median	2.17	2.09	3.59
Time to cancellation	Std	2.38	2.38	2.17
Clinical Time to Approval	Mean	8.07	8.09	7.92
Clinical Time to Approval	Median	8.88	8.88	8.95
Clinical Time to Approval	Std	1.84	1.85	1.79

Table 2: Therapeutic Classes

This table reports the breakdown of projects by therapeutic classes. The therapeutic activity codes, as defined in the PharmaProjects Therapeutic Classification System (PTCS). This classification is based on the original classification devised by the European Pharmaceutical Market Research Association (EPHRA). The PTCS differs from the EPHRA classification in that it has been considerably revised to more accurately reflect the types of products in R&D in 2004, rather than the marketed products of the 1970s. The PTCS is divided into fourteen major sections covering broad therapeutic areas, such as blood and clotting products, anticancer agents or respiratory agents. This classification is similar to the World Health Organization's (WHO) typology used by Danzon, Nicholson, and Pereira (2003).

Therapeutic Class	Total	Internal	External
Alimentary/Metabolic	348	286	62
Anti-infective	749	654	95
Anticancer	442	396	46
Antiparasitic	18	18	0
Blood/clotting	203	197	6
Cardiovascular	653	631	22
Dermatological	78	67	11
Genitourinary and Sex Hormones	112	99	13
Hormones excluding Sex Hormones	65	63	2
Immunological	71	56	15
Musculoskeletal	272	244	28
Neurological	735	636	99
Respiratory	292	250	42
Sensory	19	13	6
Total	4057	3610	447

Table 3: Control Variables

This table reports summary statistics on two control variables used in this study.

Panel A reports the distribution of the potential market size of the drug projects in this study, This distribution is based on assessments made by *PharmaProjects*.

Panel B reports data on the "novelty" of the drugs developed. Novelty is defined using the pharmacological description of the drugs. The pharmacology describes a drug's mechanism of action in the body, through which it exerts its therapeutic effect, i.e. it identifies the biological agent or process the drug stimulates or inhibits. Panel B describes the mean "rank" of each drug. Each drug is ranked in chronological order of appearance of its pharmacological mechanism for its specific therapeutic class of in the sample of all the compounds developed in the past 20 years as reported by *PharmaProjects*.

Panel A: Distribution of Estimated Project Potential Market Size

Market Size	Internal	Alliance
US\$ 0-500 million	9.2 %	7.9 %
US\$ 501-2000 million	29.8 %	33.8 %
US\$ 2001-5000 million	35.5 %	37.1 %
US\$ 5001-10000 million	16.1 %	9.1 %
Over US\$ 10000 million	8.1 %	9.4 %

Panel B: Mean Rank of drug Novelty

	Mean	Median	Std
Internal Projects	24.07	8	48.07
Alliance Projects	28.31	10	50.11

Table 4: Summary Statistics on Sample Companies

This table reports summary statistics on the firms in the sample. The numbers are averages of all the firm-year observations. Sales, assets, cash, market cap are in Millions of Dollars. Number of employees is in thousands. Q is defined as the market value of equity the book value of assets less the book value of equity divided by the book value of assets. All the numbers are deflated to year 2000 Dollars.

Statistic	Sales	Assets	Cash	R&D	Market Cap	Q	Employees
Mean	9,751.33	12,663.86	2,158.25	1,034.74	32,684.01	3.45	38.84
Median	8,030.59	10,952.53	1,314.33	934.27	27,952.96	3.51	42.22
Std	9,691.63	12,219.54	3,053.65	844.02	28,937.93	1.09	35.76
Max	35,198.89	42,362.05	12,629.75	2,945.42	106,737.00	6.71	136.61
Min	14.08	277.71	179.91	40.70	606.57	1.52	0.17
25%	1,029.82	2,318.71	485.73	279.91	5,918.06	2.74	4.25
75%	15,350.96	18,720.97	2,489.55	1,716.58	55,619.64	4.03	52.16

