

**Patents and licensing and the commercialization of academic biomedical research**

by

Richard George Wehby

Ph.D., Biomedical Sciences  
University of Connecticut Health Center, 1998

A.B., Biology  
Princeton University, 1991

Submitted to the Harvard-MIT Division of Health Sciences and Technology  
and the Sloan School of Management in Partial Fulfillment for the Degrees of

Master of Science in Health Sciences and Technology  
and  
Master of Science in Management of Technology  
at the  
Massachusetts Institute of Technology  
June 2004

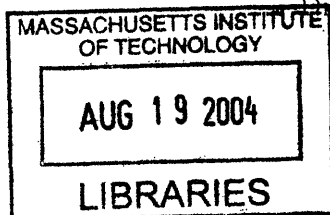
© 2004 Massachusetts Institute of Technology  
All rights reserved

Signature of Author.....  
Biomedical Enterprise Program  
Harvard-MIT Division of Health Sciences and Technology, and  
MIT Sloan School of Management  
May 7, 2004

Certified By.....  
Fiona E. Murray  
Assistant Professor in Management of Technological Innovation & Entrepreneurship,  
MIT Sloan School of Management  
Thesis Supervisor

Certified by.....  
R. Rox Anderson  
Associate Professor of Dermatology, Harvard Medical School, and  
Research Director, Wellman Laboratory of Photomedicine, Massachusetts General Hospital  
Thesis Supervisor

Accepted by.....  
Martha L. Gray  
Edward Hood Taplin Professor of Electrical and Medical Engineering, and  
Director, Harvard-MIT Division of Health Sciences and Technology



**ARCHIVES**

# **Patents and licensing and the commercialization of academic biomedical research**

by

Richard George Wehby

Submitted to the Harvard-MIT Division of Health Sciences and Technology  
and the Sloan School of Management  
on May 7, 2004 in Partial Fulfillment for the Degrees of  
Master of Science in Health Sciences and Technology  
and Master of Science in Management of Technology  
at the Massachusetts Institute of Technology

## **ABSTRACT**

This thesis is part of a larger body of research being undertaken by Dr. Fiona Murray and colleagues examining value creation and sharing between and among the three principal players in the commercialization of academic biomedical research: universities, biotech firms, and big pharma. The Recombinant Capital database provided access to contracts for biomedical technology licensed from academe to biotech, and also subsequent contracts that included that same technology from biotech to big pharma. These two contracts comprise a contract “pair”. Importantly, these contract “pairs” were unredacted, that is, all parts of the contracts, including the commercial terms, were available. This thesis will lay the foundation for later work by examining the contracts between university and biotech, from the University’s point of view. The goal is to identify factors that give the university more power in a pricing negotiation, and that predict higher economic value for the contract. The Specific Aim is to determine if certain University factors have a significant effect on predicting the economic value of the university-biotech licensing agreement. Four groups of readily quantifiable factors that contain attributes that might add power to the University in its pricing negotiation with the Biotech firm were identified: Institutional factors, Single Inventor factors, Aggregate factors, and Invention factors. The hypothesis is that at least one of these factors will have a significant effect on predicting the value of the licensing agreement, as determined using ordinary- and multiple linear regression models. In formulistic terms, the null- and test- hypotheses are: (H0) no factor has a significant effect on predicting economic value, and (H1) at least one factor has a significant effect on predicting economic value. A multiple regression model of the factors as explanatory variables for the economic value of the license revealed that two independent university factors significantly predict economic value of the contract. These combined factors account for 64% of the variance of the dependent variable (in excess of control), and have coefficients that are significant ( $p < 0.001$ ). The results are discussed in the context of its importance to university technology transfer officers, biotech firms and venture capitalists.

Thesis Supervisor : Fiona E. Murray

Title: Assistant Professor in Management of Technological Innovation & Entrepreneurship,  
MIT Sloan School of Management

Thesis Supervisor: R. Rox Anderson

Title: Associate Professor of Dermatology, Harvard Medical School, and  
Research Director, Wellman Laboratory of Photomedicine, Massachusetts General Hospital

*To my wife,  
Anne Armour Wolf*

## ACKNOWLEDGEMENTS

One of the benefits of attending the Biomedical Enterprise Program and of researching and writing this thesis is the opportunity provided to meet and interact with men and women of the highest scholarship, accomplishment and character. I extend my grateful appreciation to all who have aided me along the way, but would like to mention several particularly key people.

First, I would like to thank my advisors, Fiona Murray and Rox Anderson, for their scholarship, insights, guidance and supervision. I would also like to thank Lita Nelsen of the MIT Technology Transfer Office for serving as outside reader, and Mark Edwards of Recombinant Capital for generous access to their database of contracts. Without these individuals this thesis could not have been done.

I would also like to thank Anita Perkins, Reference Librarian at Dewey Library, and Lucretia McClure, Assistant to the Director at Countway Medical Library for their help with various databases that were instrumental for the analyses.

My appreciation must also be extended to the faculty, staff and advisors of the Biomedical Enterprise Program, for their unwavering support during the past two years. In particular, I would like to thank Joe Bonventre, Beverly Dammin, Bonnie Fendrock, Martha Gray, Cher Hugus, Stan Lapidus, Jennifer Mapes, Tony Sinsky, Jim Utterback, Marsha Warren and Dave Weber.

My colleagues in the program have been a source of camaraderie and inspiration and we have formed relationships that will last our lifetime. I would especially like to thank the members of my BEP inaugural class: Adrian Gottschalk, Jim Graham, Peter Smith and Matt Strobeck.

Finally, I would like to thank my wife, Anne Armour Wolf, for her patience, love and support throughout these past two years.

## Table of Contents

Abstract .....	2
Dedication .....	3
Acknowledgements .....	4
Chapter 1. Introduction.....	6
Chapter 2. Literature Review .....	9
Chapter 3. Methods.....	15
Chapter 4. Summary Statistics .....	32
Chapter 5. Test of the Regression Model.....	59
Chapter 6. Discussion .....	68
References .....	72
Richard G. Wehby Biographical Note .....	73

# Chapter 1

## Introduction

This thesis is part of a larger body of research being undertaken by Dr. Fiona Murray and her colleagues. They have begun an examination of value creation and sharing between and among the three principal players in the commercialization of academic biomedical research: universities, biotech firms, and big pharma (all terms are defined in Chapter 3, Methods). The first of what is expected to be a series of publications appeared last year (Edwards, Murray and Yu, 2003). That work will be discussed in depth in the Literature Review (Chapter 2). For now, the most important point is that they have access to a unique database and were able to obtain contracts for biomedical technology licensed from academe to biotech, and then a subsequent contract *that included that same technology* from biotech to big pharma. These two contracts comprise a contract “pair”.

Importantly, these contract “pairs” were unredacted, that is, all parts of the contracts, including the commercial terms, were available. The goal of Dr. Murray’s research is to use these pairs of contracts to understand in detail the process of taking an idea from its invention in the laboratory, through development and on to commercialization. In particular, it is hoped to shed light on the process of how economic value is created by the three players. Having commercial terms available will enable a measure of what the parties believed were equitable economic terms.

This thesis will lay the foundation for later work by examining the contracts between university and biotech, from the University’s point of view. The goal is to identify factors that give the university more power in a pricing negotiation, and that predict higher economic value for the contract.

*Specific Aim: To determine if certain University factors have a significant effect on predicting the economic value of the university-biotech licensing agreement.*

From prior work in the literature, interviews with licensing professionals and insight into the problem, four general groups of readily quantifiable factors (explained in further detail in Chapter 2, Methods) contain attributes that might add power to the University in its pricing negotiation with the Biotech firm:

1. Institutional factors:
  - a. Country
  - b. Carnegie Classification
  - c. NIH funding to the institution, and rank
  - d. Federal obligations for R&D to the institution

- e. Total Science and Engineering (including Life Sciences) R&D expenditures
2. Single Inventor factors
  - a. Number of articles published
  - b. Number of times papers were cited
  - c. Number of patents issued and patent applications filed
3. Aggregate factors
  - a. Total number of papers by all inventors of the licensed patent and/or patent application
  - b. Total number of citations for all inventors of the licensed patent and/or patent application
  - c. Total number of patents and/or patent applications by all inventors of the licensed patent and/or patent application
  - d. Number of patents and/or patent applications in the contract
  - e. Tangible property included in the contract
4. Invention factors
  - a. Stage of the invention
  - b. Field of the invention

Our working hypothesis is that at least some of these factors will have a significant effect on predicting the value of the licensing agreement, as determined using ordinary- and multiple linear regression models. Chapter 2 will explain why these are the correct models to use to test the hypothesis. In formulistic terms, the null- and test- hypotheses are

*H0: no factor has a significant effect on predicting economic value*

*H1: at least one factor has a significant effect on predicting economic value*

This hypothesis has either a “yes” or “no” answer, and thus is informative regardless of the result:

- If at least one factor is predictive of economic value, then this factor may be manipulated by the University or its agents, so as to further enrich the source of invention.
- If no factor turns out to be predictive, then other factors need to be considered, such as
  - the negotiating experience and tactics of the Technology Transfer professionals,
  - the need of the biotech firm for a specific enabling technology in its IP portfolio

Chapter 2 is a brief Literature Review, and Chapter 3 presents the Methods used in this work. Chapter 4 will provide sample statistics. Chapter 5 tests a multiple regression model of the factors as explanatory variables for the economic value of the license. The results show that two independent university factors can significantly predict the economic value of the contract. These combined factors account for 64% of the variance of the dependent variable, and have

coefficients that are significant ( $p < 0.001$ ). Chapter 6 is a General Discussion. It is hope that this work is not only of academic interest, but may also be useful to university technology transfer officers, biotech firms and venture capitalists.

## Chapter 2

### Literature Review

This thesis focuses on licenses between university and biotech, for university patents. As such, we here briefly review our current understanding of the factors affecting the degree of university patenting, and of license structures between both university and biotech, and also biotech to pharma.

We first show that Universities are more interested in drugs and medical technologies, and less interested in mechanical technologies, than overall inventors, and this difference has increased over time. Next we briefly discuss the Bayh-Dole Act, which provided incentives for universities and private enterprise to commercialize Federally-funded research.

Following this we present studies showing that universities are increasingly patenting inventions of less importance, and that this may be due in large measure to increased propensity of technology transfer officers to patent.

Finally, we look at an empirical study on licensing contract terms by Edwards, Murray and Yu (2003) that lays the foundation for this thesis.

#### *1. Universities have increased their “propensity to patent”*

Henderson and colleagues (1995) studied university patents from 1966 – 1988, and showed that the number of university patents grew at a faster rate than other patents. Further, since university patenting has increased faster than university research spending, the ratio of university patents to R&D more than tripled during this period. In contrast, domestic patenting was constant, but industry R&D spending increased greatly, so that the ratio of industry patents to R&D was halved over the same time period. Thus, for a given research effort the universities’ “propensity to patent” has increased significantly while the overall propensity to patent has decreased (Henderson et al., 1995).

This increase in university patenting has not been uniform: in 1988 Drug and Medical patents comprised about 35% of patents (up from less than 15% in 1966), chemical patents comprised about 25 – 30%, electronic and related were about 20 - 25%, mechanical patents about 10 – 15%, and about 5% other. In contrast, overall patenting is 30 –35% mechanical, 20 –25% each for chemical and electronics, 10 – 15 % other, and less than 10% drugs and medical. Universities are thus more interested in drugs and medical technologies, and less interested in mechanical technologies, than overall inventors, and this difference has increased over time (Henderson, et al., 1995).

## ***2. Bayh-Dole Act provides incentives to Universities for commercialization.***

There are certain committant events that may have contributed to this increase by Universities, such as increases in Federal funding for research and the rise of University technology transfer offices . Certainly among the strongest influences has been the Bayh-Dole Act passed by Congress in 1980 (P.L. 96-517), and University patenting and licensing has increased sharply since passage (Jensen and Thursby, 1998). This law gave universities and other non-profit institutions the right to own intellectual property rights to inventions funded with federal monies. The law was expanded in 1984 (P.L. 98-620), such that universities were permitted to assign these property rights to other parties.

Most university patents are for very basic research, and substantial additional investment is typically required to transform these research results into marketable products. Proponents of P.L. 98-620 reasoned that no private firms would invest in development of university technologies unless they could own the property rights of the underlying . The sharp rise in University patents after passage of the act is consistent with this interpretation but it is hard to rule out that several mutually reinforcing trends (e.g. Bayh-Dole act, increased Federal funding, rise of TTOs) have been at work (Jensen and Thursby, 1998; Henderson et al., 1995).

For example, prior to 1989 the top recipient of biotechnology patents was Merck, but by 1999 the combined University of California system held that spot (Edwards, *et al.*, 2003). Moreover, in the decade 1992 – 2002 twelve academic institutions were among the top 40 biotechnology patent generating entities, including Stanford University, MIT, MGH, and Scripps Research Institution (*ibid*). This increase in university patents was mirrored by an increase in university licenses: a survey conducted by the Association of University Technology Managers (AUTM) showed that for the period 1991 – 1996 licenses executed increased 75%, with 13,087 licenses executed during that period (Jensen and Thursby, 1998). A more recent AUTM survey shows that over 400 licenses were executed by the top 25 U.S. academic hospitals and research institutes in the year 2000. These licenses were granted to startups, as well as to small and large companies, and running royalty income in year 2000 reached \$110 MM (AUTM licensing survey, FY 2000; Edwards *et al.*, 2003).

## ***3. Size and rank of institution affects quality of patents.***

Henderson and colleagues (1995) proposed to alternative hypotheses to explain the increased propensity to patent. Either

*Increase in university patenting is due to increased rate of invention*

or

*Increase in university patenting is due to lowered threshold on what is patentable*

The authors approached this question by looking at the distribution of the number of citations received by university patents versus a random selection of patents. They demonstrated that over time the fraction of university patents with zero citations has greatly increased, indicating that they are less important, and strongly supporting the hypothesis that there has been a decline in the threshold for what is patentable by universities.

The authors also looked at whether there had been a shift in the source of university patents, that is, if the fraction of patents from large institutions had fallen. They determined that such a decline had occurred: the fraction of patents attributed to the top 4 institutions was 50% in 1965, but only 25% in 1988. Further, the 15 top institutions had superior patents than those of the random sample, while those from lower-ranked institutions did not, so that on average those institutions that patent more were receiving patents of higher quality. Nonetheless, even the top institutions had seen a decline in quality since about 1983.

Thus the authors demonstrated that the decline in importance and generality university patents during the period up to 1988 has at least two components: (1) the top institutions are producing patents of lower quality, and (2) an increasing fraction of university patents are coming from smaller institutions, which always had patents that were not as highly cited as those from larger institutions.

#### ***4. Technology licensing officers have increased their propensity to patent***

One concern raised by the increase in university licensing is that the nature of university research has changed, from more basic to more applied. Alternative explanations suggest that there is an increased willingness on the part of faculty and administrators to license, or that business now increasingly relies on external R&D. Thursby and Thursby (2000) addressed this issue by creating a three-stage input-output model of university technology licensing, each stage modeling a particular step in the process: *Disclosure*, *Patents*, and *License Agreements*.

Their dataset consisted of 65 research universities reporting their activity over the years 1994 - 1997, and accounted for 57% of all federal research support and 61% of industry research support. The sample accounts for 68% of licenses executed, 61% of disclosures, and 67% of new patent applications by the 132 respondents of the 1997 AUTM survey.

Their results suggest that there is a fairly large increase in the propensity of Technology Transfer officers to patent, and a small increase in the propensity of faculty to disclose new inventions.

They suggest that universities are “delving more ‘deeply’ into the available pool of commercializable inventions” and that “while many more technologies are being offered and licensed to industry, the proportion of licenses executed to those offered is falling” (Thursby and Thursby, 2000).

This increase in propensity to patent by university administrators is consistent with the reasoning behind the Bayh-Dole act, and strongly supports that it has been effective in promoting more commercialization of university research.

### ***5. The structure of therapeutics licenses over time***

The work by Henderson and colleagues and Thursby et al. have discussed university factors that bear upon patents and their licenses, but the nature of those licenses has not been elucidated. Yet for the licensees and licensors contractual terms are critical to future success in transforming ideas into commercializable products. We now discuss the results of an empirical study on licenses and their terms.

Edwards, Murray and Yu (2003) have reviewed a database of over 2000 licenses executed from 1975 to the present, and selected a subset of 265 licenses from university to biotech. The basis for selection was that all were licenses for therapeutics, and that publicly-filed contracts listing complete and unredacted economic terms were available. This subset included 119 distinct universities or other such academic centers licensing patents and/or know-how to 122 distinct biotechnology companies.

Licenses from university to biotech may include the following pre-commercialization payments:

- upfront fee
- license maintenance fee
- research payments,
- milestone payments,
- sublicensing revenue sharing.

Post-commercialization payments principally include:

- royalties (including a minimum annual royalty)
- sublicense revenue sharing.

The frequencies of these elements over time have remained fairly constant, except for the frequency of clinical milestone payments, which have increased in recent years compared with the late 1970s and early '80s. (See Figure 2-1)

Figure 2-2 shows that the economic terms for these licenses has changed over time. In general, the overall trend has been for higher payments to universities over time, but principally in upfront, sponsored research and license maintenance fees .

### ***6. Universities get a small share of ultimate economic value of successful therapeutics patents.***

The same dataset that was analyzed for university-biotech licenses was then scrutinized to find all the licenses executed by the 122 companies in the dataset. From this groups of 1500 alliances were extracted those licenses that explicitly sublicensed part or all of the specific university invention. This yielded a data set of 160 contracts in which a university invention licensed to the biotechnology firm in the first dataset was the basis for one or more subsequent sublicenses by

the biotechnology firm, usually to a large pharmaceutical firm. This dataset was further reduced to 112 licenses that had been filed with the SEC, and for which at least some of the economic terms were available. From this biotech-pharma dataset a third dataset was extracted, consisting of those companies for which the full economic terms of the downstream contract were known. Recall that the second dataset had been selected from a larger dataset of 265 contracts for which full upstream economic terms were known. This third dataset thus consisted of 36 “matched” instances where both upstream and downstream full economic terms were known.

For each contract, the amount of payment owed the university by the biotech firm, both pre-commercial and post-commercial, was calculated. These amount were a percentage of the total enriching payments received by the biotech firm from its commercial partner.

Taking into account both pre- and post-commercialization payments, the value split for sharing the profits was calculated at a ratio of 7:29:64, for university: biotech: pharma. Put another way, the pharma firm captures about two-thirds of the profits, and the remainder is split by biotech and university on an 80:20 basis. The authors further noted that those universities that required licensees to submit complete and unredacted sublicenses received significantly greater pre-commercial payments, but somewhat lower royalty payments (Edwards et al., 2003).

## ***7. Implications for this thesis***

The literature has established that university factors such as larger size and larger rank are predictive of more important patents, and that citations can be used as a measure of patent importance (Henderson et al). Also, university inventor factors, such as slightly increased propensity on the part of faculty to disclose inventions, can influence the degree of patenting (Thursby and colleagues). Finally, the terms of university-biotech therapeutics licenses are known, and the share of the profits to Universities have been shown to be small (Edwards, Murray and Yu).

There has been no work, however, on the relationship between university factors and the value of the contract. Presumably there are factors, such as those discussed by Henderson et al and Thursby and colleagues, that affect the value of those contracts. Given the relatively small proportion of the profits that accrue to universities, as demonstrated by Edward, Murray and Yu, any factors that can be identified that predictably increase contract value will be of interest to universities. If these factors are manipulable by universities, then they will have practical utility, as well.

The Specific Aim of this thesis is to determine if certain University factors have a significant effect on predicting the economic value of the university-biotech licensing agreement, from the University point of view. The hypothesis to be tested is presented in Chapter 3.

