

**MODELLING STOCK MARKET PERFORMANCE OF
FIRMS AS A FUNCTION OF THE
QUALITY AND QUANTITY
OF INTELLECTUAL PROPERTY OWNED**

A Thesis Presented to the
Academic Faculty

By
Lokendra P. S. Chauhan

In Partial Fulfillment of the
Requirements for the Degree of
Master of Science in
The School of Public Policy

Georgia Institute of Technology
Atlanta GA
August, 2007

MODELLING STOCK MARKET PERFORMANCE OF FIRMS AS A
FUNCTION OF THE QUALITY AND QUANTITY OF
INTELLECTUAL PROPERTY OWNED

Approved by:

Dr. Diana Hicks, Advisor
School of Public Policy
Georgia Institute of Technology

Dr. Bill Rouse
School of Industrial and Systems Engineering
Georgia Institute of Technology

Dr. Douglas Bodner
School of Industrial and Systems Engineering
Georgia Institute of Technology

Date Approved: July 9, 2007

ACKNOWLEDGEMENTS

I am particularly indebted to my advisor, Prof. Diana Hicks and co-advisor Prof. Bill Rouse for their mentorship, guidance, and encouragement during this research. Without Prof. Rouse's generous funding support, ideas and direction to pursue this research topic, I would not have been able to focus and do something as meaningful. I owe a lot to Prof. Diana Hicks, for without her training and immense availability I would not have been able to accomplish anything significant. She is the root of my interest in this exciting and growing field of economics of innovation. Being a novice to social sciences, I learnt a lot from how to design and conduct research, to how to write thesis and academic papers, and I probably will keep learning from her. I wish to emphasize that not only they both have been remarkable as advisors to guide my professional endeavors but also as elders who have supported me in my hard times in personal life.

Many other people have also contributed a lot to whatever I learnt during this research experience. I wish to convey my heartfelt thanks to Prof. Nicoleta Serban her support on statistics related issues in this research. Prof. Bronwyn Hall and Prof. Stuart Graham have also helped me every time I turned to them with their constructive criticism and I am highly thankful to them. Finally, I wish to thank Prof. Minjae Song , Dr. Andrea Ribas, Dr. Douglas Bodner and Prof. Marco Ceccagnoli for their constructive feedbacks. I also want to thank all my friends at Georgia Tech who have been willing to talk about this research whenever I needed, and especially to my wife because without her proof-reading for typographical errors and grammatical mistakes and moral support I could not have completed my research and this thesis.

TABLE OF CONTENTS

Approval	ii
Acknowledgements	iii
List of Tables	vi
List of Figures.....	vii
Summary	viii
1. Introduction	1
1.1 Patents as Value Drivers	1
1.2 Research Questions	3
2. Literature Review	6
2.1 Overview	6
2.2 Patent Quality Indicators	6
2.3 Citations	8
2.4 Market Value and Patents	10
3. Research Design	13
3.1 Model	13
3.2 Assumptions	17
3.3 Data	23
3.4 Method	28
3.5 Threats to Validity	30
3.5.1 Internal Validity	30
3.5.2 Validity of Statistical Conclusions	32
3.5.3 Construct Validity	33
3.5.4 External Validity	33

4.	Results and Discussions	35
4.1	Preliminary Analysis	37
4.2	Detailed Results	40
5.	Conclusions, Future Directions and Policy Implications	53
5.1	Conclusions	53
5.2	Future Directions	54
5.3	Policy Implications	55
	References	57
	Appendix – I: Descriptive Statistics for Data	59
	Appendix – II: Results	65

LIST OF TABLES

Table 4.1: Description of Variables	35
Table 4.2: Correlation Matrix of the Variables Used in the Analysis	36
Table 4.3: Coefficients for the Top-level Combined Regression	37
Table 4.4: Combined Table of Observations	42
Table A1.1: Descriptive Statistics	64
Table A2.1: Analysis of Variance.....	66
Table A2.2: Coefficient Values for Independent Variables	66
Table A2.3: Values for Length of Lag between R&D Expenses and Stock Returns	67

LIST OF FIGURES

Figure 3.1: Mean Generality Score over Time	26
Figure A1.1: Patent Counts Distribution of Firm Data Observations	59
Figure A1.2: Asset Value Distribution of Firm Data Observations	59
Figure A1.3: Industry-wise Distribution of Firm Data Observations	60
Figure A1.4: Year-wise Distribution of Firm Data Observations	60
Figure A1.5: Statistical Summary for Natural Log of Tobin's Q of Firms	61
Figure A1.6: Statistical Summary for Natural Log of Patent Counts of Firms	61
Figure A1.7: Statistical Summary for Natural Log of Assets of Firms	62
Figure A1.8: Statistical Summary for Ratio of R&D Expenses and Assets of Firms	62
Figure A1.9: Statistical Summary for Natural Log of Average Number of Citations Received by the Firm's Patents of that Year	63
Figure A1.10: Statistical Summary for Natural Log of Average Number of Claims Made by the Firm's Patents of that Year	63
Figure A2.1: Residual Plots for LnQ as Dependent Variable	65
Figure A2.2: Residual Plots for R&D/Assets as Dependent Variable	66

SUMMARY

This thesis attempts to analyze a part of the big and complex process of how intellectual property ownership and technological innovation influence the performance of firms and their revenues. Here I analyze firm's stock market performance as a function of the quantity and quality of intellectual property (patents) owned by the firm in the context of the three US high-technology sectors, Pharmaceuticals, Semiconductors and Wireless. In these sectors value of a firm is predominantly driven by the technologies which firm owns. I use citation based indicators and number of claims to measure the quality of patents. This research presents empirical evidence for the hypothesis that in high-tech sectors, companies which generate better quality intellectual property perform better than average on the stock market. I also find that firms which are producing better quality technologies (good R&D) invest more in R&D regardless of their market performance. Furthermore, though smaller firms get relatively lesser returns on quality and quantity of R&D they tend to invest a bigger fraction of their total assets in R&D when they are generating high quality patents. Larger firms enjoy the super-additivity effects in terms of market performance as the same intellectual property gives better returns for them and returns to R&D are relatively more in the pharmaceutical industry than semiconductor or wireless industries.

Chapter – 1

Introduction

In this knowledge-based economy of our times, intellectual property rights have become important strategic resources for firms in the high-technology industries. A patent right is not just a means of protection of knowledge about a new technology or products discovered by such firms but it also has the potential to improve the market performance of the organization. Firms get a short term legal monopoly power in the market when their product is protected by intellectual property laws (usually a patent in this case). Therefore, such firms have greater assurance of making larger profits, which in turn increases the confidence of investors. Thus the advantage of having good technologies protected also gets reflected in stock prices.

1.1 Patents as Value Drivers

Patents are intangible assets for any company, and researchers studying economics of technological change have been working on how patents affect the productivity and valuation of the firms or what are the returns to the R&D investments. While returns in productivity come slowly distributed over time (Bloom and Van Reenen, 2002), stock market returns are almost immediate. Better stock performance reflects the market's (external to the company) belief in the success and increased revenues for the company due to which the market decides to invest more in the company. On the other hand, R&D investments are internal decisions which occur primarily due to the belief of a company

in itself or its research capabilities for pursuing the market opportunities while seeking and often leading to subsequent buy in from the market into this decision, which would further lead to the growth in revenues of the company. Another factor which influences R&D spending is the amount of profits made by the firm in any year.

From our knowledge of stock markets, we can conveniently assume here that the stock market absorbs each new piece of information and has intelligence to react to it depending upon the importance the market gives to that type of information and other factors. Moreover, the market is able to provide this response almost immediately without a significant lag in time. Hence company announcements about their further investments in R&D or new patents being granted etc, are absorbed fast and generate the intelligent market reaction for that piece of information.

In this paper we present our research to find empirical evidence for the hypothesis that in high-tech sectors, companies which generate better quality intellectual property perform better than average on the stock market. Furthermore, we seek to empirically identify any other trends or patterns which may be observable at the industry level in the three selected US industry sectors with high innovation and patenting activity.

To interpret and measure the performance of the firm, we use the Tobin's Q, which is the ratio of the market value of a firm's assets (as measured by the market value of its outstanding stock and debt) to the replacement cost of the firm's assets. Innovation or knowledge stock is the independent variable along with the asset value of the firm. Patent

indicators, like patent counts, citation counts, number of claims etc are used to measure the knowledge stock or innovative technologies owned by the firm. As it is difficult to calculate the replacement cost of firm's assets, typically replacement cost of the intangible assets like intellectual property and brand are not included in the calculation of Tobin's Q. Due to this Q values would tend to be greater than one for the firms which have significant amount of these intangible assets. Therefore using Tobin's Q to measure the impact on market value of firm due to these intangible assets is a good approach.

1.2 Research Questions

This paper attempts to analyze a part of the big and complex process of how technological innovation and intellectual property influences the performance of firms and their revenues. Here we analyze firm's stock market performance as a function of the quantity and quality of patents owned by it for the three US high-technology sectors, where value of a firm is predominantly driven by the technologies which firm owns. These industry sectors are semiconductor equipment manufacturing (NAICS 334413), wireless and communications equipment manufacturing (NAICS 334220) and pharmaceutical and medical drugs manufacturing (NAICS 32541).

Six dimensions used in this research for exploring and measuring the effect of innovative activity on market performance are: quantity or number of patents, quality of patents measured using patent indicators used commonly by researchers, size of company measured by the value of its assets, R&D investments, industry and time. The specific

research questions explored in this study in the context of these three industry sectors with high levels of innovation activity mentioned above are as follows:

- Do companies with more patents perform better than their competitors?
- Do companies with better average quality of patents perform better?
- If such an effect of quality or quantity of patents on market value of firms exists, then how does it vary across the three different industry sectors considered here?
- What are the trends of this phenomenon over time?
- How does it change with value of assets of the company and what are the interaction effects among these dimensions?
- Do companies which are successful in market or are able to do better in R&D subsequently invest more in R&D?

Answers to these questions should give a somewhat clearer picture of how technological innovation or intellectual property can impact market performance of firms, in what ways and how does this impact vary with increasing firm size, amount of innovative activity and quality of innovations. Currently we do not know much about how exactly shareholders/ investors value a firm's intellectual property but in this study we intend to determine whether there is an effect due to intellectual intangible assets or not, and if yes then we discuss and speculate what could be the possible explanations for the observations made. A reasonable guess would be that decision makers in market (buyers of stock) or in the company (R&D managers) do not use specific indicators for quality of patents which we use in our study. Therefore what we are measuring is probably a proxy

for some other information which is readily used by them to support their decision making under the surrounding uncertainties.

Further elaborate discussion on these issues will be done later in this paper in the results section. The remaining sections of this paper are organized as following - next section is the literature review, followed by the section on research design where the model used, assumptions made and data used will be described with the methodology of analysis and discussion over threats to internal and external validity of this research. After that results will be discussed, followed by conclusions, suggestions for future work and policy implications.

Chapter – 2

Literature Review

2.1 Overview

In past two decades, research efforts relating to the use of patent data for market valuation of firms has gained significant momentum, providing growing evidence that R&D output as an intangible capital has become an important determinant of the market value of firms. Patents are a good proxy for R&D output of firms, therefore use of number of patents and other patent based indicators to evaluate and measure the R&D output of firms has become quite popular amongst researchers. Patent data is published by the patent offices, so there is very little chance to manipulate it because patent examiners make their decisions objectively. Moreover, patent data is structured, hence it is relatively easy to use for the purpose of analysis than the data on journal articles, though language of patent data is relatively more ambiguous than an average journal article (Porter, 2005).

2.2 Patent Quality Indicators

Though patent data is authentic and structured, its weakness is that all patents are not of the same quality and hence the same value. Griliches (1981, 1990) presented evidence that the distribution of values for the patents is highly skewed and thus concluded that patent counts or quantity of patents is not a good indicator of R&D output, rather it should be the quality of R&D or patents which determines the value of patents. By then

Pakes and Schankerman (1984, 1986) had developed a model to measure the patent quality and hence value by the observed renewal decisions for these patents.

Trajtenberg (1990) found correlation between the patent citations and independent measures of the social value of innovations. Later Putnam (1996) said that family size of a patent is also a good indicator of its value. Citations from other patents started getting more acceptance as good indicators for patent value when Thomas (1999) analyzed the relationship between patent citations, and renewal decision made by patent owners. He observed a significant correlation between these across a number of time periods.

Another similar paper by Harhoff et al (1999) used a survey to obtain the private economic value estimates for 964 US and German patents for which German patent renewal fees were paid to their full-term expiration in 1995. He noted that the patents which were renewed to their full-term were much more heavily cited than the patents expiring before the full-term, thus establishing that higher the patent's economic value, the more that patent would be subsequently cited. In other words, citations received from other patents are a good indicator of the quality of patents. Then to counter this notion Jaffe, Trajtenberg and Fogarty (2000) conducted a survey of patentees and concluded that though patent citations can tell us a lot about the quality of innovation, they are a noisy measure. In the same year Hall, Jaffe and Trajtenberg (2000) published that citation-weighted patent stocks are more highly correlated with Tobin's Q ratio of the firms that own those patents than patent counts by themselves. This occurs mainly due to the high valuations placed by the market on the firms that hold very highly cited patents.

2.3 Citations

Citation count, though noisy, is a well established indicator of patent quality because only a fraction of patents receive any citation ever. It is a well chosen construct to measure patent quality and value. One big issue with use of citation count as quality indicator for patents is that it is available only later in the life span of a patent. Hence in order to identify some early stage quality indicators of patents, the idea of using number of claims as an indicator of patent quality was proposed by Tong and Frame (1994). They showed that patent counts weighted by number of claims are more highly correlated with R&D spending at national level. Lanjouw and Schankerman (1997) showed high worth patents are litigated more often, and the number of claims correlate well with the probability of litigation for the patent. Later Reitzig (2004) found that number of claims was a highly significant indicator for patent value and quality at an early stage in the life of a patent, where patent value and quality was measured in terms of number of litigation suits faced by the patent.