Table 5: Regression of the Probability of Advancing from Phase I to Phase II

The model estimated is a linear probability regression. The dependent variable is the probability of advancing from phase I to phase II in the 2 years following phase I. It takes the value of 1 if the drug advanced to phase II and 0 if not. An alliance project is defined as a project that was originated by another company and an alliance contract was signed. There are several controls: 1) Novelty - Each compound has a pharmacological description (a drug's mechanism of action in the body, through which it exerts its therapeutic effect). Compounds are ranked by the number of drugs developed for the same therapeutic class with the same pharmacological description over time. Novelty is the log of the inverse of this rank. 2) Bio - a dummy that receives a value of 1 if the compound is based on a biologic agent and 0 if not. 3) Market Size - There are three dummies as defined by *The PharmaProjects*: market size is up to 2000\$m, between 2000 and 5000\$m and more than 5000\$m. 4) Three fixed effects: company fixed effects, therapeutic class fixed effects and year fixed effects. The t-statistics, reported in parentheses, are based on robust standard errors that are corrected for firm-level clustering.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Alliance	0.1591 (7.71)***	0.1621 (5.69)***	0.1763 (5.43)***	0.197 (5.02)***	0.1951 (5.19)***	0.1969 (5.12)***	0.2074 (4.68)***
Novelty		0.0255 (2.24)**	0.0256 (2.28)**	0.0272 (2.45)**	0.0304 (2.61)**	0.0297 (2.51)**	0.0289 (2.55)**
Bio			0.0778 (2.46)**	0.0817 (2.51)**	0.0700 (2.32)**	0.0735 (2.34)**	0.0772 (2.37)**
Market Size - Large				0.0397 (0.30)	0.0666 (0.50)	0.0582 (2.33)**	0.1535 (2.42)**
Market Size - Small				-0.0532 (2.20)**	-0.045 (1.96)*	-0.0426 (1.72)*	-0.0279 (1.32)
Observations	2985	2985	2985	2985	2985	2985	2985
R^2	0.10	0.11	0.11	0.11	0.14	0.15	0.25
Company FE	No	No	No	No	Yes	Yes	Yes
Therapeutic Class FE	No	No	No	No	No	Yes	Yes
Year FE	No	No	No	No	No	No	Yes

Table 6: Regression of the Probability of Advancing from Phase II to Phase III

The model estimated is a linear probability regression. The dependent variable is the probability of advancing from phase II to phase III in the 3 years following phase I. It takes the value of 1 if the drug advanced to phase II and 0 if not. An alliance project is defined as a project that was originated by another company and an alliance contract was signed. There are several controls: 1) Novelty - Each compound has a pharmacological description (a drug's mechanism of action in the body, through which it exerts its therapeutic effect). Compounds are ranked by the number of drugs developed for the same therapeutic class with the same pharmacological description over time. Novelty is the log of the inverse of this rank. 2) Bio - a dummy that receives a value of 1 if the compound is based on a biologic agent and 0 if not. 3) Market Size - There are three dummies as defined by *The PharmaProjects*: market size is up to 2000\$m, between 2000 and 5000\$m and more than 5000\$m. 4) Three fixed effects: company fixed effects, therapeutic class fixed effects and year fixed effects. The t-statistics, reported in parentheses, are based on robust standard errors that are corrected for firm-level clustering.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Alliance	-0.0894 (2.49)**	-0.1021 (2.52)**	-0.0956 (2.96)***	-0.1048 (3.47)***	-0.1276 (3.45)***	-0.1462 (3.54)***	-0.1742 (3.90)***
Novelty		0.0563 (2.29)**	0.0744 (2.37)**	0.0545 (2.42)**	0.1001 (2.52)**	0.1313 (2.96)***	0.1520 (2.60)**
Bio			-0.0571 (1.21)	-0.0457 (1.12)	-0.0605 (1.08)	-0.0617 (1.10)	-0.0562 (1.05)
Market Size - Large				0.0239 (1.25)	0.0325 (1.27)	0.0386 (1.32)	0.0297 (1.28)
Market Size - Small				-0.0282 (0.90)	-0.0179 (0.84)	-0.0025 (0.10)	-0.0095 (0.50)
Observations	2108	2108	2108	2108	2108	2108	2108
R^2	0.08	0.08	0.09	0.09	0.13	0.23	0.28
Company FE	No	No	No	No	Yes	Yes	Yes
Therapeutic Class FE	No	No	No	No	No	Yes	Yes
Year FE	No	No	No	No	No	No	Yes

Table 7: Regression of the Probability of Getting FDA Approval if Moved to Phase III