Terms of Agreement	Pre 1980 - 1986	1987 - 1990	1991 - 1994	1995 - Present
<i>Post commercial payments</i>				
Royalties	4% (n = 25)	5.1% (n = 43)	4.2% (n = 62)	3.9% (n=24)
Minimum Annual Royalty	\$13,438 (n = 8)	\$38,212 (n = 22)	\$50,392 (n = 34)	\$53,479 (n = 11)
Sublicense Revenue Sharing	37.4% (n = 9)	34.3% (n = 17)	28.4% (n = 27)	28.4% (n = 14)
<i>Pre-commercial payments</i>				
Upfront fee	\$20,085 (n=21)	\$40,655 (n=35)	\$48,649 (n = 53)	\$87,942 (n = 24)
Reserch Payments	\$409,321 (n=14)	\$434,467 (n = 22)	\$1,159,941 (n = 31)	\$585,323 (n = 18)
Maintenance fee (5 years)	\$39,041 (n=8)	\$53,333 (n = 15)	\$90,495 (n = 22)	\$183,909 (n=11)
Milestone payments	\$16,250 (n=2)	\$324,359 (n = 12)	\$445,017 (n = 25)	\$1,585,679 (n = 11)
Sublicense revenue sharing	46.6% (n = 8)	27% (n = 11)	23.4% (n = 21)	25.4% (n = 12)
Total number of Deals	n = 40	n = 70	n = 110	n = 45

Figure 2-1. Average economic terms of university-biotechnology company deals. Source: Fiona Murray

## Average Financial Terms of Downstream Commercialization Alliances

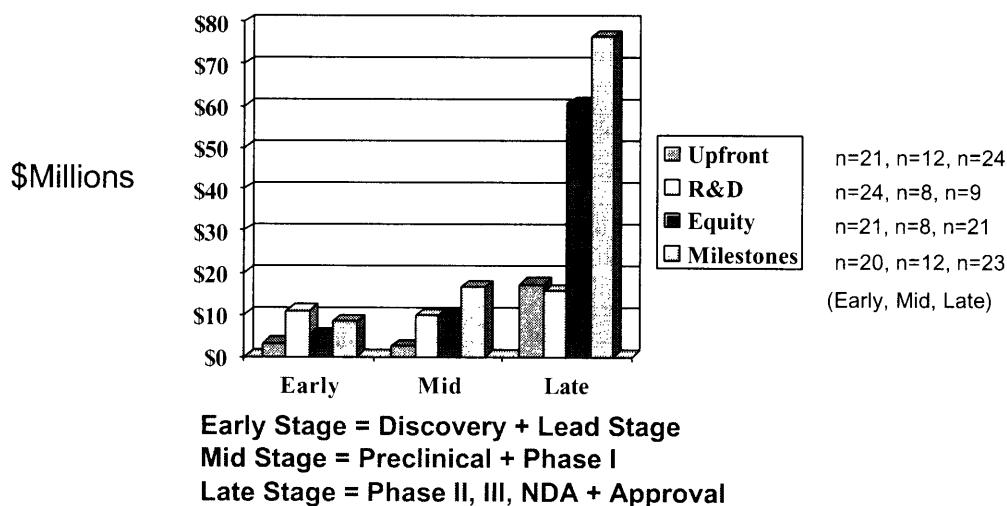


Figure 2-2. Average Financial Terms of Downstream commercialization alliances. Source: Fiona Murray

## Chapter 3

### Methods

The Specific Aim of this work is to determine if certain University factors have a significant effect on predicting the economic value of the university-biotech licensing agreement, from the University point of view. In formulistic terms, the null- and test hypothesis are, respectively,

*H0: no factor has a significant effect on predicting the economic value of the contracts*

*H1: at least one factor has a predictive effect on predicting the economic value of the contracts*

This hypothesis is tested with a multiple linear regression model. The dependent variable is the Expected Value of the contracts. The independent variables are the University factors.

This chapter will outline the methods used, including,

- terminology,
- how and why the sample was selected
- the source of contracts, and features analyzed
- the model used to calculate contract expected value
- the multiple regression model and its dependent and independent factors
- database sources
- statistical package used

#### 3.1 Terminology

##### *i. “University”, “Biotech”, and “Pharma”*

This work is part of a larger body of work by Dr. Murray and colleagues that looks at intellectual property flow, and deal flow in terms of three parties:

- University that invents and licenses a technology
- The Biotech firm that licenses the technology from the University and develops it to the point that
- The Pharma firm that licenses the technology from the Biotech firm (where development subsequently take place jointly or in pharma only).

In this and similar analysis these terms are defined as follows:

1. “University” is understood to mean a university, medical center, medical school, academic-affiliated hospital, research institute or other such public or private institution of higher learning that includes among its missions the discovery and public dissemination of new knowledge.

2. “Biotechnology” firms” is used loosely, and can refer to either biotechnology firms in the pure sense (*i.e.*, business focused on research and development of recombinant DNA, antibodies, genomics, proteomics, or other applications of “large molecule”), or a small-molecule firm *i.e.*, a relatively small research and development pharmaceutical company (capitalization < \$ 5 B) , that is, not a large-cap multinational pharmaceutical, nor a generics manufacturer or distributor.
3. “Pharma” is used to describe large-cap (capitalization > \$ 5 B) multinational pharmaceutical firms, with fully-developed research, development, manufacturing, sales, marketing forces, as well as experience in conducting clinical trials (either by itself, or through a CRO) and successfully bringing new drugs to market. It does not include companies dedicated solely to the manufacture and sale of generic drugs.

***ii. “Upstream” and “Downstream”***

This IP-flow and deal-flow can be thought of from the point of view of the biotech firm: deals with the University, to in-license the technology, are “upstream”. Deals with Pharma firms, where they out-license the technology, are “downstream”. In this work we will focus on deals between Universities and Biotech firms, or what are called “upstream” deals.

***iii. “Early”, “Mid”, and “Late” stages of development.***

To remain consistent with prior analyses, we follow the usage of Edwards , Murray and Yu (2003) who defined the stages as:

Early: Discovery and Lead Generation  
 Mid: Preclinical and Phase I  
 Late : Phase II, III and approval

These stages are shown in the drug-development pipeline in Figure 3-1.

**3.2 How and why the sample was selected.**

In this work we focus on the University factors affecting upstream contract value. Subsequent work will want to examine downstream contracts, to see what the Biotech firms do to add value. Having a matched set of contracts is essential for comparisons between upstream and downstream contracts. A sample of 36 such matched sets was identified by Edwards, Murray and Yu (2003). In this work we will restrict our analysis to the upstream contracts in this dataset.

There are certain restrictions in the dataset:

- (1) The deals are restricted to what reasonably appear to be the development of therapeutic products. In general,

therapeutics have a higher valuation than diagnostics or test kits, so valuations based upon this data set can not be extrapolated to categories other than therapeutics.

- (2) The information is from documents the companies decided were material, and so disclosed to the SEC. Thus, there exists the possibility that many deals that were not considered material were not filed. This may skew the data towards larger values, as deals that are small may be below the threshold of “material”. Further, “material” is defined by each company with reference to itself.
- (3) From samples of 265 upstream and 112 downstream contracts they found only 36 contracts that met the criteria of revealing complete and unredacted terms. This is less than 18% of the over 2000 contracts contained in their university-biotech company data set.

### **3.3 Source of contracts and features analyzed.**

Through the collaboration of Dr. Murray of MIT Sloan and Mr. Mark Edwards of Recombinant Capital, we have gained access to a web-accessed database of over 2000 contracts filed with the Securities and Exchange Commission of the United States of America. This has permitted us to review and analyze the contracts described above.

The contracts are all unique, but nonetheless share features common to all such as, precise definitions of terms used, identification of the licensed product(s) and the field, territory and term for which the license will be valid, monetary consideration in exchange for the grant of license, and conditions for termination of the agreement.

Contract features bearing upon the hypothesis were analyzed, including:

1. Contract Structure
  - a. Type of contract (e.g., license, research collaboration, co-marketing, co-development, etc.)
  - b. Licensed products, processes, and know-how
  - c. Field, territory, term and exclusivity of use
  - d. Tangible property transferred
  - e. Number of patents or other intellectual property in the contract
  - f. State of intellectual property at the time of contract signing (i.e., patent application, patent issued)
2. Monetary aspects of the agreement
  - a. Upfront fees
  - b. License maintenance fees
  - c. Research payments
  - d. Milestone payments, and their conditions
  - e. Royalty payments, and their conditions
  - f. Sublicense royalty payments, and their conditions
3. Names of inventors

These contract features were used to value the contract (see below) and to identify institutions, inventions and inventors for further analysis.

### **3.4 The Contract valuation model, from the University point of view.**

The model for payments expected by the university, and hence value of the deal for the university, are based on the standard drug pipeline model for therapeutics (See Figure 3-1; Murray, 2003). The model uses decision trees (Figure 3-2) to calculate the value at each node. This is the standard valuation method in the industry, and takes into account the cash flow at each node and the probability of successfully passing each stage to calculate the expected value at each node (Figure 3-3). The expected values are discounted back to calculate the present value of the deal. The model contains certain simplifying assumptions:

1. Assume patents are valid and do not infringe, and that patent applications issue. If minimum annual royalty is based on issuance of patent, assume it issues four years after filed. Further assume that, if necessary, licensed compound is further modified and protected by additional patents that permit a total of 14 years of sales (See sales model, below).
2. License maintenance fees terminate when royalties begin
3. Sales only in country where patent valid, and of product based on the licensed patent(s).
4. One therapeutic product for one therapeutic indication
5. Royalty payments to university as per contract: until expiry of the last of the patent rights (typically 20 years from filing date), or until such time as specified in the contract.
6. One sublicensee, who pays an upfront fee of \$1 MM at Phase IIA (development year = 6). This subject to revenue sharing by university, only if not credited against royalties.
7. Sales evenly between licensee (50%) and sublicensee (50%)
8. Royalty rate from sublicensee to licensee assumed to be university royalty rate + 2%.
9. No payments to university on milestone payments made by sublicensee to licensee.
10. Industry averages (See Figure 3-2) for:
  - a. For phases of development
  - b. For probability of success in each phase (Figure 3-3)
11. Assume probability of launching sales, once registration is achieved, is 100%
12. Typical industry sales curve (Murray, 2003 ):
  - a. Log rise to peak sales, 3 years of peak sales, log fall to zero sales upon entry of generic
  - b. Peak sales in year 6, 7,8
  - c. Rise in sales from year 0 – 6, fall in sales from year 9-14
13. Assume moderately successful peak sales of \$300 million per year (contrast with blockbuster sales of > \$1 billion per year; Murray, 2003).
14. Only one discount rate for the project, 12%. This is the historical risk-free rate (5%) plus a 7% market risk premium. We believe this is a reasonable discount rate given that additional project risk is taken into account by the probability of success at each phase.

15. Assume net sales are proportional to gross sales. Many contracts are written such that royalties are payable on net sales. We have no way of calculating sales net of such items as trade or quantity discounts, credits or allowances, sales taxes and other taxes and duties, transportation, insurance, and bad debts, all of which are commonly excluded from gross sales to define net sales. Our royalty model is based on gross sales, and we assume that a simple shift in baseline (effectively, a change in the value of the constant in the regression formula) would relate gross and net sales.
16. Equity provided to the University at the signing of the contract is valued at the price stated in the contract. If no price is stated (only number of shares stated), assume a price of \$1/share. If equity is promised at a future financing date, equity is priced at the maximum possible contracted price, and said future financing is assumed to occur one year from signing of contract. Presumably these shares are priced as discounted future earnings, so now further discounting or modeling is required.
17. If the biotech firm assumes "all patent costs", this is allocated as an upfront fee of \$50,000.
18. If the contract is for an option, assume that biotech firm further contracts for the license at one year after signing the option.

The model incorporates the payment amounts (upfront and maintenance fees, milestones, royalties from licensees and sublicensees) and schedules, as stated in the contracts, to calculate the value of the deal to the University. An example of one such calculation is (for the University of Florida/ Advanced Tissue Sciences contract) is shown in Figure 3-4, and the predicted sales curve is shown in Figure 3-5.

**A drug company has a drug (drugs) in development at distinct stages in its life**

**To quantify the value of, one must:**

- Define output of each stage
- Assign a value to each output at every stage

**Result: all molecules, from early hits and leads to late lifecycle drugs have an NPV value**

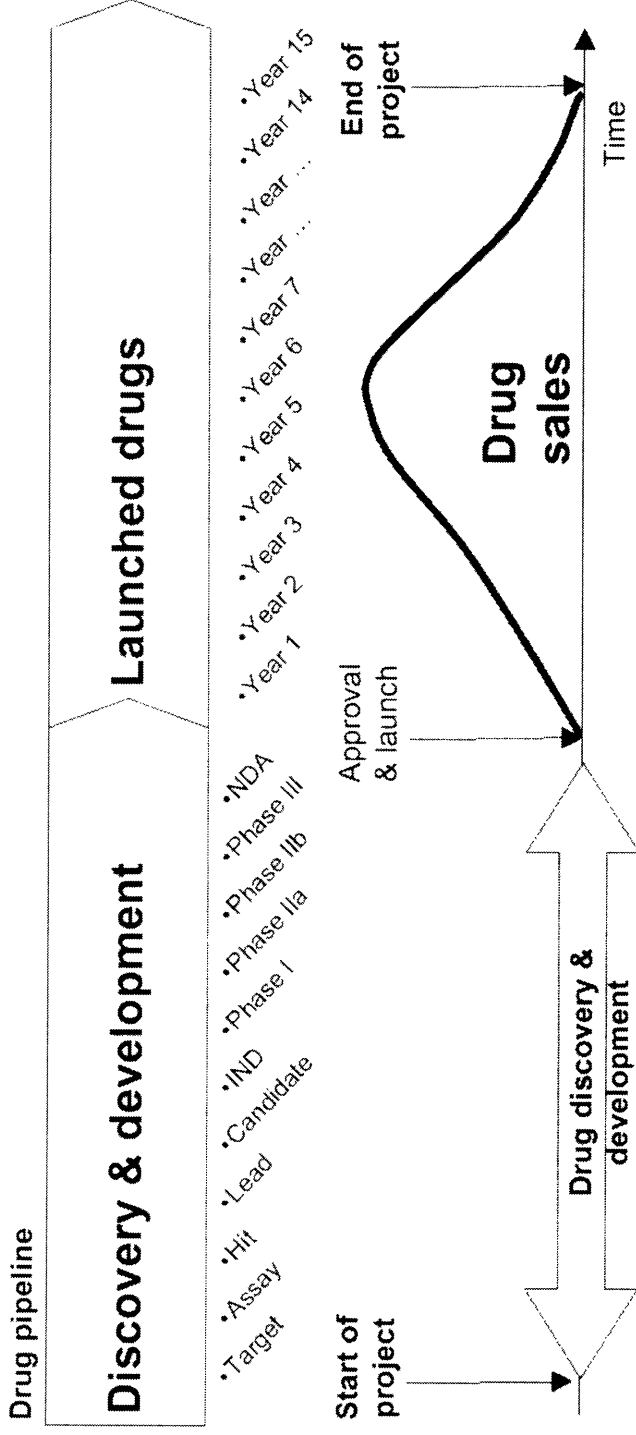


Figure 3-1. Drug pipeline model used for valuation. Source: Murray, 2003.

## Expected Value for each Phase of Development

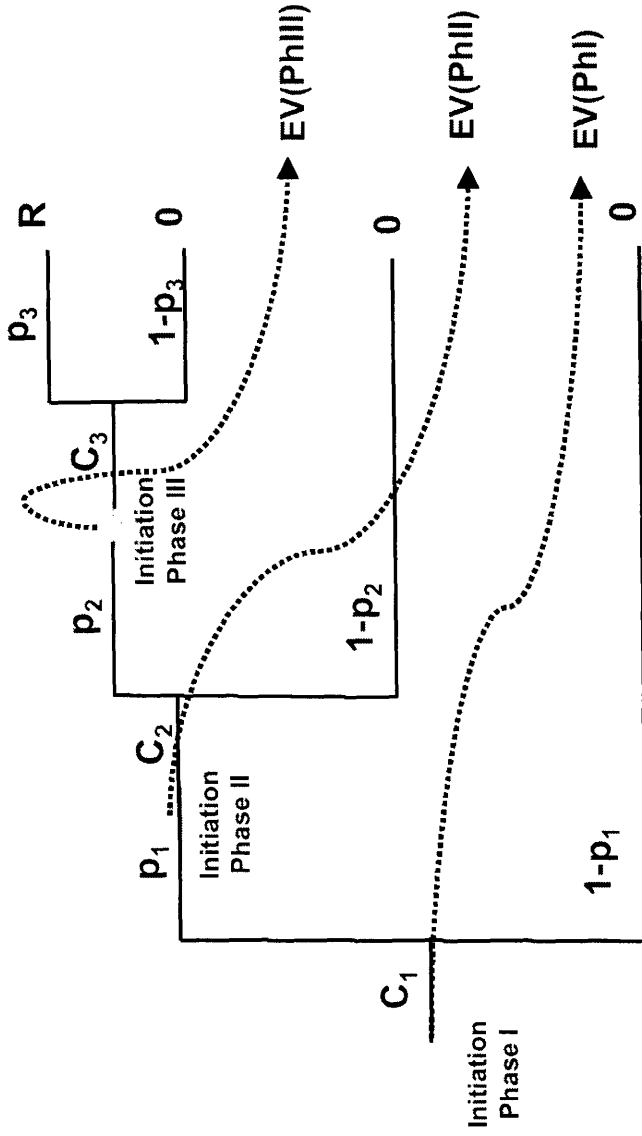


Figure 3-2. Using decision trees to calculate expected value. Each node has a probability of success (" $p_n$ ") and a probability of failure (" $1-p_n$ "). The product of the CashFlow (" $C_n$ ") and the probability of success is the expected value (EV) at each node. Source: Murray, 2003.

# Current Drug Development is Inefficient with low POS

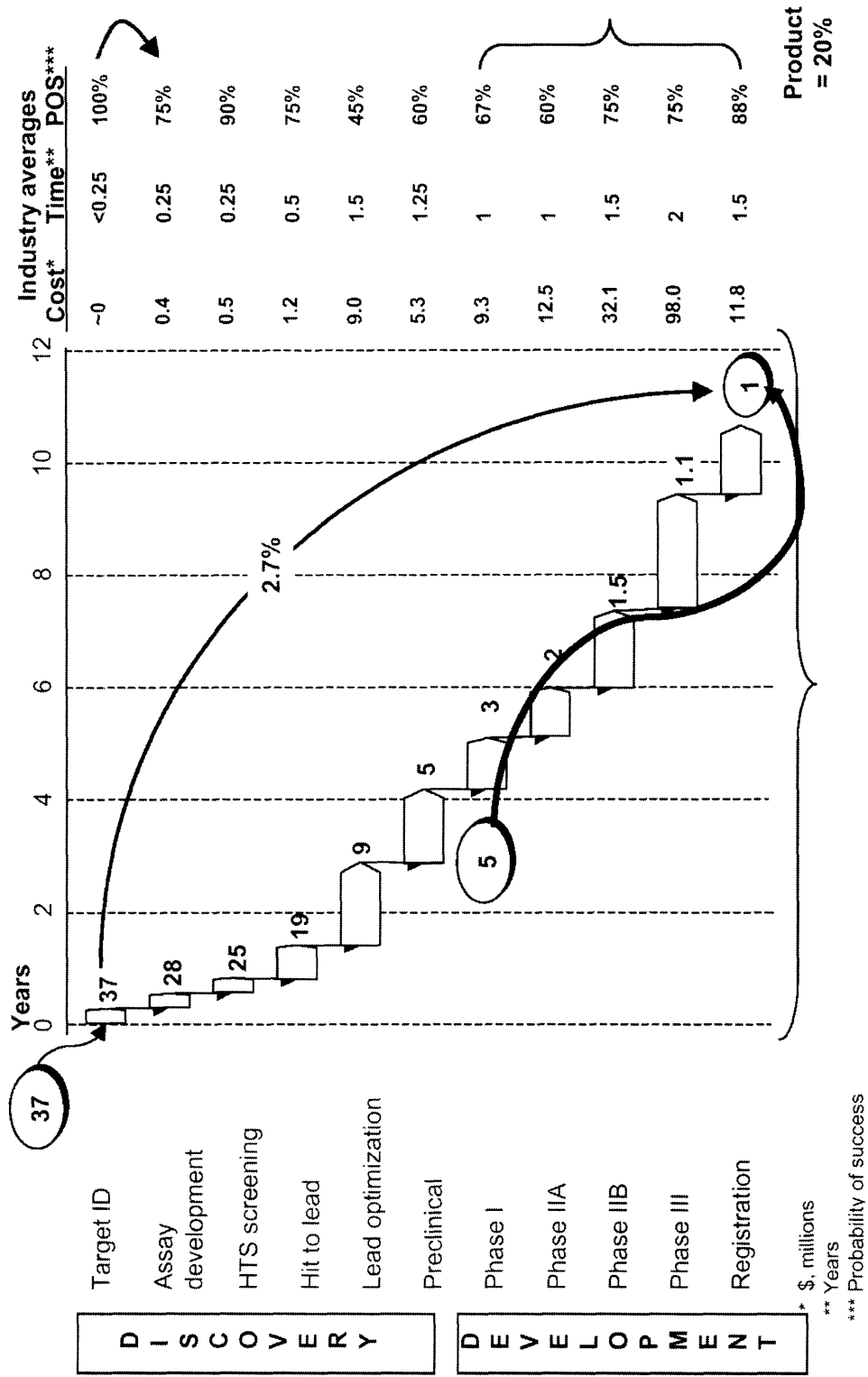


Figure 3-3. Industry-average time in phases, costs, and probability of success. Data for time and POS are used in the present model. Source: Murray, 2003.