In 2004, using a composite latent variable Lanjouw and Schankerman (2004) showed that family size, backward citations and claims are all significant indicators of patent quality for which they used patent renewal information as a proxy. It is posited that patents with larger family size, full term renewals, and more litigation suits have these characteristics because they are more valuable in the market.

Around the same time, Bloom and Van Reenen (2002) considered different options possible for a firm regarding the use of its intellectual property and used a real options framework to show that patents have an economically and statistically significant effect on the firm level productivity. They also showed that the higher market uncertainty reduces this impact of new patents on the productivity of the firms because this effect appears slowly distributed over time, unlike the market performance which is immediate and driven by the market sentiments.

Two citation based indicators “generality” and “originality” were discussed by Hall, Jaffe and Trajtenberg (2001) in their description file for NBER patent database. These are used for examining the impact linkages with the other innovations or patents.

$$Generality_i = 1 - \sum_j^{n_i} s_{ij} \quad \dots\dots\dots \text{equation 2.1}$$

where s_{ij} is the percentage of citations received by a patent i that belongs to the patent class j , out of n_i patent classes. The sum is also called Herfindahl concentration index. A higher generality score would mean that the patent was cited by subsequent patents which belong to a wide range of fields. Thus generality is a measure of breadth of a patent. Similarly originality scores are calculated using the references made by the patent. Therefore if a patent references other old patents from a wide range of patent classes then the originality score of the patent would be higher. This is also a measure of breadth of

the patent. In this paper we have used these two indicators (generality and originality) for measuring the quality of innovation of the firms.

In this study we attempt to improve the measure of the knowledge stock or quality of innovation (patents) for the publicly listed US firms in pharmaceutical, semiconductors and wireless industries during the period 1986-1999. To explore whether we can measure the patent quality better we use multiple indicators together - number of patents and their citation counts along with number of claims, generality and originality indices.

2.4 Market Value and Patents

While most of the research work mentioned above deals with exploring and relating how patent information can be used to analyze the industries as a whole and to identify patterns and trends from the data. Narin, Breitzman and Thomas (2004) have developed methods (two US patents) and successfully demonstrated the use of patent information to stock portfolio selection for getting better returns on investments.

They used three citation based indicators for assessing the quality of technologies which a firm owns. These indicators, introduced previously by Deng, Lev and Narin (1999), were the following: Current Impact Index (CII), which measures that relative to all US patents in the industry sector, how frequently the company's patents from previous five years are cited by patents issued in the most recent year; Science Linkage (SL), which is the average number of references a company's patents make to scientific papers; Technology Cycle Time (TCT) which is the median age of the patents cited by a firm's

patents. They regressed these indicators, patent counts, patent growth and R&D intensity as the independent variables with the Market to Book (MTB) value ratio of firms as the dependent variable. MTB is an indicator quite similar to the Tobin's Q which is used most often in research. This highlights the critical importance of use of patent data in valuation and investment decision making for R&D, stocks or mergers and acquisitions.

Deng, Lev and Narin (1999) used these indicators to determine the extra effect of intangible assets on market value of firm, over and above the effect due to earnings or profits of the company. An issue which could be raised in their approach is that current profits, which are most often considered a proxy for the firm's future performance in the market, would partially be the return on intangible assets like intellectual property owned by the firm. Therefore one would not be able to extract out the full effects of intellectual property in analysis.

Beginning with Hausman, Hall and Griliches (1984), until Hall (2006), many researchers in economics of innovation have used R&D expenses, R&D intensity (i.e. R&D/Sales ratio), R&D/Asset ratio or R&D stocks¹ as one of the measures of innovative activity inside the firm. For this study also, conducting a regression using Negative Binomial distribution would have been a better and more insightful idea if we had used the R&D stocks distributed over a time period equal to the expected lag for patent outputs from the R&D investments, after which these returns truncate. Citations data with truncation would provide a good proxy for the quality of patents in that case. The whole stream of

¹ R&D stocks for any firm are calculated using declining balance formula and the past R&D spending history of the firm.

literature starting from Hausman, Hall and Griliches (1984) uses this approach for answering research questions on links between R&D investments, product innovation and firm's market performance. They use Poisson Pseudo Maximum Likelihood Estimator (PMLE) estimates treating data as one long cross-section with a fixed effects model. However, our research question is different; we assume perfect absorption of information by the market and then explore the immediate market response to the information releases about new patents granted and R&D investments by the firms. Moreover we intend to find comparative strategic insights into the relationship between innovation and market performance across three different US industry sectors considered here.

Chapter – 3

Research Design

The research model in this study attempts to find evidence for two forms of decision making about the quality and market potential of technology innovations produced by any given firm. These forms are the decision of stock market, and that of the firm's internal R&D investments. We hypothesize that both of these decisions are favorable when the quality and market potential of technology innovations are perceived good or optimistic. Internal R&D investment decision making symbolizes a firm's belief in the potential of technologies it plans to build and its scientists; whereas stock market's decision in the form of Tobin's Q represents market's sentiments about the expected growth of the firm at any point of time. We examine how and what information the market uses for valuing a company's prospects. This study is an exploratory study only. We approach the issue under consideration in a hierarchical manner for an overview of causal relations between the quality and quantity of innovations and these forms of decision making without delving to the deepest levels in order to find out the exact mechanics involving many other variables, in which we are not interested here.

3.1 Model

The regression model we use in this study has long been in use by researchers in economics of technological change (Hausman, Hall and Griliches, 1984). It is derived using the standard Cobb-Douglas production function for the market value of firm in

terms of the value of its capital assets and knowledge capital. This Cobb-Douglas function is then transformed to measure the impacts of these independent variables on market performance measured in terms of Tobin's Q ratio.

In our analysis, at first we assume knowledge capital (K) as a function of number of patents, ratio of R&D expenses to the capital assets C (R&D/C) and time, and run the regression. After that we include the quality aspect also and assume patent quality as a Cobb-Douglas function with patent indicators used in this study and time as inputs. We then compare the results from both of these assumptions.

$$K = \eta * Pats^\alpha CiteF^\beta Cite^\lambda Claims^\delta \left(R \& D / C \right)^\eta T^\mu \dots \text{equation 3.1}$$

The final functional specification we use for our regression analyses is the following:

$$\ln Q_{i,t} = \ln A + \beta_0 \ln C_{i,t} + \beta_1 \ln Pats_{i,t} + \beta_2 \ln CiteF_{i,t} + \beta_3 Cite_{i,t} + \beta_4 \ln Claims_{i,t} + \beta_5 \ln \left(R \& D / C \right)_{i,t} + \beta_6 \ln T_t \dots \text{equation 3.2}$$

We also substitute forward citations by “generality” and domestic references (citations) by “originality”, which are the other quality indicators for explaining the breadth of innovation thereby allowing for a variation in the regression. We make an assumption for the relation between forward citations and generality scores, and between citations and originality scores. We verified this assumption by performing a regression, which was a

highly significant regression with an F-statistic value = 417.24 for the analysis. The following relationship was assumed:

$$Generality = B * CiteF + c \quad \dots \text{equation 3.3}$$

For the factor A, we make an assumption that it is derived by multiplying a generic constant and an industry constant. Therefore we get

$$A = \alpha \kappa \quad \dots \text{equation 3.4}$$

To account for this industry based contribution to the constant, we add dummy variables for different industries in our regressions. If we use the additively separable linear specification for firm level market value function, which Griliches (1990) and then later others have used in previous research work, then again we get the same final model equation for our regression analysis. So though his specification makes a good assumption that the marginal shadow value of assets is equalized across firms, but later one has to assume that the contribution due to the knowledge capital is relatively very small, which we are reluctant to assume for the dataset used here.

Therefore by using this model, we investigate effects of the six different factors identified from previous literature on the market performance of a firm which is measured by Tobin's Q ratio. These factors are: quantity or number of patents, quality of patents

measured by commonly used patent indicators, size of company measured by the value of its assets, R&D investments, industry and time.

We conduct regression analyses on the dataset used here with this model in a hierarchical manner, by carrying out regressions using independent variables for six different factors mentioned above while testing different assumptions or scenarios. To begin with, we use number of patents as the only indicator of knowledge capital and then later we add more variables which measure the quality of patents. We also carry out the same analysis using categorical variables for company size and patenting activity in order to determine if there are some differences in patterns observed. We conduct this analysis with a combined dataset built for our regressions by using the independent variables and dummy variables for industries together. Thereby we assume the same variance in the error term across all observations (homoskedastic). As this is time series data, expecting heteroskedasticity is highly probable, so we check whether assuming homoskedasticity was correct by running separate regressions for industries, separately for each year and also by using dummy variables for all years. This is important because the data used is from 1986-1999 and we seek to identify patterns or insights which could be valued even now. To delve deeper, one could even use firm level dummies to isolate any firm specific effects, which do exist.

Other phenomenon of interest, which we explore in this study, is whether firms which are performing better in the market or creating good quality intellectual property, invest more in R&D or not. For this we regress R&D as a dependent variable with Tobin's Q and

indicators of quantity and quality of intellectual property. Earlier when we had Tobin's Q as our dependent variable we were analyzing the market's decision on a firm's potential performance or revenues and how it was influenced by the indicators used in the analysis. On the other hand, by using R&D investments as a dependent variable we analyze how R&D managers, board members or executives of the firm assess or make decisions regarding the firm's potential revenue streams in the form of new products created by their R&D divisions. It is assumed that they take their decisions in the best interest of the company based on the potential they see in the technology and their firm's chances of capitalizing that potential.

3.2 Assumptions

In this analysis we make some assumptions which are discussed below with the reasons and logic behind doing so, along with their possible consequences in the analysis:

1. We assume that patents represent the intangible assets, due to which stock market values the firm more than its book value. This assumption can be justified by the fact that patents are exclusive rights to exclude others from the inventions, hence treating them as intangible assets is reasonable. Patents have value in themselves even when the option of commercialization has been delayed or not considered by the firms which own them. This could also be due to strategic reasons like licensing revenues or blocking competitors. The literature has provided support for this assumption since the pioneering approach from Griliches (1981) by finding empirical evidence for it.

2. Next we assume that any patent's economic value is a good indicator of the quality of patent and vice versa. This assumption is also supported by the literature because it holds true statistically, but it need not be true always. Good inventions can and do fail in market due to many reasons and relatively not so novel technologies can also succeed in the market. As we are using a sample population large enough, we can expect this assumption to hold true.

3. Another assumption we make is, while considering the contribution of intellectual property to the intangible assets of the firms, we do not consider the contribution of licensing (both licensing in and out) of technologies. We use only the technologies created by in-house R&D. We do so because in this study we are interested in the effect of intellectual property ownership alone. Licensing activities can have a significant effect on the productivity of the firm and thereby on its performance in stock market. We do not have data on licensing or cross-licensing of patents amongst companies, so in order to concentrate on effects due to intellectual property owned by the firms, we assume that the effect of licensing activity is same for all the companies, which may not be true and could lead to either erroneous results or bad fit. As we analyze only the firms from R&D intensive sectors and ones pursuing in-house R&D, hence our assumption should not distort the results much.

4. Normal distributions have been assumed for the variables used in the regression analysis. Measures of most of the factors used here are highly skewed but because we use natural logarithms of their values in the analysis, so their distributions can be approximated to normal (Cameroon and Trivedi, 1998). We checked this assumption for all the variables by plotting a histogram for the log values used in the analysis and all the distributions appeared to be near normal, which we assume is good enough here.

5. We also use Generality and Originality scores of patents in place of Forward Citations and Backward Citations (domestic references) in some regressions to see how these indicators for breadth of technology correlate with the market performance of firms. For verifying whether the substitution relations assumed by us were correct, we regressed these new indicators with the respective citation counts and validated them.

6. Effects of any Mergers and Acquisitions activity have been assumed to be negligible over the patterns identified in the study. These effects are highly significant for any individual company and will definitely lead to some error in the results, but we do not have any data on the changes in the company structure over time. To verify this assumption we conduct our regression analyses with the dummy variables for years in place of using one variable for time and also separate regressions for each year. In effect, it means we take an industry level snap-shot for each year and then compare those snap-shots. This way, if the

general patterns observed in the normal analysis appear in these regressions as well, then we can be confident that our assumption is reasonable and empirically supported by the data used here.

7. We assume that the effect of historical events like the dotcom boom or any other industry or economy specific events can be ignored without getting spurious results in terms of patterns identified. Use of dummy variables for time should be able to isolate any such effects, and if our overall conclusions do not change then this assumption holds true.
8. It is assumed that each innovation which results in improved performance, better productivity or a new product is being patented. Thereby we ignore the effects of trade secrets and business methods. Business method patents came into the picture in late 1990s, but even before that innovative business methods were valuable to firms. This will lead to an increase in error like the other assumptions but we do not have any data on these forms of intellectual properties for companies.
9. We do not consider the effects of labor or human capital in the performance of companies because we do not have data on that. Again the assumption is that the effect of labor is independent and will not change our conclusions though it will decrease our R^2 of the analysis.

10. Assuming a common model for all the firms in all industries is an obvious oversimplification, but we suppose analyzing these three R&D intensive industry sectors (Semiconductors, Pharmaceuticals and Wireless) only should control for that. To investigate inter-industry differences in our analysis, we use dummy variables for the semiconductors and wireless industries, treating pharmaceutical as our base sector.
11. Though log values of the variables have been used for analysis on the basis of the model assumed but not all non-linear relations in the data could be accounted for here. We assume that it will not lead to much change in the significance levels of the variables in the results of analysis though standard error will increase due to this leading to a lower R^2 .
12. In our analysis, we do not account for the temporal correlations between the patenting activity, patent quality and R&D expenses across different years. Firms which invest more in R&D or the ones which have very large patent portfolios will tend to get more number of patents and average citations. These effects would contribute to and can be determined using the R&D productivity of firms, which further leads to better market performance relatively slowly over time. We do not seek to identify these effects as they are not apparent before a significant amount of lag time, whereas patent indicators used in this study have an immediate effect over our dependent variable of interest, market performance.