The model estimated is a linear probability regression. The dependent variable is the probability of receiving FDA approval in the 4 years following Phase III. It takes the value of 1 if the drug received FDA approval and 0 if not. An alliance project is defined as a project that was originated by another company and an alliance contract was signed. There are several controls: 1) Novelty - Each compound has a pharmacological description (a drug's mechanism of action in the body, through which it exerts its therapeutic effect). Compounds are ranked by the number of drugs developed for the same therapeutic class with the same pharmacological description over time. Novelty is the log of the inverse of this rank. 2) Bio - a dummy that receives a value of 1 if the compound is based on a biologic agent and 0 if not. 3) Market Size - There are three dummies as defined by *The PharmaProjects*: market size is up to 2000\$m, between 2000 and 5000\$m and more than 5000\$m. 4) Three fixed effects: company fixed effects, therapeutic class fixed effects and year fixed effects. The t-statistics, reported in parentheses, are based on robust standard errors that are corrected for firm-level clustering.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Alliance	-0.0929 (2.09)**	-0.0961 (2.06)**	-0.0913 (2.18)**	-0.0953 (2.21)**	-0.105 (2.54)**	-0.1133 (2.34)**	-0.1089 (3.10)***
Novelty		-0.0753 (0.63)	-0.0779 (0.31)	-0.0716 (0.57)	-0.0787 (0.61)	0.0364 (0.27)	-0.1016 (1.16)
Bio			-0.1179 (1.76)*	-0.1238 (1.79)*	-0.1061 (1.61)	-0.1114 (1.30)	-0.1170 (1.19)
Market Size - Large				0.0479 (1.95)*	0.0759 (1.27)	0.1715 (1.35)	0.0476 (1.21)
Market Size - Small				0.017 (0.55)	0.0181 (0.58)	0.0138 (0.40)	0.0244 (0.79)
Observations	997	997	997	997	997	997	997
R^2	0.08	0.10	0.11	0.11	0.13	0.15	0.16
Company FE	No	No	No	No	Yes	Yes	Yes
Therapeutic Class FE	No	No	No	No	No	Yes	Yes
Year FE	No	No	No	No	No	No	Yes

Table 8: Matching Estimators - Probability of Advancing from Phase I to Phase II
This table reports the results of a matching estimator as specified by Heckman, Ichimura, and Todd (1998) and Abadie and Imbens (2002). Observations are matched based on all observable characteristics: Novelty, Bio, Market Size, Therapeutic Class, and Year. I use 4 matchings per observation in order to improve the statistical reliability. Then the probability of advancing from Phase I to Phase II is compared between the matched observations. Panel A reports average treatment effects and Panel B reports average treatment effects for the treated. All the estimators are bias corrected following Rubin (1973a and 1973b) and Abadie and Imbens (2002).

Panel A: Average Treatment Effect			
Coef.	Std. Err.	z stat	$P > z $
0.2944	0.300	9.81	0.0000

Panel B: Average Treatment Effect for the Treated			
Coef.	Std. Err.	z stat	$P > z $
0.2034	0.318	6.40	0.0000

Table 9: Regression of the Probability of Advancing from Phase I to Phase II - Acquisitions

The model estimated is a linear probability regression. The dependent variable is the probability of advancing from phase I to phase II in the 2 years following phase I. It takes the value of 1 if the drug advanced to phase II and 0 if not. An acquired project is a project that belonged to a smaller firm (that had in the past a contractual alliance with the larger firm) but subsequently got acquired by the larger firm. There are several controls: 1) Novelty - Each compound has a pharmacological description (a drug's mechanism of action in the body, through which it exerts its therapeutic effect). Compounds are ranked by the number of drugs developed for the same therapeutic class with the same pharmacological description over time. Novelty is the log of the inverse of this rank. 2) Bio - a dummy that receives a value of 1 if the compound is based on a biologic agent and 0 if not. 3) Market Size - There are three dummies as defined by *The PharmaProjects*: market size is up to 2000\$m, between 2000 and 5000\$m and more than 5000\$m. 4) Three fixed effects: company fixed effects, therapeutic class fixed effects and year fixed effects. The t-statistics, reported in parentheses, are based on robust standard errors that are corrected for firm-level clustering.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Alliance Acquisition	-0.0104 (0.49)	-0.0085 (0.39)	-0.0099 (0.45)	-0.0098 (0.45)	-0.0119 (0.48)	-0.0147 (0.60)	-0.0202 (0.96)
Novelty		0.0403 (3.40)***	0.0401 (3.43)***	0.0425 (3.75)***	0.0468 (3.95)***	0.0454 (3.80)***	0.0322 (3.61)***
Bio			0.0738 (2.41)**	0.0775 (2.56)**	0.0664 (2.42)**	0.0697 (2.30)**	0.0732 (2.25)**
Market Size - Large				0.4150 (0.50)	0.05987 (0.72)	0.06875 (1.33)	0.0824 (1.74)*
Market Size - Small				-0.0732 (2.75)***	-0.0677 (2.66)**	-0.068 (2.43)**	-0.0468 (2.16)**
Observations	2567	2567	2567	2567	2567	2567	2567
R^2	0.00	0.01	0.02	0.02	0.05	0.07	0.12
Company FE	No	No	No	No	Yes	Yes	Yes
Therapeutic Class FE	No	No	No	No	No	Yes	Yes
Year FE	No	No	No	No	No	No	Yes