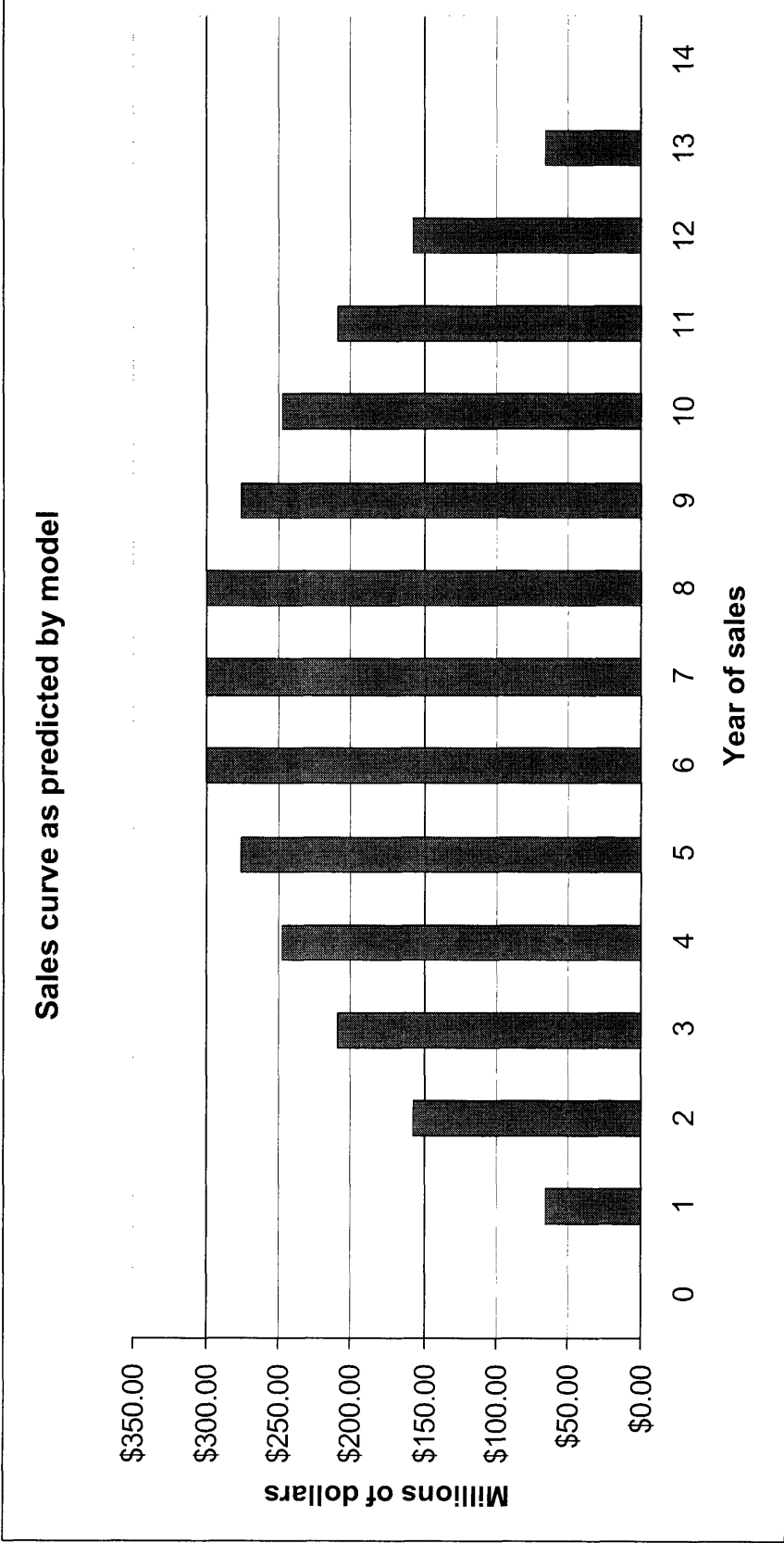


Figure 3-5. Sales curve predicted by model. See text for model assumptions.

### 3.5 Linear regression model

Our hypothesis is that certain factors can significantly predict contract value. The appropriate model for such a hypothesis is a linear regression model (Rosner, 2000). In our model the contract value is the dependent variable and the University factors are the independent variables:

$$\text{Contract value} = (\text{coefficient} \times \text{UniversityFactor 1}) + (\text{coefficient} \times \text{UniversityFactor 2}) + \dots + \text{constant}$$

The contract value is the Expected Value (EV), as determined from the Deal Valuation Model presented above.

The independent variables are certain factors that might give bargaining power to the University in a pricing negotiation. Prior work in the literature (Henderson et al., 1995; Thursby and Thursby, 2000; Edwards, Murray and Yu, 2003) interviews with licensing professionals and insight into the problem have identified four general categories in which the factors may be grouped:

1. Institutional factors:
  - a. Country
  - b. Carnegie Classification
  - c. NIH funding to the institution, and rank
  - d. Federal obligations for R&D to the institution
  - e. Total Science and Engineering (including Life Sciences) R&D expenditures
2. Single Inventor factors
  - a. Number of articles published
  - b. Number of times papers were cited
  - c. Number of patents issued and patent applications filed
3. Aggregate factors
  - a. Total number of papers by all inventors of the licensed patent and/or patent application
  - b. Total number of citations for all inventors of the licensed patent and/or patent application
  - c. Total number of patents and/or patent applications by all inventors of the licensed patent and/or patent application
  - d. Number of patents and/or patent applications in the contract
  - e. Tangible property included in the contract
4. Invention factors
  - a. Stage of the invention
  - b. Field of the invention

These factors were used in the following way:

1. *Country*: This factor was coded as 1= U.S., 0=non-US.

2. *Carnegie Classification*: The size or focus of the institution may bear upon the value of the contract. The Carnegie Classification is a classification of institutions based upon descriptive data. The Classification was developed by the Carnegie Commission on Higher Education in the early 1970s to aid in its policy studies. The classification was first published for public use in 1973, and has been revised in 1976, 1987, 1994, and 2000. The institutions are classified at the time the deal was structured (i.e., if a deal was signed in 1992, the classification in effect at that juncture, in this case based on the 1987 publication, is used.

Most of the institutions in the sample are classified as Research Institution I, and these definitions were slightly different in 1987 and 1994:

1987 These institutions offer a full range of baccalaureate programs, are committed to graduate education through the doctorate degree, and give high priority to research. They receive annually at least \$33.5 million in federal support and award at least 50 Ph.D. degrees each year.

1994 These institutions offer a full range of baccalaureate programs, are committed to graduate education through the doctorate degree, and give high priority to research. They award 50 or more doctoral degrees each year. In addition, they receive annually at least \$40 million or more in federal support.

Partly as a consequence of the definition changes, the number of institutions classified as R1 and R2 by the Carnegie Foundation changed over time as the definition changed. The institutions in our study have remained stably in the same category over this time period. This factor was coded as R1=1, MED = 0.

3. *NIH Funding Levels and Ranking*: Federal grants specific for biomedical research may influence the size of the contracts. These rankings are for what the NIH calls “not-for-profit” institutions, with one exception. This exception is the University of Medicine and Dentistry of New Jersey, which is ranked in the category of “medical schools”. Total number of NIH recipients varies per year, but is on the order of 500, so the lowest rank can be considered to be 500. We looked not only total NIH funds, but also subsets, such as funds for Research Grants, Training Grants, Fellowships, and Contracts. The factors used in the regression were both the actual expenditures (dollars) and the NIH ranking of those expenditures in the year of the contract.

4. *Federal R&D obligations*: In addition to NIH funds, institutions can receive other Federal funds, such as from the National Science Foundation. These other types of grants may bear upon the value of the contracts. The factors used in the regression were the actual Federal obligations (dollars) in the year of the contract.

5. *Academic R&D Expenditures:* In addition to NIH and other Federal grants and contract, institutions have other sources of funds for R&D, including State and local funding, alumni and other donations, and interest from their endowments. Thus, the total Science and Engineering (including Life Sciences) R&D expenditures by institutions can be larger than total Federal R&D obligations. The R&D expenditure categories of interest are total Science and Engineering (including Life Sciences) , as well as breakdowns for Life Sciences (defined by NSF as agriculture + biological + medical + other), Biological Sciences- and Medical Sciences. The factors used in the regression were both the actual expenditures (dollars) and the NSF ranking of those expenditures in the year of the contract.

6. *Number of articles published by a single inventor:* The number of publications may reflect several inventor traits: intellectual fecundity, sufficient funding levels to carry out novel work and to hire assistants, diligence, as well as age and relative career development. Instead of looking at the total number of publications by an inventor, however, we looked at the total number of published articles. “Total publications” includes book chapters, invited reviews, abstracts from meetings, Letters to the Editor, etc., which although indicative of other academic duties and status, do not necessarily indicate that are actively producing a new body of work. The bulk of articles, however, are peer-reviewed and transmit the discovery of new knowledge. It is an assumption of this work that articles, containing new knowledge, may more likely lead to new inventions for which patent protection can be sought. As such, we determined the number of articles published 20 years prior and up to and including the year of the upstream contract. We restricted the time period to 20 years to avoid inadvertently including publications from different authors with the same names. We nonetheless noticed that some authors were listed with last name and first initial only, and that the publications were in widely disparate fields and journals, suggesting that several authors’ work was being pooled into one ambiguous listing. As such, we eliminated four authors with ambiguous SCI listing (K. Kurachi, S. Sinha, F. Takaku, and S. Watanabe) and this revised list includes 50 different authors. The factor used in the regression was the number of articles (up to the contract date) by the inventor on the patent with the single greatest number of published articles.

7. *Number of citations of published work in previous five years.* Although the number of articles might be a measure of the intellectual fecundity of an inventor, it provides little insight into the relative merit of those publications. Some authors are known to produce work in “packets” called “least-publishable units”, or LPUs. This is because the academic promotion system often depends on having a certain number of publications per year, as axiomized in “publish or perish”. Relative importance of the work, however, can be assessed by how often an author’s work is cited. A search for total number of citations is impractical, as electronic databases only provide aggregate citations up to the present day, and do not let you find aggregate citations within a given time interval. As such, the searching must be done by hand, using the Thomson ISI Science Citation Index catalogs for the citations in the years of interest, a laborious process. We decided to restrict our

search to the five years prior to the upstream contract, so as to determining if being “highly-cited” in this time period would affect the value of the contract.

For example “Total number of times that publications by Anthony E. Pegg were cited between 1997 and 1993” is a type search that must be done by hand, using either the annual and/or cumulative five-year volumes of the print version of Science Citation Index. For a highly-cited author like Dr. Pegg, it would have taken too long to individually count all citations. Counting the number of columns of print that cited the author approximated the number of citations. The number of columns was then multiplied by the approximate number of citations per column of print (determined by sampling and averaging), where one column of printed citations = 160 citations. Thus, if Dr. Pegg has citations adding up to  $18\frac{3}{4}$  columns, then  $18.75 \times 160 = 3000$  citations.

The factor used in the regression was the number of citations in the previous 5 years of the inventor on the patent with the single greatest number of citations in this time period.

*8. Number of prior patents issued or applications filed per inventor:* Might the relative “inventiveness” affect contract value? This might be a measure of commercial sophistication on the part of the inventor, or it may signal ability to reduce abstract ideas to practice, a skill necessary for effective patent prosecution and which could increase the commercial value of a patent. For each inventor we determined the number of U.S. patents issued, filed, or claiming priority from a filing (foreign or domestic) prior to the upstream contract date. The factor used in the regression was the number of patents and/or patent applications filed (up to the contract date) by the inventor on the patent with the single greatest number of such patents and/or applications.

*9. Total number of patents, citations, and articles, by inventors per contract.* The value of the contract may depend on the aggregate value of having a patent from a highly successful team, with many notable partners, rather than only one single star player. In these cases, the factors for the regression model are

- a. sum of all patents and/or patent applications by inventors, which we take as a measure of commercial experience, or a “contract commercialization index”
- b. sum of all citations in the prior 5 years by the inventors, which we take as a measure of the scientific impact of the inventors, or a “contract impact index”
- c. sum of all articles published by inventors, here taken to measure the scientific fecundity, novelty or originality of the inventors, or a “contract novelty index”

*10. Total number of patents and/or patent applications in a contract.* The value of the contract may depend on how many different pieces of intellectual property are included in the contract. The factor used in the regression model was the total number of patents and patent applications included in the contract

*11. Tangible property included in the patent.* Receipt of tangible property, such as hybridoma cell lines that produce a certain monoclonal antibody, might increase contract

value. This factor was coded as 0=no tangible property in contract, 1=tangible property in contract.

*12. Stage of the invention.* An invention that is further in its development has mitigated some of the risk, and so should be more highly valued. This factor was coded as 0=Early, 1=Mid. There were no late-stage inventions in the contract, see Chapter 4: Sample Statistics.

*13. Field of the invention.* Certain fields might have greater needs or larger markets, and so should be more highly valued. This factor was coded with two Dummy variables :

Dummy1 = 1 if cancer, 0=otherwise, and

Dummy2 = 1 if immune, 0 = otherwise.

These factors are by no means exhaustive, but are meant to provide a first attempt at explanatory variables for hypothesis testing.

### **3.6 Comparison with a random sample for representativeness.**

This work is part of a larger research task by Dr. Murray and colleagues whose work requires comparison between upstream and downstream unredacted contracts. Thus this sample of contracts was chosen because they form the upstream component of a matched pair of upstream and downstream contracts. Our upstream contracts (n=28 unique contracts) represent about 11% (28 / 265) of the entire unredacted upstream contract in the database. As a first attempt to determine the representativeness of our sample we will compare certain characteristics of our sample with a control set of 28 randomly-selected upstream contracts:

1. Distribution of signing dates of contracts
2. Proportion of foreign institutions
3. Proportion of Carnegie-classified institutions
4. Proportion of R1 institutions
5. Royalty rate, mean and median.

The random sample was selected from among the 237 remaining unredacted upstream contracts that did not have a matched, unredacted downstream contract. The list of these contracts also provided the royalty rates charged to the licensee.

### **3.7 Databases.**

Much of the information gathered to perform this analysis was obtained from web-based databases. These include:

**Stock prices:** Historical stock market prices, outstanding shares, etc. were obtained from Center for Research in Securities Prices (CRSP), available through the Wharton Research Data Services (WRDS) of the University of Pennsylvania Wharton School of Business, via an MIT Dewey Library subscription (<http://wrds.wharton.upenn.edu>). Market capitalization was determined by multiplying the price per share by the number of shares outstanding, that is, market cap = \$/share \* # shares outstanding.

**Company and business information:** Company information was obtained from company annual reports and press releases (via Lexis-Nexus database), SEC filings (via the SEC web site: <http://edgar.sec.gov>), as well as Thompson Research (via an MIT Dewey Library subscription), and Wharton's WRDS (<http://wrds.wharton.upenn.edu>).

**Academic Typology:** Classification of academic institutions was based *Carnegie Foundation for Teaching Classification*. The latest version is available on line at <http://www.carnegiefoundation.org/Classification/>. They provide additional research resources through their links: <http://www.carnegiefoundation.org/Classification/links.htm>.

**NIH Funding Levels and Ranking:** Data on NIH funding of Extramural research are publicly available on the NIH website (<http://grants1.nih.gov/grants/award/trends>).

**Federal R&D obligations:** The National Science Foundation's Division of Science Resources has compiled a database (WebCASPAR) that is available through their web site (<http://www.nsf.gov/sbe/srs/>).

**Academic R&E Expenditures:** These data are also available through NSF's WebCASPAR (<http://www.nsf.gov/sbe/srs/>).

**U.S. Patents and Patent Applications:** Information about U.S. patents and patent applications were obtained via the United States Patent and Trademark Office web site (<http://www.uspto.gov>).

**Published citation search:** Thomson ISI's Science Citation Index provides online information on papers published from 1945 – 2004. Available information includes numbers of papers published by an author, as well as number of times a particular paper has been cited by other authors. Unfortunately, the citation search is cumulative to the present day, so one cannot limit the search to a particular time period. This latter type of search must be done by hand using the print edition. Both the online and print versions were available HMS Countway libraries through the ISI "Web of Science" (<http://isi1.isiknowledge.com.ezp3.harvard.edu/portal.cgi/wos/>).

### **3.8 Statistical software package.**

The STATA statistical software package for Windows, Release 8 (Intercooled) was used for statistical analyses (StataCorp., 2003). Unless otherwise stated, mean values are expressed as “mean  $\pm$  s.e.m.”, not “mean  $\pm$  s.d.”

## Chapter 4

### Sample Statistics

The regression model presented in Chapter 3 tests a dependent variable (EV) against several independent explanatory variables (University factors). This chapter provides sample statistics of these dependent and independent variables.

#### 4.1 Sample Breakdown

The sample of 36 pairs of contracts were drawn between

- Number of unique universities (or equivalent) = 25
- Number of unique biotech firms = 23
- Number of unique pharma firms = 28

These parties contracted 28 unique upstream contracts and 33 unique downstream contracts. This thesis focuses exclusively on the 28 unique upstream contracts.

#### 4.2 Dependent Variable: Expected Values

Using the EV model that was proposed in Chapter 2 of what the University might expect, we evaluated these contracts from the University point of view. The EV for Development Phase, Sales Phase and Total EV are shown in Figure 4-1. Some contracts did not pay anything during the Development Phase: Univax paid nothing to Walter Reed during Development, and neither did Immulogic to UNC. The largest Development Phase EV was in the Sugen/NYU contract, with an EV of \$1.5 MM. For the dataset as a whole, the mean EV for Development and Sales were about the same, but the median EV was higher during the Sales than during the Development Phase. These differences between mean and median values suggest the distributions are not normal, which will be an important consideration for the linear regression model.