13. Any firm specific effects on the market performance and R&D innovation have been neglected. These would include the organizational culture, leadership, brand value etc. It could have been captured by using dummy variables for each firm but we are not interested in those effects.
14. We have assumed that Tobin's Q is a good indicator to measure the firm's market performance for our research because it measures the growth expectations of a firm. Moreover, we seek to determine the effect of intellectual property on the market performance and thus on the market value of firm. As Tobin's Q greater than one indicates the presence of intangible assets, therefore using it as an indicator to measure the effect of intangible assets like intellectual property is a valid assumption.
15. We assume that in the selected industry sectors, value of the firm is primarily driven the intellectual property owned by the firm. This assumption is in congruence with the popular beliefs about these industries, and high patenting activity observed in these sectors testifies for this belief.

We seek independent effects due to only the variables we are interested in, so obviously our model does not take into account many other different and independent variables which could have affected the performance of firms (e.g, locational advantage or disadvantage, industry cycles, firm's life in market, etc.). If data is made available for any more such variables then a more rigorous and insightful analysis can be done.

3.3 Data

The dataset used in this study has been made by combining two datasets. Company performance data was taken from ATIVO's Research LLC's database. We take yearly data on Tobin's Q-ratio values, capital assets and R&D expenses from this database. The second database used is the publicly available NBER patent database, which is available online from the NBER website. This database was used to get the data on all the patent based indicators used in this study.

We first select all the companies which were present in the stock market for five or more years during the period 1986-1999 in semiconductor equipment manufacturing (NAICS-334413, 56 firms), wireless and communications equipment manufacturing (NAICS-334220, 43 firms) and pharmaceutical and medical drugs manufacturing (NAICS-32541 69 firms). Then we looked for the patents belonging to each of these companies in the NBER patent database.

Company name matching in NBER database is not accurate, firstly because company structure data over time is not available in it and secondly, not all patents have their assignee/owner names matched. Still the database is of great value because it has information on all the patents for the period 1963-1999. It has a numeric variable "assignee-id" for each patent, which when searched in another table listed in the database gives the name of the company. There can be many assignee-ids allotted to each company due to different names of subsidiaries or variations in their company names.

Hence, in order to find more matches, a MS Excel macro was made to download the patent numbers for each company from USPTO website. We then used these patent numbers to extract the other information from the NBER database. After combining both set of patents and keeping the unique records only, the number of records in the dataset for our use got increased by around 8-10%. Both approaches missed some patents. Patent numbers downloaded from USPTO website were less than what one could get from the name search in the NBER data. Still 8-10% patents numbers downloaded from USPTO had not been matched with the same company in NBER data. All possible name variations for the company names were not used while downloading the patent numbers from USPTO website using our macro, so these numbers were less than what one can get from NBER. Finally, with this two-pronged approach we got 13,831 patent records for pharmaceutical companies, 41,828 patent records for semiconductor firms and 13,732 patent records for the companies in wireless sector.

In the NBER patent database we found many missing entries for some fields, especially in the number of claims field. Many times number of claims (which can not be zero) was found missing in the NBER data. Most records with missing claims were from year 1999 but there was a large number of missing entries for claims in other years too. So we substituted these blanks by the mean value of claims for that company's remaining patents in that class in that year. This was done assuming that the same patent-attorney wrote those patent applications for that company in that year, so the number of claims in those patents should have the same mean value. For those records where number of

claims entry could not be filled using this method, we relaxed the same year assumption at first, and then for the ones still remaining the assumption of same class was relaxed. After that only about hundred or less records in each industry category had the number of claims missing, so those values were filled in by manually looking up for those patents on the USPTO website.

For some records, the entry for number of citations made (domestic references) was also missing and an approach similar to claims was used with random manual checking for a fraction of these entries to confirm that this process is reasonably correct. There is one update available for the NBER data for patents till 2002 at Prof. Bronwyn Hall's personal website at University of California Berkeley. This update does not have values for number of claims so we could not use the whole data for analyzing patents till 2002. But we used this data to find the updated number of citations received and hence generality scores for all the patents in our dataset.

Another issue we faced was that number of citations received and generality scores for the patents in later years were obviously not a true value for those. Moreover we used only the number of citations received or generality scores for building our dataset, so truncating the citation values at a four or five year time period was not possible. To tackle this we first plotted the distribution of these values for all patents in the NBER database, all patents in our data set and then for the patents for each industry sector in our dataset (Figure 1). From a cursory examination of the plot, we could see that the mean values for the first four years 1986-89 were roughly the same, and then it started decreasing slowly

initially and approached zero rapidly in last few years. We estimated a good fit for this distribution $(1 - x^3 + c)$, where x is the ratio of time passed from base year to total number of years and c is the mean value for the last year).

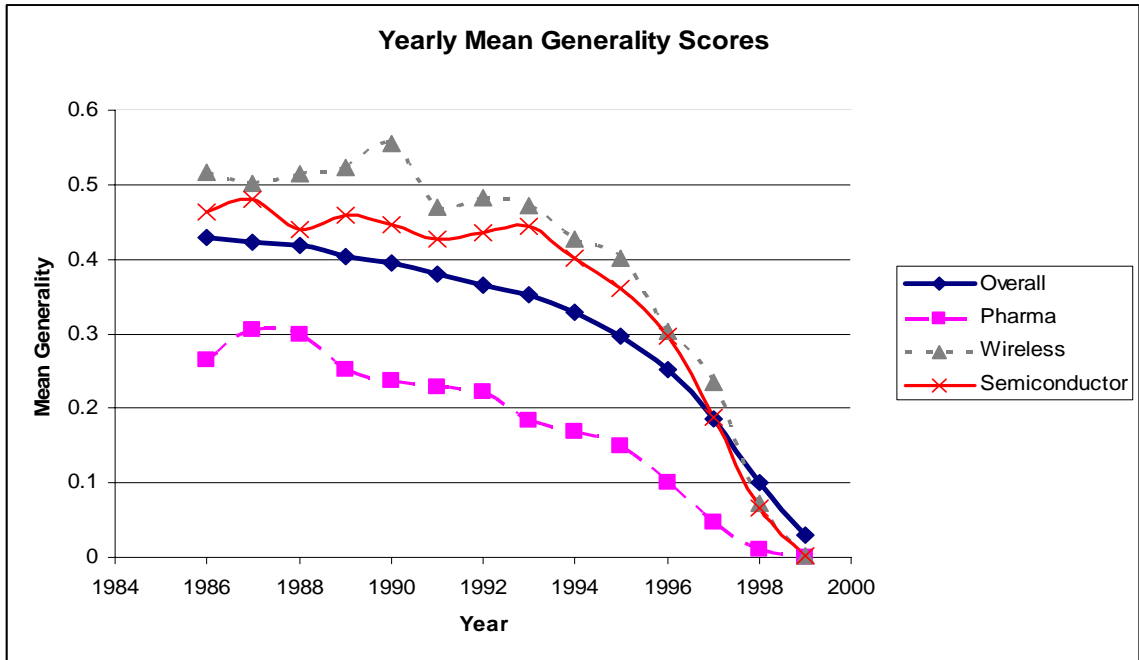


Figure 3.1: Mean Generality Score over Time

Citations received and generality scores do not always have a value because not all patents receive a citation in their lifespan, hence any substitution to tackle the problem of having erroneous values for these fields in later years could be of multiplicative form only. Additive substitution would give non-zero values to even those patents which probably will never get a citation. So we calculate the mean value for first four years and then assume that yearly mean values will tend to reach this value. Thus, we determine a yearly multiplying factor for the patents in that industry by dividing the mean of first four years by yearly mean of the value calculated from the data. Multiplying

generality/citation values in all the records from 1990-99 with the yearly industry multiplication factor determined, we get the corrected values to be used further in our analysis methodology.

From the plot, we also found that yearly means of generality scores for patents tend to reach a uniform distribution given enough time. Another notable observation was that yearly means of generality scores in the pharmaceutical industry were significantly less than those for the patents from semiconductors industry, and these mean values for wireless patents was significantly higher than for semiconductor patents. As generality is a measure of breadth of use of technology, obviously wireless application will be more generic in nature, followed by semiconductors and then pharmaceutical patents which have very narrow applications in only the field of medicine. Hence the pattern in generality scores observed is as expected for the industry sectors considered. It shows the usefulness of IP protection in these industries and also the relatively higher importance of breadth and thus of licensing activities in the wireless and semiconductor sectors.

The next step to be taken was the fixed effects rescaling for the citations, references and claims. Mean values for all these fields has been increasing over time due to either the sheer propensity to cite or claim more, or other similar reasons. For carrying out fixed effects rescaling, we divide each such entry by the yearly mean value for that group or patent class. Thus we need not make any assumptions about the underlying reasons due to which differences in yearly mean values occur for each of these counts. Though the disadvantage here is that by assuming no structured reasons for the difference, we can not

distinguish between the real and artificial differences (Hall, Jaffe and Trajtenberg, 2000). For claims, the scaling we did was a little more sophisticated, as we used the mean values for the same company's other patents in that class, and relaxing this assumption successively as we did in substituting values for the missing claim entries.

Having done all these rescaling and corrections, we calculate the yearly mean values for all fields of interest for each company and then in a combined dataset add that year's observation for Tobin's Q, capital assets and R&D expenses. Next, we create new columns for the natural log values of these variables. The next step was to create dummy variables for semiconductors and wireless companies with pharmaceutical as the base sector, and for each year with 1986 as the base year. After this we created interaction terms of interest here and categorical dummy variables according to the asset value or company size and patent counts. We also create another datasheet of only those observations where Tobin's Q was greater than one. This is to compare the difference between analysis results for all observations and observations with Tobin's Q value greater than one.

3.4 Method

For this analysis, we perform our regression analyses in the following hierarchical manner while noting down results, key observations and differences at each step. To begin with, we do the regression with just the dummy variables for industries and the other main independent variables in the dataset by considering number of patents as the knowledge stock. We do the same analysis for all the observations in the data and then

observations with Q-ratio greater than one only. Then in place of the natural log values for number of patents and assets, we use the categorical variables created for size and patent counts. Next step was to include the indicators for patent quality into the analysis, where the natural log of citations and forward citations were used first, and in later analyses those were substituted by the originality and generality scores respectively.

Subsequently, we drop the variables which were consistently appearing as totally insignificant with high variance inflation factors and include the interaction terms for the remaining independent variables. Then we conduct the same analysis separately for observations from each industry sector and then separately for the observations from each year. We note changes in significance, variance inflation factors of the main independent variables, dummy variables and interaction terms for each analysis and the changes in F-statistic and R^2 of the regression for each of these analyses. Cut-off value for variance inflation factor used was 25, and for checking significance of variables we consider 1%, 5% and 10% significance level in our analyses.

Furthermore, we analyze effects of Q-ratio, number of patents and patent quality indicators over R&D, in a similar manner as above. This stage-wise approach makes sure that the conclusions we draw are consistent. Additionally in order to find the length of lag period for returns to R&D in these three sectors we use cross correlation method between the time series for Q-ratio and R&D expenses. The cross correlation analysis was done individually for only those companies for which data for all 14 years was available. The lag values once determined will be compared with expected values for these industry sectors to validate the results.

3.5 Threats to Validity

In order to be able to make suitable and fair measurements for drawing right conclusions we need to consider the four threats to validity, proposed by Cook and Campbell (1979) and discussed by Porter (2005) in perspective of technical data. These four general concerns are following:

- Internal validity
- Validity of statistical conclusions
- Construct validity
- External validity

In the context of our research following is the discussion over each of these threats to validity mentioned above:

3.5.1 Internal Validity

This is to check whether we are able to distinguish if the analysis performed has made any real difference by bringing out correct conclusions or we interpret incorrectly because the data itself had some sort of bias in it, which mislead the conclusions derived from the analyses.

Hence in order to establish internal validity of the research design we have to explore whether the data used is a correct measure of what we want to measure and there are no

distortions, events or other reasons due to which data may be potentially biased. As we have discussed in the previous section describing the data, though imperfect, the indicators used here for measuring innovation are the best available right now according to the popular school of thought in the literature on economics of innovation or technological change. Apart from that, we have taken full care in replacing missing values in a careful manner that does not distort the overall distributions and patterns in data. The only issue here is that by doing so we tend to decrease the error terms because we favor the averages. This may increase our R^2 for different analyses a little but as we do not use R^2 as the main criteria for selection or rejection of variables, or deriving the conclusions, rather we just report it, so it should not affect our conclusions.

Secondly, in order to take care of any historical events or industry specific reasons which may have altered the patterns and trends in data, we also perform our analysis with dummy variables for different industry sectors and different years. Additionally, we split the data year-wise and industry-wise, and conduct separate regressions on these sub-sets of data. A third factor which we have ignored is if there are some company specific reasons for difference, like patenting behavior etc. We do not use separate dummy variables for each company to isolate such effects but we create categorical dummy variables according to firm-size and the extent of patenting activity. This should take care of such issues to a reasonable extent.

The fourth aspect to be considered here is the real probability distributions of these variables. Patent data is count data, which by nature follows Poisson distribution, or more

precisely it belongs to the standard generalized form of Poisson, Negative Binomial distribution. Thus, an Ordinary Least Squares (OLS) estimator is clearly inappropriate for the underlying distribution of patent data, because it specifies a conditional mean function which may take negative values and a variance function which is homoskedastic, so Cameron and Trivedi (1998) suggest that OLS with log transformation of the dependent variables should be used to transform this skewed distribution into a normal distribution. That way one can use OLS with patent data; hence our log transformation for the model serves this purpose also, along with linearizing the regression model used.