Table 10: Regression of the Probability of Advancing from Phase I to Phase II - Expertise

The model estimated is a linear probability regression. The dependent variable is the probability of advancing from phase I to phase II in the 2 years following phase I. It takes the value of 1 if the drug advanced to phase II and 0 if not. There are several controls: 1) Novelty - Each compound has a pharmacological description (a drug's mechanism of action in the body, through which it exerts its therapeutic effect). Compounds are ranked by the number of drugs developed for the same therapeutic class with the same pharmacological description over time. Novelty is the log of the inverse of this rank. 2) Bio - a dummy that receives a value of 1 if the compound is based on a biologic agent and 0 if not. 3) Market Size - There are three dummies as defined by *The PharmaProjects*: market size is up to 2000\$m, between 2000 and 5000\$m and more than 5000\$m. 4) Three fixed effects: company fixed effects, therapeutic class fixed effects and year fixed effects. *Expertise* is a dummy variable that receives the value of 1 if the firm is an "expert" in the therapeutic class of the project and 0 otherwise. A firm is defined as an expert if: 1) It had at least one drug approved in the 3 years prior to the project's Phase I date, or 2) It had at the time of the project's Phase I date *at least 2* other projects in Phase III in the same therapeutic class. The t-statistics, reported in parentheses, are based on robust standard errors that are corrected for firm-level clustering.

	(1)	(2)	(3)	(4)	(5)
Alliance	0.1591 (7.71)***	0.197 (5.02)***	0.2074 (4.68)***	0.2011 (5.76)***	0.2232 (5.24)***
Novelty		0.0272 (2.45)**	0.0289 (2.55)**	0.0216 (2.32)**	0.0317 (2.51)**
Bio		0.0817 (2.51)**	0.0772 (2.37)**	0.0764 (2.38)**	0.0744 (2.29)**
Market Size - Large		0.0397 (0.30)	0.1535 (2.42)**	0.0378 (1.86)*	0.1425 (2.28)**
Market Size - Small		-0.0532 (2.20)**	-0.0279 (1.32)	-0.0485 (2.08)**	-0.0212 (1.21)
Expertise				0.0192 (3.75)***	0.0011 (2.05)**
Expertise*Alliance				-0.0447 (2.93)***	-0.0446 (2.55)**
Observations	2985	2985	2985	2985	2985
R^2	0.10	0.11	0.25	0.12	0.25
Company FE	No	No	Yes	No	Yes
Therapeutic Class FE	No	No	Yes	No	Yes
Year FE	No	No	Yes	No	Yes

Table 11: Regression of the Probability of Advancing from Phase I to Phase II - Diversification

The model estimated is a linear probability regression. The dependent variable is the probability of advancing from phase I to phase II in the 2 years following phase I. It takes the value of 1 if the drug advanced to phase II and 0 if not. There are several controls as described in Table 5. *Concentration* is a dummy variable that receives the value of 1 if the ratio of projects in a specific therapeutic class to the total number of projects the firm is undertaking. Panel A reports summary statistics for the concentration variable. Panel B reports the regression. The t-statistics, reported in parentheses, are based on robust standard errors that are corrected for firm-level clustering.