Deal Date	Contract	EV, Development Phase, \$	EV Sales Phase, \$	EV, Total, \$
Jan-85	Univ. Ill. / Alliance Pharma	\$ 182,332	\$ 435,294	\$ 617,625
Feb-87	Dana Farber / Procept	\$ 643,746	\$ 233,156	\$ 876,902
Apr-87	MIT /Immulogic	\$ 425,493	\$ 133,165	\$ 558,659
Jun-87	SUNY / Cortech	\$ 10,000	\$ 366,909	\$ 376,909
Jul-87	B&W / Athena	\$ 236,438	\$ 670,940	\$ 907,378
Dec-87	SRI / US BioScience	\$ 2,277	\$ 551,130	\$ 553,407
Jan-88	UBC/ QLT	\$ 50,619	\$ 145,098	\$ 195,716
Aug-88	Children's /Enzytech	\$ 118,761	\$ 11,826	\$ 130,587
Oct-88	Salk /Ligand	\$ 379,503	\$ 1,958,821	\$ 2,338,324
Apr-89	Baylor /Cephalon	\$ 578,049	\$ 290,196	\$ 868,245
Jun-89	NTIS / Genaera	\$ 48,922	\$ 465,166	\$ 514,088
Jul-89	UNC / Immulogic	\$ -	\$ 153,652	\$ 153,652
Jul-90	Princess Margaret / Immulog	\$ 78,882	\$ 276,574	\$ 355,456
Aug-91	Max Planck / Sugen	\$ 1,218,953	\$ 335,460	\$ 1,554,412
Aug-91	NYU / Sugen	\$ 1,472,553	\$ 335,460	\$ 1,808,013
Aug-91	Walter Reed / Univax	\$ -	\$ 336,428	\$ 336,428
Dec-91	UF / SunPharm	\$ 887,930	\$ 206,956	\$ 1,094,886
Dec-91	UWisc / Viagene	\$ 19,130	\$ 31,996	\$ 51,125
Jan-92	NDRC / Genetic Therapy	\$ 183,498	\$ 470,999	\$ 654,496
Jul-92	MIT /ATS	\$ 1,406,038	\$ 326,863	\$ 1,732,901
Nov-92	UF /ATS	\$ 428,493	\$ 317,548	\$ 746,041
Feb-93	B&W / NPS	\$ 402,814	\$ 163,896	\$ 566,710
Apr-93	UCSD / Imclone	\$ 17,309	\$ 71,886	\$ 89,195
Jan-94	ARI / Alexion	\$ 87,433	\$ 217,647	\$ 305,080
Sep-95	UMDNJ / BDS	\$ 21,683	\$ 174,139	\$ 195,822
Mar-96	Ohio State / AVI	\$ 146,485	\$ 244,853	\$ 391,337
Oct-96	Stanford / Rigel	\$ 119,660	\$ 40,974	\$ 160,634
Mar-98	Penn State / Procept	\$ 552,343	\$ 883,815	\$ 1,436,158
	<i>n</i>	28	28	28
	<i>min</i>	\$ -	\$ 11,826	\$ 51,125
	<i>max</i>	\$ 1,472,553	\$ 1,958,821	\$ 2,338,324
	<i>Mean</i>	\$ 347,119	\$ 351,816	\$ 698,935
	<i>SEM</i>	\$ 82,512	\$ 71,313	\$ 113,775
	<i>SD</i>	\$ 428,744	\$ 370,555	\$ 591,191
	<i>Median</i>	\$ 164,408	\$ 283,385	\$ 556,033

*Figure 4-1 Expected values of the contracts, along with their summary statistics.*

#### *4.2.1 No difference in EV across time.*

The work by Edwards, Murray and Yu (2003) demonstrated that contract terms have changed over time. This suggested that EV might have also changed over time. To determine if our sample showed changes in EV over time, we segregated the contracts into five-year bins (1985 – 1989, n=12; 1990-1994, n=12; and 1991-1995, n=4) and examined the three

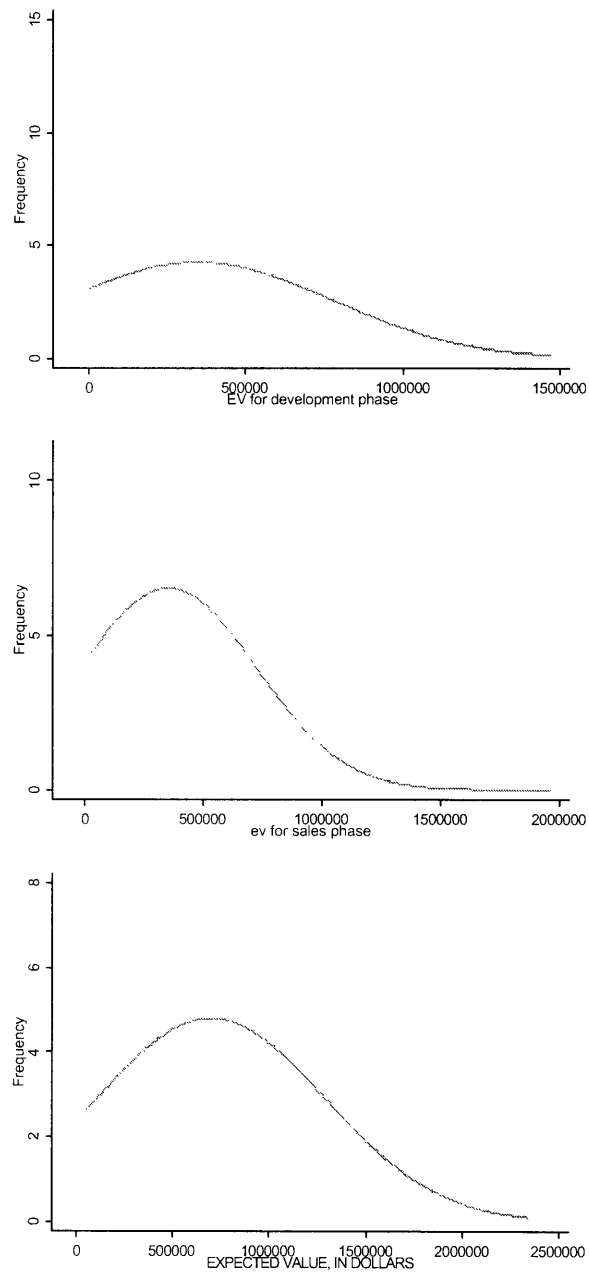
different EV categories: EV during the Development Phase, EV during the Sales Phase, and the Total EV. Note from Figure 4-2 that in 6/9 of the cases the SD was larger than the mean, indicating a large variance. For any given category (e.g., EV in Development Phase) there was no significant difference in EV across the three 5-year bins (ANOVA). As such, all the EV of a given category (e.g., total EV) can be aggregated and treated as one dataset of n=28.

5-yr bin	EV, Sales Phase	EV, Development Phase	EV, Total
1985-1989			
<i>mean</i>	\$ 451,279	\$ 223,012	\$ 674,291
<i>sd</i>	\$ 512,449	\$ 230,635	\$ 591,071
<i>sem</i>	\$ 154,509	\$ 69,539	\$ 178,215
<i>n</i>	12	12	12
1990-1994			
<i>mean</i>	\$ 257,643	\$ 516,919	\$ 774,562
<i>sd</i>	\$ 124,666	\$ 572,651	\$ 627,894
<i>sem</i>	\$ 37,588	\$ 172,661	\$ 189,317
<i>n</i>	12	12	12
1995-1999			
<i>mean</i>	\$ 335,945	\$ 210,043	\$ 545,988
<i>sd</i>	\$ 374,899	\$ 234,420	\$ 602,061
<i>sem</i>	\$ 216,448	\$ 135,342	\$ 347,600
<i>n</i>	4	4	4

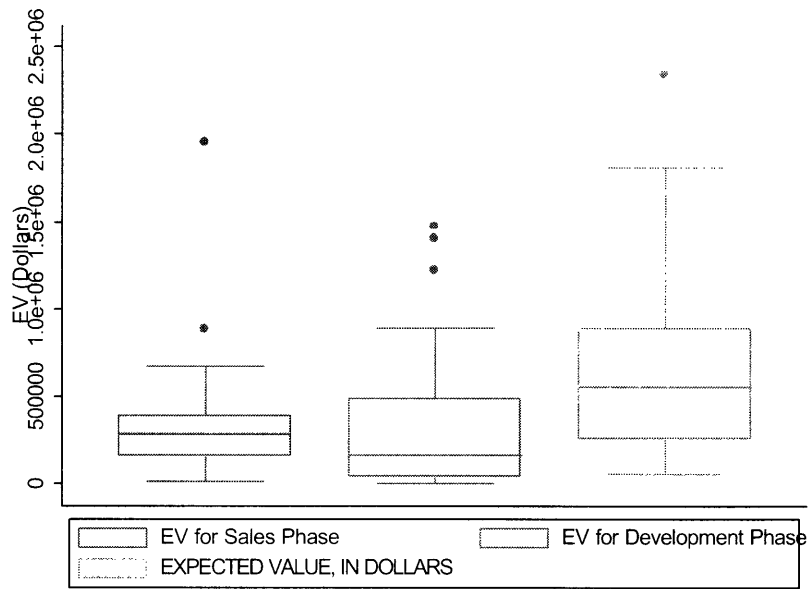
**Figure 4-2** EV for Sales- and Development Phases, and Total EV, by 5-year bin. For a given category (e.g., EV, Sales Phase) there was no difference in EV across time (ANOVA).

#### 4.2.2 EV Distribution, mean, quartiles

As suggested by the summary statistics of Figure 4-1, the distributions of EV during Sales- and Development Phases, as well as Total EV were not normal (Figure 4 –3). The box plots of Figure 4-4 show the quartile divisions, and indicate the presence of outliers. Both these figures indicate that the data must be normalized so as to not violate the requirements of a linear regression model. This is done in Chapter 5.



**Figure 4-3** Histograms showing distributions of EV for Development (top), Sales (mid) and Total EV (bottom). Each panel shows a normal curve in addition to the bars of the histogram. Note the distributions are not normal.



**Figure 4-4. Box plot for EV of Sales and Development Phases, and Total EV.** The bottom of the box is the 25<sup>th</sup> percentile, and the top of the box is the 75<sup>th</sup> percentile. A horizontal line inside the box is the 50<sup>th</sup> percentile (median). The vertical line connects the 75<sup>th</sup> percentile to the largest non-outlying value, and another vertical line connects the 25<sup>th</sup> percentile to the smallest non-outlying value. Outliers, defined as  $x > \text{upper quartile} + 1.5 * (\text{upper quartile} - \text{lower quartile})$ , are identified by a dot.

### 4.2.3 Types of contracts.

Not all contracts were of the same type. Most (26/28) included a license (the two contracts that were for options included the licensing terms, and were subsequently followed by licenses, so are included in the dataset). Many (11/28) included sponsored research (Figure 4-5). The most frequent type of contract was for a license only.

Number of contracts	Components
1	option
1	option + sponsored research
1	license + option
5	license + sponsored research + option
4	license + sponsored research
12	license only
2	license + equity
1	license + equity + sponsored research
1	license + cooperative research

**Figure 4-5. Contract components for total n = 28 contracts**

We divided all contracts into whether or not they included a sponsored research component: 11/28 included sponsored research, and 17/28 did not. Contracts that included a sponsored research component had significantly higher Developmental Phase EV than contracts that did not (\$680,000 vs. \$130,000, respectively, Student's t-test, one-tailed,  $p < 0.001$ ). The effect of the time-value of money and the probability of successfully reaching later stages resulted in Total EV that were twice as large for contracts containing sponsored research, than those that did not (\$991,000 vs. \$509,000, respectively, Student's t-test, one-tailed,  $p < 0.001$ ).

#### ***4.2.4 Diversity and Ranges of some contract terms.***

Contracts were not only of different types, but also differed in the terms they included and the value assigned to those terms. The most common items were royalties from licensees (28/28), expressed as a percentage of net sales.<sup>1</sup> Equally common were sublicensee royalties (28/28), which were specified in one of two ways: as a percentage of sublicensees net sales (13/28) or as a percentage of the royalty the sublicensee pays to the licensee (15/28). Only half the contracts charged an upfront licensing fee (14/28), but this was more than the number of contracts that included sponsored research (11/28).

These and other contract terms had higher or lower values, as shown in Figure 4-6. It is the hypothesis of this thesis that certain university factors can predict the expected value of the sum of all these terms. We now turn to these University Factors.

---

<sup>1</sup> Net sales are defined in each contract. A representative definition is "Licensee's billings for Licensed Products produced hereunder less the sum of the following: (a) Discounts allowed in amounts customary in the trade; (b) Sales, tariff duties and/or use taxes, excise taxes and all other applicable taxes directly imposed and with reference to particular sales; (c) Insurance and outbound transportation prepaid or allowed; and (d) Amounts allowed or credited on returns. No deduction shall be made for commissions paid to individuals whether they be with independent sales agencies or regularly employed by Licensee and on its payroll, or for cost of collections. Licensed products shall be considered "sold" when billed out or invoiced." (MIT/Immologic contract, April 3, 1987).

Contract Item	n / 28	min	max	Mean	SEM	SD	Median
Royalty from licensee, percent of net sales	28	0.50%	7.50%	4.19%	0.38%	1.95%	4.00%
Royalty from sublicensee							
Percent of sublicensee net sales	13	0.50%	7.50%	4.10%	0.69%	2.38%	4.00%
Percent of sublicensee royalty to licensee	15	5.00%	50.00%	25.37%	3.77%	14.11%	22.50%
Upfront licensing fee	14	\$ 10,000	\$ 150,000	\$ 55,000	\$ 13,620	\$ 49,107	\$ 27,500
Additional licensing fees (total)	4	\$ 10,000	\$ 950,000	\$ 316,875	\$ 247,772	\$ 429,153	\$ 153,750
License maintenance fees (total)	6	\$ 50,000	\$ 705,000	\$ 293,500	\$ 129,304	\$ 289,132	\$ 173,000
Sublicense issuance fee	1	\$ 10,000	\$ 10,000	\$ 10,000	\$ -	\$ -	\$ 10,000
Minimum royalty/payment prior to sales (total)	9	\$ 135,000	\$ 1,530,000	\$ 625,556	\$ 170,586	\$ 482,490	\$ 570,000
Pay All patent prosecution costs	5	-	-	-	-	-	-
Past patent prosecution costs	4	\$ 5,475	\$ 200,000	\$ 91,691	\$ 47,326	\$ 81,970	\$ 80,644
Patent Issuance fee	1	\$ 750,000	\$ 750,000	\$ 750,000	\$ -	\$ -	\$ -
Purchase of licensed property from inventor	2	\$ 50,000	\$ 250,000	\$ 150,000	\$ 141,421	\$ 141,421	\$ 150,000
Milestone	5	\$ 50,000	\$ 1,975,000	\$ 558,000	\$ 407,165	\$ 814,329	\$ 150,000
Sponsored Research	11	\$ 35,000	\$ 8,311,091	\$ 1,998,270	\$ 794,824	\$ 2,513,455	\$ 1,171,000
Equity	4	\$ 7	\$ 400,000	\$ 170,002	\$ 100,442	\$ 173,971	\$ 140,000
Revenue share (percent of non-royalty revenue)	8	2.0%	25.0%	18.1%	3.4%	9.0%	22.5%
Cooperative research agreement (no funding)	1	-	-	-	-	-	-

Figure 4-6. Terms present in the contracts in our dataset of university-biotech licenses, and their representative values.

### **4.3 Independent Variables: University factors**

In this section we provide sample statistics for the factors that might provide universities with increased negotiation power that were proposed in Chapter 3 (Methods). Recall that these were grouped into (1) University factors, (2) Inventor factors, (3) IP factors, and (4) invention factors.

#### **4.3.1 Institutional Factors**

##### ***a. Country***

Most institutions (20/25) were located in the United States, but the sample includes five from outside the U.S.: two from Australia (Austin Research Institute, Victoria; Princess Margaret Hospital, Perth), as well as one each from the U.K. (National Research Development Corporation, now British Technology Group), Canada (University of British Columbia, Vancouver) and Germany (Max Planck Institute of Biochemistry, Martinsreid).

##### ***b. Carnegie Classification***

Figure 4-7 shows the institutions in the sample set and their classification using the Carnegie criteria relevant at the time the upstream deal was made. Note that almost half the sample are not U.S. academic institutions as included in the Carnegie Classification, so that our sample of 25 unique “universities” by Carnegie standards contains only 13 academic institutions. Using this system, most (n=11/13) academic institutions were classified as “R1”, or “Research Institution I”.

As we shall see, however, the federal support received by the R1 institutions in our sample was in excess of the baselines quoted above, and so all institutions were stably classified as R1. Our sample of 13 constitutes 15 – 19 % of all R1 institutions during the time period of the upstream contracts.

Two institutions in the sample were classified as Med: Baylor College of Medicine and University of Medicine and Dentistry of New Jersey.

Most of the academic institutions in our sample were “Public” (10/13), and the three remaining (MIT, Stanford, Baylor) were “Private Not for Profit” institutions.

Institution	upstream deal date	FICE	Carnegie Classification	Control
University of Illinois at Urbana Champagne	Jan-85	1775	R1	Public
Dana Farber Cancer Institute	Feb-87			
MIT	Apr-87	2178	R1	PNP
State University New York Stony Brook	Jun-87	9555	R1	Public
Brigham & Women's Hospital, Boston	Jul-87			
Southern Research Institute (appears to be for the whole institute)	Dec-87			
University of British Columbia	Jan-88			
Children's Hospital Boston	Aug-88			
Salk Institute for Biological Studies	Oct-88			
Baylor College of Medicine	Apr-89	4949	MED	PNP
National Technical Information Service	Jun-89			
University of North Carolina, Chapel Hill	Jul-89	2974	R1	Public
Princess Margaret Hospital, Perth, Australia	Jul-90			
Max Planck Institute Biochemistry, Martinsreid, Germany	Aug-91			
New York University	Aug-91	2785	R1	Public
U.S. Walter Reed Army Inst. Of Research	Aug-91			
University of Florida	Dec-91	1535	R1	Public
University of Wisconsin at Madison	Dec-91	3895	R1	Public
National Research Development Corporation, UK	Jan-92			
MIT	Jul-92	2178	R1	PNP
University of Florida	Nov-92	1535	R1	Public
Brigham & Women's Hospital, Boston	Feb-93			
University of California at San Diego	Apr-93	1317	R1	Public
Austin Research Institute, Victoria, Australia	Jan-94			
University of Medicine & Dentistry of New Jersey, Newark	Sep-95	2620	MED	Public
Ohio State University, all campuses	Mar-96	8802	R1	Public
Stanford University	Oct-96	1305	R1	PNP
Pennsylvania State University (University Park)	Mar-98			

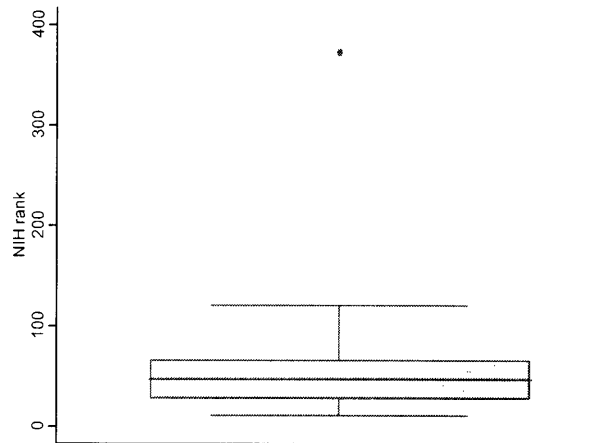
*Figure 4-7. Carnegie classification of the sample, sorted by date of contract.* Note the many empty spaces : these are not “academic institutions” by Carnegie criteria. FICE = Federal Interagency Committee on Education Code, a unique identification code for each academic institution. R1 and MED are defined in the Chapter 3. “Public” = public institution, “PNP” = private not for profit institution. Source: NSF webCASPARS.

***c. NIH funding rank of institutions.***

Figure 4-8 shows the institutions by rank of total NIH funds received, as ordered by the NIH in the year of the contract. In our sample, ranks range from a high of 10 (Stanford University) to a low of 372 (University of British Columbia). These ranks represented total NIH funding ranging from \$153 MM to \$ 0.9 MM in their contract year (actual dollars; not adjusted for inflation). Three institutions were not listed by the NIH (Austin Research Institute, National Research Development Corporation, and National Technical Information Service). Three additional institutions have received funds in other years but did not receive NIH funding in the year of the contract (Max Planck Institute of Biochemistry , Princess Margaret Hospital, Walter Reed Army Institute of Research). If we remove these six



institutions that were not ranked in their contract year, and also remove UMDNJ because it was ranked in the separate “Medical” category, this leaves a sample of 21 “not-for-profit institutions”. For this sample, the mean rank was 60, with median 46, and the spread is shown in Figure 4-9.



**Figure 4-9. Box plot of NIH Funding Rank for the 21 institutions in the reduced sample (see text).** Rank is that given by the NIH to the institution, based on total NIH funding to that Institution, in the year of the contract. Total number of recipients vary per year, but is on the order of 500, so lowest rank can be considered to be 500.

***d. Federal Science and Engineering obligations to institutions.***

The total of federal funds for R&D can be substantially higher than the total of NIH funds received. For example, UMDNJ received \$16MM in NIH funds in 1995, but total Federal obligations in the same year were \$49MM (Figure 4-10). During this period annual Federal R&D funds received by institutions ranged from \$16 MM (SRI) to almost \$320 MM (Stanford).

Institution	upstream deal date	Federal R&D obligations in the upstream year, in thousands of dollars, not inflation-adjusted
University of Illinois at Urbana Champagne	Jan-85	83,122
Dana Farber Cancer Institute	Feb-87	31,391
MIT	Apr-87	205,792
State University New York Stony Brook	Jun-87	40,134
Brigham & Women's Hospital, Boston	Jul-87	47,779
Southern Research Insitute.(appears to be for the whole institute)	Dec-87	16,475
University of British Columbia	Jan-88	not listed
Children's Hospital Boston	Aug-88	21,688
Salk Institute for Biological Studies	Oct-88	27,986
Baylor College of Medicine	Apr-89	72,715
National Technical Information Service	Jun-89	not listed
University of North Carolina, Chapel Hill	Jul-89	105,334
Princess Margaret Hospital, Perth, Australia	Jul-90	not listed
Max Planck Institute Biochemistry, Martinsreid, Germany	Aug-91	not listed
New York University	Aug-91	92,019
U.S. Walter Reed Army Inst. Of Research	Aug-91	not listed
University of Florida	Dec-91	79,467
University of Wisconsin at Madison	Dec-91	205,012
National Research Development Corporation , UK	Jan-92	not listed
MIT	Jul-92	275,279
University of Florida	Nov-92	109,895
Brigham & Women's Hospital, Boston	Feb-93	86,394
University of California at San Diego	Apr-93	217,412
Austin Research Institute, Victoria, Australia	Jan-94	not listed
Universty of Medicine & Dentistry of New Jersey, Newark	Sep-95	49,695
Ohio State University, all campuses	Mar-96	111,039
Stanford University	Oct-96	317,932
Pennsylvania State University (University Park)	Mar-98	203,633

**Figure 4-10. Total Federal R&D obligations to institutions in the upstream year.** Note this figure can be substantially higher than NIH funds received (compare with Figure 4-8). Source: NSF.