3.5.2 Validity of Statistical Conclusions

This issue involves validation of our results to assure they are statistically significant and are not merely chance observations. Normally this validation is done over two dimensions - sample size and reliability of measurements. Sample size is not relevant here as we use a population and not a random sample for our analysis, and we seek to understand the patterns and relations in this population. We also conduct and prefer the combined analysis with dummy variables for industries and interaction terms in this study because it assumes the same error distributions all throughout thereby reducing chances or erroneous conclusions.

Additionally to increase the robustness of results, we conduct our analysis in a hierarchical manner and consider only consistent results as conclusions and report the other results. Our use of a commercial database for financial data of firms and patent data

(which is structured and reviewed by the USPTO before granting patents) from NBER to construct the dataset used in the study enhances the reliability of measurements by decreasing noise.

3.5.3 Construct Validity

This section deals with assuring that the constructs and causal relationships used for making the model are proper and there are no omitted variable biases present. We build our model based on a popular school of thought from the literature in the field of economics of innovation. Hence the constructs and causal relationships used to derive this model are well in agreement with the theory on this topic. Furthermore, our use of industry specific data may isolate some industry specific local patterns in the data.

3.5.4 External Validity

In this section we discuss whether the conclusions derived from this study can be generalized. We conduct a hierarchical analysis; carrying out the whole analysis combined and then in parts, and finally with the use of dummy variables and interaction terms to isolate the local patterns and trends (industry or year-wise) from the global patterns in the data. Hence, though our data is old (1986-1999), we can generalize our main conclusions (global patterns) which appear consistent in analyses and are in accordance with theory, for these three industry sectors with high patenting propensity or any other similar sectors like biotechnology, chemicals and nanotechnology in the US. Any industry specific or time dependent conclusions can not be generalized.

For firms outside US, due to the different patenting systems, generalizing the conclusions about the extent of correlation between the patent quality indicators with the firm performance could be misleading. However, the generic patterns and causal relationships should remain similar reflecting the positive impact of intellectual property on firm performance. Hence generalizing these results should not lead to misleading conclusions in the context of other countries.

Chapter – 4

Results and Discussion

Results of this study are interesting and align well with the findings of previous research work in this field. As mentioned earlier, we attempt to track and measure the changes in performance of firms with respect to changes in six basic variables namely: quantity or number of patents, quality of patents measured using patent indicators used commonly by researchers, size of company measured by the value of its assets, R&D investments, industry and time. Following is the table describing the variables used in our research:

Table 4.1: Description of Variables

Variable Used	Description
LnQ	This represents the logarithm of yearly mean of Tobin's Q of a firm
LnAsset	Logarithm of estimated value of Capital Assets of a firm for the year
LnPatents	Logarithm of number of patents granted to the firm in any given year
LnCiteF	Logarithm of the corrected average number of forward citations received by the patents of a firm in a particular year
LnCite	Logarithm of the corrected average number of citations made by the patents of a firm in a particular year
LnClaims	Logarithm of the corrected average number of claims made by the patents of a firm in a particular year
Generality	Corrected average generality score for all the patents of a firm in a particular year
Originality	Corrected average originality score for all the patents of a firm in a particular year
LnT	Logarithm of time passed starting from the base year, T = 1,2,3,....,14 for 1986-99.
Sdummy	Dummy variable for firms belonging to the semiconductor manufacturing industry
Wdummy	Dummy variable for firms belonging to the wireless manufacturing industry

Table 4.2: Correlation Matrix of the Variables Used in the Analysis

	LnQ	LnAsset	LnPatents	R&D/Assets	LnCite-F	LnCite	LnClaims	Generality	Originality
LnAsset	-0.136 (0.000)								
LnPatents	-0.047 (0.059)	0.785 (0.000)							
R&D/Assets	0.103 (0.002)	-0.196 (0.000)	-0.153 (0.000)						
LnCite-F	0.047 (0.016)	-0.033 (0.328)	0.101 (0.003)	-0.178 (0.000)					
LnCite	-0.024 (0.479)	0.015 (0.665)	0.049 (0.143)	0.70 (0.038)	-0.044 (0.190)				
LnClaims	0.057 (0.088)	-0.105 (0.002)	0.011 (0.747)	0.119 (0.000)	0.183 (0.000)	0.158 (0.000)			
Generality	0.031 (0.087)	-0.080 (0.017)	0.054 (0.105)	0.011 (0.734)	0.602 (0.000)	0.162 (0.000)	0.215 (0.000)		
Originality	-0.019 (0.578)	-0.093 (0.006)	-0.051 (0.128)	0.084 (0.012)	0.013 (0.688)	0.601 (0.000)	0.115 (0.001)	0.206 (0.000)	
LnT	0.028 (0.404)	0.117 (0.000)	0.060 (0.075)	0.398 (0.000)	-0.468 (0.000)	0.271 (0.000)	-0.003 (0.932)	-0.164 (0.000)	0.128 (0.000)

()Values in parentheses are the p-values from significance test of corresponding correlations

4.1 Preliminary Analysis

We began our analysis with calculating the correlation matrix (presented in Table 4.2) for the variables used in our analysis. A first look at the correlation matrix suggests that we can expect to get the variables representing assets, patent counts, R&D, forward citations to appear significant in our regression analyses. Also, we can expect all patent indicators to correlate with R&D/Assets at a high degree of significance in our regressions with R&D/Assets as dependent variable. Next we conduct our first preliminary regression, with number of patents as the measure of knowledge stock or innovative activity inside the firms. Following are the results of this preliminary analysis (Table 4.3):

Table 4.3: Coefficients for the Top-level Combined Regression

Predictor	Coef	SE Coef	T	P	VIF
Constant	1.09289	0.08198	13.33	0.000	
LnAsset	-0.08432	0.01265	-6.66	0.000	2.9
LnPatents	0.07427	0.01606	4.62	0.000	2.7
Sdummy	-0.30262	0.04153	-7.29	0.000	1.3
Wdummy	-0.36244	0.05428	-6.68	0.000	1.3
LnT	0.09217	0.02733	3.37	0.001	1.4
RDbyAsset	-0.111	0.1182	-0.94	0.348	1.4

As we can notice, “p” values and Variance inflation factors (VIF)² for all the coefficients except R&D/Assets ratio are very low, showing that all these variables are highly significant and have independent effects over the performance of the company. Though the correlation analysis suggested so, R&D/Assets do not appear significant in the regression. This may have happened because the causal link we are looking for between R&D/Assets and Tobin’s Q may be opposite in direction for the immediate time-frame in the context of dataset used here. Theoretically firms can not get immediate returns to R&D in terms of revenues. and if Q-ratios do not depend on recent R&D expenses, then we interpret it as: the market does not react on announcements by firms about which new R&D projects they will start or how much money they are investing in their R&D for any cutting edge technologies. Rather the market absorbs the information about patenting activity as number of patents granted is positively correlated with Tobin’s Q of the firms. According to the literature, we can expect the returns to R&D to be distributed over time and we discuss it later in this section. For this analysis R² was low (10.6%), but F-statistic value was quite high (17.43), thereby making this regression highly significant.

Following the methodology described earlier, we add the quality aspect of the patents granted with the interaction terms. We note that the quality indicators - citations (domestic references), originality and claims – turn out to be statistically totally insignificant in this analysis along with R&D/Assets. At the same time, the interaction terms which were significant in this analysis extracted the significance effect from the variables representing assets and number of patents. As expected, VIFs also increased

² VIF are used to test for multi-collinearity in multiple regression analysis. High values of VIF indicate multicollinearity problems in the analysis thereby suggesting that results may be spurious.

due to the use of interaction terms, but in terms of inference we come to know that interaction effects amongst independent variables exist in our study, hence not just the quantity and quality of patents are important but having both of these together enhances the value of firm even more, making it a perfect metaphor for “whole is greater than the sum of parts”.

Furthermore, these results tell us that bigger companies with more patents are perceived as better investment opportunities by the market. Also for large firms, effect of better quality of patents is less which probably hints that larger organizations become institutions and their priority is to patent every innovation in order to manage R&D well. Further insights on this about how it is different for small and big firms will be discussed later in the text where results from the categorical analysis will be described.

We also note that the market performance of firms in these sectors is improving over time, thus suggesting that market was valuing high-technology companies more during 1986-99. Another interesting observation here is, forward citations and generality are the quality indicators of patents which are determined at a later stage in the life of a patent, but they emerge out as highly significant in our analyses. This possibly means that the market has intelligence to foresee which technologies will be more successful in the long run. As generality depends on the citations received by the patent from the patents which are granted later, we can speculate that it indicates later developments on the same technology are market driven. If market analysts of market research predict good revenues from any stream of technology, more R&D will be carried out in that

technology field, hence these initial patents will receive more citations later. Because investments in the stock market are driven by such revenue forecasts, these patents and the companies owning them are relatively valued more, even at an early stage in the life of patent.

4.2 Detailed Results

Moving ahead following the analysis methodology, we perform regressions for Q-ratio greater than one, for categorical variables for capital size and number of patents, separate regressions for industries and years and finally repeat these analysis after dropping the variables R&D expenses, originality, domestic references (backward citations) which were consistently insignificant in the analyses. We note down all the results and following is the table (Table 4.4) of our observations which were consistent in all or most of the regressions done in this study.

We aimed to explore answers for some specific research questions in this study along with any other patterns or trends which can emerge from our analyses of this data. Our findings from this research are following:

- Results of this study indicate that companies with more patents being granted in any specific year perform better on the stock market in that year. No difference was found in this observation even when categorical variables were used. When interaction terms for interaction amongst the variables LnPatents, LnAsset and

- generality were used, then the significance of LnPatents gets transferred into the interaction terms, but that does not change our interpretation.
- Our results show that firms with better average quality of patents consistently perform better on the stock market. In our analysis, only one of our patent quality indicators, citations received (also its variant generality) came out significant (rather highly significant). This is a strong indicator of quality (other strong indicators are family size, renewal data, international patent families, litigation data) which can be extracted out from patent data and which we have in our dataset. Other indicators used are domestic references and number of claims, which are not as strong, and were not found significant. But if patent quality appears as highly significant in a macro level analysis like this one, then a more sophisticated patent data analysis should be able to extract out the significance of these variables as well.
 - Regarding the question, “if such an effect of quality or quantity of patents on market performance of firms exists, then how does it vary across the three different industry sectors considered?” Our answer is: Dummy variables for industry sectors were found highly significant consistently in all analyses conducted. The coefficients for the variables were negative. Therefore we can say that semiconductor and wireless equipment manufacturing industries give lesser returns on intellectual property than the pharmaceutical industry. Our interpretation of this is, firstly semiconductor and wireless industries are more

Table 4.4: Combined Table of Observations

Parameter	Causal Link	Expected effect	Observed Effect	Significance level
Capital Assets	Market performance of firm is measured in terms of its capital and intangible assets	Highly Positive	Negative	***
Number of Patents	Patents are a good measure of intangible assets of the firm, hence they should affect market performance	Highly Positive	Highly Positive	***
R&D Expenses (R&D/Assets)	R&D expenses are the investments to develop intangible assets	Positive	None	No Significance
Citations Received	Most reliable indicator for quality of a patent and its inventiveness, hence should affect the market performance	Positive	Highly Positive	***
References made (Citations to other patents)	Indicator for quality of a patent, hence should affect the market performance	Positive	None	No Significance
Generality	Most reliable indicator for breadth and quality of a patent, hence should affect the market performance of firm	Positive	Highly Positive	***
Originality	Indicator for quality of a patent and its breadth, hence should affect the market performance	Positive	None	No Significance
Number of Claims	Indicator for quality of a patent and its inventiveness, hence it should affect the market performance	Positive	None	No Significance
Time	Represents the time dimension of market performance of firms. On average this has been improving over time	Positive	Highly Positive	***
Industry: Semiconductor	Different industry sector than Pharmaceuticals	Unknown	Negative	***
Industry: Wireless	Different industry sector than Pharmaceuticals	Unknown	Negative	***

competitive and second, most products of the companies in these sectors did not fall into the category “vital necessities”, if not in the “consumer luxury” segment in the period 1986-99. Both of these reasons explain why these sectors gave lesser returns on the quantity and quality of intellectual property owned.

- Another observation is that the time variable in our regression was always found highly significant with a positive coefficient. This means on average market value of high-technology firms in this time period (1986-99) was increasing with time. It could be due to two reasons, markets were growing as surplus income of people was increasing during this period, and secondly, value of intellectual property was growing in general. Though both of these reasons are also quite correlated but it would mean that firms started giving more importance to intellectual property in their business during this period.

- Along the company size dimension, which is measured by capital assets, we found that coefficient for it was consistently highly significant but contrary to our expectation it was negative. But in the categorical analysis where we used dummy variables for different company sizes, the coefficients for small company sizes were coming negative and for large companies they were positive. Our interpretation of this is, smaller companies are not the most popular ones on the stock market, even when they have a good patent portfolio. Obviously small companies are relatively risky investments in market because they normally do

not have long history, muscle power (capital assets), brand which can make the intellectual property more valuable. Additionally, small firms face the risk of failure during expansion or scaling up. This is well in line with the common notion that value of intellectual property also depends on who owns it, along with other variables. Another noteworthy observation is that most often small companies have small patent portfolios, and average patent value for larger portfolios is more than the average patent value for small portfolios (Reitzig, 2004 and Pitkethly 1997).