Panel A: Summary Statistics

	Mean	Median	StD
Internal	0.1264	0.9097	0.6310
Alliance	0.1241	0.1111	0.2431

Panel B: Regression

	(1)	(2)	(3)	(4)	(5)
Alliance	0.1591 (7.71)***	0.197 (5.02)***	0.2074 (4.68)***	0.2145 (5.44)***	0.2381 (5.30)***
Novelty		0.0272 (2.45)**	0.0289 (2.55)**	0.0208 (2.29)**	0.0322 (2.41)**
Bio		0.0817 (2.51)**	0.0772 (2.37)**	0.0761 (2.41)**	0.0722 (2.35)**
Market Size - Large		0.0397 (0.30)	0.1535 (2.42)**	0.0377 (1.86)*	0.1521 (2.32)**
Market Size - Small		-0.0532 (2.20)**	-0.0279 (1.32)	-0.0485 (2.08)**	-0.0224 (1.89)*
Concentration				-0.0875 (3.21)***	-0.0675 (3.32)***
Concentration*Alliance				-0.0021 (0.43)	-0.0038 (0.51)
Observations	2985	2985	2985	2985	2985
R^2	0.10	0.11	0.25	0.12	0.25
Company FE	No	No	Yes	No	Yes
Therapeutic Class FE	No	No	Yes	No	Yes
Year FE	No	No	Yes	No	Yes

Table 12: Regression of the Probability of Advancing from Phase I to Phase II - Resource Allocation Across Different Projects

The model estimated is a linear probability regression. The dependent variable is the probability of advancing from phase I to phase II in the 2 years following phase I. It takes the value of 1 if the drug advanced to phase II and 0 if not. An alliance project is defined as a project that was originated by another company and an alliance contract was signed. 1) Cash is the amount of cash (deflated to year 2000 Dollar terms) the company has scaled by the total number of projects the company had 3 years earlier. 2) All Projects - is the total number of projects in the firm scaled by the firm's sales (deflated to year 2000 Dollar terms). 3) Same Therapy Projects - is the total number of projects in the firm targeting the same therapeutic class scaled by the firm's sales (deflated to year 2000 Dollar terms). The same controls are used as described in Table 5. The t-statistics, reported in parentheses, are based on robust standard errors that are corrected for firm-level clustering.

	(1)	(2)	(3)	(4)	(5)
Alliance	0.0860 (2.61)**	0.1081 (2.32)**	0.1191 (2.71)**	0.1635 (3.18)***	0.1771 (3.10)***
Cash	0.0012 (8.02)***	0.0019 (6.24)***	0.0019 (6.21)***	0.0013 (4.45)***	0.0023 (6.08)***
Cash*Alliance	-0.0010 (7.15)***	-0.0016 (6.01)***	-0.0015 (5.88)***	-0.0010 (4.10)***	-0.0020 (5.77)***
All Projects		-0.0039 (1.95)**	-0.0030 (1.98)**	-0.0070 (2.92)***	-0.0089 (2.96)***
All Projects*Alliance		0.0028 (1.90)**	0.0034 (1.95)**	0.0068 (2.81)***	0.0081 (2.72)***
Same Therapy Projects		-0.0192 (1.84)*	-0.0190 (1.80)*	-0.0137 (2.01)**	-0.0134 (2.05)**
Same Therapy Projects*Alliance		-0.0015 (1.99)**	-0.0017 (1.97)**	-0.0074 (2.52)***	-0.0094 (2.85)***
Novelty	0.0132 (1.90)*	0.0121 (1.97)**	0.0110 (2.12)**	0.0106 (2.23)**	0.0376 (2.65)**
Bio	0.0635 (1.95)*	0.0743 (2.17)**	0.0760 (2.32)**	0.0734 (2.41)**	0.0980 (2.60)**
Market Size - Large	0.0669 (1.01)	0.0139 (1.17)	0.0036 (1.04)	0.0954 (2.11)**	0.1250 (2.03)**
Market Size - Small	-0.0275 (1.76)*	-0.0360 (1.39)	-0.0316 (1.15)	-0.0361 (1.26)	-0.0241 (0.81)
Observations	2985	2985	2985	2985	2985
R^2	0.13	0.14	0.15	0.16	0.27
Company FE	No	No	Yes	Yes	Yes
Therapeutic Class FE	No	No	No	Yes	Yes
Year FE	No	No	No	No	Yes