#### *e. Science and Engineering R&D expenditures by Institutions*

Figure 4-11 shows these figures for total Science and Engineering R&D expenditures, as well as breakdowns for Life Sciences-, Biological Sciences- and Medical Sciences R&D expenditures. There can be substantial differences in the level of spending among these categories. For example, in 1992 MIT had total Science and Engineering R&D expenditures of \$334 MM, but of this only \$38 MM was spent on Life Sciences. This is supported by looking at the rankings: in 1992 MIT was ranked #5 in Total Science and Engineering R&D expenditures, but #21 in Biological Sciences R&D expenditures and #96 in Total Life Sciences R&D expenditures. This held for the sample as a whole, such that there was no correlation between ranks in Life Sciences R&D expenditures and Total Science and Engineering expenditures. There was a strong correlation between ranks in Life Sciences

R&D expenditures and Medical R&D expenditures ( $r^2=0.86$ ), and a modest negative correlation between ranks in Medical R&D expenditure and Biological R&D expenditures ( $r^2 = -0.23$ )

Institution	UpStream Deal Date	All Amounts in thousands of dollars (Actual dollars; not inflation-adjusted)									
		TOTAL S/E/M/LS* R&D expenditures for given upstream year		LIFE SCIENCES ** Total R&D Expenditures for given upstream year		BIOLOGICAL SCIENCES Total R&D Expenditures for given upstream year		MEDICAL SCIENCES Total R&D Expenditures for given upstream year			
		Rank	Amount	Rank	Amount	Rank	Amount	Rank	Amount	Rank	Amount
University of Illinois at Urbana Champagne	Jan-85	13	139,563	53	37,468	55	12,255	83	6,415		
Dana Farber Cancer Institute	Feb-87										
MIT	Apr-87	2	264,416	74	32,844	11	32,552	164	292		
State University New York Stony Brook	Jun-87	74	54,850	87	26,096	44	17,230	85	8,517		
Brigham & Women's Hospital, Boston	Jul-87										
Southern Research Institute, Birmingham	Dec-87										
University of British Columbia	Jan-88										
Children's Hospital Boston	Aug-88										
Salik Institute for Biological Studies	Oct-88										
Baylor College of Medicine	Apr-89	29	134,681	10	134,681	4	57,948	12	76,733		
National Technical Information Service	Jun-89										
University of North Carolina, Chapel Hill	Jul-89	38	122,097	22	91,710	11	43,086	33	40,375		
Princess Margaret Hospital, Perth, Australia	Jul-90										
Max Planck Institute Biochemistry, Martinsreid, Germany	Aug-91										
New York University	Aug-91	52	112,106	32	87,436	30	30,958	41	41,624		
U.S. Walter Reed Army Inst. Of Research	Aug-91										
University of Florida	Dec-91	37	140,257	36	83,456	52	23,319	53	31,984		
University of Wisconsin at Madison	Dec-91	4	326,489	5	187,200	6	58,766	9	101,919		
National Research Development Corporation, UK	Jan-92										
MIT	Jul-92	5	333,908	96	38,101	21	38,101	not listed this year			
University of Florida	Nov-92	40	140,189	37	87,227	51	23,846	55	34,611		
Brigham & Women's Hospital, Boston	Feb-93										
University of California at San Diego	Apr-93	10	307,051	16	153,495	85	13,701	5	139,794		
Austin Research Institute, Victoria, Australia	Jan-94										
University of Medicine & Dentistry of New Jersey, Newark	Sep-95	81	96,365	39	96,365	32	39,349	39	57,016		
Ohio State University, all campuses	Mar-96	19	262,147	28	133,876	38	36,717	49	49,981		
Stanford University	Oct-96	9	351,526	18	160,399	90	11,941	8	148,022		
Pennsylvania State University (all campuses)	Mar-98	13	362,643	53	103,348	30	46,249	75	26,030		

\* TOTAL S/E/M/LS = Science & Engineering + Medical and other life sciences = Total of all academic disciplines R&D for these institutions

\*\* LIFE SCIENCES = Agricultural + Biological + Medical + Other

Figure 4-11. Science and Engineering R&D expenditures, and relevant sub-categories. Ranks and expenditures are for the year of the upstream contract. Expenditures are in thousands of actual dollars, not inflation-adjusted. Blank spaces are for institutions not listed in the NSF database.

### **4.3.2 Single Inventor factors**

#### ***a. Number of published articles per inventor in 20 years prior to upstream date.***

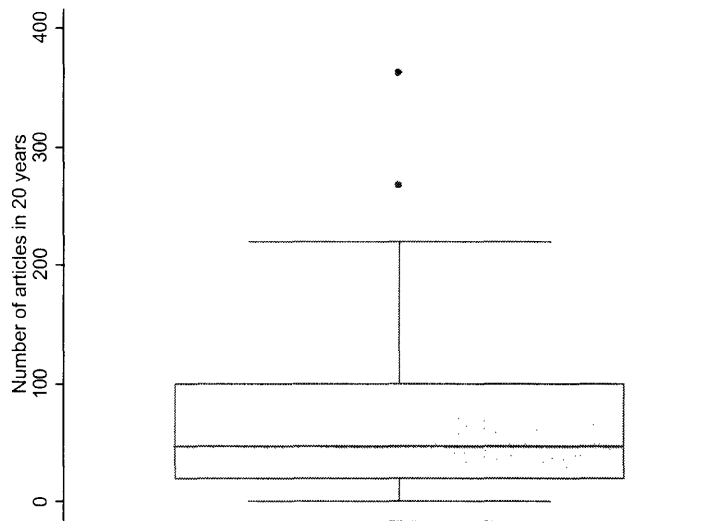
Figure 4-12 shows that the number of articles per inventor published 20 years prior and up to and including the upstream contract ranged from 368 (Ian F.C. McKenzie, Austin Research) to 1 (Patricia Lawman, Univ. Florida). . There was no relationship between year of upstream contract and number of articles published. Mean number of publications was 71 with median of 46, and the spread as shown in Figure 4-13.

upstream date	Inventor	No. articles in 20 yrs	No. articles /yr	No. citations in 5 yrs
Jan-94	Ian F.C. McKenzie	362	18	262
Mar-98	Anthony E. Pegg	267	13	3000
Feb-93	Edward M. Brown	219	11	976
Aug-91	Joseph Schlessinger	191	10	820
Aug-91	Axel Ullrich	180	9	2426
Oct-88	Ronald M. Evans	174	9	944
Feb-87	Ellis L. Reinherz	139	7	5800
Dec-91	Howard M. Temin	132	7	776
Apr-93	John Mendelsohn	128	6	232
Aug-88	(M.) Judah Folkman	116	6	3688
Jan-92	Earl W. Davie	111	6	280
Apr-89	Stanley H. Appel	101	5	376
Jan-85	David M. Long, Jr.	100	5	650
Apr-93	Gordon Sato	94	5	111
Mar-98	Stanton L. Gerson	80	4	700
Jun-87	Aaron Janoff	80	4	1148
Apr-87	Malcolm L. Gefter	78	4	300
Dec-91	Raymond J. Bergeron	72	4	520
Mar-98	M. Eileen Dolan	70	4	640
Jun-87	Sanford R. Simon	68	3	86
Mar-98	Robert C. Moschel	64	3	168
Jul-90	Wayne R. Thomas	60	3	280
Jul-89	David G. Klapper	52	3	312
Nov-92	Michael J. P. Lawman	48	2	154
Jan-94	Mauro S. Sandrin	47	2	95
Mar-96	Vernon C. Stevens	45	2	185
Aug-91	Avigdor Shaffer	42	2	128
Feb-93	Steven C. Hebert	40	2	474
Jul-87	Dennis J. Selkoe <sup>A</sup>	39	2	460
Jun-89	Michael A. Zasloff <sup>B</sup>	34	2	280
Oct-96	Garry P. Nolan	31	2	670
Jul-92	Joseph P. Vacanti	29	1	91
Dec-87	James R. Piper <sup>C</sup>	26	1	156
Dec-87	Thomas P. Johnston <sup>C</sup>	24	1	144
Jul-87	Keneth S. Kosik <sup>A</sup>	21	1	243
Apr-89	Warren J. Strittmatter	20	1	219
Jun-87	Micha Vered	19	1	64
Apr-93	J. Denny Sato	18	1	78
Apr-87	Jean-Gerard Guillet	17	1	209
Sep-95	Raphael J. Mannino	16	1	87
Nov-92	Charles E. Bagwell	16	1	22
Apr-89	James L. McManaman	13	1	38
Jan-88	David H. Dolphin	12	1	56
Jul-92	Robert S. Langer	9	0	89
Sep-95	Susan Gould-Fogerite	8	0	47
Aug-88	Robert S. Langer	8	0	64
Oct-88	David J. Mangelsdorf	7	0	105
Jun-87	Qi Long Ying	7	0	11
Oct-88	Anthony E. Oro	3	0	24
Nov-92	Patricia Lawman	1	0	1

upstream date	Inventor	No. articles in 20 yrs	No. articles /yr	No. citations in 5 yrs
Feb-87	Ellis L. Reinherz	139	7	5800
Aug-88	(M.) Judah Folkman	116	6	3688
Mar-98	Anthony E. Pegg	267	13	3000
Aug-91	Axel Ullrich	180	9	2426
Jun-87	Aaron Janoff	80	4	1148
Feb-93	Edward M. Brown	219	11	976
Oct-88	Ronald M. Evans	174	9	944
Aug-91	Joseph Schlessinger	191	10	820
Dec-91	Howard M. Temin	132	7	776
Mar-98	Stanton L. Gerson	80	4	700
Oct-96	Garry P. Nolan	31	2	670
Jan-85	David M. Long, Jr.	100	5	650
Mar-98	M. Eileen Dolan	70	4	640
Dec-91	Raymond J. Bergeron	72	4	520
Feb-93	Steven C. Hebert	40	2	474
Jul-87	Dennis J. Selkoe <sup>A</sup>	39	2	460
Apr-89	Stanley H. Appel	101	5	376
Jul-89	David G. Klapper	52	3	312
Apr-87	Malcolm L. Gefter	78	4	300
Jan-92	Earl W. Davie	111	6	280
Jul-90	Wayne R. Thomas	60	3	280
Jun-89	Michael A. Zasloff <sup>B</sup>	34	2	280
Jan-94	Ian F.C. McKenzie	362	18	262
Jul-87	Keneth S. Kosik <sup>A</sup>	21	1	243
Apr-93	John Mendelsohn	128	6	232
Apr-89	Warren J. Strittmatter	20	1	219
Apr-87	Jean-Gerard Guillet	17	1	209
Mar-96	Vernon C. Stevens	45	2	185
Mar-98	Robert C. Moschel	64	3	168
Dec-87	James R. Piper <sup>C</sup>	26	1	156
Nov-92	Michael J. P. Lawman	48	2	154
Dec-87	Thomas P. Johnston <sup>C</sup>	24	1	144
Aug-91	Avigdor Shaffer	42	2	128
Apr-93	Gordon Sato	94	5	111
Oct-88	David J. Mangelsdorf	7	0	105
Jan-94	Mauro S. Sandrin	47	2	95
Jul-92	Joseph P. Vacanti	29	1	91
Jul-92	Robert S. Langer	9	0	89
Sep-95	Raphael J. Mannino	16	1	87
Jun-87	Sanford R. Simon	68	3	86
Apr-93	J. Denny Sato	18	1	78
Jun-87	Micha Vered	19	1	64
Aug-88	Robert S. Langer	8	0	64
Jan-88	David H. Dolphin	12	1	56
Sep-95	Susan Gould-Fogerite	8	0	47
Apr-89	James L. McManaman	13	1	38
Oct-88	Anthony E. Oro	3	0	24
Nov-92	Charles E. Bagwell	16	1	22
Jun-87	Qi Long Ying	7	0	11
Nov-92	Patricia Lawman	1	0	1

Figure 4-12. Inventors, their published articles in the 20 years up to and including the upstream contract year, and their citations in the five years prior to the upstream contract year. Left grouping: sorted by articles. Right grouping: sorted by citations. See text for details.

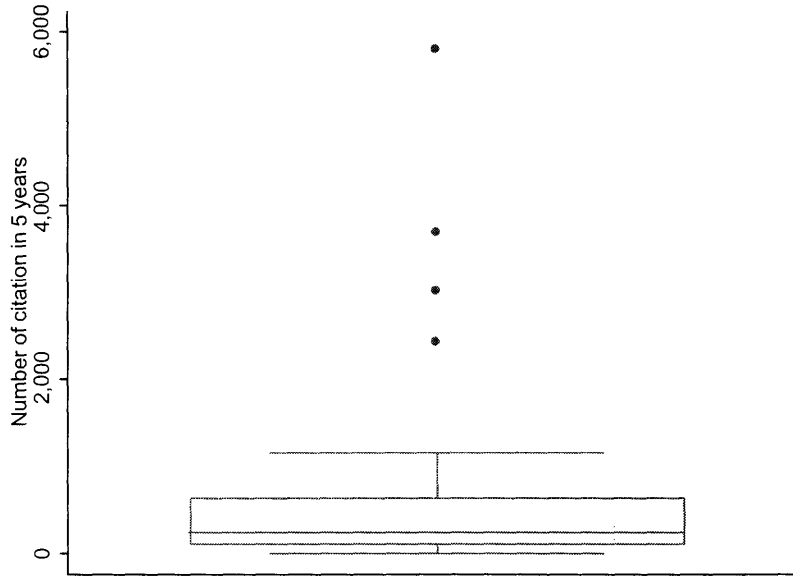


**Figure 4-13. Box plot of number of article published by inventors in the 20 years up to and including the year of the upstream contract.** See text for details.

***b. Number of citations of published work in previous five years.***

As shown in Figure 4-12 the number of citations in this five year window ranged from 5800 (Ellis L . Reinherz, Dana-Farber Cancer Institute) to 1 (Patricia Lawman, Univ. Florida). There was some correlation between the number of articles and the number of citations ( $r^2=0.47$ ). It would be hard to say which is the explanatory variable in this case, as these variables are to some degree interrelated.

The mean number of citations was 574, with median 225 and distribution as shown in Figure 4-14.



*Figure 4-14. Boxplot of number of citations in 5 years prior to the contract date. See text for details.*

***c. Number of prior patents issued or applications filed per inventor***

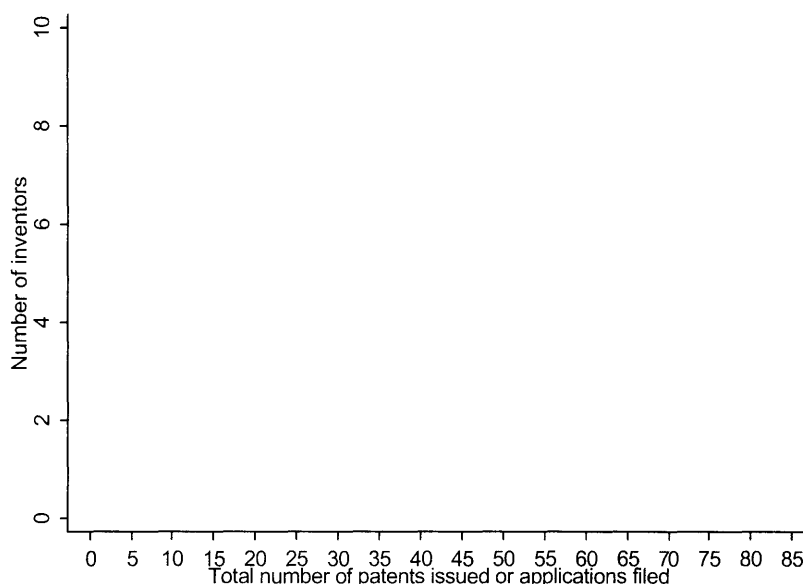
Figure 4-15 shows that there is wide disparity among inventors: some have no patents issued or filed prior to the upstream contract, while Robert Langer of MIT had 82 patents prior to the MIT / Advanced Tissue Sciences upstream contract in 1992 (he is currently listed on 149 patents and/or patent applications).

UpStream Deal Date	Upstream Contract	Inventors	# Articles in 20 years prior and upto upstream date	Approx # citations in five years prior to upstream deal year	# of U.S. IP with date prior to upstream date
Jul-92	MIT/ATS	Robert S. Langer	9	89	82
Aug-88	Children's/Enzytech	Robert S. Langer	8	64	76
Dec-91	Univ. Florida/SunPharm	Raymond J. Bergeron	72	520	30
Mar-96	Ohio State/ AVI	Vernon C. Stevens	45	185	18
Mar-98	Penn State/Procept	Anthony E. Pegg	267	3000	16
Jul-90	incess Margaret/Immulo	Wayne R. Thomas	60	280	14
Mar-98	Penn State/Procept	Robert C. Moschel	64	168	13
Jan-94	Austin Res/Alexion	Ian F.C. McKenzie	362	262	10
Aug-91	MaxPlanck/Sugen	Axel Ullrich	180	2426	8
Aug-88	Children's/Enzytech	(M.) Judah Folkman	116	3688	8
Aug-91	NYU/Sugen	Joseph Schlessinger	191	820	7
Oct-88	Salk/Ligand	Ronald M. Evans	174	944	7
Jan-92	NRDC/Genetic Therapy	Earl W. Davie	111	280	7
Jul-92	MIT/ATS	Joseph P. Vacanti	29	91	7
Dec-87	SRI/USB	James R. Piper	26	156	7
Feb-87	Dana Farber / Procept	Ellis L. Reinherz	139	5800	6
Mar-98	Penn State/Procept	Stanton L. Gerson	80	700	5
Jun-89	NTIS/Genaera	Michael A. Zasloff	34	280	5
Oct-96	Stanford/Rigel	Garry P. Nolan	31	670	5
Apr-89	Baylor/Cephalon	Stanley H. Appel	101	376	4
Jul-89	UNC/Immulogic	David G. Klapper	52	312	4
Nov-92	Univ. Florida/ATS	Michael J. P. Lawman	48	154	4
Sep-95	UMDNJ / BioDelivery	Raphael J. Mannino	16	87	4
Jan-88	UBC/QLT	David H. Dolphin	12	56	4
Sep-95	UMDNJ / BioDelivery	Susan Gould-Fogerite	8	47	4
Feb-93	B&W/NPS	Edward M. Brown	219	976	3
Dec-91	U.Wisconsin /Viagene	Howard M. Temin	132	776	3
Jan-85	Univ. Illinois/ Alliance	David M. Long, Jr.	100	650	3
Apr-87	MIT/Immulogic	Malcolm L. Gefter	78	300	3
Feb-93	B&W/NPS	Steven C. Hebert	40	474	3
Apr-93	UCSD/Imclone	J. Denny Sato	18	78	3
Apr-87	MIT/Immulogic	Jean-Gerard Guillet	17	209	3
Nov-92	Univ. Florida/ATS	Patricia Lawman	1	1	2
Apr-93	UCSD/Imclone	John Mendelsohn	128	232	1
Apr-93	UCSD/Imclone	Gordon Sato	94	111	1
Jun-87	SUNY/Cortech	Sanford R. Simon	68	86	1
Jan-94	Austin Res/Alexion	Mauro S. Sandrin	47	95	1
Aug-91	Walter Reed/ Univax	Avigdor Shafferman	42	128	1
Dec-87	SRI/USB	Thomas P. Johnston	24	144	1
Apr-89	Baylor/Cephalon	James L. McManaman	13	38	1
Jun-87	SUNY/Cortech	Aaron Janoff	80	1148	0
Mar-98	Penn State/Procept	M. Eileen Dolan	70	640	0
Jul-87	B&W/Athena	Dennis J. Selkoe	39	460	0
Jul-87	B&W/Athena	Keneth S. Kosik	21	243	0
Apr-89	Baylor/Cephalon	Warren J. Strittmatter	20	219	0
Jun-87	SUNY/Cortech	Micha Vered	19	64	0
Nov-92	Univ. Florida/ATS	Charles E. Bagwell	16	22	0
Oct-88	Salk/Ligand	David J. Mangelsdorf	7	105	0
Jun-87	SUNY/Cortech	Qi Long Ying	7	11	0
Oct-88	Salk/Ligand	Anthoy E. Oro	3	24	0

**Figure 4-15. Inventors whose work is in the upstream contract, and the number of prior U.S. patents issued and/or applications filed or claiming priority from a filing (foreign or domestic) prior to the upstream contract.** Also listed are number of inventors articles in the previous 20 years, and number of citations of their published work in the 5 years prior to the upstream contract.