- We also tested for the interaction effects amongst these variables in our analyses. Interaction between assets and industry sectors was not consistently significant in different analyses, but coefficients were positive, meaning that bigger companies are in a relatively advantageous position in the semiconductor or wireless sector than pharmaceutical firms. Interaction between patent quality and industry sectors was never significant, and between patent counts and industry sectors was significant with negative coefficients. This would mean that in semiconductor and wireless sectors, each new patent is of relatively lesser value than it is in pharmaceutical industry. A reasonable explanation for this is, semiconductor and wireless industries are very consumer oriented (demand driven) and fast changing (also converging) industries hence more competition will lead to lower margins, whereas in the pharmaceutical sector products remain relevant for longer time periods and suitable alternatives are rare (supply driven) so patents are expected to be more valuable.

When we consider the interactions between the variables for time, assets, patent counts and quality (forward citations or generality), then results show that with time the importance of assets has grown, whereas importance of patent counts and quality has decreased over time. These results are significant but variance inflation factors for the interaction terms are quite high (35, 59, 44) which means a significant amount of multi co-linearity is present when these interaction terms are included in the regression analysis; hence it is highly probable that these results could be spurious.

Interaction terms amongst assets, patent counts and generality were also tested, and results show that interaction between assets and patent counts, and the interaction between assets and generality was highly significant, whereas interaction between patent counts and generality was significant at 5% level some times and 10% level some other times. Coefficient for interaction between assets and generality was negative and other coefficients were positive. This would mean that for larger firms in any of these sectors, patent quality is relatively not as important as for small firms, in order to perform better. Whereas patent counts are definitely more important for large firms, alternatively it may also mean that large firms and their investors do not give as high a weight to quality of patents as small firms or their investors. For the third interaction term between patent counts and generality we have a positive coefficient, thereby meaning that better quality patents are more valuable in general.

- Regressions conducted separately and with year dummy variables did not show much variance in terms of results and significance levels of the variables, but some of these regressions were not significant by themselves (low F-statistic), though number of observations was enough to provide statistically significant results. Separate regression for those years for which year-dummies were appearing significant had low F-statistic. Regression for some other years also had low F value. We speculate that this means our assumption of same variance everywhere for error term is right, though we may be missing out on accounting for any general or industry specific events or other causes which affected the performance of firms in the time period 1986-99. Probably at the cost of goodness of fit (here R^2) we were able to get good significance level (high F-statistic) for our analyses in this study.
- Analysis of all the observations v/s observations with $Q > 1$ showed that we get better fit and improved significance levels when we use data for $Q > 1$ only, hence we decided to do remaining regressions with observations where Q was greater than one. An explanation for this would be that we are exploring how intellectual property owned by a firm can explain value of its intangible assets, which are apparent only for observations with $Q > 1$. There could be many reasons due to which firms do not perform well in the market and we are not using any variables which could address that.

- As we noticed before, R&D expenses were not at all significant. This means market does not give value to the announcements about future R&D direction and spending. Literature says that returns to R&D are lagged and distributed over time in years after investment. We therefore attempted to determine the length of lag time between R&D investments and returns in the form of company performance or high Q-ratio. For this we took only those 35 firms for which we had data for all fourteen years. We then performed a cross correlation analysis between the time series for Q-ratio and R&D expenses to determine the length of lag period individually for each of these firms. In this analysis one has to calculate the correlation between the two time series vectors by sliding over other (correlation of $F(t)$ and $G(t+k)$ where $k = \{-n, n\}$). The point where one observes a peak in the correlation, that value of “k” is taken as the length of lag period in years.

The lag period could be determined for only 16 firms, for rest of the firms lag values determined either did not make any sense (e.g. negative or zero values, very high values) or there was no peak observed in the cross-correlation, so we dropped such observations. Average and median length of lag period for pharmaceutical firms was 7 years (varied between 6-9 years), for semiconductor and wireless firms it was 3 years (varied between 2-6 years for wireless firms and 1-6 years for semiconductor firms). In all observations, years near the peak correlation point also showed high cross correlation. Another observation was that all pharmaceuticals firms, for which positive reasonable lag were found, were the firms with a low or lower-middle level of patenting activity. For firms with high

patenting activity (most often the large firms) these values were zero, and for one firm it was found to be one year. As not all lag values could be determined, external validity of these results is highly questionable.

Values determined for length of lag period for each of these sectors are quite reasonable and as expected, but number of companies for which lag could be found was low. Hence we can merely speculate but not claim that the values determined are a representative for the respective industry sectors. One reason for not being able to determine lag length for all firms could be: for this method to work “stationarity” (no trend) for the time series is assumed and both the time series for Q-ratio and R&D expenses have growing trends, which would make it difficult for this method to work. Cross correlation after differencing was tried but it did not yield better results.

Another plausible reason for this is, if annual increase of R&D expenditures is marginal, as it normally happens, then there is not a strong reason to expect lagged effects on returns, or to be able to isolate the effect of R&D on performance which most often would be distributed over years. High correlations observed for the years near the peak correlation year would support this distributed returns explanation. Though the classical assumption is that firms invest in R&D considering expected rates of return of their investments, it is quite hard to isolate this effect. Other variables may count here, e.g. market conditions,

advertising or business processes. Thus zero or no lag length can be interpreted as distributed or continuous returns of R&D on performance.

- The other research question we explored was whether firms which perform well on stock market or have good R&D outputs (new products and innovations) in turn invest more in R&D than the firms which are not performing so well. In other words, do firms try to get into a virtuous cycle of innovation driven growth leading to more innovation. As we know, better stock performances are the reflection of market's (external to the company) belief in the success and increased revenues for the company due to which market decides to invest more in the company. On the other hand, R&D investments are the internal decisions which occur primarily due to belief of a company in itself or its research capabilities for pursuing the market opportunities, leading to subsequent buy in from the market into this decision and finally towards the growth in revenues of the company. We analyze R&D investments as dependent variable (R&D/Assets is the variable used) within the same year with performance of firm (Tobin's Q) in that year and quality and quantity of patents.

Results for this analysis were highly significant (F-statistic = 48) and show that R&D/Asset ratio is not related to the performance of firms on stock market in any given year. Variable "number of patents granted" in the same year was significant in the continuous regression but not in the categorical regression with categories for patenting activity and company size. Assets, time, industry dummies,

generality and number of claims were found to be highly significant in the both regressions. Coefficient for the time variable used was positive showing that over the years R&D/Asset ratio has increased on an average for all successful firms (Only firms with $Q > 1$ were taken in analysis here). Coefficient for assets was negative, indicating that as firms grow their R&D/assets ratio does not grow as fast, though overall R&D expenses are growing.

In the categorical regression this coefficient for small firms was found positive and negative for the large firms. This is interesting and a reasonable observation, because we consider the R&D/Assets ratio and it should grow faster for firms less than a couple of hundred million dollars worth capital assets than it would for firms with assets in billions, even though the increase in absolute R&D expenses amount may be much more for the larger firms.

Coefficient for industry dummies was found to be negative, which indicates that because pharmaceutical industry products are vital for life and the products from semiconductors and wireless sectors are consumer products, hence R&D has relatively less importance in these sectors than in pharmaceutical sectors. Probably advertising, distribution channels, network size and other aspects of business are relatively more important in those sectors.

- Another observation was that the coefficients for generality and number of claims were both highly significant and positive. This translates into more R&D

investments for firms which are already conducting good R&D. Explanation for generality's significance can be that it denotes the market research and market feedback or reviews, and that is the reason R&D managers or board members pump in more money into further research. For claims, the explanation becomes more speculative. It is a weak indicator of quality and no R&D investment decision maker would look into the patent applications to see how many claims were made in order to decide further investments in that research stream.

Number of domestic references (represented by originality here) is an equally good indicator of quality as number of claims, but that was not found significant. We therefore speculate that number of claims is indirectly related to or acting as a proxy for the performance or assessment of the capabilities of researchers in the firm. If we assume that final claims in a patent application are quite similar to whatever claims about scope and use of technology that were made in the proposals by these researchers while requesting funds, then number of claims in a patent can be expected to behave as a proxy for the performance or capabilities of the inventors. This would justify more investments in the research to expand on the patented technology, because R&D managers can believe the claims inventors are making for their future research. It may also suggest that internal R&D decision making of firms can not be called fully rational if belief in what inventors claim could affect investment decisions. As R&D investment decision making is itself a large and growing research area in decision sciences, this speculation could be worth exploring further.

Finally, in this study we found that in all our regressions adjusted R^2 values were quite low (between 10-25%). This suggests that the relationships observed are quite noisy, thereby reflecting the complexity of stock market valuations which lead to high levels of noise in modeling. However our regressions were highly significant with F-statistic values greater than 10, which suggests that these causal relationships posited are empirically well supported by the dataset used here, and one can rely on these insights.

Chapter – 5

Conclusions, Future Work and Policy Implications

4. 1 Conclusions

The results of this paper show that the immediate market performance, and thus market value of high technology firms with yearly mean of Tobin's Q-ratio greater than one correlates significantly with both quantity and quality of patents granted to the company. Furthermore, the market gives better returns to technology innovation for pharmaceutical firms than for semiconductor or wireless companies. Additionally, we show that strong interaction effects amongst the variables - assets, patent quantity and patent quality exist in determining the market value of a firm. Except for the interaction term between assets and quality of patents, the coefficients for the interactions were positive, which implies super-additivity in effects of intellectual property over market performance because combined effects of patent quality and quantity, and firm size and patent counts are more than the sum of these effects individually on a firm's market performance. This super-additivity phenomenon puts small firms on a very disadvantageous position relative to the large firms.

The other finding was that R&D investments by firms do not depend on the performance of these firms in market. But it does depend on the quality of patents which also influences the market performance of firms. Therefore both market value and the R&D expenses of a firm in any given year get influenced by the quality of patents granted to it in that year. The relationship between market performance and R&D investments is not

apparent in the same year, but lagged and distributed or continuous returns to R&D on market value were observed in a subset of data. Results and the model used in this research do not account for all innovation activity happening in the industries and any licensing activities taking place amongst firms.

4.2 Future work

This study provides us with some good and interesting results, which lead to further research questions to be explored in depth. Here we present some of these questions which are definitely worth further research:

- We use the OLS estimator in this study which involves patent and citations count data. This leads to a bad fit in our results though significance levels of the analyses and variables were quite high. Hence if the research question demands, one can use another regression model which would be more suitable for count data (negative binomial distribution based) to increase the fit.
- Companies do not always take up the option of commercialization for each patent granted to them. If data can be gathered then exploring the same research questions with a real-options based model considering all options over any patent which a company has after getting the patent granted, would definitely provide a much better understanding of R&D investment decision making within the firms and its impacts on the market performance of firms.

- As we noticed, dummy variables for semiconductors and wireless industries were found negative consistently in all our regressions. This could be well worth the research effort to explore and empirically support any explanations for this phenomenon.
- Assuming a linear relationship between firm performance and independent variables used itself is an assumption which could be verified using an “artificial neural networks” based estimator, which could determine the nonlinear relationships.
- We could not determine lag values for all the firms from our dataset but that is a very interesting research question worth much more detailed investigation. Using a larger and more detailed dataset this could be achieved.

Like any other research, this study also leads to many more questions to be answered subsequently in further research. Yet it provides some useful insights into how market performance of firms is related to its intellectual property, size, time and industry characteristics in any given year.

4.3 Policy Recommendations

Along with other challenges of growth and expansion, small firms have to also face the fact that though intellectual property gives them a lot of advantage, but proportionately it gives more advantage to large firms. This provides ample opportunities for the policy makers to attempt making the competition fair or leveling the field by putting in place

policy initiatives or scaffolding to support smaller firms. Furthermore, policy makers can work on ways to decrease different forms of uncertainties associated with the businesses in the economy, as it would be helpful to all firms. Another area for policy initiatives could be to standardize and mandate the information disclosures about finer details of R&D expenses reporting and information about technologies licensed in or out from the company, to make the business activities in the industries more transparent for competitors and investors.

References

- Bloom, N. and Van Reenen, J. (2002) "Patents, Real Options and Firm Performance", *The Economic Journal* 112 (478), C97–C116.
- Cameron, C. and Trivedi, P. K. (1998), "*Regression Analysis of Count Data*", Econometric Society Monograph No.30, Cambridge University Press, New York NY.
- Cook, T. D. and Campbell, D. T., "*Quasi-Experimentation: Design and Analysis Issue for Field Settings*", Houghton Mifflin Company, Boston MA.
- Deng, Z., Lev, B. and Narin, F. (1999), "Science & Technology as Predictors of Stock Performance", *Financial Analysts Journal*, 55, 3, 20-32.
- Griliches, Z. (1981), "Market Value, R&D and Patents." *Economic Letters* 7: 183-87
- Griliches, Z. (1990), "Patent Statistics as Economic Indicators: A Survey", *Journal of Economic Literature*, XXVIII (Dec.): 1661-1707.
- Hall, B. H., Jaffe A. B. and Trajtenberg, M., (2000). "Market Value and Patent Citations: A First Look," *NBER Working Papers* 7741
- Hall, B. H., Jaffe, A. B. and Trajtenberg, M. (2001). "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools." *NBER Working Paper* 8498.
- Hall, B. H., (2006) "R&D, Productivity, and Market Value", *Annales, Revised version of a paper presented at the International Conference in memory of Zvi Griliches*, Paris, August 2003.
- Harhoff, D., Narin, F., Scherer, F. M. and Vopel, K. (1999), "Citation Frequency and the Value of Patented Inventions", *Review of Economics and Statistics* 81(3): 511-15
- Hausman, J., Hall, B. and Griliches, Z. (1984), "Econometric models for count data with an application to the patents-R&D relationship." *Econometrica*, 52: 909-938.
- Jaffe, A. and Trajtenberg, M. (1998), "International Knowledge Flows: Evidence from Patent Citations", *NBER Working Paper* 6507.
- Jaffe, A., Trajtenberg, M., and Fogarty, M. (2000): "The Meaning of Patent Citations: Report on the NBER/Case-Western Reserve Survey of Patentees", *NBER Working Paper* 7631.
- Lanjouw, J.O., Schankerman, M., (1997), "Stylized facts of patent litigation: value, scope, ownership", *NBER Working Paper No.* 6297.