There was no correlation between the number of articles they had published and the number of patents and/or applications which they had issued or filed. Nor was there any correlation between the number of citations in the previous 5 years and the number of patents and/or applications which they had issued or filed.

Fully 20% of all inventors whose work was included in upstream contracts had no prior patents issued and/or application filed. A further 15% had only one patent and/or application filed. The median was 4, and the distribution may be seen to be highly skewed (Figure 4-16). At the high end were contracts with IP by Robert Langer of MIT, who had 76 and subsequently 82 patents issued or filed prior to these particular upstream contracts.



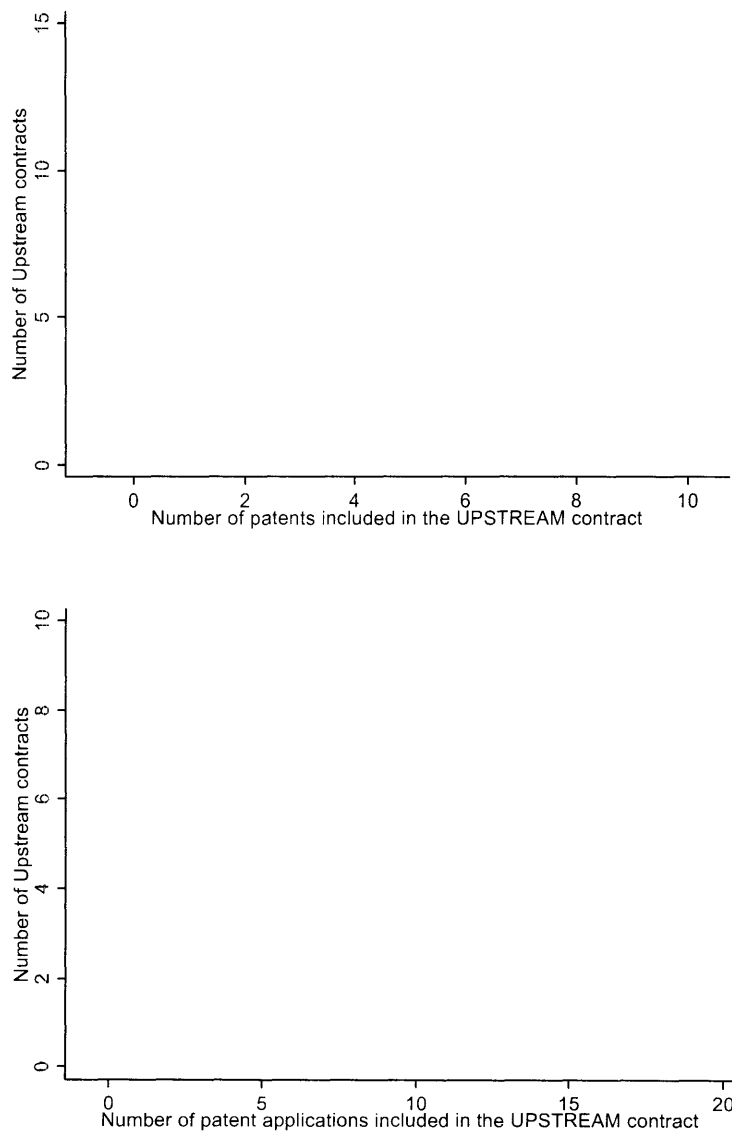
*Figure 4-16. Distribution of Number of U.S. patents and/or patent applications issued, filed or claiming priority from a filing (foreign or domestic) prior to upstream deal.*

### 4.3.3 Aggregate factors

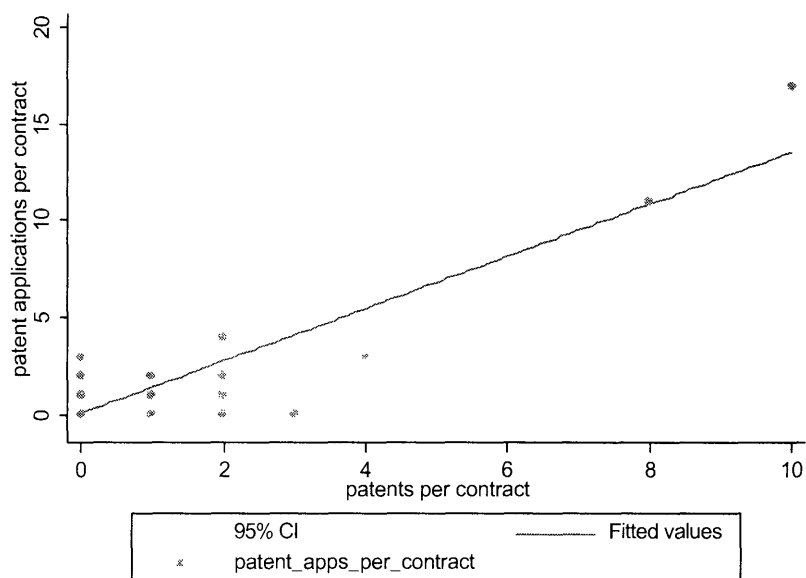
#### *a. Number of U.S. patents and patent applications per contract*

In our sample of upstream contracts, the most common finding was that the contract did not contain a patent that had issued (mode = 0), and contained one patent application that had been filed (Figure 4-17). There were two contracts (Max Planck/Sugen and NYU/Sugen) where all patents and know-how of the university were licensed to the biotech. If these two contracts are removed and the sample size reduced to 26 upstream contracts, there is a very

strong correlation between number of patents and number of patent applications in an upstream contract ( $r^2 = 0.89$ ). If we assume that patents are more desirable than patent applications (after all, patents have issued and give the right to exclude others from making, using or selling the technology) we can run a linear regression with number of patents as the explanatory variable. This model showed that the number of patent applications included in the upstream contract was significantly dependent on the number of patents ( $F[1,24] = 91.5$ ,  $p < 0.0001$ ,  $r^2 = 0.79$ ; Figure 4-18). Thus, biotech companies were willing to invest in ideas before they were protected, or that might subsequently be shown to infringe on other IP, and if they contracted for many patents, they concurrently contracted for many patent applications.



**Figure 4-17. Number of U.S. patent applications filed or claiming priority from either a foreign or domestic prior filing). See text for details.**



**Figure 4-18. Fit of linear regression model.** The number of patent applications is significantly dependent on the number of patents. See text for details.

**b. Contract “Commercialization” index.**

The most commercial experience was provided by inventors in the MIT/Enzytech contract, and the least in the Brigham&Women’s/ Athena Neuroscience contract (Figure 4-19).

**b. Contract “Impact” index.**

The most team impact, as measured by greatest aggregate number of citations of their published work, was for the Penn State/ Procept contract, with 4508 citations.

**c. Contract “Novelty” index.**

The greatest team novelty, as measured by greatest aggregate number of articles of was for the Penn State/ Procept contract, with 481 articles published.

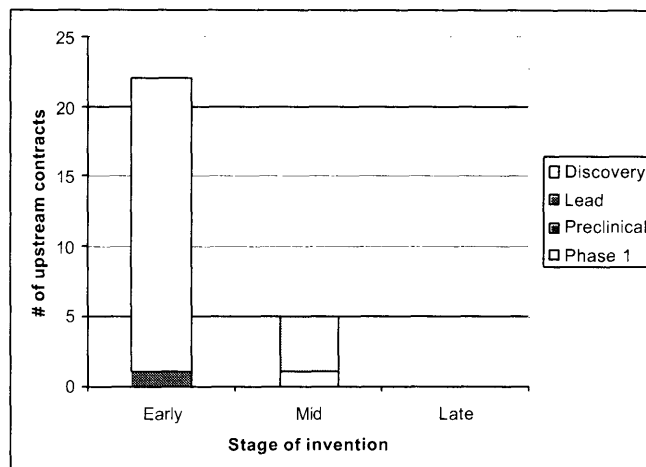
Date	Contract	Contract "Novelty" Index	Contract "Impact" Index	Contract "Commercialization" index
Jan-85	Univ. Ill. / Alliance Pharma	100	650	3
Feb-87	Dana Farber / Procept	139	3	6
Apr-87	MIT /Immulogic	95	509	6
Jun-87	SUNY / Cortech	174	1309	1
Jul-87	B&W / Athena	60	703	0
Dec-87	SRI / US BioScience	50	300	8
Jan-88	UBC/ QLT	12	56	4
Aug-88	Children's /Enzytech	124	3752	84
Oct-88	Salk /Ligand	184	1073	7
Apr-89	Baylor /Cephalon	134	633	5
Jun-89	NTIS / Genaera	34	280	5
Jul-89	UNC / Immulogic	52	312	4
Jul-90	Princess Margaret / Immulogic	60	280	14
Aug-91	Max Planck / Sugem	180	2426	8
Aug-91	NYU / Sugem	191	820	7
Aug-91	Walter Reed / Univax	42	128	1
Dec-91	UF / SunPharm	72	520	30
Dec-91	Uwisc / Viagene	135	776	3
Jan-92	NDRC / Genetic Therapy	111	280	7
Jul-92	MIT /ATS	38	180	89
Nov-92	UF /ATS	65	177	6
Feb-93	B&W / NPS	259	1450	6
Apr-93	UCSD / Imclone	240	421	5
Jan-94	ARI / Alexion	409	357	11
Sep-95	UMDNJ / BDS	24	134	8
Mar-96	Ohio State / AVI	45	185	18
Oct-96	Stanford / Rigel	31	670	5
Mar-98	Penn State / Procept	481	4508	34

Figure 4-19. Indices for commercialization, impact, and novelty, for each upstream contract. See text for details.

#### 4.4 Invention Factors

##### a. Stage of the invention

Most inventions by far (21/28) are in discovery phase, and so are “early”, following the usage of Edwards et al (Figure 4-20). Recall that “Early” includes those therapeutics in discovery or that have a lead generated (1/28). Five inventions are “Mid” stage, which includes those with Preclinical- or Phase 1 data. There were no “Late” stage projects.



*Figure 4-20. Stage of the invention in Upstream contracts, following the terminology of Edwards et al (See Chapter 2, Methods).*

**d. Field of the invention**

The licensed technologies are described briefly in Figure 4-21. They range from novel antimicrobial peptides, to modifications of specific enzymes, to cloned genes for blood clotting factors, to hybridomas for the production of specific monoclonal antibodies. These technologies were then grouped into larger “themes” of general therapeutic categories (Figure 4-23). Some licensed technology was so broad that it was simultaneously categorized in two areas. For example, “modified peptides and methods for making vaccines against colorectal and pancreatic cancer” (the Ohio State/AVI BioPharma contract of March 1996) can be grouped in with “cancer” and with “immune control”. On the other hand “neurotrophic hormones for ALS, Parkinson's, Alzheimer's” (the Baylor/Cephalon contract of April 1989) is unambiguously categorized as CNS. In some cases the patent was so general that the downstream contract was used to help elucidate a specific therapeutic area.

The bulk of the contracts were in the general therapeutic areas of immune control (n=9), which includes allergy and transplantation, and cancer (n=8).

CONTRACT PAIRS	# issued U.S. patents in the contract	# U.S. patent apps in the contract	stage	E M L	description of the the licensed technology	therapeutic theme
Austin Research--> Alexion --> US Surgical	0	1	discovery	E	modification, use or blockade of transferase enzyme or their products in xenotransplantation	immune
Baylor --> Cephalon --> Schering-Plough	2	4	discovery	E	neurotrophic hormones for ALS, Parkinson's, Alzheimer's	cns
Brigham & Women's --> Athena Neuroscience --> Eli Lilly	0	0	discovery	E	research agreement (with option to license) for development of diagnostic and therapeutic products for Alzheimer's Disease	cns
Brigham & Women's -->NPS Pharma --> SmithKline	0	0	discovery, will apply for patent upon signing contract	E	bovine parathyroid calcium receptor genes for use in therapeutics directed at calcium receptors	bone (osteoporosis)
Children's Hospital --> Enzyte <sup>TM</sup> --> Schering-Plough	2	0	to start clinical trials in 6 months	M	controlled release of macromolecules from erodable and non-erodable polymers	drug delivery
Dana-Farber Cancer Inst --> Procept --> BMS	0	0	discovery	E	research to study regulatory and effector molecules involved in control of T-cell activation, with goal of developing therapeutics that interact with these targets, so as to regulate T-cell activation, up or down	immune
Max Planck --> Sugen --> Amgen	Any and all of Max Planck's patents	2	discovery	E	any product for the diagnosis, treatment of or prevention of human disease that acts via a Tyrosine-kinase, tyrosine-phosphatase- or opiate receptor, or the DNA encoding said receptors	hematology, CNS
MIT--> Advanced Tissue Sciences --> Smith & Nephew	8	11	discovery	E	various matrix materials and methods for tissue repair and regeneration	bone
MIT--> Immulogic --> Merck; MMD	0	2	discovery	E	immunomodulating drugs and synthetic vaccines, and methods for creating them	immune
NRDC--> Genetic Therapy --> Genetics Inst.	1	0	discovery	E	Factor IX gene clone for human gene therapy using retroviral vectors	hematology
NTIS --> Genaera (Magainin) --> Colgate; Sandoz	1	2	discovery	E	magainins, novel antimicrobial polypeptides	infectious disease
NYU --> Sugen --> Amgen	Any and all of NYU's patents	2	discovery	E	any product for the diagnosis, treatment of or prevention of human disease that acts via a Tyrosine-kinase, tyrosine-phosphatase- or opiate receptor, or the DNA encoding said receptors	hematology, CNS
Ohio State --> AVI BioPharma --> SuperGen	10	17	preclinical -ready to start Phase 1 for CRC, planned Phase 1 for pancreatic ca	M	modified peptides and methods for making vaccines against colorectal and pancreatic cancer	cancer, immunology
Penn State --> Procept --> Access Oncology	4	3	obtaining preclinical and clinical data from NCI; upon receipt of data ready to start Phase 2, so must be wrapping up Preclinical and Phase 1	M	06 benzyl quanine (BG) and related technologies as adjunct iv chemotherapeutic agent	cancer
Princess Margarete Hospital --> Immulogic --> MMD	0	1	discovery	E	cloned dust mite allergens, and uses thereof in treatment of immunological disorders, including allergy	immune
Salk Institute --> Ligand Pharma --> Glaxo	0	3	discovery	E	intracellular steroid- and steroid-like receptors that modulate gene expression, including glucocorticoid, aldosterone, thyroid hormone and retinoic acid receptors	cardiovascular regulation
Southern Res. Inst --> US Bioscience --> SmithKline; Schering-P	1	0	discovery	E	use of WR-151327 to protect tissues against toxic effects of chemotherapy and radiation therapy	cancer
Stanford --> Rigel Pharma --> Janssen Pharma, Pfizer, Novartis	0	1	discovery	E	cell lines, plasmids, and vector libraries useful for developing methods to screen for transdominant effector peptides and their targets	cancer; immune
SUNY--> Cortech --> MMD	0	1	discovery	E	synthetically-produced human leukocyte elastase inhibitors	pulmonary
UBC --> QLT --> Baxter, Ciba-Geigy	0	1	discovery	E	certain porphyrin derivatives that are preferentially taken up by tumors, and then light-activated to make them cytotoxic	cancer
UCSD --> Imclone --> BMS	1	0	lead (had two leads, 528 and 225 chimeric antibodies)	E	hybridomas that produce monoclonal antibodies to epidermal growth factor receptor, for use in cancer therapeutics	cancer
UF-->SunPharm --> Parke-Davis, Nippon Kayyaku	1	1	preclinical	M	polyamine analogs for antineoplastic, antiviral and antiretroviral applications	cancer
UF--> Advanced Tissue Sciences --> Kinn	0	1	discovery	E	stem cell proliferation factor, protein, sequence, antibodies against SCPF, receptors and their ligands	cancer, immune
Univ Illinois --> Alliance Pharma --> Boehringer Ingelheim	3	0	preclinical ready to start Phase 1 in six months	M	development of certain perfluorocarbon compounds for radiographic imaging, MRI contrast agents, and synthetic blood oxygen carrier	infectious disease; ophthalmology
UNC --> Immulogic --> MMD	0	1	discovery	E	allergic proteins from ragweed, and uses thereof for treatment of immunological disorders, including allergy	immune
Univ Wisc, Madison --> Viagene --> Bayer, Chiron	2	1	discovery	E	use of helper cells, target cells, retroviruses, retrovirus vectors in eukaryotic recombinant DNA	immune
UMDNJ--> BioDeliverySciences-->Wyeth	2	2	discovery	E	liposomes and cochleates for drug and vaccine delivery	drug delivery
Walter Reed Army Inst--> Univax --> Genentech	0	1	discovery	E	polypeptides that selectively bind to antibodies made against HIV, and vaccines derived from those polypeptides	immune

**Figure 4-21. For each upstream contract, this figure shows the Number of patents and patent applications licensed, stage of the invention , Code for stage E = early, M =mid, L = late (see Chapter 2, Methods), description of the licensed technology, and general therapeutic theme.**

General therapeutic theme	# contracts with this theme
immune control (includes allergy and transplantation)	9
cancer	8
cns	4
hematology	3
drug delivery & reformulation	2
bone (orthopedics and osteoporosis)	2
cardiovascular	1
ophthalmology	1
disease	2
pulmonary	1

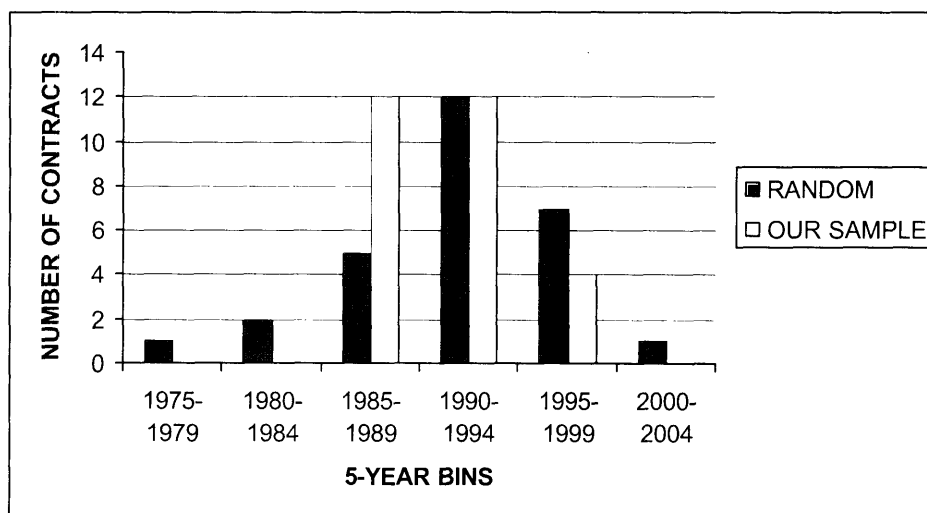
**Figure 4-22. General therapeutic themes of the contracts.** A contract could contain more than one theme, so there are 33 items in the table, although there were only 28 contracts.

#### 4.5 Comparison with a randomly-selected sample

A random selection of 28 university-biotech contracts was made from the 237 upstream contracts that were not part of our sample. This group of contracts is called the “control” group.

There were no significant differences between sample and control for the features that were examined (Figure 4-23, 4-24)

From this preliminary analysis it appears that the results from our sample should generalize to the population from which the random sample was taken.



**Figure 4-23. Distribution of signing dates of contracts in our sample, and in a random sample.** Both distributions are normal, and are not significantly different from each other. See Figure 4-24.

<b>Feature compared</b>	<b>Sample</b>	<b>Control</b>	<b>Statistical test</b>	<b>Significance</b>
Distribution of contract signing dates, in 5-yr bins (n=28)	see figure 4-23		Kolmogorov-Smirnoff Test	N.S.
Normality of signing date distributions (n=28)	Normal	Normal	Shapiro-Francia Test	N.S.
Proportion of Foreign institutions	5/28	4/28	Fisher's Exact Test	N.S.
Proportion Listed as Carnegie institution	14/28	19/28	Fisher's Exact Test	N.S.
Proportion of R1 insititutions	12/14	18/19	Fisher's Exact Test	N.S.
Royalty rate, mean $\pm$ sem	4.19 $\pm$ 0.38 %	4.45 $\pm$ 0.53 %	Student's t-test	N.S.
Royaly rate, median	4.00%	4.25%	-	-

Figure 4-24. There were no significant differences between our sample and a randomly-selected control along a number of different dimensions.

## Chapter 5

### Test of the Regression Model

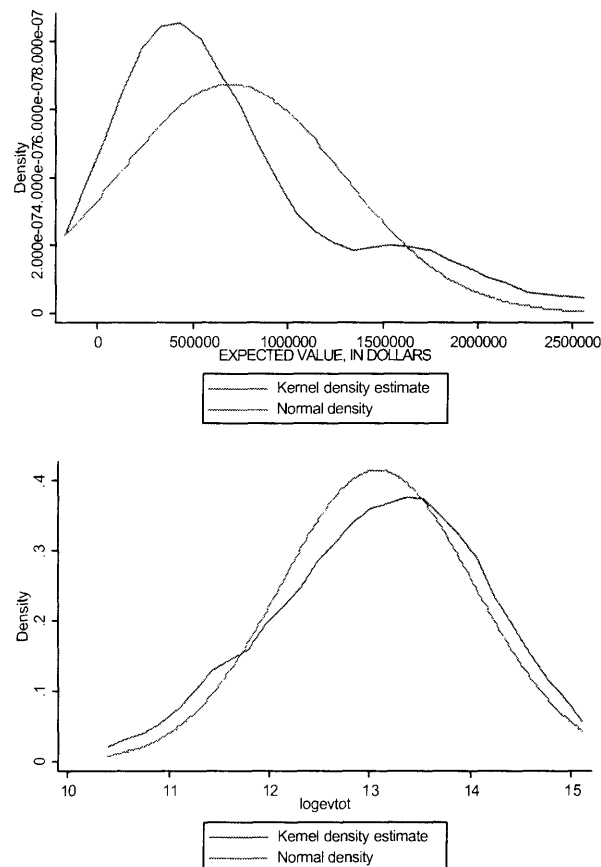
Our hypothesis proposes that contract economic value (dependent variable) can be significantly predicted from university factors (independent variables). The sample statistics demonstrated that many of the variables are not normally distributed, and normal or near-normal distributions are required to have valid tests for the significance of the coefficient weights.

For example, the dependent variable (EV) was found to be strongly skewed: so a natural-log transformation was performed, to obtain a more normal distribution ( Figure 5-1) .

Most of the independent variables were also found to be non-normally distributed so these were also ln-transformed for a more normal distribution. Dummy variables were left as coded and were not ln-transformed. As such the regression tested the natural log of the independent variables (except normally-distributed independent variables and dummy variables, both left untransformed) as significant predictors of the natural log of the dependent variable.<sup>2</sup>

---

<sup>2</sup> Independent variables were tested using the “Forward selection” method, which introduces variables into the model one at a time. At each step the variable is tested to see if it improves the model, i.e. explains more of the dependent variable variance. If it does it is kept. If not, it is not kept (Finkelstein and Betensky, 2004).



**Figure 5-1** *The distribution of the expected value was skewed (top panel; normal distribution has lower peak value), with a tail to the right, so a ln transform was performed. Transformed data (bottom panel; normal distribution has higher peak value) are more normally -distributed, and were used as the dependent variable.*

## 5.1 Results

All the University factors (except for dummy variables) presented in Chapter 3 and summarized in Chapter 4 were ln-transformed and tested as explanatory variables for the ln-transformation of the dependent variable. (Figure 5-2). Two different control variables were tested: presence/absence of sponsored research (we know from Sample Statistics in Chapter 4 that contracts with sponsored research have higher EV) and years of patent life remaining (which is reflected in the years of discounting in the Expected Value formula, and thus increases as EV increases). Several factors were correlated (Figure 5-3), and of these only the factor with the highest  $R^2$  was chosen for a multiple regression. Of the non-correlated factors, only two had sufficient explanatory power and had coefficient weights that were significantly different from zero. These factors were ln-transformations of “Life Sciences R&D

Independent Variables	Control Regression				Regression 2 = Control + Explanatory variable			
	R <sup>2</sup>	Adjusted R <sup>2</sup>	p value for coefficient weight	# Of observations	R <sup>2</sup>	Adjusted R <sup>2</sup>	p value for Exp.var. coefficient weight	# Of observations
<b>Controls</b>								
sponsored research (yes/no)	0.2624	0.2340	0.005	28				
# years left on patent rights	0.1556	0.1232	0.038	28				
<b>Explanatory Variables, simple linear regression using "sponsored research" as control</b>								
# years left on IP					0.3113	0.2562	0.195	28
country (US/non US)					0.2627	0.2037	0.921	28
Carnegie Classification (R1/MED)					0.5484	0.4663	0.886	14
Public vs. Private-Not-for-Public					0.5823	0.5064	0.359	14
ln(total NIH funding)					0.3135	0.2413	0.396	22
total NIH funding					0.4065	0.3440	0.065	22
ln(total NIH funding Rank)					0.3135	0.2413	0.396	22
ln(Number of NIH Research grants)					0.3441	0.2750	0.211	22
Amount of NIH Research Grant \$ received)					0.4046	0.3419	0.067	22
ln(Number of NIH Training grants received)					0.4600	0.3964	0.032	22
ln(Amount of NIH training Grant \$ received)					0.4245	0.3568	0.059	20
ln(Number of NIH Fellowships received)					0.2951	0.2209	0.632	22
ln(Amount of NIH Fellowship \$ received)					0.2873	0.2122	0.873	22
ln(Number of NIH contracts received)					0.3310	0.2474	0.320	19
ln(Amount of NIH contract \$ received)					0.3391	0.2565	0.278	19
total Federal science R&D obligations					0.2928	0.2143	0.448	21
ln(total Federal science R&D obligations)					0.3028	0.2253	0.364	21
ln(Science & Engineering Expenditures)					0.3973	0.2968	0.679	15
ln(Science & Engineering Rank)					0.3884	0.2865	0.955	15
ln(Biological Sciences Expenditures)					0.3904	0.2888	0.841	15
ln(Biological Sciences Expenditures Rank)					0.3943	0.2933	0.737	15
ln(Life Sciences Expenditures)					0.6351	0.5743	0.015	15
ln(Life Sciences Expenditures Rank)					0.7310	0.6862	0.002	15
ln(Medical R&D Expenditures)					0.5625	0.4829	0.034	15
ln (Medical R&D expenditures Rank)					0.6708	0.6109	0.006	14
ln(number of articles published by inventor on patent that had the most articles)					0.3208	0.2665	0.155	28
ln(number of total citations in previous 5 years by inventor on patent that had the most such citations)					0.3989	0.3508	0.025	28
ln(number of patents and/or patent applications by inventor on patent with greatest number of such patents and/or applications)					0.3020	0.2438	0.336	27
ln(commercialization index)					0.3031	0.2450	0.327	27
ln(impact index)					0.3598	0.3081	0.063	28
ln(novelty index)					0.3337	0.2804	0.114	28
ln(total number of patents and/or patent applications included in the contract)					0.1200	-0.0755	0.475	27
tangible property (yes/no)					0.2627	0.2037	0.917	28
stage of the invention (early/mid)					0.2842	0.2269	0.391	28
therapeutic area, coded as Dummy1 (1=cancer, 0 =other)					0.3122	0.2572	0.190	28
Dummy2 (1=immune, 0=other)					0.3286	0.2749	0.129	28

**Figure 5-2. Regression Table for Regressions 1 and 2.** The dependent variable was ln(contract expected value). Two different control variables were tested: presence/absence of sponsored research (we know from Sample Statistics that contracts with sponsored research have higher EV) and years of patent life remaining (which is reflected in the years of discounting in the expected value formula, and thus increases as EV increases). See text for details.

	SP_RES	NIH_sT	NIH_RE~T	loglsst	loglsr~k	logume~t	logume~k	logmaxc~	Dummy2
SP_RES	1								
NIH_sT	0.0287 (0.8990)	1							
NIH_RES_sT	0.0428 (0.8500)	0.9987 (0.0000)	1						
loglsst	0.0686 (0.8081)	0.5938 (0.0196)	0.5988 (0.0184)	1					
loglsrank	0.0382 (0.8924)	-0.564 (0.0285)	-0.5636 (0.0287)	-0.8476 (0.0001)	1				
logumedst	0.201 (0.4909)	0.4955 (0.0716)	0.4965 (0.0709)	0.8493 (0.0001)	-0.6994 (0.0054)	1			
logumedrank	0.0638 (0.8285)	-0.7882 (0.0008)	-0.7841 (0.0009)	-0.7981 (0.0006)	0.8605 (0.0001)	-0.8013 (0.0006)	1		
logmaxcite~v	0.0657 (0.7397)	0.1175 (0.6026)	0.117 (0.6040)	-0.0008 (0.9976)	0.0056 (0.9842)	-0.2888 (0.3167)	0.1152 (0.6950)	1	
Dummy2	0.2293 (0.2405)	-0.1374 (0.5419)	-0.1343 (0.5512)	0.163 (0.5616)	-0.0663 (0.8143)	0.105 (0.7210)	0.1349 (0.6456)	-0.1868 (0.3412)	1

**Figure 5-3. Pairwise correlations between explanatory variables (statistical significance is in parentheses, underneath the correlation).** If statistical significance  $p < 0.05$ , then the two variables are correlated, and only one may be used in the model. Note that loglsrank ["ln(Life Sciences R&D Expenditures, Rank)"] is correlated with many other variables, but not with logmaxcite~v ["ln(number of citations in previous 5 years by inventor on patent that had the most such citations)"] or Dummy2 (1=immune therapeutic, 0 = other therapeutic).

Regression 3 = control + 2 Explanatory variables				
	$R^2$	Adjusted $R^2$	$p$ value for Exp. var. coefficient weights, in order listed	# Of observations
<b>Regression #3 Test of uncorrelated variables, cutoff for single correlation is significance = 0.05</b>				
control+ln(Life Sciences Expenditures Rank)+ln(number of total citations in previous 5 years by inventor on patent that had the most such citations)	0.9000	0.8727	.000,.001	15
control+ln(Life Sciences Expenditures Rank)+ln(Number of NIH Training grants received)	0.7316	0.6584	.010,.879	15

**Figure 5-4. Regression table for Regression 3.** The dependent variable was ln(contract expected value). Control as in Figure 5-2. See text for details.

Expenditures, Rank” (an institutional factor) and “number of citations in previous5 years by inventor on patent that had the most such citations” (a single inventor factor). The multiple regression output is shown in Figure 5-5.

Source	SS	df	MS			
Model	15.3104065	3	5.10346885	Number of obs =	15	
Residual	1.70122333	11	.154656666	F( 3, 11) =	33.00	
Total	17.0116299	14	1.21511642	Prob > F =	0.0000	
				R-squared =	0.9000	
				Adj R-squared =	0.8727	
				Root MSE =	.39326	

logevtot	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
SP_RES	1.579397	.2169077	7.28	0.000	1.101986	2.056807
loglsrank	.7846942	.1294068	6.06	0.000	.4998717	1.069517
logmaxcite~v	.4991454	.115785	4.31	0.001	.2443043	.7539865
_cons	6.707626	.8379377	8.00	0.000	4.863337	8.551914

**Figure 5-5. Regression output in Stata for the dependent variable (logevtot), the control variable (SP\_RES) and the two independent variables (loglsrank, logmaxcite~v).** See text for details.

The F-test is significant at  $p < 0.0001$ , indicating that the factors are predicting the dependent variable.

Note that 90% of the dependent variable variance is explained by the control and independent factors ( $R^2 = 0.9000$ , Adjusted  $R^2 = 0.8727$ ), compared with  $R^2 = 0.2624$  (Adjusted  $R^2 = 0.234$ , Figure 5-2) for the control factor by itself. Thus, the two explanatory factors contribute an additional 64% explanatory power to the model.

Also note that the coefficient weights of the factors are both significantly different from zero at  $p \leq 0.001$ , and that all confidence intervals exclude zero.

The least squares regression line is given by:

$$\text{logevtot} = (1.57x \text{ SP\_RES}) + (0.78 x \text{ loglsrank}) + (0.50 x \text{ logmaxcite~v}) + 6.7$$

That is, when SP\_RES and loglsrank are held constant and logmaxciteinv increases by one unit, logevtot will increase by 0.50 units. Similarly, when SP\_RES and logmaxciteinv are held constant and loglsrank increases by one unit, logevtot will increase by 0.78 units.

In other words, assuming a contract starts at \$500,000, an increase in rank by 10 units would increase the value of the contract to \$2.4 million (Figure 5-6, top panel).

In like fashion, again assuming a contract starts at \$500,000, an increase in maximum citations for the inventor with the most citations by 550 additional citations results in an increase in the value of the contract to \$830,000 (Figure 5-6, bottom panel).

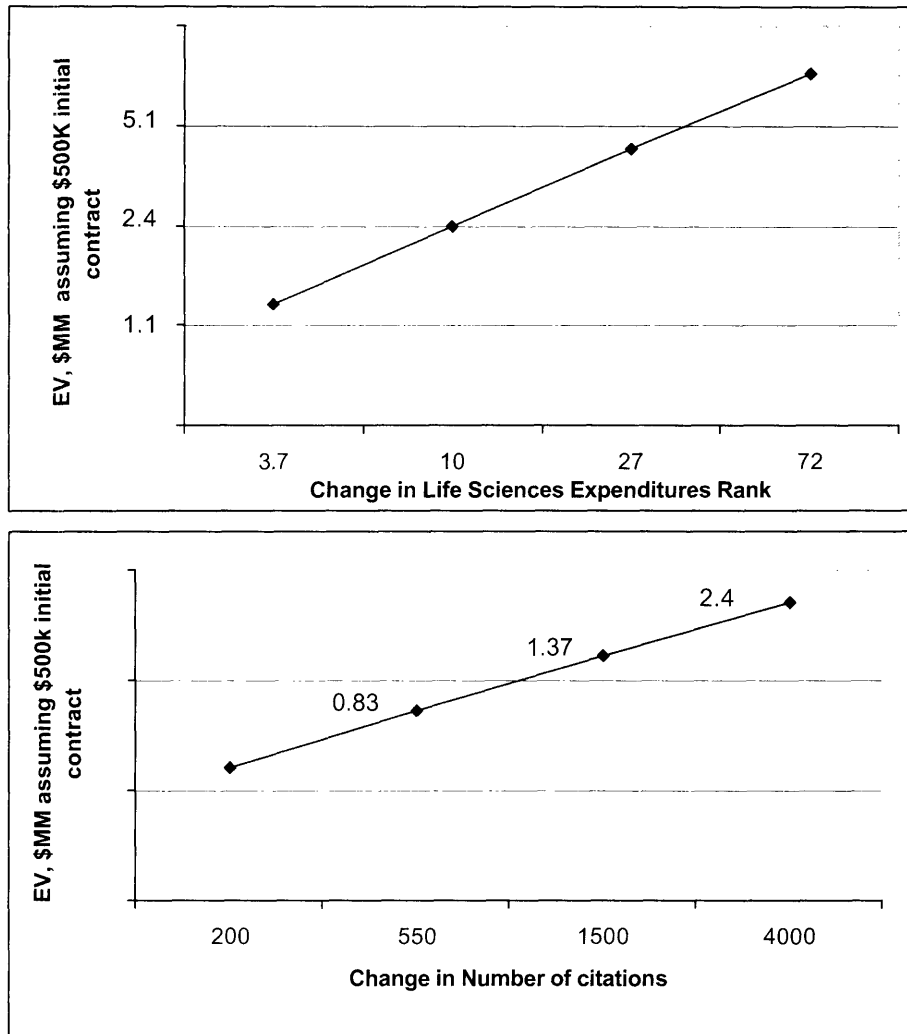
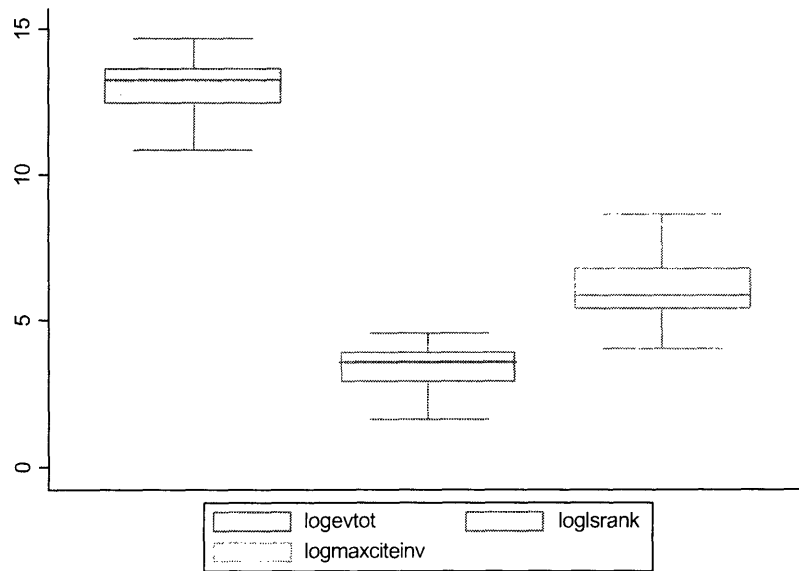


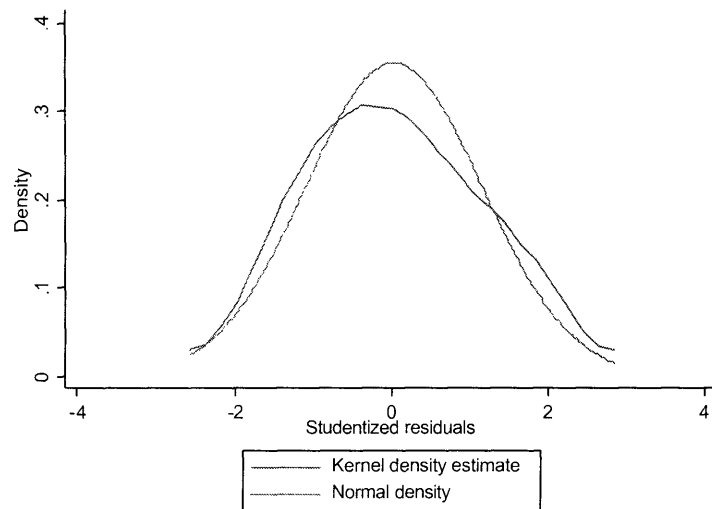
Figure 5-6 Predicted contract EV (assuming a baseline value of \$500,000) for a contract, for a change in Life Sciences Expenditure Rank (top panel) and change in Number of Citations (bottom panel). See text for details.

## 5.2 Goodness of fit

In the ln-transformations no observations were outliers, or had undue influence or leverage (Figure 5-7). Residuals were approximately normally distributed (Figure 5-8), and the Cook-Weisberg test was not significant, indicating that residuals are homoscedastic ( $\chi^2=.25$ ,  $p<0.62$ ).



**Figure 5-7. Box plots of ln-transformed dependent (logevtot) and independent (loglrank, logmaxciteinv) variables.** There are no outliers. Left box = logevtot, middle box = loglrank, right box = logmaxciteinv.