- Lanjouw, J. O. and Schankerman, M. (2004), "Patent Quality and Research Productivity: Measuring Innovation with Multiple Indicators.", *Economic Journal* 114(495): 441-465.
- Narin, F., Thomas, P. and Breitzman, A. (2004), "Using Patent Citation Indicators to Manage a Stock Portfolio," in Moed H. et al. (ed.) *Handbook of Quantitative Science and Technology Research*, Kluwer Academic Publishers, Dordrecht
- Pakes, A. and Schankerman, M. (1984), "The Rate of Obsolescence of Knowledge, Research Gestation Lags, and the Private Rate of Return to Research Resources," pp. 73-88, in Z. Griliches (Ed.), *Patents, R&D, and Productivity*, The University of Chicago Press, Chicago IL.
- Pitkethly., R. (1997), "The Valuation of Patents: A Review of Patent Valuation Methods With Consideration of Option Based Methods and the Potential for Further Research", *Judge Institute Working Paper* WP 21/97, University of Oxford.
- Porter, A.L. and Cunningham, S. W. (2005). *Tech Mining: Exploiting New Technologies for Competitive Advantage*, Wiley, New York NY.
- Putnam, J. (1996), "The value of international patent rights." PhD Thesis, Yale University.
- Reitzig M, (2004), "Improving patent valuations for management purposes: validating new indicators by analyzing application rationales", *Research Policy* 2004 Sep Vol 33:6-7 p 939-957.
- Schankerman, M., Pakes, A., (1986). "Estimates of the value of patent rights in European countries during the post-1950 period", *Economic Journal* 97, 1-25.
- Thomas, P. (1999), "The Effect of Technological Impact upon Patent Renewal Decisions", *Technology Analysis & Strategic Management*, Volume 11(2), 181-197.
- Tong, X. and Frame, J. D. (1994), "Measuring national technological performance with patent claims data", *Research Policy*, 23(2), 133-141.
- Trajtenberg, M, (1990) "A Penny for Your Quotes: Patent Citations and the Value of Innovations", *The Rand Journal of Economics*, 21(1), 172-187.

Appendix – I: Description of Data Statistics

Following are the plots describing the characteristics of the data used in this study in terms of different categorizations and statistics:

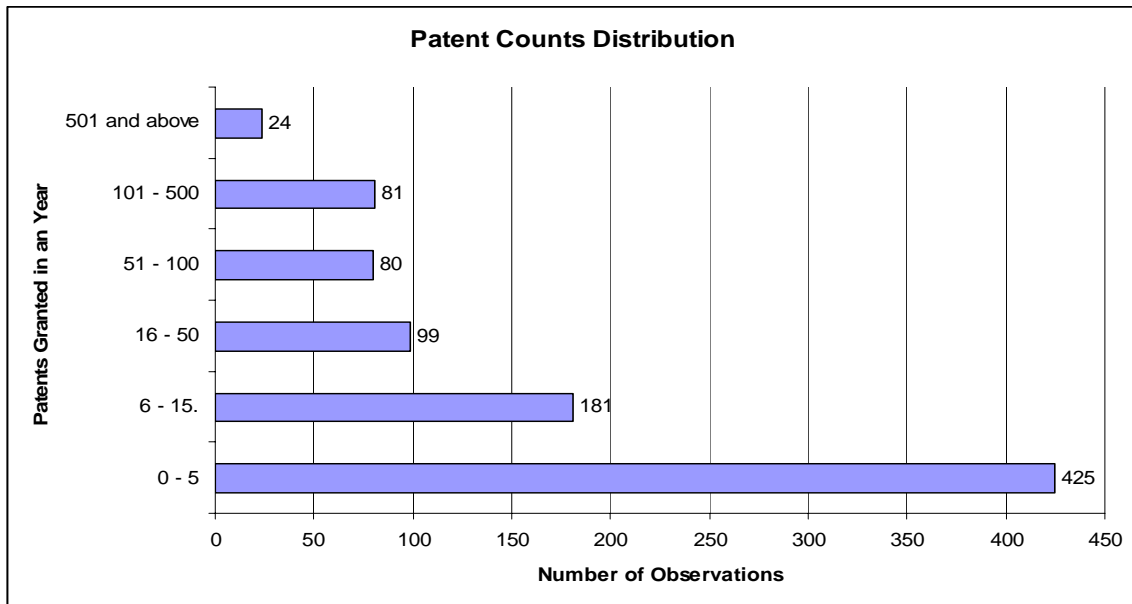


Figure A1.1: Patent Counts Distribution of Firm Data Observations

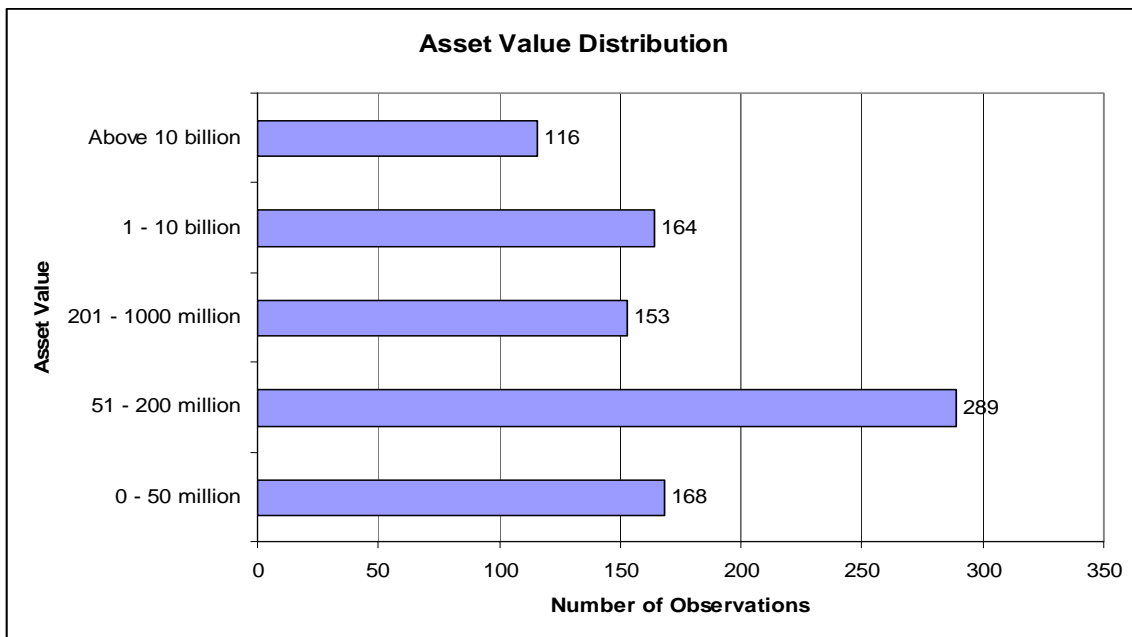


Figure A1.2: Asset Value Distribution of Firm Data Observations

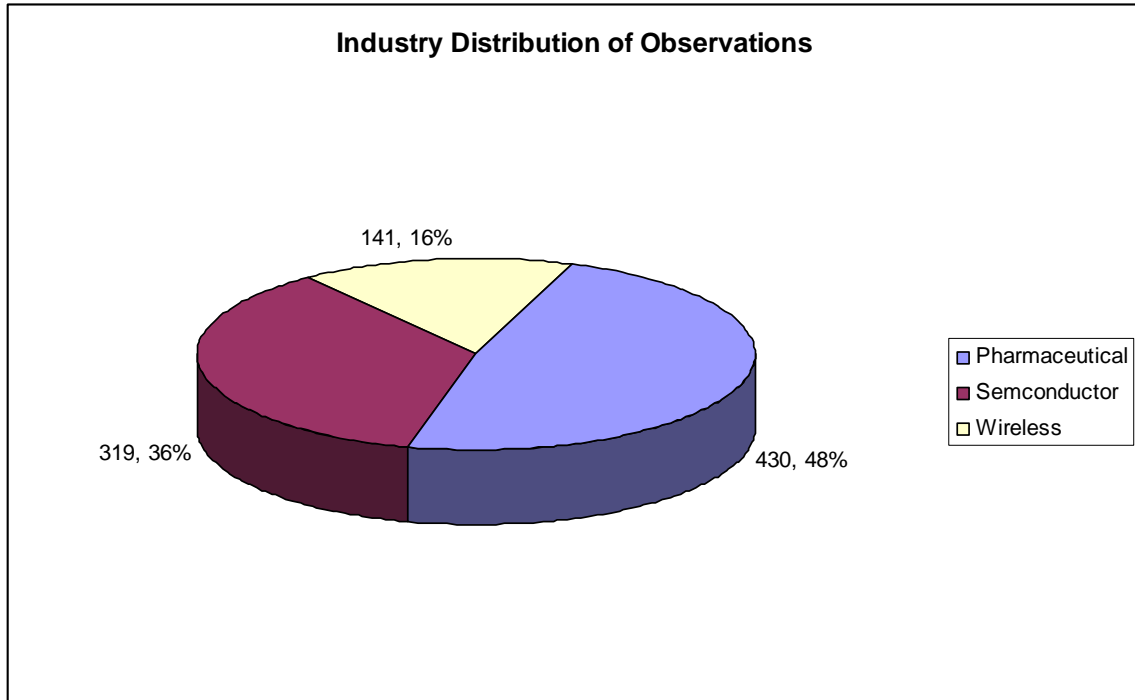


Figure A1.3: Industry-wise Distribution of Firm Data Observations

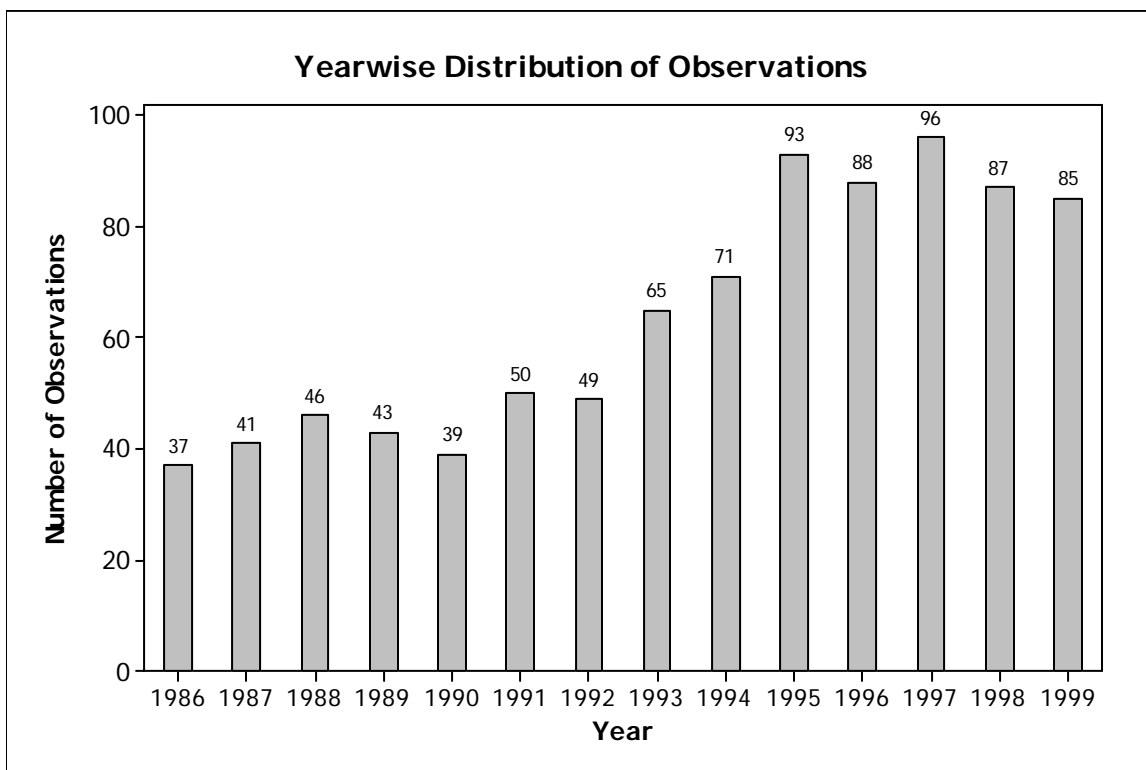


Figure A1.4: Year-wise Distribution of Firm Data Observations

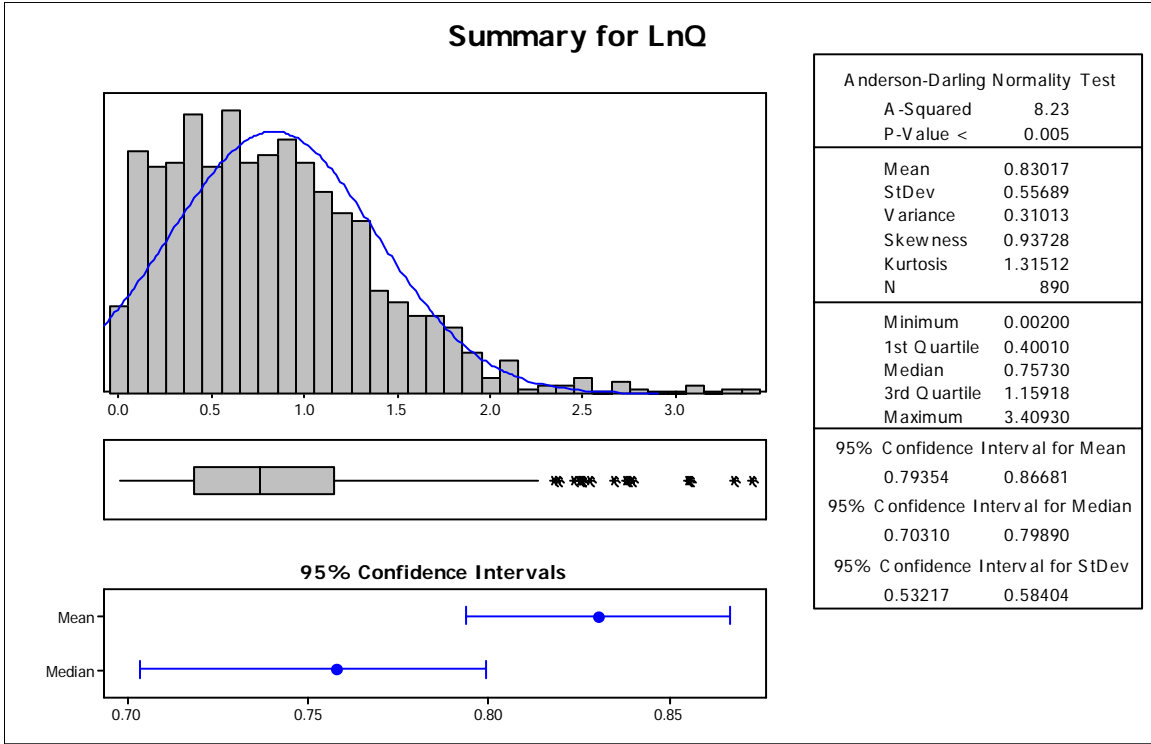


Figure A1.5: Statistical Summary for Natural Log of Tobin's Q of Firms in a Year

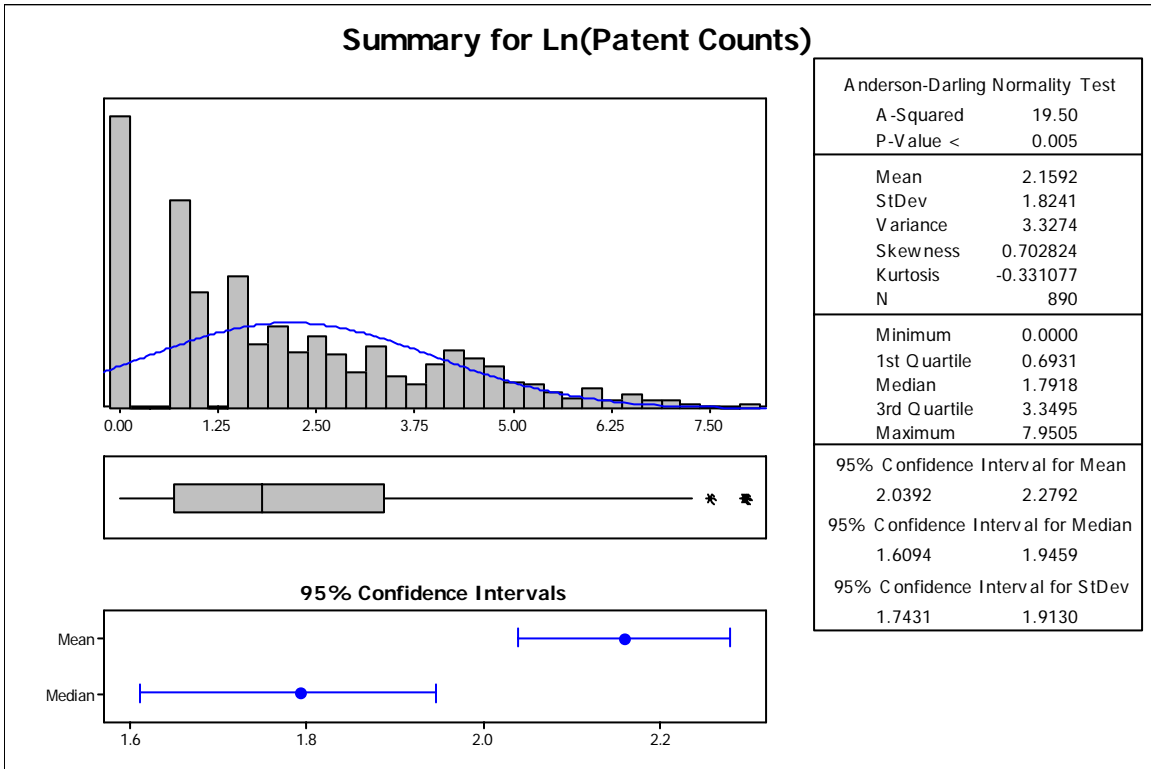


Figure A1.6: Statistical Summary for Natural Log of Patent Counts of Firms in a Year

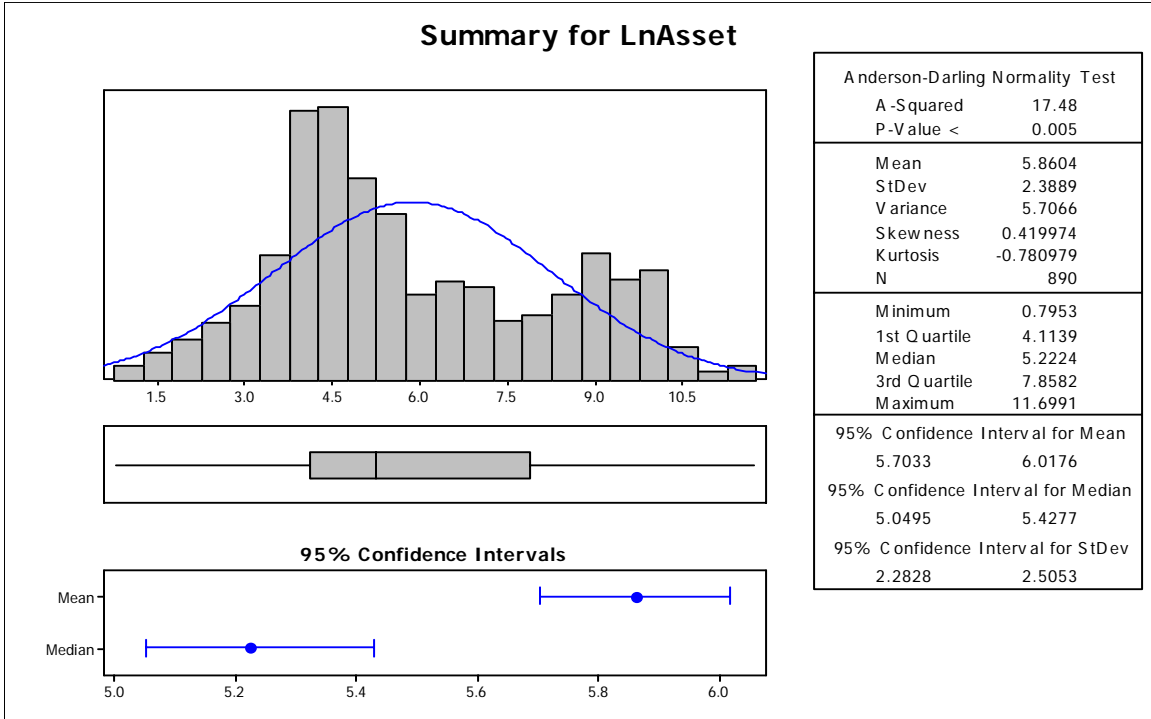


Figure A1.7: Statistical Summary for Natural Log of Assets of Firms in a Year

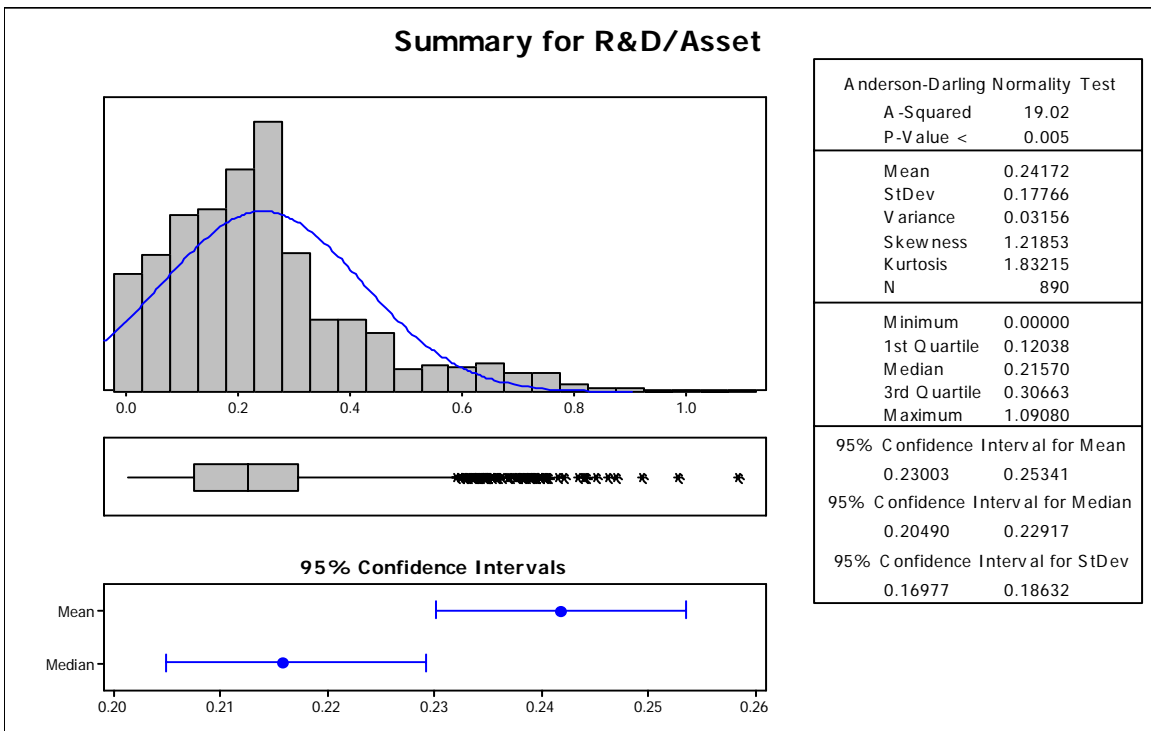


Figure A1.8: Statistical Summary for Ratio of R&D Expenses and Assets of Firms in a Year

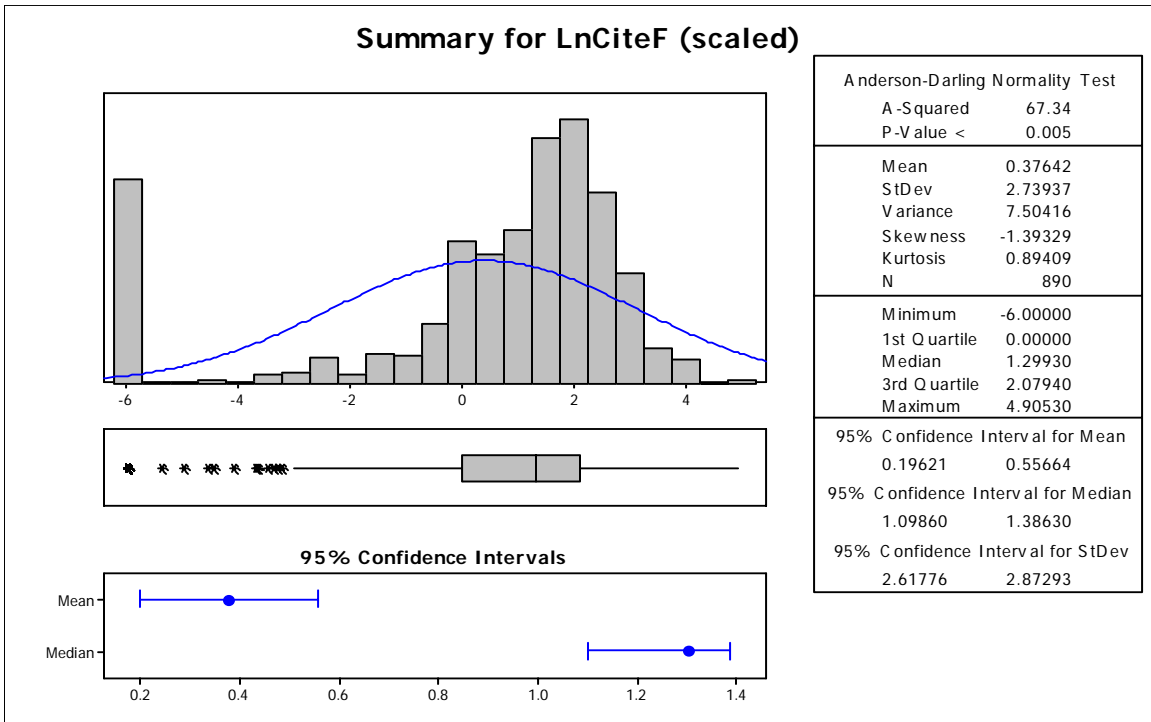


Figure A1.9: Statistical Summary for Natural Log of Average Number of Citations Received by the Firm's Patents of that Year