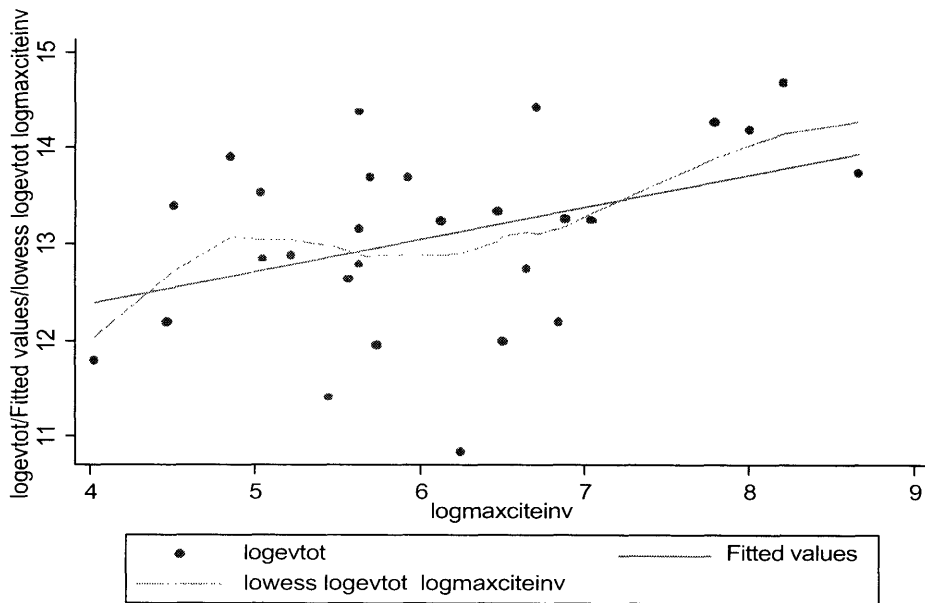
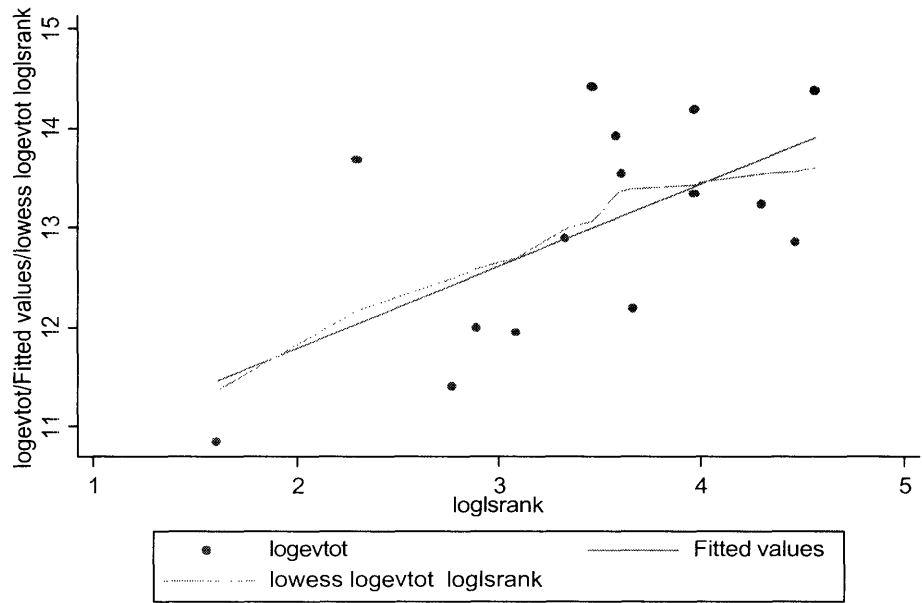
**Figure 5-8. Distribution of residuals is approximately normal.** Normal density is the curve with the higher peak.

The independent Explanatory variables  $\log\text{lsrank}$  and  $\log\text{maxciteinv}$  were not collinear. There was fairly good linearity between  $\log\text{lsrank}$  and the dependent variable except for the rightmost end of the range. There was some non-linearity between  $\log\text{maxciteinv}$  and the dependent variable, but it consisted of a broad oscillation about the linear plot, and so is an acceptable model of a linear response (Figure 5-9).

Using the Ramsey RESET test there was no specification error, that is, there were no omitted variables that should have been included in the model, nor were there irrelevant variables included in the model [ $F(3,8)=0.4$ ,  $p<0.7583$ ].

Finally, we determined if any of the outcomes of the dependent variable were dependent on others. This might arise because there are two MIT contracts at different time points (MIT/ Immulogic and MIT/ ATS) and two University of Florida contracts, also at two different time points (UF/ SunPharm and UF/ ATS). As such, at each of these institutions the first contract might have influenced the second contract. The Durbin-Watson test for correlated residuals was used to test this possibility, and the outcomes were found to be independent [ $d\text{-statistic}(4,15) = 1.19$ ].

Thus, we have no reason to reject the proposed model on goodness-of-fit grounds.



*Figure 5-9. Plots of the dependent variable (logevtot) vs. the independent variables (loglrank, top, and logmaxciteinv, bottom), show fairly good linearity.*

## Chapter 6

### General Discussion

This work has tested the hypothesis that certain university factors could significantly predict the value of university-biotech licensing contracts for therapeutics. This was tested with a multiple regression model, in which the ln transformation of the expected value of the contract was the dependent variable. Two independent factors were found to explain 64% of the variance of the dependent variable (in excess of a control independent variable). These two explanatory factors were ln transformations of

1. the Ranking of the institution in Life Sciences R&D Expenditures, and
2. the number of citations in the previous 5 years of the inventor on the patent that had the most citations.

#### 1. General methodological issues

This work relied on electronic and printed database searches. In the process of doing the research it became clear that there was some ambiguity in listings and citations of authors. Some authors, were listed only as first initial and surname. An examination of some of the listed and cited publications showed publications in widely disparate fields and journals. In these cases it was easy to discern that two or more separate authors had been placed in one listing. It is impossible to say in how many cases authors with same names in similar fields were lumped together, so that inflated measures of publications or citations have inadvertently been reported. A related issue is authors who have published under more than one name. For example, Moses Judah Folkman has published as “M. J. Folkman”, but by far the bulk of his work is published under “J. Folkman”. Reasonable attempts were made to verify that the articles and citations belonged to one and only one author, but error may nonetheless be present.

It is also acknowledged that the variables selected as representative of publications and citations have a certain arbitrariness about them. It is certainly possible that other specific metrics may have given different results, in particular, it is possible that our measure of publications was insufficiently sensitive. For example, perhaps a different measure, say total publications vs. only articles, may have given different results. Similarly, a different measure of citations, say, total citations for the most-cited article, may have produced different results.

Finally, it must be recognized that a multiple regression study may find factors that are significant in a particular study, but it doesn't claim to find unique factors, or the only factors. It is certainly possible that other factors may be significantly predictive of contract value.

## **2. Relationship to prior scholarship**

Our findings are in keeping with prior studies. For example, Henderson and colleagues (1995): demonstrated that size and ranking of a university are directly proportional to the importance of a patent. We here equate importance of a patent with the Expected value of a license contract based on that patent. In our sample neither country of origin nor Carnegie classification were predictive of contract value (= patent importance), but total Life Sciences R&D expenditures was significantly predictive. Further, these Life Sciences expenditures were significantly correlated with NIH funding levels, so these, too, were predictive. We assume that total expenditures are proportional with size and rank, and in order to raise funds through traditional NIH and NSF granting mechanisms universities must demonstrate both that the work is important and novel and that the scientists that will carry out the work have a record of producing results. Both of these are measures of rank. Additionally, Henderson and her colleagues showed that citations could be used as measure of importance, and we demonstrated that the number of citations by the inventor on the patent with the most citations was predictive of Expected value (= importance).

Our findings also bear upon the work of Thursby and Thursby (2000). Their work shows that faculty have a propensity to disclose inventions, and that tech licensing officials have a propensity to commercialize that technology, and although more patents are being issued fewer commercialization deals are actually getting done. This may relate to our finding that the number of prior patents by the inventor with the most patents, as well as total number of patents by all inventors, and total number of patents per contract had no significance upon contract value. It appears that prior “commercialization experience” (as measured by its proxy, number of patents issued and/or patent applications filed) is not relevant for contract value.

Finally, this work can be related to the findings of Edwards, Murray and Yu (2003) who found that universities are getting a very small fraction of the total profits from commercialization of academic research. This work shows that certain university factors are predictive of higher value. If universities can manipulate these factors, then they may increase their share of captured value.

## **3. Practical implications for Universities.**

The university factors tested in this thesis were selected based on findings in the literature, as discussed above. Importantly, by the measures that we *a priori* selected, they are generalizable to the greater body of university-biotech therapeutic contracts.

What might these factors mean in a practical sense? Can they be manipulated to benefit the University?

Total “Life Sciences” is defined by the NSF as “agriculture + biological + medical + other”. This factor was highly correlated with Medical Expenditures Rank, but not Biological Sciences Expenditures Rank, so presumably medical expenditures forms a large portion of this variable. In addition to the NIH- and NSF funding vehicles, the university has several sources to draw on for increasing its expenditures.

- Targeted Alumni donations
- Philanthropic gifts from high net worth individuals
- Corporate sponsored research
- Local and state government funding
- Direct Federal budget line-item funding (Congressional “pork”).

The motivation to pursue these extra-granting funding routes is significant: for every 4 units increase in expenditures, the value of a \$500k contract (somewhat less than the mean EV in our sample) increases to \$1 MM. The question for the universities will be: how much must be invested in order to reap the benefits of increased value of contracts, and is this an economically viable tradeoff?

As to citations, these can be increased by

- Doing novel and important work
- Publishing in journals that are covered by Thomson ISI’s Citation Index
- Publishing “least publishable units”, so that the work is broken up into separate entities, each of which may then gain independent recognition and be cited separately
- Having an active and effective university public affairs office that dispenses timely press releases about novel and potentially important biomedical work, particularly if it has commercial applications.

The motivation here is perhaps not as great as in the expenditures case, because it may be more difficult to attain: citations must increase by 500 in order to increase the value of a contract from \$500k to \$1MM.

These attempts at manipulating the expected value of a contract are not measures that can be effected by the Technology Transfer Office, but TTO officers should make known these findings, and their financial implications, to the University officials responsible for setting of University policy. Perhaps policy directive from the top may be effective.

#### **4. Practical implications for Biotech firms and Venture Capitalists.**

Biotech firms and venture capitalists can manipulate the factors that have no bearing on university Expected Value. For example, country of origin and type of institution, numbers of publications by the inventors, therapeutic area, and stage of the invention (at least through Phase 1) seem to make no difference in Expected Value. An effective strategy for biotech firms and VCs may be to seek the most innovative and potentially important work coming

from the smaller, less well-ranked institutions and from less well-known inventors. This way they could pay less for the contract (in EV terms) and reap more profits for themselves. This is not an easy strategy to implement, however, for it demands that biotech firms and VCs be constantly on top of the relevant biomedical literature, and that the literature be read by those with sufficient knowledge to understand the findings, interpret them in the context of the present biomedical and business climate, and to make an informed prediction on the business viability of pursuing the particular line of research. These skill sets in both biomedical science and business are rare, and will demand a new group of “biomedical scientist / entrepreneur / financier”. Our work predicts that there will be a growing need for these types of individuals.

## References

- Edwards, M.G., Murray, F., and Yu, R. (2003) Value creation and sharing among universities, biotechnology and pharma, *Nature Biotechnology*, 21(6):618-624.
- Finkelstein, D. and Betensky, R. (2004) Lectures delivered at the Harvard-MIT Division of Health Sciences Technology class in Statistical Planning and Analysis of Biomedical Investigations, Course HST-190/191, January, 2004, Boston , Mass.
- Henderson, R., Jaffe, A.B., Trajtenerg, M. (1995 = 1998) Universities as a source of commercial technology: a detailed analysis of university patenting 1965-1988, *NBER Working Paper No. 5068*, Cambridge, Mass., and *Review of Economics and Statistics*, 119 - 127.
- Jensen, R. and Thursby, M. (1998) Proofs and prototypes for sale: the tale of university licensing, *NBER Working Paper No. 6698*, Cambridge, Mass.
- Murray, F. (2003) Lectures delivered at MIT-Sloan School of Management, Course 15.968 Building a Biomedical Business, Fall 2003, Cambridge, Mass.
- Rosner, B. (2000) *Fundamentals of BioStatistics, 5<sup>th</sup> Edition*, Duxbury Press, Pacific Grove, California.
- StataCorp. 2003. *Stata Statistical Software: Release 8.0*. College Station, TX: StataCorp LP
- Thursby, J.G. and Thursby, M.C. (2000) Who is selling the ivory tower? Sources of growth in university licensing, *NBER Working Paper No. 7718*, Cambridge, Mass.

## **RICHARD G. WEHBY**

### **- BIOGRAPHICAL NOTE -**

#### **EMPLOYMENT HISTORY**

##### **SCION PHARMACEUTICALS**

Medford,, MA

June - December, 2003 Business Development Consultant. Determine next therapeutic indications for the firm; conduct sales and marketing forecasts of firms' and competitor's therapies; scientific, business and strategic assessment of competitors and possible partners; meet with CROs for planning of clinical trials; advise on in-licensing opportunity.

##### **BIOMEDICAL ENTERPRISE CONSULTANTS**

Cambridge, MA

2002 – Present Founding Partner. Provide consulting services in marketing- and competitive strategy to early-stage venture-backed biotechnology- and biomedical device companies, in conjunction with Navigator Technology Ventures (Cambridge, MA).

##### **MASSACHUSETTS INSTITUTE OF TECHNOLOGY**

Cambridge, MA

1998-2002. Postdoctoral Associate, Department of Brain and Cognitive Sciences and Center for Learning and Memory. Determined that information about identity of complex objects is present in part of the brain where it was thought to be absent. Results presented at Society for Neuroscience Annual International Meeting, San Diego, CA, November 2000.

##### **UNIVERSITY OF CONNECTICUT HEALTH CENTER**

Farmington, CT

1991-1998. Graduate Research Assistant, Neuroscience Program. Discovered OP-1 effects on control of neural crest development, investigated potassium ion channel properties, and identified electrophysiological properties of taste-responsive neurons. Taught Medical- and Dental Students as Graduate Instructor in Human Neuroanatomy laboratory (1997).

#### **EDUCATION**

**MASSACHUSETTS INSTITUTE OF TECHNOLOGY & HARVARD MEDICAL SCHOOL** Cambridge, MA  
Dual Degree Candidate, June 2004 Biomedical Enterprise Program, S.M from MIT Sloan School of Management in Management of Technology, and S.M. from Harvard-MIT Division of Health Sciences and Technology in Health Sciences and Technology.

##### **UNIVERSITY OF CAMBRIDGE**

Cambridge, England, U.K.

January 2003 and October 2002. Tech Transfer and Bioscience Enterprise Symposium, and Epidemiology . through affiliation with Bioscience Enterprise Program at the Institute of Biotechnology, and the Cambridge- MIT Institute.

##### **UNIVERSITY OF CONNECTICUT HEALTH CENTER**

Farmington, CT

Ph.D. Biomedical Sciences 1998. Thesis titled "Insular cortex of the Syrian golden hamster". Discovered that conventional histochemical test for nitric oxide synthase yielded false positive caused by previously unidentified NADPH diaphorase. Identified types and locations of cortical neurotransmitters using immunohistochemistry. Determined morphology of nitric-oxide-containing neurons and their interconnections with other parts of the brain. "Distinguished Dissertation Award" nominee. NIH Training Grant Recipient. Coursework included medical neuroanatomy, neurophysiology, immunology, developmental neurobiology, cell physiology.

##### **PRINCETON UNIVERSITY**

Princeton, NJ

A.B. Biology 1991. Senior Thesis "Towards a Neurobiological Understanding of Selective Attention". Proposed neurobiological mechanisms for voluntary control of attention. Coursework included physics, biology, microbiology, inorganic and organic chemistry, calculus.

#### **RECENT HONORS**

- **Finalist**, Venture Bowl 2004 Business Plan Competition; **Semi-finalist**, 2003 & 2004 MIT 50K Entrepreneurship Competition ,(3) **Semi-finalist**, 2004 New Hampshire Business Plan Competition,
- Elected Member of MIT Chapter Sigma Xi Science Honor Society

## PUBLICATIONS

- Wehby, R.G., Barry, M.A., and Frank, M.E. Specific and non-uniform distribution of NOS- and non-NOS NADPH diaphorases in the insular cortex of the Syrian golden hamster. *Manuscript in preparation*.
- Wehby, R.G. and Frank, M.E. (1999) NOS and non-NOS NADPH diaphorases in the insular cortex of the Syrian golden hamster, *Journal of Histochemistry and Cytochemistry*, 47(2):197
- Wehby, R.G. (1998) Insular cortex of the Syrian golden hamster: architectonics and characterization of the nNOS-containing neurons, *Doctoral Dissertation, University of Connecticut Health Center, Farmington, Conn., USA*.
- Varley, J.E., Wehby, R.G., Rueger, D.C., and Maxwell, G.D. (1995) Number of adrenergic and Islet-1 immunoreactive cells is increased in avian trunk neural crest cultures in the presence human recombinant osteogenic protein-1 (OP-1), *Developmental Dynamics*, 203(4): 434.
- Wehby, R.G. (1990) "Towards a Neurobiological Understanding of Selective Attention", *Senior Thesis, Princeton University, Princeton, N.J., USA*.
- Valentino, R.J., Curtis, A.L., Parris, D.G., and Wehby, R.G. (1990) Antidepressant actions on brain noradrenergic neurons, *Journal of Pharmacology and Experimental Therapeutics*, 253(2): 833.
- Valentino, R.J. and Wehby, R.G. (1989) Locus ceruleus discharge characteristics of morphine-dependent rats: effects of naltrexone, *Brain Research*, 488: 126 - 134
- Valentino, R.J. and Wehby, R.G. (1988) Corticotropin-releasing factor: Evidence for a neurotransmitter role in the locus ceruleus during hemodynamic stress. *Neuroendocrinology*, 48: 674.
- Valentino, R.J. and Wehby, R.G. (1988) Morphine effects on locus ceruleus neurons are dependent on the state of arousal and availability of external stimuli: Studies in anesthetized and unanesthetized rats, *Journal of Pharmacology and Experimental Therapeutics*, 243(3): 1178.
- Wehby, R. G. (1982) Creationism clashes with nature, *Miami Herald*, Sunday, January 17, p. 3E.

## SELECTED PRESENTATIONS

- Wehby, R. G., Kiddoo, C.E., and Miller, E.K. (2000) Neuronal selectivity to objects and their locations in the Macaque posterior parietal cortex. *Soc. for Neuroscience Abstracts*, 26(1): 975.
- Wehby, R. G., and Frank, M.E. (1999) NOS- and non-NOS diaphorases in the insular cortex of the Syrian golden hamster, *Soc. Neuroscience Abstracts*, 25(1):173
- Wehby, R. G. (1996) Characterization of nitric oxide synthase (NOS) neurons in the insular cortex of the Syrian golden hamster, *Soc. Neuroscience Abstracts*, 22(2):1565.
- Wehby, R. G., Burger, G. E., and London, J. A. (1996) Nitric oxide synthase (NOS) in the insular cortex of the Syrian golden hamster, *Chemical Senses*, 21(6):803.
- Wehby, R.G. and London, J.A. (1995) Architectonics of the insular cortex of the Syrian golden hamster, *Chemical Senses*, 20(6):799.
- Wehby, R.G., Zeiger, J.deB., and London, J.A. (1994) Forebrain projections from the gustatory cortex of the Syrian golden hamster. *Soc. Neuroscience Abstracts*, 20: 978.
- Wehby, R.G., Varley, J.E., and Maxwell, G.D. (1993) Recombinant human osteogenic protein-1 (OP-1) increases adrenergic cell number in avian neural crest cultures. *Soc. Neuroscience Abstracts*, 19: 1727.
- Fink, C., Wehby, R., Donta, T., London, J., and Davidson, R. (1993) Temperature-dependent effects on astrocytic potassium currents in rat organotypic brain cultures. *Soc. Neuroscience Abstracts*, 19: 1871.
- Wehby, R.G. and London, J.A. (1992) Inhibition of neuronal responses in hamster gustatory cortex, *Chemical Senses*, 17(5): 716.
- Barry, M. A., Wehby, R.G., Savoy, L.D., and Frank, M.E. (1991) Change in expression of acetylcholinesterase in terminal fields of gustatory primary afferents following peripheral nerve damage, regeneration, and reinnervation of taste buds. *Soc. Neuroscience Abstracts*, 17(1): 50.
- Aulisi, E.F., Wehby, R.G., Katz, J.L., and Valentino, R.J. (1989) Selective proconflict effect of corticotropin-releasing factor (CRF). *Soc. Neuroscience Abstracts*, 15(2):1068.
- Valentino, R.J. and Wehby, R.G. (1988) Corticotropin-releasing factor (CRF) mediates locus ceruleus (LC) activation by hemodynamic stress. *Soc. Neuroscience Abstracts*, 14(1): 445.
- Aulisi, E.F., Wehby, R.G., and Valentino, R.J. (1988) Naltrexone (NTX) - precipitated opiate withdrawal increases locus ceruleus (LC) spontaneous discharge and disrupts LC sensory responses. *Soc. Neuroscience Abstracts*, 14(2): 1155
- Valentino, R.J. and Wehby, R.G. (1987) Morphine effects on locus ceruleus neurons are dependent on the state of arousal and availability of external stimuli: Studies in anesthetized and unanesthetized rats. *Soc. Neuroscience Abstracts*, 13(2): 766.