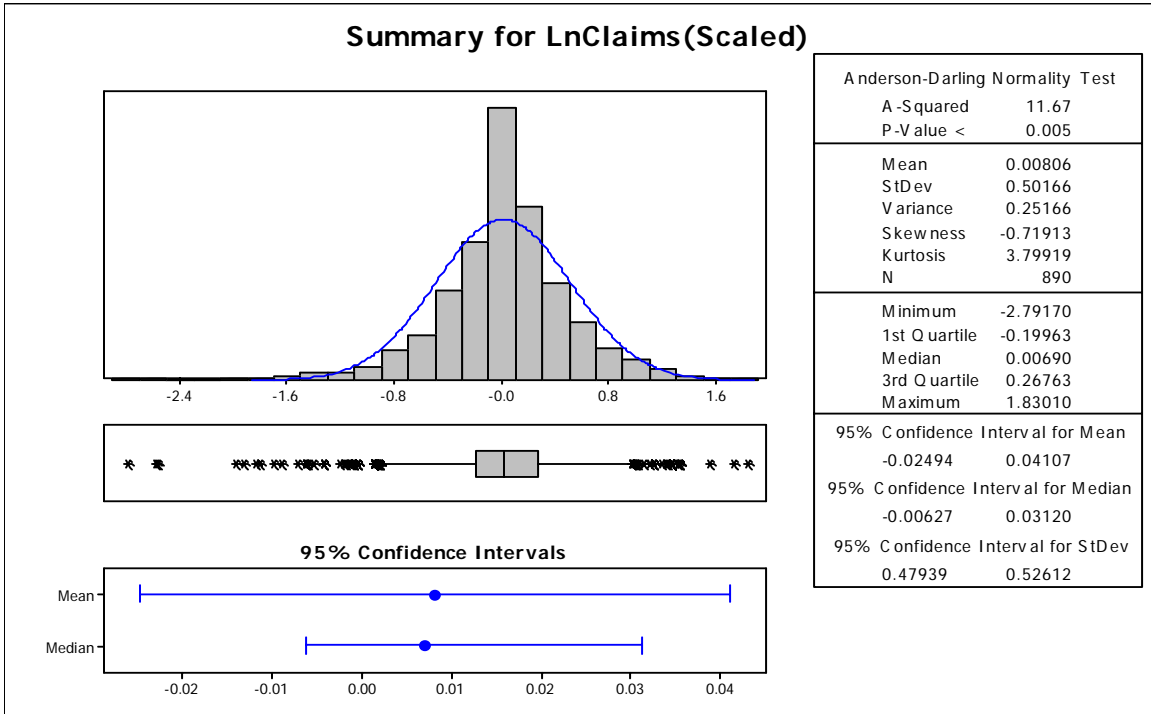


Figure A1.10: Statistical Summary for Natural Log of Average Number of Claims Made by the Firm's Patents of that Year

Descriptive Statistics

Table A1.1: Descriptive Statistics

Variable	Mean	SE Mean	TrMean	StDev	Min	Q1	Median	Q3	Max
Patent Counts	59.0800	6.7200	25.7700	200.3500	1.0000	2.0000	6.0000	28.5000	2837.0000
LnPatents	2.1592	0.0611	2.0443	1.8241	0.0000	0.6931	1.7918	3.3495	7.9505
Assets	4301.0000	389.0000	2416.0000	11605.0000	2.2200	61.2000	185.0000	2587.0000	120463.0000
LnAssets	5.8604	0.0801	5.8259	2.3889	0.7953	4.1139	5.2224	7.8582	11.6991
Q Ratio	2.7684	0.0806	2.4434	2.4045	1.0020	1.4920	2.1325	3.1873	30.2440
LnQ	0.8302	0.0187	0.7953	0.5569	0.0020	0.4001	0.7573	1.1592	3.4093
R&D Expenses	854.4000	73.4000	469.0000	2190.7000	0.0000	10.2000	46.5000	284.1000	20123.0000
R&D/Asset	0.2417	0.0060	0.2278	0.1777	0.0000	0.1204	0.2157	0.3066	1.0908
Sales	2604.0000	249.0000	1381.0000	7430.0000	0.0000	16.0000	102.0000	1233.0000	87548.0000
CiteF	6.6770	0.3550	5.1680	10.5820	0.0000	1.0000	3.6670	8.0000	135.0000
LnCiteF	0.3764	0.0918	0.5545	2.7394	-6.0000	0.0000	1.2993	2.0794	4.9053
Cite	9.6910	0.3710	8.2070	11.0780	0.0000	4.6970	7.1230	10.9540	137.4000
LnCite	1.9673	0.0254	1.9674	0.7561	-1.3863	1.5751	1.9652	2.3956	4.9229
Claims	1.1348	0.0202	1.0770	0.6013	0.0613	0.8190	1.0070	1.3069	6.2344
LnClaims	0.0081	0.0168	0.0204	0.5017	-2.7917	-0.1996	0.0069	0.2676	1.8301
Generality	0.3334	0.0090	0.3202	0.2690	0.0000	0.0000	0.3362	0.5123	0.9000
Originality	0.3825	0.0064	0.3825	0.1897	0.0000	0.2771	0.3836	0.5000	0.8711

Appendix – II: Results

Following are the representative set of results from this study which have not been presented in the report in this form:

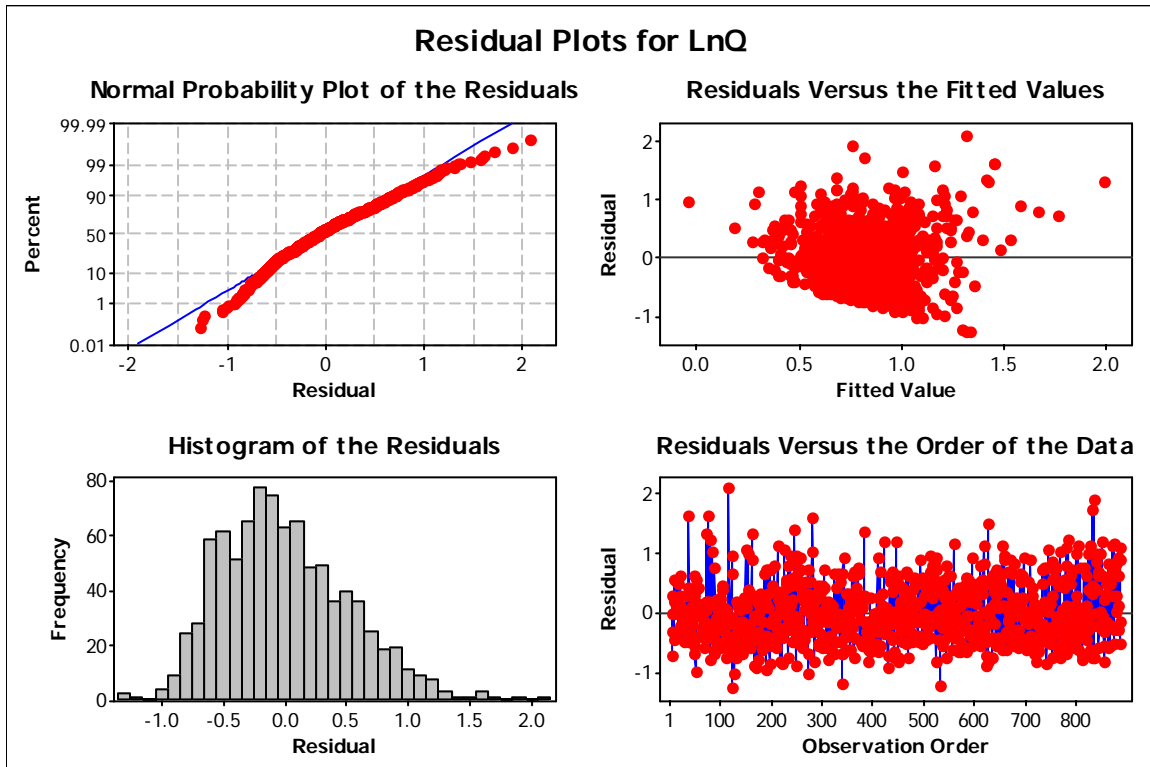


Figure A2.1: Residual Plots for LnQ as Dependent Variable

Results for Regression Analysis with R&D/Asset as a Dependent Variable

The regression equation determined is:

$$\begin{aligned} \text{RDbyAsset} = & 0.160 - 0.00666 \text{ LnQ} + 0.00857 \text{ LnPatents} - 0.0244 \text{ LnAsset} \\ & + 0.0268 \text{ Inclaims-scaled} + 0.00226 \text{ LnCiteF} - 0.0888 \text{ sdummy} \\ & - 0.125 \text{ wdummy} + 0.130 \text{ LnT} \end{aligned}$$

Standard Error = 0.149082 R-Sq = 30.2% R-Sq(adj) = 29.6%

Table A2.1: Analysis of Variance

Source	DF	SS	MS	F-statistic	p-value
Regression	8	8.4785	1.0598	47.68	0
Residual Error	881	19.5806	0.0222		
Total	889	28.0591			

Table A2.2: Coefficient Values for Independent Variables

Predictor	Coef	SE Coef	T	P	VIF
Constant	0.15995	0.025	6.4	0	
LnQ	-0.006659	0.009478	-0.7	0.483	1.1
LnPatents	0.008568	0.004675	1.83	0.067	2.9
LnAsset	-0.024366	0.003609	-6.75	0	3
Inclaims-scaled	0.02681	0.01036	2.59	0.01	1.1
LnCiteF	0.002258	0.002184	1.03	0.031	1.4
Sdummy	-0.0888	0.01181	-7.52	0	1.3
Wdummy	-0.12502	0.01502	-8.32	0	1.2
LnT	0.130449	0.008862	14.72	0	1.4

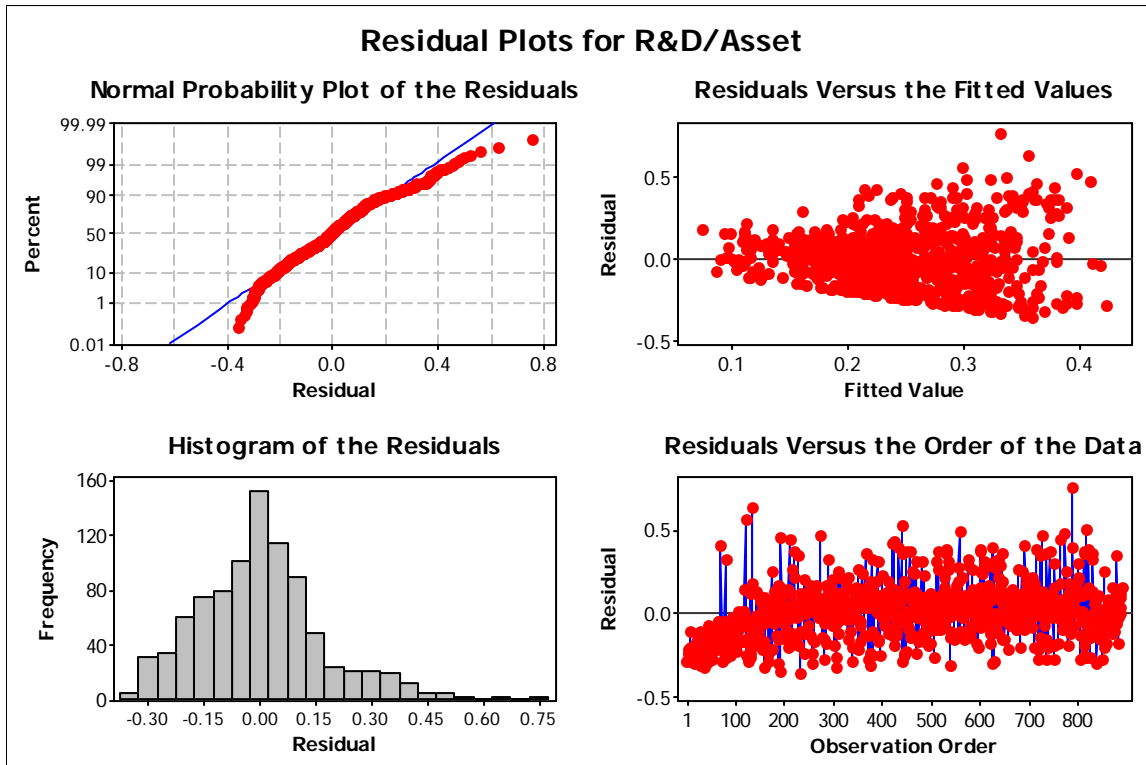


Figure A2.2: Residual Plots for R&D/Assets as Dependent Variable

Table A2.3: Values for Length of Lag between R&D Expenses and Stock Returns

Industry	Stock Symbol	Highest Correlation Year and Value	Second Highest Correlation Year and Value	Total Patents in any Year	Lag Found
Pharma					
	XOMA	7 0.531		Low	Yes
	WYE	7 0.246		Mid	Yes
	RGEN	6 0.416	7 0.401	Low	Yes
	PFE	0 0.902	1 0.742	High	
	NVO	0 0.852	1 0.733	High	
	MRK	0 0.677	1 0.499	High	
	LLY	0 0.640	1 0.530	High	
	JNJ	0 0.861	1 0.747	High	
	IMMU	7 0.307		Low	Yes
	GSK	2 0.628	1 0.591	Mid	
	CYTO	0 0.246	1 0.213	Low	
	BOL	8 0.419	9 0.411	Mid	Yes
	BMY	0 0.620	1 0.567	High	
	AMGN	0 0.281	1 0.200	Mid	
Semiconductor					
	ADI	2 0.569	3 0.529	Mid	Yes
	AMD	3 0.337	2 0.327	High	Yes
	CY	4 0.238	5 0.194	Mid	Yes
	ENER	4 0.553	5 0.491	Mid	Yes
	IBM	6 0.622	7 0.583	High	
	IDTI	3 0.299	4 0.283	Mid	Yes
	INTC	0 0.935	1 0.721	High	
	LLTC	0 0.836	1 0.686	Low	
	LSCC	1 0.280	2 0.274	Mid	Yes
	LSI	0 0.512	1 0.478	High	
	MSCC	5 0.668	4 0.578	Low	Yes
	MU	0 0.287	1 0.272	High	
	NSM	2 0.562	1 0.531	High	Yes
	SMSC	10 0.221	11 0.205	Low	
	TXN	0 0.687	1 0.602	High	
	XLNX	0 0.562	1 0.388	Mid	
Wireless					
	KYO	5 0.279	6 0.257	Mid	Yes
	MOT	2 0.365	1 0.302	High	Yes
	QCOM	3 0.324	4 0.256	Mid	Yes

Though these results are as expected, but they can not be taken as valid for generalization to the dataset used in this study and obviously for any other similar dataset.